

Discrete Choice with Endogenous Peer Selection*

Nail Kashaev

Natalia Lazzati

Western University

UC Santa Cruz

nkashaev@uwo.ca

nlazzati@ucsc.edu

June 2, 2026

Abstract We develop a continuous time discrete choice model of peer effects. The distinctive feature of the model is that agents might not consider all peers at the moment of making a decision. Instead, they select *some* of them on the basis of a mechanism that depends on recent choices. We characterize the equilibrium behavior and study the empirical content of the limited attention peer effect model. We allow changes in the choices of peers to affect both the set of peers to which the agent pays attention and her preferences over the alternatives. We exploit variation in choices together with variation in the size of the set of potential peers (or reference groups) to recover the preferences of the agents and the peer selection mechanisms. We apply our results to model expansion and contraction decisions by fast-food restaurants and find evidence of limited attention to actions of competitors.

JEL codes: C31, C33, D83, O33

Keywords: Peer Effects, Random Network, Continuous Time Markov Process, Bounded Rationality

*For their useful comments we thank Victor Aguiar, Roy Allen, Tim Conley, and David Rivers. We would also like to thank the audience at Network Science and Economics conference (Miami Herber BS) and Triangle Econometrics Conference (UNC), and the seminar presentations at Duke University. Kashaev gratefully acknowledges financial support from Social Sciences and Humanities Research Council Development and Insight Development Grants.

1. Introduction

Social interactions are a cornerstone of many economic and social outcomes (Durlauf and Young, 2001). Consumers purchase products after seeing their friends buy them, and firms open new stores in response to the actions of their competitors. A large empirical literature measures these peer effects in discrete choice settings under the assumption that agents observe or pay attention to all their peers.¹ This means that, for instance, consumers take into account the purchasing decisions of all their friends and firms consider the actions of all their competitors at the moment of making a decision. However, in many settings, different constraints or limitations might lead agents to pay attention to the choices of a smaller subset of peers, e.g., agents could be choosing actions using a subjective model of peer effects that may be misspecified. This paper offers a model of peer selection in which the peer selection mechanism is allowed to depend on recent choices.

Similarly to Kashaev and Lazzati (2019) and Kashaev, Lazzati and Xiao (2026), we study a continuous-time discrete choice model with social interactions. The network specifies the set of potential peers (referred to as reference group) of each agent. These connections are known by the agents, but not by the researcher. At random times governed by independent Poisson alarm clocks, an agent wakes up and selects an alternative from a finite menu. When the clock rings, the agent first draws a subset of peers, the *active peer set* or *active peers*, from her reference group. The probability of selecting any given peer depends on the current choices of the agent and her peers, allowing agents to pay more attention to peers that have selected similar alternatives. The agent evaluates the alternatives under the influence of the active peers and then selects an option. The resulting profile of choices evolves according to a continuous-time Markov process. We show that (under some restrictions) if the researcher observes the choices of all agents in the population over a long period of time, then she can recover the social network (i.e., the reference group for each agent), the peer selection mechanism, and the random preferences of each agent captured by the distribution over the choice set conditional on realization of the choices of the active peers.

The modeling restrictions ensure the existence of a unique equilibrium. This equilibrium is

¹This paper uses a behavioral definition of peers. We think the peers of an agent are all other agents that can directly affect her choices.

characterized by an invariant distribution over the set of choice configurations, which describes the long-run frequencies of each possible choice profile. For the identification of the model, the important feature is that the invariant distribution has full support. Since the probability of selecting a peer is affected by previous choices of connected agents, which are determined by the equilibrium distribution, the realized distribution over random active peer sets is endogenous.

We illustrate key features of the model through a simple example with a closed-form solution. The example shows that allowing agents to include peers who choose similar alternatives more often in the active peer set increases the correlation of choices among agents with similar preferences. In specific applications, this result can be interpreted as polarization and our model can be seen as offering a micro-foundation for echo chambers (Hu, Li and Tan, 2025).

Our contribution is threefold: First, we connect bounded rationality with peer effects by embedding endogenous peer selection into a discrete choice model of peer effects in preferences. In a simple two-agent example, we show how endogenous attention can change the comparative statics: paying attention to peers that make similar choices can either reinforce or dampen coordination relative to a standard model depending on whether peer effects are positive or negative. Second, we offer a simple setting with transparent identification arguments that can be incorporated (as we explain below) into a larger model of either correctly specified or misspecified network structure. Similarly to the use of menu variation in consideration set models, we exploit variation in the sizes of the reference groups to identify the network, the peer-selection mechanism, and preference parameters from a single long panel of choices while imposing only mild exclusion and dimension-reducing conditions. Third, we show how the limited attention to firms operating in the same market can shape expansion and contraction decisions in our empirical illustration.

The starting point of the identification strategy is the set of Conditional Choice Probabilities (CCPs). Each CCP specifies the frequency of choices made by an agent given the choice configuration at the time of making the decision. We assume the analyst can consistently estimate these CCPs from the data. We first show that, under a set of mild restrictions, variation in the choices of potential peers generates variation in the frequency of choices of a given agent. Thus, the set of potential peers, or reference groups, can be recovered from the data. We then implement a thought

experiment: We show that if the researcher could manipulate the reference groups of the agents by creating or deleting links, then she could recover the rest of the model via a recursive strategy. However, in the applications that motivate this paper, the researcher has no control over reference groups and cannot arbitrarily alter them. We therefore replicate the same logic using two models in which the necessary variation arises endogenously within the model. The first model assumes there is variation in the sizes of reference groups across similar agents. If agents were people, this is the same as saying that some people have more friends than others. The second model, instead, assumes no peer effects in the outside option. This modeling restriction has been motivated by many empirical applications, where the outside option means *do nothing*.

We recover the peer-selection mechanism and the choice rule of each agent nonparametrically as a function of the choice configurations which, in our dynamic setting, correspond to the states of the system. These two parts of the model are left quite general and can be easily microfounded with various existing results in the literature. For instance, we could generate the peer-selection mechanism as the product of random meetings and the probability of including the met agents in the active peer set based on relative payoffs—that would capture the advantage of doing so. We could also view the choice rule as the reduced-form value from selecting an alternative. This reduced-form could capture forward-looking behavior without explicitly modeling it. In this case, the choice rule would implicitly reflect the immediate payoff of selecting the option and the discounted value of payoffs that occur in the continuation state. In this context, agents might display correct beliefs about the social environment or a rather misspecified model.

We illustrate our model by studying the expansion and contraction decisions of five major fast-food restaurant chains operating in geographically distinct markets in China. In this setting, each firm-market pair is an agent and firms respond to the actions of their competitors when deciding whether to open or close restaurants. We allow managers to ignore the presence of some competitors. This can happen for different reasons. For example, the manager may not think that a particular restaurant is a close substitute at a specific moment in time. In addition, monitoring competitors is costly in terms of time, data collection, and analysis, and managers may be unable to process all relevant market information continuously. This idea builds on previous theories of

bounded rationality in firms offered by [Simon \(1955\)](#). Consistent with this view, we find substantial heterogeneity in attention to competitors across firms and markets. Limited attention leads firms to overexpansion, and for some firms it primarily induces distortions in the probabilities of opening stores in different cities, while for others it fundamentally changes the way they respond to the actions of competitors.

We finally connect our work to the existing literature. First, the ideas are related to the consideration set models where agents do not pay attention to all available alternatives (see, for instance, [Aguiar, Boccardi, Kashaev and Kim, 2023](#), and references therein). While this literature focuses on *consideration sets of alternatives*, we study *consideration sets of peers* in an interactive framework. In doing so, our work also contributes to the growing econometric literature on network formation (see [Graham, 2015](#), [De Paula, 2017](#), [Chandrasekhar, 2016](#), for an overview). Within this literature, the studies by [Leung \(2015\)](#), [Menzel \(2015\)](#), [Miyauchi \(2016\)](#), [Boucher and Mourifie \(2017\)](#), [Mele \(2017\)](#), [de Paula, Richards-Shubik and Tamer \(2018\)](#), [Thirkettle \(2019\)](#), [Ridder and Sheng \(2020\)](#), [Sheng \(2020\)](#), [Badev \(2021\)](#) and [Gualdani \(2021\)](#) have analyzed game-theoretic models of creating links. Another related literature models social decision-making as based on observations from a sampled subset of network neighbors (e.g. [De Groot, 1974](#), [Bala and Goyal, 1998](#), [Golub and Jackson, 2010](#)) or from a small set of peers with its size constrained by the cost of processing the information ([Park, Sarangi and Tzavellas, 2016](#)). We complement these approaches by offering a general peer-selection mechanism in which the selection probabilities for effective network formation depend on recent choices. This distinction matters for identification, as endogenous active peer set formation can generate correlated behaviors.

There is a large econometric literature on the identification of models of social interactions where the choices of peers shape agents' preferences over alternatives (see, for example, [Blume, Brock, Durlauf and Ioannides, 2011](#), [Bramoullé, Djebbari and Fortin, 2020](#), [De Paula, 2017](#) and [Graham, 2015](#), for comprehensive reviews of this literature). On the theoretical side, [Chambers, Cuhadaroglu and Masatlioglu \(2025\)](#) recover the impacts that agents have on each other using variation in group composition. Although existing social interaction models typically assume correct beliefs about peer behavior, a growing literature documents biased social learning and misperceived

peer behavior, e.g., formation of peer-induced beliefs (Boucher, Dedewanou and Dufays, 2022) and misclassification in social interactions (Lin and Hu, 2024, Meyer and Mittag, 2017, Lewel, Qu and Tang, 2025). We can think of our model as a discrete choice framework with endogenous peer selection that incorporates misspecified beliefs about peers.²

Finally, our empirical application relates to the literature on misoptimizing (e.g., boundedly rational) firm behavior in the behavioral industrial organization (Simon, 1945, Armstrong and Huck, 2010, Heidhues and Kőszegi, 2018). Most of this literature introduces deviations from rationality to the demand side of the economy (i.e. behavioral consumers). In contrast, we study a setting in which firms themselves may misoptimize (Bloom and Van Reenen, 2007, DellaVigna and Gentzkow, 2019, Goldfarb and Xiao, 2011, Hortacsu and Puller, 2008, Cho and Rust, 2010, Tourek, 2022, Almunia, Hjort, Knebelmann and Tian, 2024, Dube, Manning and Naidu, 2025, Galdon-Sanchez, Gil and Uriz-Uharte, 2025, Song and Stern, 2025, Hjort, Li and Sarsons, 2026, Kashaev et al., 2026).

The remainder of the paper is organized as follows. Section 2 presents the model and its assumptions. Section 3 elaborates on equilibrium behavior. We establish identification of the model parameters in Section 4. Section 5 presents our empirical application and we conclude in Section 6. All proofs can be found in Appendix A.

2. Model

This section describes the model and the main assumptions used in the paper.

2.1. Network, Peer Selection, and Preferences

Network There is a finite set of agents $\mathcal{A} = \{1, 2, \dots, A\}$, $A \geq 2$, choosing from a finite set of alternatives $\mathcal{Y} = \{0, 1, \dots, Y\}$, $Y \geq 1$. In some contexts, 0 will indicate the default alternative.

²See Esponda and Pouzo (2016) for an equilibrium set-up of misspecified models.

We refer to $\mathbf{y} = (y_a)_{a \in \mathcal{A}} \in \mathcal{Y}^{\mathcal{A}}$ as a choice configuration. Agents are connected through a social network. The network is described by a set of edges. Each edge identifies two connected agents and the direction of the connection. For each Agent $a \in \mathcal{A}$, her reference group is defined as follows:

$$\mathcal{N}_a = \{a' \in \mathcal{A} : a' \neq a \text{ and there is an edge from } a \text{ to } a'\}.$$

Example 1. Suppose that there are four agents and three alternatives. That is, $\mathcal{A} = \{1, 2, 3, 4\}$ and $\mathcal{Y} = \{0, 1, 2\}$. The reference groups are

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

This means that, for instance, Agents 2 and 3 may affect the choices of Agent 1. □

Peer Selection and Preferences The revision of choices follows a standard continuous-time Markov process. Agents have independent Poisson alarm clocks with rates $\lambda = (\lambda_a)_{a \in \mathcal{A}}$. The alarm of Agent a is triggered at exponentially distributed moments with mean $1/\lambda_a$. When this happens, the agent first selects a subset of peers $\mathcal{N} \subseteq \mathcal{N}_a$ to whom she will pay attention (i.e., the active peers) and then makes a choice under their influence (i.e., only the active peers affect her current choice). We do not model the cognitive process that leads to peer selection. We take it as given and identify it nonparametrically. Many existing models can be used to motivate the peer-selection mechanism we invoke (see Section 2.3.a for details). (All identification results in Section 4.1 are valid if agents make simultaneous decisions, i.e., they have the same clock.)

For each agent in the population, the described setting introduces a two-stage decision process that depends on the current configuration of choices (\mathbf{y}):

Step 1 Agent a picks a set of active peers $\mathcal{N} \subseteq \mathcal{N}_a$. The ex-ante probability that set \mathcal{N} is picked is $S^a(\mathcal{N} \mid \mathbf{y}, \mathcal{N}_a)$ with $S^a(\mathcal{N} \mid \mathbf{y}, \mathcal{N}_a) \geq 0$ and $\sum_{\mathcal{N} \subseteq \mathcal{N}_a} S^a(\mathcal{N} \mid \mathbf{y}, \mathcal{N}_a) = 1$.

Step 2 After selecting peers $\mathcal{N} \subseteq \mathcal{N}_a$, the agent picks an alternative according to a choice rule $R^a(\cdot \mid \mathbf{y}, \mathcal{N})$ that satisfies $R^a(v \mid \mathbf{y}, \mathcal{N}) > 0$ for all v and $\sum_{v \in \mathcal{Y}} R^a(v \mid \mathbf{y}, \mathcal{N}) = 1$.

The choice rule summarizes the decision process after the peer selection is completed. We

associate choice rules with preferences, since, in many instances, the preference parameters can be recovered from the choice rules. If agents are forward-looking, the choice rule will also contain information about beliefs on future actions of other agents and discounted values of future payoffs (see Section 2.3.b for further details).

Example 1 (continued). Suppose that, at the moment of making a choice, Agent 1 only selects Agent 2 from her reference group $\mathcal{N}_a = \{2, 3\}$. That is $\mathcal{N} = \{2\}$. Hence, her probability of selecting alternative 1 is $R^1(1 \mid \mathbf{y}, \{2\})$. Assume Agent a 's indirect utility from v given the active peer set \mathcal{N} is $u_{a,v}(\mathbf{y}, \mathcal{N}) + \xi_{a,v,\mathcal{N}}$, where $u_{a,v}$ captures the mean utility from the alternative. The vector of agent- and peer-group-specific taste shocks $(\xi_{a,v,\mathcal{N}})_{v \in \mathcal{Y}}$ is continuously distributed with conditional cumulative distribution function (c.d.f.) $F_{a,\xi}(\cdot \mid \mathbf{y}, \mathcal{N})$. Then, for $v \in \mathcal{Y}$

$$R^1(v \mid \mathbf{y}, \{2\}) = \int \mathbf{1} \left(v = \arg \max_{v' \in \mathcal{Y}} \{u_{a,v'}(\mathbf{y}, \{2\}) + \xi_{a,v',\{2\}}\} \right) dF_{a,\xi}((\xi_{a,v,\{2\}})_{v \in \mathcal{Y}} \mid \mathbf{y}, \{2\}).$$

If the $\xi_{a,v,\mathcal{N}}$ s are independent and identically distributed (i.i.d.) shocks, distributed according to the standard Type I extreme value distribution, then the above expression simplifies to

$$R^1(v \mid \mathbf{y}, \{2\}) = \frac{\exp(u_{1,v}(\mathbf{y}, \{2\}))}{\sum_{v' \in \mathcal{Y}} \exp(u_{1,v'}(\mathbf{y}, \{2\}))}. \quad \square$$

As the example demonstrates, our framework allows dependence between preferences and selection of peers. That is, the peer effect in preferences can appear both via mean utility (e.g., direct effect from playing the same video game with your friends) and latent preference shocks (e.g., learning the latent quality of the product via choices of peers).

At the moment of making a choice, the ex-ante probability for Agent a of selecting each alternative is a finite mixture given by

$$P_a(v \mid \mathbf{y}) = \sum_{\mathcal{N} \subseteq \mathcal{N}_a} R^a(v \mid \mathbf{y}, \mathcal{N}) S^a(\mathcal{N} \mid \mathbf{y}, \mathcal{N}_a).$$

We refer to $\mathbf{P} = (P_a)_{a \in \mathcal{A}}$ as the set of Conditional Choice Probabilities (CCPs).

2.2. Main Assumptions

This section adds more structure to the peer selection process and preferences. Regarding the selection process, we assume that each peer is selected independently from the others.

Assumption 1 (Independent Selection). *The probability of considering a subset of peers $\mathcal{N} \subseteq \mathcal{N}_a$ is given by*

$$S^a(\mathcal{N} | \mathbf{y}, \mathcal{N}_a) = \prod_{a' \in \mathcal{N}} Q^a(a' | \mathbf{y}) \prod_{a' \in \mathcal{N}_a \setminus \mathcal{N}} (1 - Q^a(a' | \mathbf{y})),$$

where $Q^a(a' | \mathbf{y})$ is the probability that Agent a' is selected by Agent a given \mathbf{y} such that $Q^a(a' | \mathbf{y}) = 0$ for $a' \notin \mathcal{N}_a$ and $0 < Q^a(a' | \mathbf{y}) \leq 1$ for $a' \in \mathcal{N}_a$.

Assumption 1 reduces the dimensionality of the identification problem. Without it, for each \mathbf{y} , the researcher would need to learn S^a over $2^{|\mathcal{N}_a|}$ points. Assumption 1 reduces this number to $|\mathcal{N}_a|$. This assumption may hold in some settings, but not in others. It is violated if, for example, adding Agent a' to the active peer set affects the likelihood that Agent a'' also be added. Assumption 1 also requires that every peer in each reference group be selected with positive probability.

We also impose restrictions on R^a to make the problem more tractable. Let $N_a^v(\mathbf{y}, \mathcal{N})$ be the number of peers of Agent a in the nonempty set $\mathcal{N} \subseteq \mathcal{N}_a$ who pick v in \mathbf{y} . That is,

$$N_a^v(\mathbf{y}, \mathcal{N}) = \sum_{a' \in \mathcal{N}} \mathbf{1}(y_{a'} = v).$$

We let $N_a(\mathbf{y}, \mathcal{N}) = (N_a^v(\mathbf{y}, \mathcal{N}))_{v \in \mathcal{Y}}$ be the vector of numbers of active peers selecting each alternative. We let $N_a(\mathbf{y}, \emptyset) = \mathbf{0}$ since, for any non-empty \mathcal{N} , $N_a(\mathbf{y}, \mathcal{N})$ has at least one non-zero component.

Assumption 2. *For each a , \mathbf{y} , $a' \in \mathcal{N}_a$, $\mathcal{N} \subseteq \mathcal{N}_a$, $\mathcal{N} \neq \emptyset$, and v*

(i) $Q^a(a' | \mathbf{y}) = Q_a(a' | y_a, y_{a'})$; and

(ii) $R^a(v | \mathbf{y}, \mathcal{N}) = R_a(v | y_a, N_a(\mathbf{y}, \mathcal{N}))$.

Assumption 2(i) states that the selection probability of a specific peer depends on the current choice of the agent and the one of the target peer. It also depends on the identity of the target peer.

This specification allows agents to pay more (or less) attention to peers who have chosen similar alternatives in the recent past. The choices of other peers do not affect this selection probability.

Example 1 (continued). Assume that Agent a follows a threshold rule when selecting peers to pay attention to: Agent a' is selected (at the moment of making a choice) if and only if $c_{a,a'}(y_a, y_{a'}) \geq \varepsilon$, where $c_{a,a'}(y_a, y_{a'})$ is an agent-pair-specific attention index that depends on the previous choices of both agents and ε is a random attention shock independent of \mathbf{y} . Then, the probability of selecting Agent a' is $Q_a(a' | y_a, y_{a'}) = F_\varepsilon(c_{a,a'}(y_a, y_{a'}))$, where F_ε is the c.d.f. of ε , and Assumption 2(i) is satisfied. \square

Assumption 2(ii) requires the choice rules to depend on the number of active peers selecting each alternative and not on the identity of peers. However, we allow for unrestricted dependence of choice rules on agent's own choice y_a .

Example 1 (continued). Suppose that indirect utilities of alternatives are linear-in-means or linear-in-aggregates: $u_{a,v}(\mathbf{y}, \mathcal{N}) = \alpha_{a,v}(y_a) + \beta_{a,v}(y_a) N_a^v(\mathbf{y}, \mathcal{N}) / |\mathcal{N}|$ or $u_{a,v}(\mathbf{y}, \mathcal{N}) = \alpha_{a,v}(y_a) + \beta_{a,v}(y_a) N_a^v(\mathbf{y}, \mathcal{N})$, where $\alpha_{a,v}(y_a)$ and $\beta_{a,v}(y_a) \neq 0$ are unknown parameters that depend on the current choice of Agent a . Then Assumption 2(ii) is satisfied. \square

The last assumption in this section is a mild regularity condition that rules out certain knife edge cases. Specifically, it ensures that when a peer switches her choice, the resulting effect on peer selection does not exactly offset the effect on preferences. This assumption allows us to identify the set of potential peer groups for each agent. Let $\mathbf{0}_{a'}^v$ be the configuration obtained from the zero vector by replacing the a' -th component by v . That is, this is a configuration where every one but Agent a' picked the default and Agent a' picked alternative v .

Assumption 3 (Regularity). *For all $a \in \mathcal{A}$, $a' \in \mathcal{N}_a$, there exists v such that*

$$\frac{Q_a(a' | 0, v)}{Q_a(a' | 0, 0)} \neq \frac{\sum_{\mathcal{N} \subseteq \mathcal{N}_a \setminus \{a'\}} [R^a(v | \mathbf{0}, \mathcal{N} \cup \{a'\}) - R^a(v | \mathbf{0}, \mathcal{N})] S^a(\mathcal{N} | \mathbf{0}, \mathcal{N}_a \setminus \{a'\})}{\sum_{\mathcal{N} \subseteq \mathcal{N}_a \setminus \{a'\}} [R^a(v | \mathbf{0}_{a'}^v, \mathcal{N} \cup \{a'\}) - R^a(v | \mathbf{0}, \mathcal{N})] S^a(\mathcal{N} | \mathbf{0}, \mathcal{N}_a \setminus \{a'\})}.$$

The left hand side of Assumption 3 captures the change in the probability that a peer is selected when she switches from the default option 0 to option v . The right hand side captures

the corresponding change in the preference component of the choice rule induced by the peer's action. Assumption 3 requires that these two effects are not equal. The restriction is imposed only at one choice configuration where everyone picks the default 0. This is not necessary, we only need the existence of a point where peer effects in preferences and in selection do not offset each other. Moreover, this choice configuration can vary with a and a' . For example, if switching Agent a' to an alternative increases or decreases both R^a and Q^a , then Assumption 3 is automatically satisfied.

2.3. Microfoundations for the Peer Selection Mechanism and Random Choices

In the paper, we take the process of peer selection and the choice rules as given and recover them nonparametrically. Next, we show that these processes can be generated by some well-known empirical models.

2.3.a. Peer Selection Mechanism.— Let us say that Agent a creates a link (at the moment of choosing) to Agent $a' \in \mathcal{N}_a$ if she includes this other agent in her active peer set. The selection of active peers is then governed by the probabilities of links between Agent a and all other agents in \mathcal{N}_a . There are many standard models for link probabilities based on random utility, limited attention, or continuous-time matching frictions (see, e.g., Duffie, Garleanu and Pedersen, 2005, Jackson, 2008, and Jackson and Asher, 1996). In these models, the links often result from random meetings and payoffs: When two agents randomly meet, forming a link yields a deterministic surplus that depends on agents' states and past behavior. In addition, agents face idiosyncratic costs or frictions that are privately observed. The link forms when the surplus is larger than the cost. We could interpret our model as summarizing this two-outcome process using a reduced-form selection probability. We allow the components of the link probability to depend on the previous actions of Agent a and the other agents in her reference group.

To illustrate this idea, let

$$\delta_{a,a'}(y_a, y_{a'}) \in (0, 1)$$

be the probability that Agent a meets Agent $a' \in \mathcal{N}_a$ at the moment of decision making. When Agents a and a' meet, the acceptance decision is governed by a random utility representation.

Specifically, if Agent a meets Agent a' , then the former obtains utility

$$U_{a,a'}(y_a, y_{a'}) = u_{a,a'}(y_a, y_{a'}) + \varepsilon$$

where $u_{a,a'}$ is a deterministic surplus function and ε is an idiosyncratic Logistic random shock. Agent a decides to add Agent a' to the set of active peers if $U_{a,a'}(y_a, y_{a'}) \geq 0$. This implies the selection probability

$$Q_a(a' | y_a, y_{a'}) = \delta_{a,a'}(y_a, y_{a'}) \frac{\exp(u_{a,a'}(y_a, y_{a'}))}{1 + \exp(u_{a,a'}(y_a, y_{a'}))}.$$

In this setup, the probability of paying attention to peers in $\mathcal{N} \subseteq \mathcal{N}_a$ is given by

$$S^a(\mathcal{N} | \mathbf{y}, \mathcal{N}_a) = \prod_{a' \in \mathcal{N}} \frac{\delta_{a,a'}(y_a, y_{a'}) \exp(u_{a,a'}(y_a, y_{a'}))}{1 + \exp(u_{a,a'}(y_a, y_{a'}))} \prod_{a' \in \mathcal{N}_a \setminus \mathcal{N}} \frac{1 + (1 - \delta_{a,a'}(y_a, y_{a'})) \exp(u_{a,a'}(y_a, y_{a'}))}{1 + \exp(u_{a,a'}(y_a, y_{a'}))}.$$

Randomness in the peer selection mechanism might also come from agents having a wrong model of who they should pay attention to or how the network works. Among other possibilities, they could misperceive who is relevant, use a wrong rule for sampling peers, or misspecify the selection probabilities. In these setups, peer selection reflects errors in beliefs, not just noise.

2.3.b. Choice Rule.— The random choice rule in our setting could simply reflect best choices for preference rankings with peer effects. Alternatively, while we do not model forward-looking behavior explicitly, one could view the choice rule as the reduced-form value from selecting an alternative. This reduced-form would implicitly incorporate the immediate payoff of selecting the option and the discounted value of payoffs that occur in the continuation state. An explicit forward-looking setup would also require adding beliefs (about future choices) to the model. Although this alternative formulation can lead to multiple equilibria, the empirical literature on dynamic games often assumes that there is a unique equilibrium in the data-generating process (Blevins, 2026).

In our model, it is also possible that agents would be choosing actions using a subjective model of peer effects that may be misspecified. In this case, the random components of the utility represent the prediction errors generated by the misspecification.

3. Equilibrium Behavior

The independent Poisson processes of the alarm clocks of the agents lead the selection process of alternatives through time. They guarantee that the transition rate from choice configuration \mathbf{y} to any different one \mathbf{y}' is as follows

$$w(\mathbf{y}' | \mathbf{y}) = \begin{cases} 0 & \text{if } \sum_{a \in \mathcal{A}} \mathbb{1}(y'_a \neq y_a) > 1 \\ \sum_{a \in \mathcal{A}} \lambda_a P_a(y'_a | \mathbf{y}) \mathbb{1}(y'_a \neq y_a) & \text{if } \sum_{a \in \mathcal{A}} \mathbb{1}(y'_a \neq y_a) = 1 \end{cases}. \quad (1)$$

These transition rates are the off-diagonal terms of the *transition rate matrix*.³ The diagonal terms are constructed from them in a simple way

$$w(\mathbf{y} | \mathbf{y}) = - \sum_{\mathbf{y}' \in \mathcal{Y}^A \setminus \{\mathbf{y}\}} w(\mathbf{y}' | \mathbf{y}).$$

We will indicate by \mathcal{W} the transition rate matrix. As the number of choice configurations is $(Y + 1)^A$, it follows that \mathcal{W} is a $(Y + 1)^A \times (Y + 1)^A$ matrix. An equilibrium in this model is an invariant distribution $\mu : \mathcal{Y}^A \rightarrow [0, 1]$, with $\sum_{\mathbf{y} \in \mathcal{Y}^A} \mu(\mathbf{y}) = 1$, of the dynamic process with transition rate matrix \mathcal{W} . It indicates the likelihood of each choice configuration \mathbf{y} in the long run. This equilibrium behavior relates to the transition rate matrix in a linear fashion

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

The assumptions in the paper guarantee the existence and uniqueness of an invariant distribution. Importantly, this invariant distribution has full support (i.e., any choice configuration \mathbf{y} realizes with positive probability).

Proposition 3.1. *If Assumptions 1 and 2 hold, there exists a unique, full support equilibrium μ .*

Example 2. Suppose that the network consists of two identical agents that select between two alternatives, option 1 and the default option 0. Let us also assume that the random preferences

³This transition rate matrix has many zeros in known locations. Blevins (2017, 2026) offers a nice discussion of this feature and its advantage for identification over discrete time models.

only depend on the choice of the other agent. As agents are identical, we will drop a from some terms. Thus, for $a = 1, 2$, the CCPs can be written in a rather succinct form

$$P_a(v | y_1, y_2) = Q(y_a, y_{-a}) R(v | 1 - y_{-a}, y_{-a}) + (1 - Q(y_a, y_{-a})) R(v | 0, 0).$$

The rates for their Poisson alarm clocks are 1. The transition rate matrix \mathcal{W} is as follows.

	(0, 0)	(0, 1)	(1, 0)	(1, 1)
(0, 0)	*	$P_2(1 0, 0)$	$P_1(1 0, 0)$	0
(0, 1)	$P_2(0 0, 1)$	*	0	$P_1(1 0, 1)$
(1, 0)	$P_1(0 1, 0)$	0	*	$P_2(1 1, 0)$
(1, 1)	0	$P_1(0 1, 1)$	$P_2(0 1, 1)$	*

The diagonal terms, *, are such that the elements in each line add up to 0. Note that, by symmetry, we have $P_1(1 | 1, 1) = P_2(1 | 1, 1)$, $P_1(0 | 0, 1) = P_2(0 | 1, 0)$, $P_1(0 | 1, 0) = P_2(0 | 0, 1)$, and $P_1(0 | 0, 0) = P_2(0 | 0, 0)$. After a simple calculation, the invariant distribution of choices, or steady-state equilibrium, is given by:

$$\begin{aligned} \mu_{(0,0)} &= \frac{P_1(0 | 1, 0) P_1(0 | 1, 1)}{P_1(1 | 0, 0) P_1(1 | 0, 1) + P_1(0 | 1, 0) P_1(0 | 1, 1) + 2 P_1(1 | 0, 0) P_1(0 | 1, 1)}, \\ \mu_{(1,0)} = \mu_{(0,1)} &= \frac{P_1(1 | 0, 0) P_1(0 | 1, 1)}{P_1(1 | 0, 0) P_1(1 | 0, 1) + P_1(0 | 1, 0) P_1(0 | 1, 1) + 2 P_1(1 | 0, 0) P_1(0 | 1, 1)}, \\ \mu_{(1,1)} &= \frac{P_1(1 | 0, 0) P_1(1 | 0, 1)}{P_1(1 | 0, 0) P_1(1 | 0, 1) + P_1(0 | 1, 0) P_1(0 | 1, 1) + 2 P_1(1 | 0, 0) P_1(0 | 1, 1)}. \end{aligned}$$

At equilibrium, the probability that the agents pick the same alternatives (i.e., coordinate) is

$$\mu_{(1,1)} + \mu_{(0,0)} = \frac{1}{1 + \frac{P_1(1 | 0, 1)}{P_1(0 | 1, 1)} + \frac{P_1(0 | 1, 0)}{P_1(1 | 0, 0)}}.$$

Let us also suppose that each agent pays attention to the other agent if (and only if) their choices coincide. That is, $Q(0, 0) = Q(1, 1) = 1$ and $Q(0, 1) = Q(1, 0) = 0$. Then the probability that

agents pick the same alternatives simplifies to

$$\Pr_{\text{same}} = \frac{1}{1 + \frac{\frac{R(1|0,0)}{R(0|0,1)} + \frac{R(0|0,0)}{R(1|1,0)}}{2}}.$$

If, in contrast, agents select each other only when their choices are different (i.e., $Q(0,0) = Q(1,1) = 0$ and $Q(0,1) = Q(1,0) = 1$), then

$$\Pr_{\text{diff}} = \frac{1}{1 + \frac{\frac{R(1|0,1)}{R(0|0,0)} + \frac{R(0|1,0)}{R(1|0,0)}}{2}}.$$

Suppose that agents are indifferent between alternatives if the peer is not selected (i.e., $R(1|0,0) = 0.5$). Then, since $f(x) = x(1-x)$ achieves its maximum at 0.5, we have that

$$\begin{aligned} R(1|0,0)R(0|0,0) &= R(1|0,0)(1 - R(1|0,0)) \\ &\geq \max\{R(1|0,1)(1 - R(1|0,1)), R(1|1,0)(1 - R(1|1,0))\} \\ &= \max\{R(1|0,1)R(0|0,1), R(1|1,0)R(0|1,0)\}. \end{aligned}$$

Thus,

$$\frac{R(1|0,0)}{R(0|0,1)} > \frac{R(1|0,1)}{R(0|0,0)} \text{ and } \frac{R(0|0,0)}{R(1|1,0)} > \frac{R(0|1,0)}{R(1|0,0)}$$

and

$$\Pr_{\text{same}} > \Pr_{\text{diff}}.$$

We can interpret this result by saying that when the peer selection is based on choice similarity, then agents with similar preferences select each other more often, and the equilibrium frequency by which they select the same alternatives is higher.

Let us finally compare these results with the standard model where agents always pay attention to all their peers. In this case, we have $P_1(1|1,1) = P_1(1|0,1) = R(1|0,1)$ and $P_1(1|0,0) =$

$P_1(1 | 1, 0) = R(1 | 1, 0)$. Thus, the probability that agents pick the same alternatives is

$$\Pr_{\text{std}} = \frac{1}{1 + \frac{\frac{R(1 | 0, 1)}{R(0 | 0, 1)} + \frac{R(0 | 1, 0)}{R(1 | 1, 0)}}{2}}.$$

Interestingly, if the peer effects on preferences are positive (i.e., $R(1 | 0, 1) > R(1 | 0, 0) > R(1 | 1, 0)$), then

$$\frac{R(1 | 0, 1)}{R(0 | 0, 1)} > \frac{R(1 | 0, 0)}{R(0 | 0, 1)} \text{ and } \frac{R(0 | 1, 0)}{R(1 | 1, 0)} > \frac{R(0 | 0, 0)}{R(1 | 1, 0)}$$

and

$$\Pr_{\text{std}} > \Pr_{\text{same}} > \Pr_{\text{diff}}.$$

However, if the peer effects are negative (i.e., $R(1 | 0, 1) < R(1 | 0, 0) < R(1 | 1, 0)$), then

$$\frac{R(1 | 0, 1)}{R(0 | 0, 1)} < \frac{R(1 | 0, 0)}{R(0 | 0, 0)} \text{ and } \frac{R(0 | 1, 0)}{R(1 | 1, 0)} < \frac{R(0 | 0, 0)}{R(1 | 0, 0)}$$

and

$$\Pr_{\text{same}} > \Pr_{\text{diff}} > \Pr_{\text{std}}.$$

□

4. Identification of the Model

4.1. Identification of the Model from the CCPs

This section shows that all parts of the model can be recovered from a long sequence of choices. These parts include the network structure, the random preferences and the peer selection mechanism. That is, we will recover

$$(\mathcal{N}_a)_{a \in \mathcal{A}}, (R_a)_{a \in \mathcal{A}}, (Q_a)_{a \in \mathcal{A}}.$$

The starting point of our identification argument is the set of CCPs $\mathbf{P} = (P_a)_{a \in \mathcal{A}}$ that can be calculated from the long sequence of choices made by the group members. Recall that, for each $v \in \mathcal{Y}$ and $a \in \mathcal{A}$, we have that

$$P_a(v | \mathbf{y}) = \sum_{\mathcal{N} \subseteq \mathcal{N}_a} R^a(v | \mathbf{y}, \mathcal{N}) S^a(\mathcal{N} | \mathbf{y}, \mathcal{N}_a)$$

where

$$R^a(v | \mathbf{y}, \mathcal{N}) = R_a(v | y_a, N_a(\mathbf{y}, \mathcal{N}))$$

and

$$S^a(\mathcal{N} | \mathbf{y}, \mathcal{N}_a) = \prod_{a' \in \mathcal{N}} Q_a(a' | y_a, y_{a'}) \prod_{a' \in \mathcal{N}_a \setminus \mathcal{N}} (1 - Q_a(a' | y_a, y_{a'})).$$

Proposition 4.1. *Under Assumptions 1, 2, and 3, \mathcal{N}_a is identified from P_a for every a .*

Example 1 (continued). Let us consider Agent 2. (The identification strategy is similar for the other agents.) The probability that Agent 2 selects alternative 1 given choice configuration \mathbf{y} is

$$P_2(1 | \mathbf{y}) = R_2(1 | y_2, \mathbf{1}(y_1 = 0), \mathbf{1}(y_1 = 1), \mathbf{1}(y_1 = 2)) Q_2(1 | y_2, y_1) + R_2(1 | y_2, 0, 0, 0)(1 - Q_2(1 | y_2, y_1)).$$

Take, to simplify the exposition, $\mathbf{y} = \mathbf{0}$. After some manipulation, the changes in P_2 when we change the choices of Agents 1, 3, and 4 from 0 to 1 are, respectively,

$$\begin{aligned} P_2(1 | 1, 0, 0, 0) - P_2(1 | 0, 0, 0, 0) &= (R_2(1 | 0, 0, 1, 0) - R_2(1 | 0, 0, 0, 0)) Q_2(1 | 0, 1) \\ &\quad - (R_2(1 | 0, 1, 0, 0) - R_2(1 | 0, 0, 0, 0)) Q_2(1 | 0, 0), \end{aligned}$$

$$P_2(1 | 0, 0, 1, 0) - P_2(1 | 0, 0, 0, 0) = 0,$$

$$P_2(1 | 0, 0, 0, 1) - P_2(1 | 0, 0, 0, 0) = 0.$$

Proposition 4.1 identifies the peers of Agent 2 as those for which these changes differ from 0. Assumption 3 guarantees that the two terms in the right-hand side of the first line are not of equal size. Following this idea, we correctly recover $\mathcal{N}_2 = \{1\}$. \square

To recover the choice rules and the peer selection mechanisms, we implement two variants of an

idea that we describe next by example. The idea relies on manipulating the size of the set \mathcal{N}_a .

Example 1 (continued). Recall that (by Proposition 4.1) we can recover the peer group of each agent. Thus, for instance, we can learn that $\mathcal{N}_1 = \{2, 3\}$. Consider the following thought experiment. Take Agent 1 and assume that we can *remove* agents from her reference group and recover her CCPs. In doing so, we will vary $|\mathcal{N}_1|$ along 0, 1, and 2. The probabilities by which Agent 1 select alternative v if we remove Agents 2 and 3, only Agent 3, only Agent 2, and none of them from her reference group are, respectively, are given by

$$P_1(v | \mathbf{y}) = R_1(v | y_1, 0, 0, 0)$$

$$P_1(v | \mathbf{y}) = R_1(v | y_1, \mathbf{1}(y_2 = 0), \mathbf{1}(y_2 = 1), \mathbf{1}(y_2 = 2)) Q_1(2 | y_1, y_2) + R_1(v | y_1, 0, 0, 0)(1 - Q_1(2 | y_1, y_2))$$

$$P_1(v | \mathbf{y}) = R_1(v | y_1, \mathbf{1}(y_3 = 0), \mathbf{1}(y_3 = 1), \mathbf{1}(y_3 = 2)) Q_1(3 | y_1, y_3) + R_1(v | y_1, 0, 0, 0)(1 - Q_1(3 | y_1, y_3))$$

$$\begin{aligned} P_1(v | \mathbf{y}) = & R_1(v | y_1, \sum_{a'=2,3} \mathbf{1}(y_{a'} = 0), \sum_{a'=2,3} \mathbf{1}(y_{a'} = 1), \sum_{a'=2,3} \mathbf{1}(y_{a'} = 2)) Q_1(2 | y_1, y_2) Q_1(3 | y_1, y_3) \\ & + R_1(v | y_1, \mathbf{1}(y_2 = 0), \mathbf{1}(y_2 = 1), \mathbf{1}(y_2 = 2)) Q_1(2 | y_1, y_2)(1 - Q_1(3 | y_1, y_3)) \\ & + R_1(v | y_1, \mathbf{1}(y_3 = 0), \mathbf{1}(y_3 = 1), \mathbf{1}(y_3 = 2))(1 - Q_1(2 | y_1, y_2)) Q_1(2 | y_1, y_3) \\ & + R_1(v | y_1, 0, 0, 0)(1 - Q_1(2 | y_1, y_2))(1 - Q_1(3 | y_1, y_3)). \end{aligned}$$

The above system relates four observed CCPs and six unknowns (i.e., the choice rule and the peer selection probabilities). It allows the researcher to uniquely recover one of the unknowns ($R_1(v | y_1, 0, 0, 0)$) and partially identify the other ones. \square

The example shows that if the researcher *could* manipulate the reference groups (by removing some links) the observed data would be informative about both the choice rules and the peer selection mechanism. In the applications we have in mind, however, the researcher has no control over reference groups and cannot arbitrarily modify them. To address this issue, we introduce two models in which the required variation arises endogenously through different channels. In each case, we also add restrictions to restore full (rather than partial) identification.

Different Sizes of Reference Groups across Agents.— In the first setup, we assume that agents can be partitioned into a finite set of types $\mathcal{H} = \{1, 2, \dots, H\}$, $H \leq A$. The type of each agent is given by a known mapping $h : \mathcal{A} \rightarrow \mathcal{H}$ where $h(a)$ encodes the type of Agent a . These types might be

related to covariates or other individual characteristics that the researcher observes. Each type incorporates two dimensions: a random choice rule that captures the preferences of the agent over the set of alternatives; and a limited attention mechanism that connects the preferences of each agent to the choices of her active peers. These types do not impose any restrictions on the reference groups of peers. That is, two agents of the same type might have very different reference groups.

Example 1 (continued). Suppose that Agents 1 and 2 have a college degree while Agents 3 and 4 do not. Thus, we can define h as $h(1) = h(2) = 1$ and $h(3) = h(4) = 2$ with $\mathcal{H} = \{1, 2\}$. \square

The main assumption is as follows.

Assumption 4. For each a , \mathbf{y} , $a' \in \mathcal{N}_a$, $\mathcal{N} \subseteq \mathcal{N}_a$, and v

(i) $Q_a(a' | y_a, y_{a'}) = Q_{h(a)}(y_a, y_{a'})$; and

(ii) $R_a(v | y_a, N_a(\mathbf{y}, \mathcal{N})) = R_{h(a)}(v | y_a, N_a(\mathbf{y}, \mathcal{N}))$.

Assumption 4 imposes that agents belonging to the same type share identical functional forms for both selection probabilities and choice rules. That is, if two agents have the same type, they use the same peer-selection rule to determine which peers are considered and the same choice rule to choose among alternatives, up to observable state variables. Types therefore represent clusters of agents that behave similarly in how they form attention over peers and how they translate the set of considered peers into choices.

In our empirical application, an agent is defined as a firm-market pair. The type of an agent is the firm to which it belongs. The coefficients governing how competitor (peer) presence affects attention and profitability are firm specific but do not vary across markets. Markets differ through observable characteristics and the realized competitive environment, but the mapping from these variables into attention and choice probabilities is common within firm. Thus, types partition agents by firm, and each firm defines a behavioral cluster with its own selection and payoff parameters.

The identification result we describe below requires variation in the sizes of reference groups among agents within each type. As before, this type of variation leads to partial identification of the choice rules and peer selection mechanisms. Under Assumptions 1 and 2, the CCPs of Agent

a can be presented as a binomial mixture (or convolution) model. For each $v \in \mathcal{Y}$ and $a \in \mathcal{A}$, $P_a(v | \mathbf{y})$ can be written as

$$\sum_{n_0=0}^{N_a^0(\mathbf{y}, \mathcal{N}_a)} \cdots \sum_{n_Y=0}^{N_a^Y(\mathbf{y}, \mathcal{N}_a)} R_{h(a)}(v | y_a, n_0, \dots, n_Y) \prod_{y=0}^Y \binom{N_a^y(\mathbf{y}, \mathcal{N}_a)}{n_y} Q_{h(a)}(y_a, y)^{n_y} (1 - Q_{h(a)}(y_a, y))^{N_a^y(\mathbf{y}, \mathcal{N}_a) - n_y},$$

where $\sum_{y=0}^Y N_a^y(\mathbf{y}, \mathcal{N}_a) = |\mathcal{N}_a|$ for all $\mathbf{y} \in \mathcal{Y}^A$. After learning $(\mathcal{N}_a)_{a \in \mathcal{A}}$ we recover $(R_a)_{a \in \mathcal{A}}, (Q_a)_{a \in \mathcal{A}}$ taking advantage of variation on the size of the reference groups, i.e. $|\mathcal{N}_a|$.

To obtain full identification, we invoke two other restrictions. In the next assumption, let $\mathbf{0}_v^{|\mathcal{N}|}$ be a size $|\mathcal{Y}|$ vector of aggregate choices where all active peers select option v . (That is, a vector of 0s except for the v -component that is $|\mathcal{N}|$.)

Assumption 5. For each $a, \mathbf{y}, \mathcal{N} \subseteq \mathcal{N}_a, \mathcal{N} \neq \emptyset, v$, and v'

(i) $R_a(v | y_a, \mathbf{0}_{v'}^{|\mathcal{N}|}) = R_a(v | y_a, \mathbf{0}_v^{|\mathcal{N}|})$; and

(ii) $R_a(v | y_a, N_a(\mathbf{y}, \mathcal{N})) \neq R_a(v | y_a, \mathbf{0})$.

Assumption 5(i) states that the choice rule applied when all active peers select the same alternative does not change with the number of active peers. This restriction is satisfied if, for instance, the choice rule depends on the proportion of peers selecting each alternative—an assumption that is rather standard in the peer effects literature. Assumption 5(ii) simply states that the choice rule of the empty active peer set differs from the choice rule of the non-empty set for any choice configuration.

Proposition 4.1 states that we can recover the reference group of each agent from the CCPs. We next use a recursive argument to show that in this setting we can recover the choice rules and the peer selection mechanism if we have variation on the number of peers in the reference groups of agents of the same type. Specifically, let

$$N_t \equiv \{|\mathcal{N}_a| : a \in \mathcal{A}, h(a) = t\}$$

be the set of all *sizes* of reference groups for agents of type t . We next show that having at least three agents of the same type with different numbers of potential peers provides enough variation

to identify peer selection probabilities and choice rules with no or just one active peer.

Proposition 4.2. *In addition to Assumptions 1, 2, 3, 4, and 5, suppose that $|\mathcal{N}_t| \geq 3$ for each $t \in \mathcal{H}$. Then, Q^a , $R^a(\cdot | \cdot, \emptyset)$ and $R^a(\cdot | \cdot, \{a'\})$, $a' \in \mathcal{N}_a$, are identified from P_a for every a .*

Although we may not identify the choice rule for all possible active peer groups nonparametrically, knowing it for the empty set and for singletons is enough if we further assume preferences are represented by the standard linear-in-means multinomial logit model. Below, assuming richer variation in the sizes of the reference groups, we show that random choice rules can be nonparametrically identified without restricting attention to the logit model.

Proposition 4.3. *If the choice rule satisfies*

$$R^a(v | \mathbf{y}, \mathcal{N}) = \frac{e^{\alpha_{h(a),v}(y_a) + \beta_{h(a),v}(y_a) N_a^v(\mathbf{y}, \mathcal{N}) / |\mathcal{N}|}}{\sum_{v' \in \mathcal{Y}} e^{\alpha_{h(a),v'}(y_a) + \beta_{h(a),v'}(y_a) N_a^{v'}(\mathbf{y}, \mathcal{N}) / |\mathcal{N}|}}$$

with the normalization $\alpha_{h(a),0}(y_a) = 0$, then R^a is identified from $R^a(\cdot | \cdot, \emptyset)$ and $R^a(\cdot | \cdot, \{a'\})$, $a' \in \mathcal{N}_a$.

Once the peer selection mechanism and the choice rule (for the cases of no and just one active peer) are recovered, we can recursively identify the choice rule for all other subsets of active peers if we have full variation in the size of connected agents within each type. For example, suppose Agent a has 2 peers, a_1 and a_2 (i.e., $\mathcal{N}_a = \{a_1, a_2\}$). Then

$$\begin{aligned} P_a(v | \mathbf{y}) &= S^a(\emptyset | \mathbf{y}, \mathcal{N}_a) R^a(v | \mathbf{y}, \emptyset) + S^a(\{a_1\} | \mathbf{y}, \mathcal{N}_a) R^a(v | \mathbf{y}, \{a_1\}) \\ &\quad + S^a(\{a_2\} | \mathbf{y}, \mathcal{N}_a) R^a(v | \mathbf{y}, \{a_2\}) + S^a(\{a_1, a_2\} | \mathbf{y}, \mathcal{N}_a) R^a(v | \mathbf{y}, \{a_1, a_2\}). \end{aligned}$$

Hence, since the selection mechanism and the choice rule with no and just one active peer are identified, we can recover the choice rule with two active peers, $R^a(v | \mathbf{y}, \{a_1, a_2\})$. Repeating the same argument for someone who has three peers, we can identify the choice rule with three active peers. Thus, with enough variation in the sizes of the reference groups for different agents within each type, we can identify R^a without any parametric restrictions.

Assumption 6. For every type $t \in \mathcal{H}$,

$$\{2, 3, \dots, \max_{\{a \in \mathcal{A}, h(a)=t\}} |\mathcal{N}_a|\} \subseteq \mathbb{N}_t.$$

Assumption 6 means that within each type there is some agent with two peers, someone with three peers, etc. This assumption is similar to the full menu variation (of alternatives) in the stochastic choice literature (Aguiar et al., 2023).

The proof of the following result follows from Proposition 4.2 and the above discussion.

Proposition 4.4. Suppose the assumptions of Proposition 4.2 are satisfied. If, moreover, Assumption 6 is satisfied, then Q^a and R^a are identified from P_a for every a .

Example 1 (continued). Assume the four agents in the example are of the same type. Thus, to simplify the exposition, we will drop the subindices in Q and R . Recall that, by Proposition 4.1, we can recover the peer group of each of them. Thus, we can learn that

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

After some manipulation, we get that

$$\frac{P_1(v | v', v'', v'', v'') - P_4(v | v'', v'', v'', v')}{P_2(v | v'', v', v'', v'') - P_4(v | v'', v'', v'', v')} = 2 - Q(v', v'').$$

Since the left-hand side is observed, we can recover $Q(v', v'')$ for each pair $v', v'' \in \mathcal{Y}$. That is, the peer selection mechanism can be recovered by combining (in a specific way) the CCPs of agents that differ regarding number of peers in their reference groups. We next show how to recursively recover the choice rule.

In the example, Agent 4 has no peers. It follows that

$$P_4(v | \mathbf{y}) = R(v | y_4, \mathbf{0}).$$

This means that we can recover from the CCPs of Agent 4 the choice rules for the empty set of active

peers. Let $\mathbf{0}_{v'}^1$ be a vector of three elements, all of which are 0 except for the one corresponding to (the different) alternative v' which is 1. Let us consider Agent 2

$$P_2(v \mid y_1 = v', y_2, y_3, y_4) = R(v \mid y_2, \mathbf{0}_{v'}^1) Q(y_2, v') + R(v \mid y_2, \mathbf{0})(1 - Q(y_2, v')).$$

Since we already identified two of the three terms on the right-hand side of the last expression, we can recover the choice rule for one active peer from the CCPs of Agent 2. Let $\mathbf{0}_{v',v''}^1$ be a vector of three elements, all of which are 0 except for the ones corresponding to alternatives v' and v'' which are 1. Let us next consider Agent 1

$$P_1(v \mid y_1, y_2 = v', y_3 = v'', y_4) = R(v \mid y_1, \mathbf{0}_{v',v''}^1) Q(y_1, v') Q(y_1, v'') + R(v \mid y_1, \mathbf{0})(1 - Q(y_1, v'))(1 - Q(y_1, v'')) \\ + R(v \mid y_1, \mathbf{0}_{v'}^1) Q(y_1, v')(1 - Q(y_1, v'')) + R(v \mid y_1, \mathbf{0}_{v''}^1)(1 - Q(y_1, v')) Q(y_1, v'').$$

Since we already identified five of the six terms on the right-hand side of the last expression, we can recover the choice rule for two active peers (that have chosen different alternatives) from the CCPs of Agent 1. \square

Outside Option with no Peer Effects.— In the second setting, we assume that the choice rules are not affected by active peers selecting the outside option 0. For the next assumption, let $\mathbf{N} = (\mathbf{N}^v)_{v \in \mathcal{Y}}$ indicate a vector of the number of peers that select each alternative.

Assumption 7. For each a , \mathbf{y} , and $v \neq 0$, (i) the choice rule $R_a(v \mid y_a, \mathbf{N})$ does not vary with \mathbf{N}^0 ; and (ii) $R_a(v \mid y_a, \mathbf{0}_{v'}^1) - R_a(v \mid y_a, \mathbf{0}) \neq 0$ for all $v \in \mathcal{Y} \setminus \{0\}$.

Assumption 7(i) allows us to mimic variation in reference groups for different choice configurations at the level of each agent. As we described before, this variation produces useful information to recover the model. In our case, it leads to a partially identified model where the choice rule and peer-selection mechanism can be recovered up to one element. Assumption 7(ii) just means that there is a peer effect in preferences.

The next assumption offers a condition that suffices to recover all elements of the model.

Assumption 8. For each $a \in \mathcal{A}$ with $|\mathcal{N}_a| \geq 1$, and all $v_1, v_2 \in \mathcal{Y}$, the value of either

$$\min_{a' \in \mathcal{N}_a} Q_a(a' | v_1, v_2) \quad \text{or} \quad \max_{a' \in \mathcal{N}_a} Q_a(a' | v_1, v_2)$$

is known.

Assumption 8 requires the existence of a known maximal or minimal selection probability for all possible configurations of choices that Agent a and the target Agent a' can take. It is satisfied if for every configuration of v_1 and v_2 there is a peer that is always selected: $Q_a(a' | v_1, v_2) = 1$ for some a' . (This peer does not have to be the same for all possible v_1 and v_2 .) The choice of this peer works as an exclusion restriction by shifting the choice rule without affecting the selection process. Importantly, the set of peers where the maximum or minimum in Assumption 8 is achieved does not need to be known. For example, if we assume that one peer is always selected, this peer does not need to be known and can be identified from the data.

Proposition 4.5. If Assumptions 1, 2, 3, 7, and 8 are satisfied, then Q^a and R^a are identified from P_a for every a .

Example 1 (continued). Assume that all agents are of the same type. Recall that, by Proposition 4.1, we know

$$\mathcal{N}_1 = \{2, 3\}.$$

Take Agent 1. If Assumption 7 holds we can mimic variation in the reference groups by considering choice configurations in which some of the agents in her reference group select the outside option. Doing so, we obtain

$$P_1(v | y_1, 0, 0, 0) = R(v | y_1, 0, 0, 0)$$

$$P_1(v | y_1, y_2, 0, 0) = R(v | y_1, \mathbf{1}(y_2 = 0), \mathbf{1}(y_2 = 1), \mathbf{1}(y_2 = 2)) Q(y_1, y_2) + R(v | y_1, 0, 0, 0)(1 - Q(y_1, y_2))$$

$$P_1(v | y_1, 0, y_3, 0) = R(v | y_1, \mathbf{1}(y_3 = 0), \mathbf{1}(y_3 = 1), \mathbf{1}(y_3 = 2)) Q(y_1, y_3) + R(v | y_1, 0, 0, 0)(1 - Q(y_1, y_3))$$

$$P_1(v | y_1, y_2, y_3, 0) = R(v | y_1, \sum_{a'=2,3} \mathbf{1}(y_{a'} = 0), \sum_{a'=2,3} \mathbf{1}(y_{a'} = 1), \sum_{a'=2,3} \mathbf{1}(y_{a'} = 2)) Q(y_1, y_2) Q(y_1, y_3) \\ + R(v | y_1, \mathbf{1}(y_2 = 0), \mathbf{1}(y_2 = 1), \mathbf{1}(y_2 = 2)) Q(y_1, y_2)(1 - Q(y_1, y_3))$$

$$\begin{aligned}
& + R(v \mid y_1, \mathbb{1}(y_3 = 0), \mathbb{1}(y_3 = 1), \mathbb{1}(y_3 = 2))(1 - Q(y_1, y_2)) Q(y_1, y_3) \\
& + R(v \mid y_1, 0, 0, 0)(1 - Q(y_1, y_2))(1 - Q(y_1, y_3)).
\end{aligned}$$

As in the thought experiment we described earlier, this approach generates a system of four observed probabilities that depend on six unknowns related to the choice rules and peer selection probabilities. The system allows the researcher to uniquely recover one of the unknowns and partially identify the other ones. If we also assume that either Agent 2 or 3 is always an active peer, then we can uniquely recover the choice rule and the peer-selection mechanism. \square

The two identification strategies we offer are different in the sets of assumptions they are imposing. (Table 1 summarizes them.) Most of these assumptions, are not mutually exclusive so these two strategies can be combined if a particular application requires it. Moreover, these identification results do not impose parametric functional form restrictions on peer-selection and choice rules. If one is willing to impose parametric restrictions, as we do in our empirical illustration, then some of these assumptions can be weakened or completely dropped (e.g., full support in networks sizes is not needed anymore).

Table 1 – Comparison of Identification Assumptions (Section 4.1)

	Strategy 1	Strategy 2
Types	Yes	No
Variation in $ \mathcal{N}_a $	Yes	No
Full support in network sizes	Yes	No
Invariance when all peers choose same option	Yes	No
Outside option with no peer effects	No	Yes
Known max/min selection probability	No	Yes

4.2. Identification of Conditional Choice Probabilities

This section shows that the CCPs, P , and the rates of the Poisson alarm clocks, λ , can be recovered from a long sequence of choices. We assume the researcher observes the choices of the agents at time intervals of length Δ and can consistently estimate $\Pr(\mathbf{y}^{t+\Delta} = \mathbf{y}' \mid \mathbf{y}^t = \mathbf{y})$ for each pair

$\mathbf{y}', \mathbf{y} \in \mathcal{Y}^A$, to construct a matrix $\mathcal{P}(\Delta)$.⁴ Matrix $\mathcal{P}(\Delta)$ relates to the transition rate matrix \mathcal{W} by $\mathcal{P}(\Delta) = e^{(\Delta\mathcal{W})}$. (Here $e^{(\Delta\mathcal{W})}$ is the matrix exponential of $\Delta\mathcal{W}$.) Often, the researcher observes the precise moment at which agents revise strategies and the choices at that time. In other cases, the researcher simply observes the configuration of choices at fixed time intervals —e.g., every Monday. [Kashaev et al. \(2026\)](#) refer to these two cases as Dataset 1 and 2, respectively. Formally, in Dataset 1, the researcher can consistently estimate \mathcal{W} , and, in Dataset 2, the researcher can consistently estimate $\mathcal{P}(\Delta)$. In the second case, the identification question is whether (or under what extra restrictions) we can uniquely recover \mathcal{W} from $\mathcal{P}(\Delta)$. The identification problem in Dataset 1 is straightforward. [Kashaev et al. \(2026\)](#), using Theorem 1 in [Blevins \(2026\)](#), offer a mild condition under which the transition rate matrix can be identified from Dataset 2.⁵

Proposition 4.6. *If Assumptions 1 and 4 hold, then the CCPs \mathbf{P} and the rates of the Poisson alarm clocks $(\lambda_a)_{a \in \mathcal{A}}$ are identified from Dataset 1. If, moreover, \mathcal{W} has distinct eigenvalues that do not differ by an integer multiple of $2\pi i/\Delta$, where i denotes the imaginary unit, then \mathbf{P} and $(\lambda_a)_{a \in \mathcal{A}}$ are generically identified from Dataset 2.*

The restriction on eigenvalues of \mathcal{W} is a regularity condition that is generically satisfied.⁶ The key element in proving the second statement in Proposition 4.6 is that the transition rate matrix in our model is rather parsimonious. Since, at each time, at most one person revises her selection with a nonzero probability, \mathcal{W} has many zeros in known locations.

5. Application

We study the dynamic expansion and contraction decisions of five leading national fast-food franchising chains —Wallace, KFC, McDonald’s, Burger King, and Dicos— operating across geographically distinct markets. The empirical setting consists of a panel of fast-food chains

⁴Here again, we assume that the choice configurations are ordered according to the lexicographic order when we construct $\mathcal{P}(\Delta)$.

⁵It is also known that if the researcher were 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).

⁶See [Blevins \(2017\)](#) for a discussion of this assumption.

operating in Chinese cities over several decades. Firms decide whether to open or close restaurants in each market, taking into account local market conditions and the presence of competitors. A key feature of our setting is that firms can endogenously update their set of active competitors over time. The possibility that firms might ignore some competitors reflects information frictions, bounded rationality, and the fact that only a subset of rivals could be strategically relevant.

5.1. Empirical Model

We first describe the model of firm expansion and contraction decisions and then introduce the specifications for selection of competitors and firm payoffs. Similar to [Kashaev et al. \(2026\)](#), we define a market at the level of a prefecture-level city, which ranks below a province and above a county in China’s administrative structure. Thus, markets are geographically isolated from each other. We collect all unknown primitives by θ .

Choice Set, Agents, Types, and Peers The empirical unit of observation is a firm–market pair (f, m) , where $f \in \mathcal{F} = \{1, \dots, F\}$ indexes firms and $m \in \mathcal{M} = \{1, \dots, M\}$ indexes geographic markets. Every firm f decides whether to open a restaurant ($v = 2$), do nothing ($v = 1$), or close a restaurant ($v = 0$) in market m . We call a pair $a = (f, m)$ an agent —knowing the firm and the market identifies the agent and vice versa. Thus, $\mathcal{A} = \mathcal{F} \times \mathcal{M}$ and $\mathcal{Y} = \{0, 1, 2\}$.⁷ Given that in our data, in about 77% days with any action (restaurant closure or opening), action happens in more than one city, we assume that the agents have perfectly synchronized clocks. This assumption does not affect the identification argument in [Section 4.1](#)

In line with our assumption about types, we assume that firm parameters do not vary across markets: preference and selection parameters are firm-specific but common to all markets where the firm operates. That is, in our application $h(a) = h((f, m)) = f$. Next, we assume that the set of firms perceived as (potential) competitors varies across markets for exogenous reasons (such as entry history or geography) and that the researcher does not directly observe which firms are included.

⁷We abuse notation a bit since \mathcal{A} was previously defined as a set of the form $\{1, 2, \dots, A\}$. To be consistent with the initial notation, we can take any one-to-one mapping $\tilde{a} : \mathcal{F} \times \mathcal{M} \rightarrow \{1, 2, \dots, |\mathcal{F} \times \mathcal{M}|\}$ and define an agent as $a = \tilde{a}(f, m)$.

This variation in the size of potential competitors (or reference groups) is used for identification, as in the theoretical model.

Since these fast-food restaurants are small and operate under franchise arrangements, we argue that the managers of these restaurants (i) make decisions independently across markets and (ii) do not always consider every competitor within the market. Managers might ignore the information about competitors for different reasons. For example, the number of competitors' restaurants may be small or the manager may not think that a particular restaurant is a substitute. In addition, monitoring competitors is costly in terms of time, data collection, analysis costs, and managers might not be able to process all market information at all times.

At the moment of deciding whether to open or close a restaurant, the attention that firm f pays to her competitors within market m depends on her own and competitor's past choices in market m . Moreover, since the markets are locally isolated, we follow the literature (e.g., Arcidiacono, Bayer, Blevins and Ellickson, 2016) and assume the marginal profit of firm f in market m from opening a new restaurant is only affected by both its own and its competitors' openings in market m . Formally, for any $a = (f, m)$, $a' \notin \mathcal{N}_a$ if $m' \neq m$.

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

Peer selection In our application, we assume that every firm may or may not consider any other firm that could operate in the market. Given the numbers of restaurants each firm had in the market at time t , N_t , the probability that Agent $a = (f, m)$ considers Agent $a' = (f', m')$ at time t is

$$Q_f(N_{at}, N_{a't}, S_{mt}) = \mathbb{1}(a' \in \mathcal{N}_a) F_{\bar{\varepsilon}}(\bar{\pi}_f(N_{at}, N_{a't}, S_{mt}; \theta)),$$

where $F_{\bar{\varepsilon}}$ is a known c.d.f.; θ is the vector of unknown parameters; and $\bar{\pi}_f(N_{at}, N_{a't}, S_{mt}; \theta)$ is the mean attention index of firm f , which is known up to θ . We allow the current number of restaurants of firms f and f' in market m to affect the attention index of firm f .

Payoff from Opening or Closing a Restaurant Conditional on a set of competitors being considered, the firm decides whether to open or close at least one new restaurant, or do nothing in that market based on its marginal profit π_{at} . This marginal profit captures not just the instantaneous (one period) profitability of an extra restaurant, but the expected profitability of the restaurant in the long run —see the discussion in Section 2.3.b.

This is an ordered response model, with the choice being

$$v = \mathbf{1} \left(\bar{\pi}_f \left(N_{at}, \sum_{a' \in \mathcal{N}} N_{a't}, S_{mt}; \theta \right) + \varepsilon \geq \gamma_{1f} \right) + \mathbf{1} \left(\bar{\pi}_f \left(N_{at}, \sum_{a' \in \mathcal{N}} N_{a't}, S_{mt}; \theta \right) + \varepsilon \geq \gamma_{1f} + \gamma_{2f} \right),$$

where ε is an idiosyncratic random shock with a known c.d.f. F_ε ; $\bar{\pi}_f(N_{at}, \sum_{a' \in \mathcal{N}} N_{a't}, S_{mt}; \theta)$ is the mean marginal profit of firm f in market m that depends on the own number of restaurants, the aggregate number of restaurants of considered competitors, and market characteristics. The profit function is known up to θ . The threshold parameters γ_{1f} and $\gamma_{2f} > 0$ (in θ) are unknown.

Remark 1. Formally, our application is using *all* previous choices of firms as determinants of selection and marginal profits. That is, we have infinite history dependence. However, similar to Kashaev et al. (2026), since the model we estimate in our application is fully parametric, this should not lead to any issues.

Model Implied CCP Altogether, the above parametrization leads to a model implied distribution over three ordered outcomes for every firm in every market:

$$\Pr_a(v \mid (N_{a't})_{a' \in \mathcal{N}_a \cup \{a\}}, S_{mt}; \theta),$$

which completely characterizes the probability of observing a new restaurant or a restaurant closure in a given market by a given firm conditional on the history and the market characteristics. When evaluated at the true parameter value θ_0 , \Pr_a matches the CCP P_a , i.e., $P_a(v \mid N_t, S_t) = \Pr_a(v \mid (N_{a't})_{a' \in \mathcal{N}_a \cup \{a\}}, S_{mt}; \theta_0)$.

The vector of parameters θ contains the parameters entering the mean marginal profits and mean attention that relate to both the covariates and the reference groups \mathcal{N}_a , $a \in \mathcal{A}$. Note that

we assume that agents that are representing the same firm are of the same type. That is, the mean profit and attention functions do not vary across markets for a given firm, but are allowed to be different for different firms. Finally, we do not assume anything about \mathcal{N}_a except that only agents operating within the same market might be in \mathcal{N}_a .

5.2. Estimation

Data The data we have consist of three objects: (i) the exact date of restaurant openings and closures $\{t_k\}_{k=1}^K$; (ii) the state of the market structure $\{N_{at_k}\}_{a \in \mathcal{A}, k=1, \dots, K}$ sampled from a continuous time over interval $[0, t_K]$, where N_{at_k} is the number of restaurants 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 $\{S_{m,t_k}\}_{m \in \mathcal{M}, k=1, \dots, K}$.

The dataset records the full history of store openings and closures from 1979 through 2024. Each observation corresponds to a discrete event in which a firm opens or closes one or more restaurants in a given city. In total, the data contain 39,198 such events, including 31,509 openings and 7,689 closures, observed over 6,172 distinct action dates. The action dataset was purchased from CnOpenData, a data marketing company.⁸

We use the early portion of the data, from 1979 to 2002, to construct initial conditions for the empirical analysis. These observations allow us to recover the pre-sample distribution of firm presence across markets and to initialize the state variables governing the number of restaurants operated by each firm. The main estimation sample covers the period from 2003 to 2019. Data after 2019 are excluded to avoid confounding effects associated with the COVID-19 pandemic.

To characterize market conditions, we merge the action data with a panel of city-level economic and demographic variables covering the years 2003-2019 obtained from China Data Online.⁹ These variables include population, land area, real GDP (deflated to 2019 prices¹⁰), and an indicator for provincial capital status. In the overlap period (2003–2019), 63.42% of days have at least one

⁸<https://www.cnopendata.com>

⁹<https://www.china-data-online.com>

¹⁰Source: World Development Indicators. The last updated date is 2025-12-16.

opening.

Likelihood Function Our identification argument does not require any parametric assumptions if there is full variation in the size of peer groups. However, for small and moderate-sized samples, we suggest using the parametric maximum likelihood estimator of CCPs P_a , as it is reasonably flexible in allowing market-specific directed network for competition and efficiently uses all variations across markets. We also add the network links in the parametrization of CCPs to estimate the CCPs, these network links, and the other selection and payoff parameters in one step—instead of first estimating the CCPs and then applying our identification argument to estimate \mathcal{N}_a . In particular, we construct from the data a state vector $r_{t_k} = (r_{at_k})_{a \in \mathcal{A}}$, where r_{at_k} indicates whether there is a change in the number of restaurants of firm f in market m at time t_k , i.e., $r_{at_k} = 2\mathbb{1}(N_{at_{k+1}} > N_{at_k}) + \mathbb{1}(N_{at_{k+1}} = N_{at_k})$. The probability of observing r_{t_k} , given the data and model parameters θ conditional on an alarm clock going off, is

$$p(r_{t_k}, S_{t_k}, N_{t_k}; \theta) = \prod_{v \in \{0,1,2\}} \prod_{a: r_{at_k} = v} \Pr_a(v \mid (N_{a't})_{a' \in \mathcal{N}_a \cup \{a\}}, S_{mt}; \theta).$$

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

$$p_0(S_{t_k}, N_{t_k}; \theta) = \prod_{a \in \mathcal{A}} \Pr_a(1 \mid (N_{a't})_{a' \in \mathcal{N}_a \cup \{a\}}, S_{mt}; \theta).$$

Finally, given that the arrival process is exponential, the log-likelihood of observing the data given θ and normalizing $\lambda_a = 1$ (the choice of “doing nothing” is not observed) is

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

The maximum likelihood estimator of θ , $\hat{\theta}$, is defined as the maximizer of \hat{L} on the parameter space Θ .

Since the network might be directed (e.g., Wallace ignores KFC but KFC pays attention to Wallace in market m), a no restriction network has $(5^2 - 5) \times 281 = 5,620$ links. That is, there are

$2^{5,620}$ possible networks. Checking all possible network structures is not feasible in our application. We use a variation of a greedy algorithm. In step 1 of the algorithm, we start from the network where the link between two firms in a given market is set to zero if and only if at least one of these firms never opened or closed a restaurant in the market. In step 2, we conduct optimization with respect to the rest of the parameters. In step 3, we set one link at a time to zero and find the one link which shutdown leads to the largest improvement in the objective function value. We repeat this step until no link removal leads to an improvement. In step 4, we optimize with respect to the rest of the parameters. Next we alternate between steps 3 and 4 until there are no improvements in the objective function value.

The estimator $\hat{\theta}$ leads to an estimator of CCPs

$$\hat{P}_a(v | N_t, S_t) = \Pr_a(v | (N_{a't})_{a' \in \mathcal{N}_a \cup \{a\}}, S_{mt}; \hat{\theta}).$$

The construction of confidence sets for the parameters and estimated CCPs would require taking into account the estimation error in the estimated network. We leave this difficult problem for future research. An important feature stemming from Assumptions 1 and 2, as shown by Proposition 3.1, is that our model has a single equilibrium or invariant distribution, i.e., the multiplicity of equilibria in the data-generating process is not an issue in our estimation.

Parameterization We assume that F_ε and F_ε are Logistic c.d.f. Given that our sample size is small relative to the number of agents—there are 4020 different dates in which we observed firms opening or closing a restaurant¹¹ for 281×5 agents—we use the following second-degree polynomial parameterization to flexibly approximate mean marginal profits and mean attention index:

$$\begin{aligned} \bar{\pi}_f(N_{at}, N_{a't}, S_{mt}; \theta) = & S'_{mt} \tilde{\beta}_f + \tilde{\alpha}_{1f} \log(1 + N_{(f,m)t}) + \tilde{\alpha}_{2f} \log(1 + N_{(f,m)t})^2 \\ & + \tilde{\alpha}_{3f} \log(1 + N_{(f',m)t}) + \tilde{\alpha}_{4f} \log(1 + N_{(f',m)t})^2 \\ & + \tilde{\alpha}_{5f} \left[\log(1 + N_{(f,m)t}) - \log(1 + N_{(f',m)t}) \right]^2, \end{aligned}$$

¹¹For some days, we observe multiple agents opening or closing. Our identification results are still valid if agents have perfectly synchronized clocks.

$$\begin{aligned} \bar{\pi}_f(N_{at}, \sum_{a' \in \mathcal{N}} N_{a't}, S_{mt}; \theta) = & S'_{mt} \beta_f + \alpha_{1f} \log(1 + N_{(f,m)t}) + \alpha_{2f} \log(1 + N_{(f,m)t})^2 \\ & + \alpha_{3f} \log(1 + \sum_{a' \in \mathcal{N}} N_{a't}). \end{aligned}$$

The mean marginal profit has two parts: the first one captures the impact of the observable market characteristics, and the second part captures the impact of the own number of restaurants and the number of restaurants of considered competitors.¹²

Remark 2. Since we use the total number of restaurants opened by each firm in every market as the determinant of selection and expansion probabilities, formally we have a model with infinite history dependence. This, however, does not constitute any issues in our application, since the mean attention and mean marginal profits take known parametric forms.

5.3. Evidence of Limited Attention to Competitors

To provide direct evidence that firms do not pay attention to all competitors in every market, we estimate a series of logistic regressions using the daily store-opening data.

Specification For each focal firm f operating in market m , we construct a binary outcome $y_{(f,m)t} \in \{0, 1\}$ equal to one if firm i opens at least one store in market m on day t , and zero otherwise. The explanatory variables are market characteristics S_{mt} (excluding the indicator for provincial capital status), number of own and competitors restaurants $\log(1 + N_{(f,m)t})$ and $\log(1 + N_{(f',m)t})$ for firm f' present in market m .¹³ The coefficient in front of $\log(1 + N_{(f',m)t})$ captures the extent to which firm f 's opening decision responds to the cumulative presence of competitor f' .

¹²We use a second-degree polynomial in our approximation to capture potential nonmonotonicities.

¹³This specification mirrors the preference stage of the structural model: $\log(1 + N)$ transformations for store counts and the same five market covariates. The indicator for provincial capital status is excluded because it is constant within a market throughout the sample period and therefore not identified in a market-by-market regression. For three cities (Longnan, Zhongwei, Zunyi) that lack 2003 market covariates, we use the earliest available year, following the same fallback used in the structural estimation.

Sample and estimation We estimate the above logistic regression separately for each active firm-market pair, retaining pairs with at least five observed openings during the sample period. Of the remaining pairs, we first attempt estimation with the full set of five covariates; if convergence fails—which occurs in very small markets where GDP, population, and land area are nearly collinear—we re-estimate dropping land area and retaining four covariates (GDP per capita, population density, GDP, and population). The final sample comprises 702 regressions spanning 255 markets and all five chains. Competitors that are never present in a given market are excluded from that market’s regression, following the same convention used in the structural model. The 702 successful regressions are distributed unevenly across chains, reflecting differences in market penetration: Wallace (217 markets), KFC (196), Dicos (136), McDonald’s (84), and Burger King (69). The lower coverage for McDonald’s and Burger King is consistent with their later and more geographically selective entry relative to the domestic chains and KFC.

Joint likelihood ratio test First, we test the null hypothesis of ignoring all competitors (i.e., all coefficients in front of $\log(1 + N_{(f',m)t})$ are zeros) using a likelihood ratio test. Table 2 reports the percentage of markets in which this joint null is rejected at the 5% level.

Table 2 – Joint likelihood ratio test: percentage of markets rejecting H_0 (all competitor coefficients zero) at the 5% level

Firm	% markets rejecting	# markets
Wallace	37.0%	216
KFC	38.3%	196
McDonald’s	40.5%	84
Burger King	35.3%	68
Dicos	38.0%	137

This null is not rejected in 60-65% of markets, depending on the chain. That is, across the majority of markets, a firm’s opening decisions show no statistically significant response to any competitor. However, in 35-40% of markets, there is still response to actions of competitors. Under our assumption that the firms’s parameters do not vary across markets, this indicates that the actions of competitors have effect on marginal profits and the observed variation is coming from limited attention.

Pairwise rejection rates Next, we analyze the results for individual coefficient estimates. Table 3 reports, for each pair of firms, the percentage of markets in which the null hypothesis that the competitor coefficient is zero is rejected at the 5% level.

Table 3 – Percentage of markets with p -value ≤ 0.05 for competitor coefficient (row = focal firm, column = competitor)

	Wallace	KFC	McDonald’s	Burger King	Dicos
Wallace	—	16.4%	16.7%	15.2%	9.6%
KFC	16.1%	—	16.4%	14.5%	15.2%
McDonald’s	6.0%	25.0%	—	13.7%	5.3%
Burger King	9.1%	7.4%	12.5%	—	13.4%
Dicos	14.4%	18.7%	9.1%	5.0%	—

Several patterns are consistent with limited attention. McDonald’s responds significantly to Dicos in only 5.3% of markets, and Wallace responds to Dicos in only 9.6% of markets, suggesting that these chains largely ignore Dicos. Similarly, Dicos responds to Burger King in only 5.0% of markets, and Burger King responds to KFC in only 7.4%. By contrast, KFC and McDonald’s respond to each other in roughly 16-25% of markets.

Overall, the results in this section indicate two important observations. First, the actions of competitors affect expansion decisions in a sizable fraction of markets. Second, firms seem to display limited attention to actions of their competitors in many markets. The objective of the structural model that we estimate is to separate these two channels affecting expansion decisions.

5.4. Estimation Results for the Main Model

Structure of the Estimated Selection Network We begin by summarizing the structure of original and the estimated firm-to-firm selection network.

The initial network that takes into account only firms that were active in the market had 3,874 links. Figure 1 depicts the percentage of markets as a function of the number of firms active in them. There is substantial heterogeneity in the number of firms present in the markets.

The estimated selection network has 1,547 directed links. So we use the likelihood value to close down 2,327 links. Figure 2 shows the distribution of the number of such selection links across

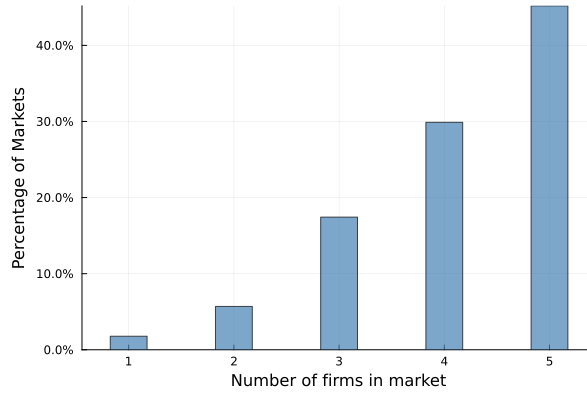


Figure 1 – Distribution of markets by number of firms present.

the markets where all 5 firms are present (there are 127 such markets). Given that the maximal number of links in such a market is 20, the networks for markets with all 5 firms present is rather sparse with the mean across market of 7.82 (standard deviation is 2.67).

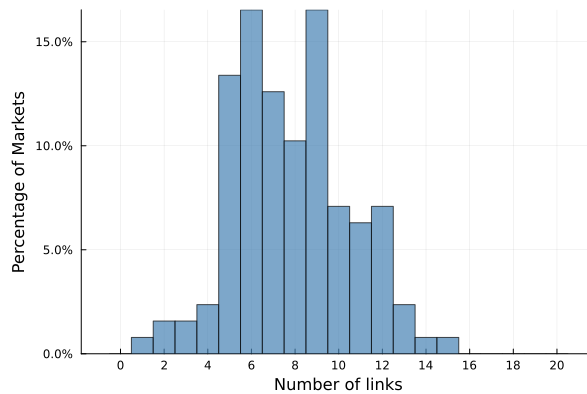


Figure 2 – Distribution of the number of selection links in markets where all five firms are present. Maximum possible is 20.

Next, we study if there is any relation between the size of the market and how attentive firms are. Figure 3 plots, for each market with at least two firms, the total number of stores at the end of the data (horizontal axis) against the ratio of the number of selection links in that market to the maximum possible (which equals $k(k - 1)$ when k firms are present). Larger markets tend to have more links in levels but the link ratio varies widely.

Next, we plot the distribution of the number of competitors considered with positive probability in markets where the firm operates. In the *initial-guess network* we assumed that every firm considers every other firm that has operated in it; the *estimated network* counts competitors based in estimated \mathcal{N}_a . Figure 4 shows both (top row: initial-guess; bottom row: estimated). The shift



Figure 3 – Total number of stores in the market (end of data) vs. ratio of selection links to maximum possible links in the market. Each point is a market with at least two firms present.

from initial to estimated reflects limited attention: firms often consider fewer competitors than the present in the market.

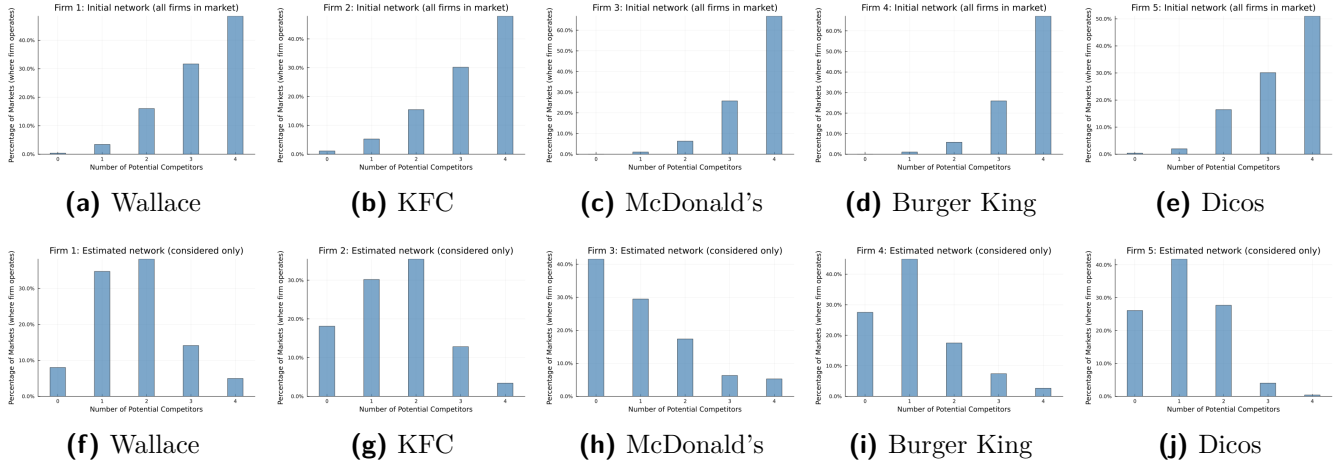


Figure 4 – The distribution of the number of competitors considered with positive probability across markets in the initial-guess network (top row) and the estimated network (bottom row).

Selection of Competitors We calculated the selection probabilities for every market and each directed firm pair (e.g, the probability of Wallace selecting KFC in Beijing) using the numbers of restaurants and covariate values observed at the end of the measurements used in estimation (December 31, 2019). Figures 5 - 8 depict the distribution of these probabilities across all markets where these firms are present (blue) and conditional on the target firm being considered (orange) for all firms except McDonald’s since its selection probabilities are almost everywhere equal to zero. These figures show that firms ignore many of their competitors in many markets. However, conditional on selecting a firm with positive probability (i.e., $a' \in \mathcal{N}_a$), all firms except McDonalds

display huge heterogeneity in paying attention to competitors. Overall, firms display striking differences: McDonald's is essentially not considering anyone as competitors, while other firms display reasonable attention with their average selection probabilities conditional on considering a firm varying in the range of 0.1-0.4 in 2019.¹⁴

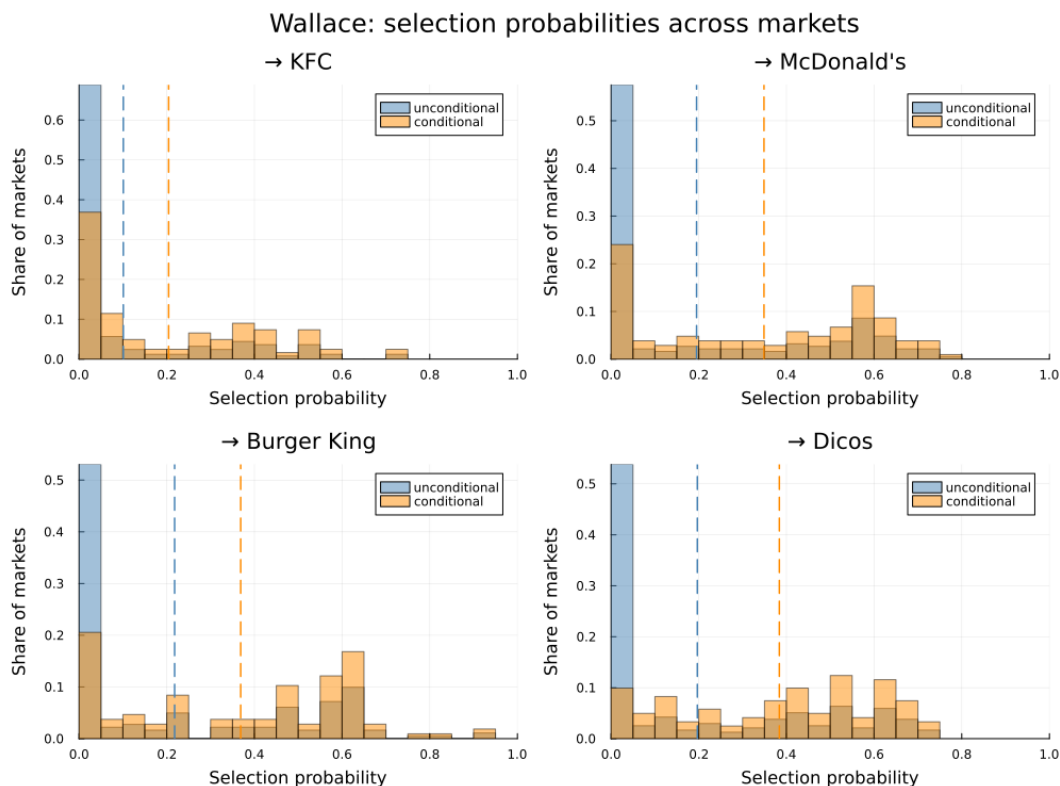


Figure 5 – Distribution of selection probabilities for Wallace across markets. Each panel shows the histogram for one competitor. Blue (unconditional): all markets where both firms are active, with ignored pairs counted as zero. Orange (conditional): markets where the competitor is not ignored. Dashed vertical lines mark the respective means.

Marginal Profits Estimates We analyze the probabilities of opening a new restaurant across 50 largest in terms of the number of restaurants markets. To quantify the effect of adding limited selection of competitors, we also estimated the marginal profit parameters assuming that all competitors that ever operated in the market are considered. We refer to the former as limited selection estimates and to the latter as full selection estimates. The results of the estimation of opening probabilities are presented in Figure 9. Table 4 reports different statistics for these scatter

¹⁴There are could be other factors affecting firms' expansion decisions that are not captured by the model. For example, a firm may operate under a different corporate strategy.

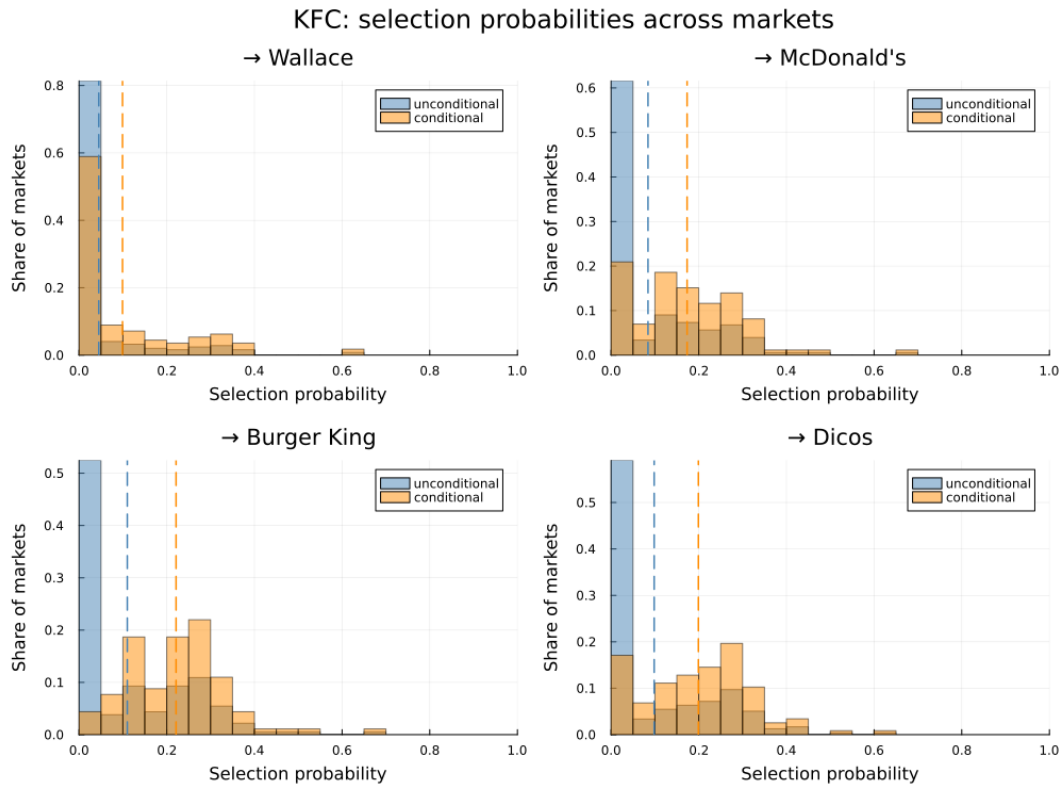


Figure 6 – Distribution of selection probabilities for KFC across markets. See Figure 5 for details.

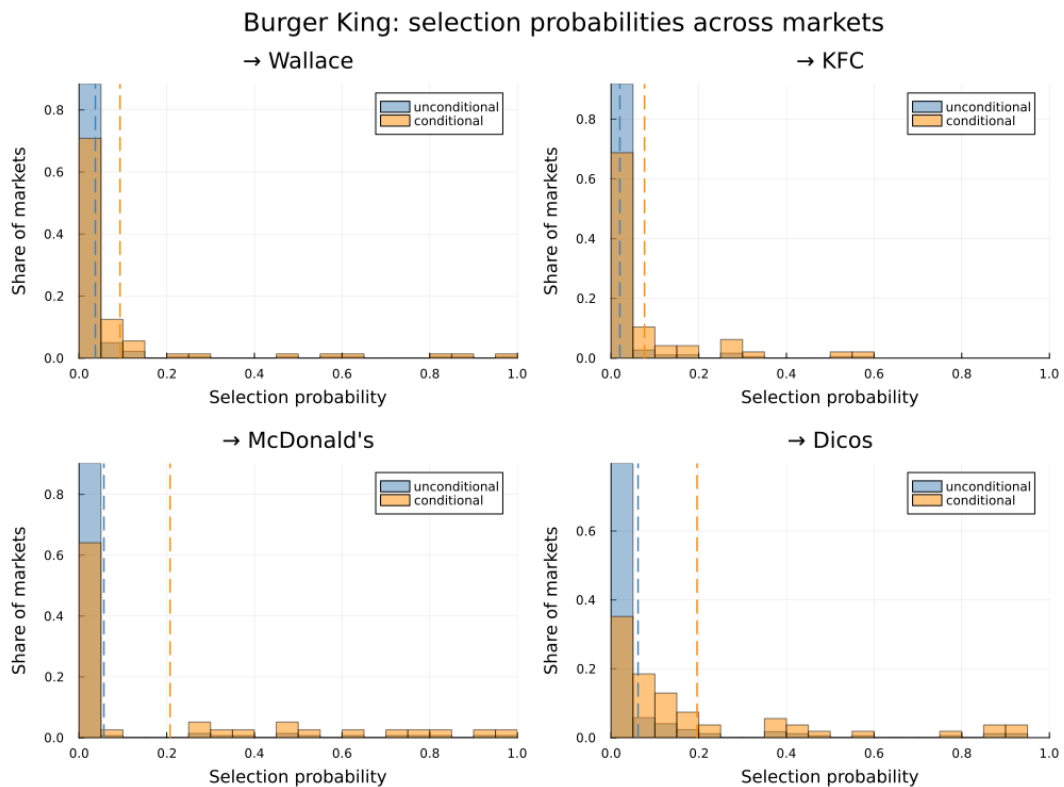


Figure 7 – Distribution of selection probabilities for Burger King across markets. See Figure 5 for details.

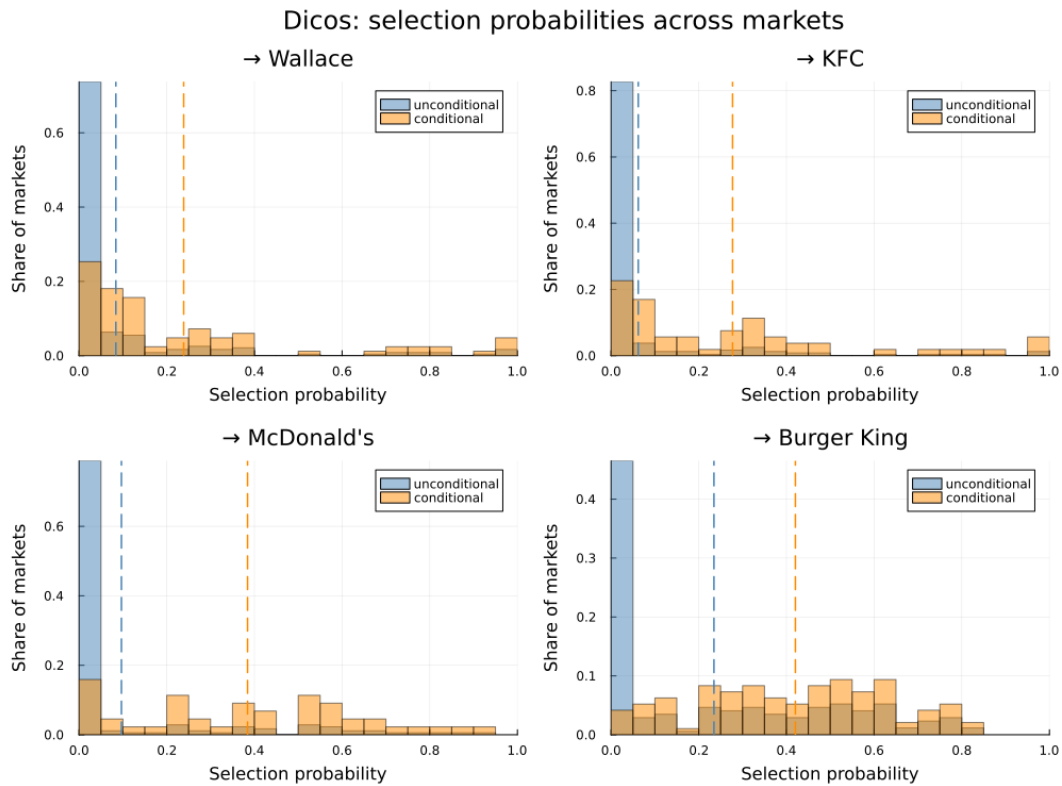


Figure 8 – Distribution of selection probabilities for Dicos across markets. See Figure 5 for details.

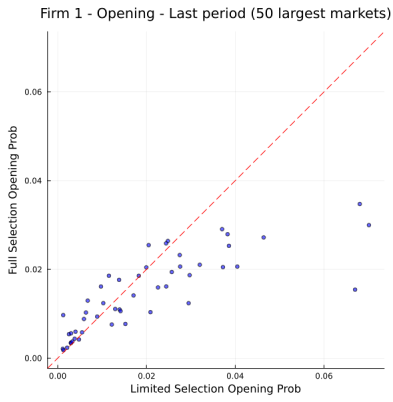
plots: the difference in average across markets probabilities, the slope of regressing full selection probabilities on limited selection probabilities and a constant, sample correlation, and the share of markets above 45-degree line. The average difference captures the average level shift induced by forcing firms to pay attention to all competitors. If it has a positive sign, then full selection increases opening probabilities on average, while a negative sign indicates the opposite. The slope measures how strongly full selection responds to cross market variation in limited selection probabilities. A slope close to one implies that market rankings (in terms of opening probabilities) are invariant and differences are mainly level shifts. A slope well below one indicates attenuation: full selection compresses dispersion, pulling high probability markets down and reducing cross market differences. The correlation coefficient captures whether the ranking of markets changes. A correlation close to one means that limited and full selection rank markets similarly, while lower correlation indicates reordering. Finally, the share of markets above the 45 degree line shows in how many markets full selection predicts more expansion than limited selection, providing a simple directional measure of competitive pressure.

The values of these statistics display several interesting patterns across firms. First, limited attention systematically inflates expansion probabilities since for all firms, the mean difference between full and limited attention is negative. McDonald’s continues to display the strongest alignment between the two specifications, with a high correlation of 0.98 and a slope close to 0.80, implying that full attention largely preserves the ranking of markets while reducing expansion levels proportionally. Burger King exhibits small attenuation with small reordering. Wallace and KFC display stronger compression effects, with slopes of 0.40 and 0.32, respectively. This means that full attention substantially reduces cross market dispersion in opening probabilities. However, correlations above 0.80 suggest that market rankings remain mostly the same. In stark contrast, Dicos has a slope of only 0.12 and a lower correlation of 0.55. This indicates both severe compression and reordering of market attractiveness under full attention. Overall, the estimates show that limited attention generates systematic overexpansion for all firms, but the magnitude and nature of the distortions vary: for most firms it primarily shifts levels and attenuates dispersion, while for Dicos it fundamentally reshapes cross market responsiveness.

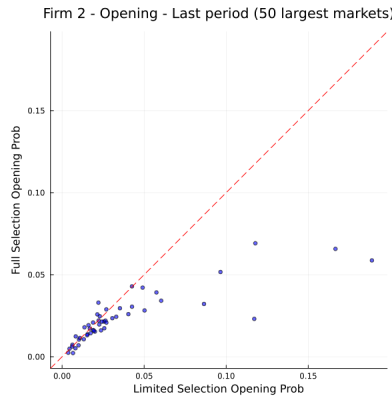
Table 4 – Opening Probabilities: Full vs Limited selection (Top 50 Markets)

Firm	$Average_{Full} - Average_{Limited}$	Slope	Correlation	Share Above 45°
Wallace	-0.00514	0.3981	0.8034	0.5
KFC	-0.01283	0.3223	0.8441	0.3061
McDonald’s	-0.00266	0.7946	0.9842	0.2979
Burger King	-0.00284	0.6465	0.9645	0.4583
Dicos	-0.00402	0.1208	0.5500	0.102

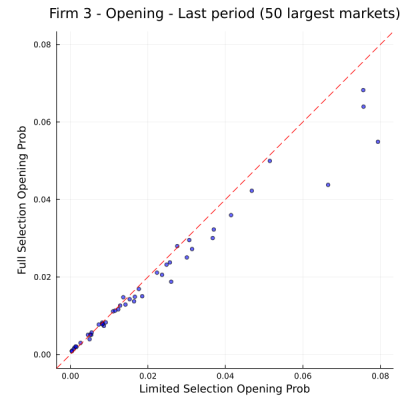
Taken together, the estimates suggest that heterogeneity in selection networks translates directly into heterogeneity in how limited selection distorts expansion behavior. For firms that do not pay much attention to their competitors, limited selection primarily shifts levels. For firms with large attention networks, limited selection alters both levels and cross market responsiveness. This findings emphasize the importance of limited selection and peer group heterogeneity in shaping expansion dynamics.



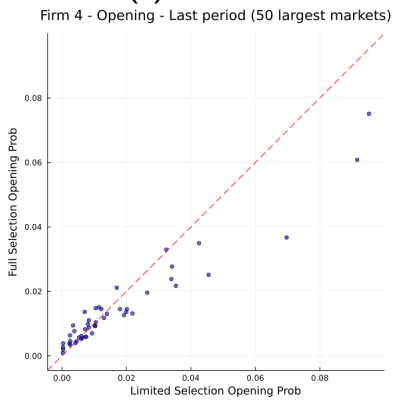
(a) Wallace



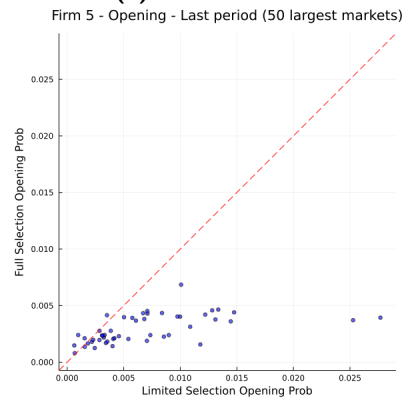
(b) KFC



(c) McDonald's



(d) Burger King



(e) Dicos

Figure 9 – Opening probability: limited vs. full selection, 50 largest markets, last period.

6. Concluding Remarks

This paper studies dynamic interactions among agents that are connected through a social network. In the model, each agent is linked to a finite set of agents and selects an option from a finite set of alternatives. The timing of the decision making of each group member follows a simple Poisson process that is independent across the agents. The distinctive feature of the model is that, at the moment of making a decision, the agent does not pay attention to all her linked agents. Instead, she first selects a subset of peers and then makes a decision under their influence. The previous choices of the linked agents affect both the probability that different agents are included in her reference group and the preferences of the agent over the alternatives once the group is formed. This model can lead to choice-based homophily in decision-making.

We exploit variation in the choices of all agents through time and variation in the size of their reference groups to recover all parts of the model. These parts include random preferences and the probability of paying attention to different agents.

The application of our framework to the analysis of the opening and closing decisions of the five largest fast-food restaurants in China shows that heterogeneity in peer selection networks translates into systematic differences in expansion dynamics between firms. Limited selection generates excess adjustment and, for some firms, substantial distortion in cross market responsiveness, emphasizing the economic importance of modeling endogenous and directed attention in competitive environments.

References

- Aguiar, Victor H, Maria Jose Boccoardi, Nail Kashaev, and Jeongbin Kim (2023) “Random utility and limited consideration,” *Quantitative Economics*, 14 (1), 71–116.
- Almunia, Miguel, Jonas Hjort, Justine Knebelmann, and Lin Tian (2024) “Strategic or confused

- firms? evidence from “missing” transactions in Uganda,” *Review of Economics and Statistics*, 106 (1), 256–265.
- 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,” Working paper.
- Badev, Anton I. (2021) “Nash equilibria on (un)stable networks,” *Econometrica*, 89 (3), 1263–1294.
- Bala, Venkatesh and Sanjeev Goyal (1998) “Learning from neighbours,” *The Review of Economic Studies*, 65 (3), 595–621.
- Blevins, Jason R (2017) “Identifying restrictions for finite parameter continuous time models with discrete time data,” *Econometric Theory*, 33 (3), 739–754.
- (2026) “Identification and estimation of continuous time dynamic discrete choice games,” *Quantitative Economics*, 17 (1), 254–296.
- Bloom, Nicholas and John Van Reenen (2007) “Measuring and explaining management practices across firms and countries,” *The Quarterly Journal of Economics*, 122 (4), 1351–1408.
- 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.
- Boucher, Vincent, Antoine Dedewanou, and Arnaud Dufays (2022) “Peer-induced beliefs regarding college participation,” *Economics of Education Review*, 90, 102307.
- Boucher, Vincent and Ismael Mourifie (2017) “My friend far, far away: a random field approach to exponential random graph models,” *The econometrics journal*, 20 (3), S14–S46.
- Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin (2020) “Peer effects in networks: A survey,” *Annual Review of Economics*, 12, 603–629.

- Chambers, Christopher P, Tugce Cuhadaroglu, and Yusufcan Masatlioglu (2025) “Revealed Social Network,” Working paper.
- Chandrasekhar, Arun G (2016) “Econometrics of network formation,” Working paper.
- Cho, Sungjin and John Rust (2010) “The flat rental puzzle,” *The Review of Economic Studies*, 77 (2), 560–594.
- De Groot, Morris (1974) “Reaching a consensus,” *Journal of the American Statistical Association*, 69 (345), 118–121.
- 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.
- DellaVigna, Stefano and Matthew Gentzkow (2019) “Uniform pricing in us retail chains,” *The Quarterly Journal of Economics*, 134 (4), 2011–2084.
- Dube, Arindrajit, Alan Manning, and Suresh Naidu (2025) “Monopsony and employer misoptimization explain why wages bunch at round numbers,” *American Economic Review*, 115 (8), 2689–2721.
- Duffie, Darrel, Nicolae Garleanu, and Lasse Heje Pedersen (2005) “Over-the-counter markets,” *Econometrica*, 73 (5), 1815–1847.
- Durlauf, Steven N and H Peyton Young (2001) *Social dynamics*, 4: Mit Press.
- Esponda, Ignacio and Damian Pouzo (2016) “Berk-Nash equilibrium: A framework for modeling agents with misspecified models,” *Econometrica*, 84 (3), 1093–1130.
- Galdon-Sanchez, Jose Enrique, Ricard Gil, and Guillermo Uriz-Uharte (2025) “The Value of Information in Competitive Markets: Evidence from Small and Medium-Sized Enterprises,” *Journal of Political Economy*, 133 (1), 252–305.
- Goldfarb, Avi and Mo Xiao (2011) “Who thinks about the competition: Managerial ability and strategic entry in US local telephone markets,” *American Economic Review*, 101 (7), 3130–3161.

- Golub, Benjamin and Matthew Jackson (2010) “Naive learning in social networks and the wisdom of crowds,” *American Economic Journal: Microeconomics*, 2 (1), 112–149.
- Graham, Bryan S (2015) “Methods of identification in social networks,” *Annu. Rev. Econ.*, 7 (1), 465–485.
- Gualdani, Chiara (2021) “An econometric model of network formation with an application to board interlocks between firms,” *Journal of Econometrics*, 224 (2), 345–370.
- Heidhues, Paul and Botond Köszegi (2018) “Behavioral industrial organization,” *Handbook of Behavioral Economics: Applications and Foundations 1*, 1, 517–612.
- Hjort, Jonas, Xuan Li, and Heather Sarsons (2026) “Across-country wage compression in multinationals,” *American Economic Review*, 116 (1), 52–87.
- Hortacsu, Ali and Steven L. Puller (2008) “Understanding strategic bidding in multi-unit auctions: a case study of the Texas electricity spot market,” *The RAND Journal of Economics*, 39 (1), 86–114.
- Hu, Lin, Anqi Li, and Xu Tan (2025) “Rationally inattentive echo chambers,” Technical report, working paper.
- Jackson, Matthew O (2008) *Social and economic networks*.
- Jackson, Matthew O and Wolinsky Asher (1996) “A strategic model of social and economic networks,” *Journal of Economic Theory*, 71 (1), 44–74.
- Kashaev, Nail and Natalia Lazzati (2019) “Peer effects in random consideration sets,” *arXiv preprint arXiv:1904.06742*.
- Kashaev, Nail, Natalia Lazzati, and Ruli Xiao (2026) “Peer effects in consideration and preferences,” *The Review of Economic Studies*, in press.
- Leung, Michael P (2015) *Econometric Methods for Network Data*: Stanford University.

- Lewel, Arthur, Xi Qu, and Xun Tang (2025) “Estimating social network models with link misclassification,” Working paper.
- Lin, Zhongjian and Yingyao Hu (2024) “Binary choice with misclassification and social interactions, with an application to peer effects in attitude,” *Journal of Econometrics*, 284 (1), 105551.
- Mele, Angelo (2017) “A structural model of dense network formation,” *Econometrica*, 85 (3), 825–850.
- Menzel, Konrad (2015) “Strategic network formation with many agents,” Working paper.
- Meyer, Bruce and Nikolas Mittag (2017) “Misclassification in binary choice models,” *Journal of Econometrics*, 200 (2), 295–311.
- Miyauchi, Yuhei (2016) “Structural estimation of pairwise stable networks with nonnegative externality,” *Journal of econometrics*, 195 (2), 224–235.
- Park, Jonghyun, Sudipta Sarangi, and Hector Tzavellas (2016) “Information overload under rational learning,” Working paper.
- de Paula, Aureo, Seth Richards-Shubik, and Elie Tamer (2018) “Identifying preferences in networks with bounded degree,” *Econometrica*, 86 (2), 643–667.
- Ridder, Geert and Shuyang Sheng (2020) “Two-step estimation of a strategic network formation model with clustering,” *arXiv preprint arXiv:2001.03838*.
- Sheng, Shuyang (2020) “Econometric analysis of large network formation models,” Working paper.
- Simon, Herbert A (1945) *Administrative behavior*: Macmillan.
- (1955) “A behavioral model of rational choice,” *The Quarterly Journal of Economics*, 99–118.
- Song, Wenting and Samuel Stern (2025) “Firm inattention and the efficacy of monetary policy: A text-based approach,” *Review of Economic Studies*, 92 (5), 3438–3469.

Thirkettle, Matthew (2019) “Identification and Estimation of Network Statistics with Missing Link Data,” Working paper.

Tourek, Gabriel (2022) “Targeting in tax behavior: Evidence from Rwandan firms,” *Journal of Development Economics*, 158, 102911.

A. Proofs

A.1. Proof of Proposition 3.1

For an irreducible, finite-state, continuous Markov chain, the equilibrium μ exists, is unique and has full support. Note that

$$P_a(v | \mathbf{y}) = \sum_{\mathcal{N} \subseteq \mathcal{N}_a} R^a(v | \mathbf{y}, \mathcal{N}) S^a(\mathcal{N} | \mathbf{y}, \mathcal{N}_a).$$

Note that $P_a(v | \mathbf{y})$ is a weighted average of $1 > R^a(v | \mathbf{y}, \mathcal{N}) > 0$. Thus, $1 > P_a(v | \mathbf{y}) > 0$ 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.

A.2. Proof of Proposition 4.1

Fix any two different Agents $a, a' \in \mathcal{A}$. Let $\mathbf{y}_{a'}^v$ be the choice configuration obtained from \mathbf{y} by replacing the a' component, $y_{a'}$, by v . First, note that if $a' \notin \mathcal{N}_a$ then

$$P_a(v | \mathbf{0}_{a'}^v) - P_a(v | \mathbf{0}) = 0,$$

where $\mathbf{0} = (0, \dots, 0)'$ is the choice configuration where everyone picks 0. We next show that if the previous difference in CCPs is different from zero, then $a' \in \mathcal{N}_a$. Note that if $a' \in \mathcal{N}_a$

$$\begin{aligned} P_a(v | \mathbf{y}) &= Q^a(a' | \mathbf{y}) \sum_{\mathcal{N} \subseteq \mathcal{N}_a \setminus \{a'\}} R^a(v | \mathbf{y}, \mathcal{N} \cup \{a'\}) S^a(\mathcal{N} | \mathbf{y}, \mathcal{N}_a \setminus \{a'\}) + \\ (1 - Q^a(a' | \mathbf{y})) &\sum_{\mathcal{N} \subseteq \mathcal{N}_a \setminus \{a'\}} R^a(v | \mathbf{y}, \mathcal{N}) S^a(\mathcal{N} | \mathbf{y}, \mathcal{N}_a \setminus \{a'\}) = \\ Q^a(a' | \mathbf{y}) &\sum_{\mathcal{N} \subseteq \mathcal{N}_a \setminus \{a'\}} [R^a(v | \mathbf{y}, \mathcal{N} \cup \{a'\}) - R^a(v | \mathbf{y}, \mathcal{N})] S^a(\mathcal{N} | \mathbf{y}, \mathcal{N}_a \setminus \{a'\}) + \\ &\sum_{\mathcal{N} \subseteq \mathcal{N}_a \setminus \{a'\}} R^a(v | \mathbf{y}, \mathcal{N}) S^a(\mathcal{N} | \mathbf{y}, \mathcal{N}_a \setminus \{a'\}). \end{aligned}$$

Moreover, by Assumption 2(ii), $R^a(v | \mathbf{y}, \mathcal{N})$ does not depend on $y_{a'}$ for any \mathcal{N} that does not contain a' . Similarly, by Assumption 2(i), $S^a(\mathcal{N} | \mathbf{y}, \mathcal{N}_a \setminus \{a'\})$ does not depend on $y_{a'}$. Hence,

$$\begin{aligned} P_a(v | \mathbf{0}_{a'}^v) - P_a(v | \mathbf{0}) &= \\ Q^a(a' | \mathbf{0}_{a'}^v) &\sum_{\mathcal{N} \subseteq \mathcal{N}_a \setminus \{a'\}} [R^a(v | \mathbf{0}_{a'}^v, \mathcal{N} \cup \{a'\}) - R^a(v | \mathbf{0}_{a'}^v, \mathcal{N})] S^a(\mathcal{N} | \mathbf{0}_{a'}^v, \mathcal{N}_a \setminus \{a'\}) \\ - Q^a(a' | \mathbf{0}) &\sum_{\mathcal{N} \subseteq \mathcal{N}_a \setminus \{a'\}} [R^a(v | \mathbf{0}, \mathcal{N} \cup \{a'\}) - R^a(v | \mathbf{0}, \mathcal{N})] S^a(\mathcal{N} | \mathbf{0}, \mathcal{N}_a \setminus \{a'\}) = \\ Q^a(a' | \mathbf{0}_{a'}^v) &\sum_{\mathcal{N} \subseteq \mathcal{N}_a \setminus \{a'\}} [R^a(v | \mathbf{0}_{a'}^v, \mathcal{N} \cup \{a'\}) - R^a(v | \mathbf{0}, \mathcal{N})] S^a(\mathcal{N} | \mathbf{0}, \mathcal{N}_a \setminus \{a'\}) \\ - Q^a(a' | \mathbf{0}) &\sum_{\mathcal{N} \subseteq \mathcal{N}_a \setminus \{a'\}} [R^a(v | \mathbf{0}, \mathcal{N} \cup \{a'\}) - R^a(v | \mathbf{0}, \mathcal{N})] S^a(\mathcal{N} | \mathbf{0}, \mathcal{N}_a \setminus \{a'\}) \neq 0, \end{aligned}$$

where the last inequality follows from Assumption 3.

A.3. Proof of Proposition 4.2

Fix some type $t \in \mathcal{H}$. Since $N_t \geq 3$, there are at least 3 agents of type t with different number of peers. Let these agents be a_1 , a_2 , and a_3 and let $N_1 < N_2 < N_3$ be the number of agents in their corresponding peer reference groups. Take any three configurations \mathbf{y}^j , $j = 1, 2, 3$, such that $y_{a_1}^1 = y_{a_2}^2 = y_{a_3}^3 = v^*$ for some $v^* \in \mathcal{Y}$ and all other components are set to v' , which may coincide

with or be different from v^* . Note that

$$\mathbb{Q}^{a_1}(a' \mid \mathbf{y}^1, \mathcal{N}_{a_1}) = \mathbb{Q}^{a_2}(a'' \mid \mathbf{y}^2, \mathcal{N}_{a_2}) = \mathbb{Q}^{a_3}(a''' \mid \mathbf{y}^3, \mathcal{N}_{a_3}) = \mathbb{Q}_t(v^*, v'),$$

for any $a' \in \mathcal{N}_{a_1}$, $a'' \in \mathcal{N}_{a_2}$, and $a''' \in \mathcal{N}_{a_3}$.

Let $q = \mathbb{Q}_t(v^*, v')$ and $t(N) = 1 - (1 - q)^N$ be the probability that at least one peer out of N peers is considered. Under Assumption 5(i), since all 3 agents pick the same option and all their peers are picking the same alternative we have that

$$\begin{aligned} \mathbb{P}_{a_j}(v \mid \mathbf{y}^j) &= \mathbb{R}_t(v \mid v^*, \mathbf{0})(1 - t(N_j)) + \mathbb{R}_t(v \mid v^*, \mathbf{0}_{v'}^1)t(N_j) \\ &= \mathbb{R}_t(v \mid v^*, \mathbf{0}) + [\mathbb{R}_t(v \mid v^*, \mathbf{0}_{v'}^1) - \mathbb{R}_t(v \mid v^*, \mathbf{0})]t(N). \end{aligned}$$

Thus, given that $\mathbb{R}_t(v \mid v^*, \mathbf{0}_{v'}^1) - \mathbb{R}_t(v \mid v^*, \mathbf{0}) \neq 0$, we deduce that

$$\frac{\mathbb{P}_{a_3}(v \mid \mathbf{y}^3) - \mathbb{P}_{a_1}(v \mid \mathbf{y}^1)}{\mathbb{P}_{a_2}(v \mid \mathbf{y}^2) - \mathbb{P}_{a_1}(v \mid \mathbf{y}^1)} = \frac{t(N_3) - t(N_1)}{t(N_2) - t(N_1)}.$$

Since the left-hand-side of the last expression is observed, if we show that the right-hand-side is a known strictly monotone function of q , then we prove that q can be identified from the data. Let $x = 1 - q$, $n_2 = N_2 - N_1$, and $n_3 = N_3 - N_1$ and note that

$$f(x) = \frac{t(N_3) - t(N_1)}{t(N_2) - t(N_1)} = \frac{1 - x^{n_3}}{1 - x^{n_2}}.$$

We next show that $f'(x) > 0$ for all $x \in (0, 1)$ and $n_3 > n_2$. After some manipulation, we get

$$f'(x) = \frac{n_2 x^{n_2-1} - n_3 x^{n_3-1} + (n_3 - n_2)x^{n_3+n_2-1}}{(1 - x^{n_2})^2}.$$

Note that the denominator is strictly positive and the numerator can be written as

$$x^{n_2-1}(n_2 - n_3 x^{n_3-n_2} + (n_3 - n_2)x^{n_3}).$$

Note that

$$\tilde{f}(x) = n_2 - n_3x^{n_3-n_2} + (n_3 - n_2)x^{n_3}$$

is such

$$\tilde{f}'(x) = -n_3(n_3 - n_2)x^{n_3-n_2-1} + (n_3 - n_2)n_3x^{n_3-1} = (n_3 - n_2)n_3x^{n_3-n_2-1}[x^{n_2} - 1] < 0$$

for all $x \in (0, 1)$. Hence, \tilde{f} is strictly decreasing on $(0, 1)$ and thus $\tilde{f}(x) \geq \tilde{f}(1) = 0$ for all x . Thus, $\tilde{f}'(x) > 0$ for all $x \in (0, 1)$ and f is strictly increasing. Since, $x = 1 - q$, then q is identified from observed CCPs.

Next, we identify $t(N)$ from q and

$$R_t(v | v^*, \mathbf{0}_{v'}^1) - R_t(v | v^*, \mathbf{0}) = \frac{P_{a_3}(v | \mathbf{y}^3) - P_{a_1}(v | \mathbf{y}^1)}{t(N_3) - t(N_1)}.$$

Finally, we identify

$$R^a(v | \mathbf{y}, \emptyset) = R_t(v | v^*, \mathbf{0}) = P_{a_3}(v | \mathbf{y}^3) - \frac{P_{a_3}(v | \mathbf{y}^3) - P_{a_1}(v | \mathbf{y}^1)}{t(N_3) - t(N_1)}t(N_3)$$

and

$$R^a(v | \mathbf{y}, \{a'\}) = R_t(v | v^*, \mathbf{0}_{v'}^1) = P_{a_3}(v | \mathbf{y}^3) + \frac{P_{a_3}(v | \mathbf{y}^3) - P_{a_1}(v | \mathbf{y}^1)}{t(N_3) - t(N_1)}(1 - t(N_3))$$

for any $a' \in \mathcal{N}_a$ and \mathbf{y} such that $y_a = v^*$ and $y_{a'} = v'$. The fact that the choice of v^* , v' , and t was arbitrary completes the proof.

A.4. Proof of Proposition 4.3

Note that

$$R^a(v | \mathbf{y}, \emptyset) = \frac{e^{\alpha_{h(a),v}(y_a)}}{1 + \sum_{v' \in \mathcal{Y} \setminus \{0\}} e^{\alpha_{h(a),v'}(y_a)}}.$$

Hence, $\alpha_{h(a),v}(y_a) = \log(\mathbb{R}^a(v \mid \mathbf{y}, \emptyset)) - \log(\mathbb{R}^a(0 \mid \mathbf{y}, \emptyset))$. To identify $\beta_{h(a),v}(y_a)$, note that

$$\mathbb{R}^a(v \mid \mathbf{y}, \{a'\}) = \frac{e^{\alpha_{h(a),v}(y_a) + \beta_{h(a),v}(y_a) \mathbb{1}(y_{a'}=v)}}{\sum_{v' \in \mathcal{Y}} e^{\alpha_{h(a),v'}(y_a) + \beta_{h(a),v'}(y_a) \mathbb{1}(y_{a'}=v')}}.$$

As a result, for $v \neq 0$,

$$\beta_{h(a),v}(y_a) = \log(\mathbb{R}^a(v \mid \mathbf{y}^*, \{a'\})) - \log(\mathbb{R}^a(0 \mid \mathbf{y}^*, \{a'\})) - \alpha_{h(a),v}(y_a),$$

for any \mathbf{y}^* such that $y_{a'}^* = v$. To identify $\beta_{h(a),0}(y_a)$, we can just flip the alternatives:

$$\beta_{h(a),0}(y_a) = \log(\mathbb{R}^a(0 \mid \mathbf{y}^*, \{a'\})) - \log(\mathbb{R}^a(v'' \mid \mathbf{y}^*, \{a'\})) + \alpha_{h(a),v''}(y_a),$$

for any \mathbf{y}^* such that $y_{a'}^* = 0$.

A.5. Proof of Proposition 4.5

Consider Agent a and let $\mathbf{0}_a^{v_1}$ be a vector where all agents select the outside option, except Agent a that is assigned option v_1 . If Assumption 7 holds, then

$$\mathbb{P}_a(v \mid \mathbf{0}_a^{v_1}) = \mathbb{R}_a(v \mid v_1, \mathbf{0}).$$

That is, we recover the choice rule for the empty set of active peers. Recall that (by Proposition 4.1) we can recover \mathcal{N}_a . Consider $a' \in \mathcal{N}_a$ and let $(\mathbf{0}_a^{v_1})_{a'}^{v_2}$ be a vector where all agents select the outside option, except Agents a and a' that are assigned options v_1 and v_2 , respectively. Note that for Agent a' that is not always selected

$$\mathbb{P}_a(v \mid (\mathbf{0}_a^{v_1})_{a'}^{v_2}) = \mathbb{R}_a(v \mid v_1, \mathbf{0}_{v_2}^1) \mathbb{Q}_a(a' \mid v_1, v_2) + \mathbb{R}_a(v \mid v_1, \mathbf{0}) (1 - \mathbb{Q}_a(a' \mid v_1, v_2)).$$

Hence,

$$P_a(v | (\mathbf{0}_a^{v_1})_{a'}^{v_2}) - P_a(v | \mathbf{0}_a^{v_1}) = \left[R_a(v | v_1, \mathbf{0}_{v_2}^1) - R_a(v | v_1, \mathbf{0}) \right] Q_a(a' | v_1, v_2).$$

Thus, if for a, v_1, v_2 , the maximum selection probability is known, then

$$\max_{a' \in \mathcal{N}_a} |P_a(v | (\mathbf{0}_a^{v_1})_{a'}^{v_2}) - P_a(v | \mathbf{0}_a^{v_1})| = \left| R_a(v | v_1, \mathbf{0}_{v_2}^1) - R_a(v | v_1, \mathbf{0}) \right| \max_{a' \in \mathcal{N}_a} Q_a(a' | v_1, v_2).$$

If the minimum selection probability is known, then

$$\min_{a' \in \mathcal{N}_a} |P_a(v | (\mathbf{0}_a^{v_1})_{a'}^{v_2}) - P_a(v | \mathbf{0}_a^{v_1})| = \left| R_a(v | v_1, \mathbf{0}_{v_2}^1) - R_a(v | v_1, \mathbf{0}) \right| \min_{a' \in \mathcal{N}_a} Q_a(a' | v_1, v_2).$$

Since the selection probability is positive, we identify

$$R_a(v | v_1, \mathbf{0}_{v_2}^1) - R_a(v | v_1, \mathbf{0}).$$

Hence, we also identify

$$Q_a(a' | v_1, v_2) = \frac{P_a(v | (\mathbf{0}_a^{v_1})_{a'}^{v_2}) - P_a(v | \mathbf{0}_a^{v_1})}{R_a(v | v_1, \mathbf{0}_{v_2}^1) - R_a(v | v_1, \mathbf{0})}.$$

Since the choice of v_1, v_2 was arbitrary, we identify the selection mechanism.

To identify the choice rule, note that since

$$R_a(v | v_1, \mathbf{0}_{v_2}^1) - R_a(v | v_1, \mathbf{0}).$$

and

$$R_a(v | v_1, \mathbf{0}) = P_a(v | \mathbf{0}_a^{v_1})$$

are identified, we also identify $R_a(v | v_1, \mathbf{0}_{v_2}^1)$. Hence, we identify the choice rule for any singleton \mathcal{N} . Then, by switching one by one all other peers to any non-default alternative recursively (similar to the proof of Proposition 4.4) we can identify the choice rule for \mathcal{N} of any size.