Online Appendix for "Peer Effects in Consideration and Preferences"

Nail Kashaev Natalia Lazzati Ruli Xiao*

January 12, 2026

The main paper contains detailed sketches of the proofs of the identification results. This Online Appendix further formalizes these arguments and presents the regularity condition. The Appendix also offers a simulation and an estimation analysis for the running example (Example 1) in the main paper. It finally adds additional details and results to the empirical application.

A. Proofs

A.1. Proof of Proposition 2.1

For an irreducible, finite-state, continuous Markov chain, the equilibrium μ exists, is unique and has full support. Thus, we only need to prove that Assumptions 2(i) and 3(i) imply that the Markov chain induced by our model is irreducible. Note that

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

Assumption 2(i) implies that given any $\mathbf{y}$, any v is always considered with a positive probability by any Agent a. Assumption 3(i) then implies that any option is picked with a positive probability if

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

considered. Thus, $0 < \mathrm{P}_a(v \mid \mathbf{y}) < 1$ for all a and $\mathbf{y}$, and we can go from one configuration to the other one in less than A steps with a positive probability.

A.2. The Regularity Condition

We formally state and discuss the regularity condition needed for the identification of the network in Section 3 of the main text of the paper. The aim of this condition is to eliminate some ties that would only arise under very unlikely situations —that we describe in the paper.

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

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

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

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

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

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

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

The next assumption describes the "regularity condition".

Assumption 5 (Regularity Condition)**.** For any $a \in \mathcal{A}$,

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

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

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

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

Assumption 5(i) ensures that peer effects in consideration and preferences do not cancel out. This guarantees that peers that affect both consideration and preferences can be distinguished from those who are not in the reference group of Agent a. Assumption 5(ii) is needed to distinguish peers who affect consideration only from those who affect preference. Specifically, for consideration-only peers, the double shifts described above are always zero. The conditions in Assumption 5(ii) ensure that the double shift in the observed CCPs is nonzero for peers who affect preference in any of the three key scenarios contemplated by Assumption 5(ii).

It is worth emphasizing that the inequality in Assumption 5(i) is only required to hold for one configuration of actions and peers. Additionally, Assumption 5(ii) allows $v_1 = v_2 = v_3$, $w_1 = w_2 = w_3$, and $(\mathbf{nr}_1, \mathbf{nc}_1) = (\mathbf{nr}_2, \mathbf{nc}_2) = (\mathbf{nr}_3, \mathbf{nc}_3)$. Furthermore, as the number of peers and/or the size of the menu grow, it gets harder to violate Assumption 5. Therefore, Assumption 5 is a mild functional form restriction that is usually generically satisfied.

The following example clarifies the scope of Assumption 5.

Example 1. Suppose that

$$\mathrm{R}_a\left(v \mid nr^{\mathcal{C}}, \mathcal{C}\right) = \frac{u_v(nr_v)}{\sum_{v' \in \mathcal{C}} u_{v'}(nr_{v'})},$$

where $u_0(nr_0) = 1$ and $u_v(\cdot)$, $v \in \mathcal{Y} \setminus \{0\}$, are strictly monotone positive functions. That is, after the consideration set is formed, Agent a picks alternatives according to a logit-type rule.

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

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

In this case, Assumption 5(i) is violated if, for all admissible values of nr^v and nc^v, the following equality holds:

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

A larger peer group makes it harder for this equality to hold for all values.

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

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

So Assumption 5(i) is violated only if, for $v \in \{1, 2\}$, $\Delta_{v,v} \ln \bar{\mathrm{P}}_a(v \mid \mathbf{nr}, \mathbf{nc}) = 0$ for all admissible values of $\mathbf{nr}$ and $\mathbf{nc}$. To illustrate how strong this condition is, note that a violation at $\mathbf{nr} = \mathbf{nc} = \mathbf{0}$ would indicate that

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

for $v \in \{1, 2\}$. Also, if Nrc_a is rich enough so that it contains $\mathbf{nr} = (0, 0)$, $\mathbf{nc} = (0, 1)$, and $\mathbf{nc}' = (1, 1)$, then we can switch from $\mathrm{Q}_a\left(v' \mid 0\right)$ to $\mathrm{Q}_a\left(v' \mid 1\right)$ without changing other parameters. Hence, we should also have that

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

The last two equalities imply that

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

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

Note that Assumption 2(iii) also establishes that

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

which guarantees Assumption 5(ii) for $i = 2$. Assumption 5(ii) is a bit more general. Specifically, violations of Assumption 5(ii) for case $i = 1$ mean that the following equations hold

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

for all combinations of the choice configuration $\mathbf{nr}$ and $\mathbf{nc}$, including $\mathbf{nr} = \mathbf{0}$ and $\mathbf{nc} = \mathbf{0}$. The set of equalities increases with the size of the peer group and/or the set of alternatives, making violations harder to arise. A similar logic carries to case $i = 3$.

A.3. Proof of Proposition 3.1

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

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

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

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

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

Note that the observed CCP can be expressed as

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

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

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

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

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

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

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

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

$$\sum_{\mathcal{C}\subseteq\mathcal{Y}\setminus\{v\}} \left[\mathrm{R}_a\left(v \mid \mathbf{0}^1_v, \mathcal{C}\cup\{v\}\right) - \mathrm{R}_a\left(v \mid \mathbf{0}, \mathcal{C}\cup\{v\}\right)\right] \mathrm{C}_a\left(\mathcal{C} \mid \mathbf{0}, \mathcal{Y}\setminus\{v\}\right) \neq 0$$

by Assumption 3(iii). Hence, the only remaining possibility is $a' \in \mathcal{NCR}_a$. But the latter contradicts Assumption 5(i), since $a' \in \mathcal{NCR}_a$ would imply that the consideration peer effect offsets the preference peer effect *everywhere* over the support. The contradiction completes the proof.

A.4. Proof of Proposition 3.2

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

$$\frac{\mathrm{P}_a\left(v \mid \mathbf{y}^w_{a''}\right)}{\mathrm{P}_a\left(v \mid \mathbf{y}\right)} = \frac{\mathrm{P}_a\left(v \mid (\mathbf{y}^v_{a'})^w_{a''}\right)}{\mathrm{P}_a\left(v \mid \mathbf{y}^v_{a'}\right)}, \tag{1}$$

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

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

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

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

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

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

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

or equivalently

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

If $a'' \in \mathcal{NR}_a \setminus \mathcal{NC}_a$, then let $i = 1$. If $a'' \in \mathcal{NC}_a \setminus \mathcal{NR}_a$, then let $i = 2$. Finally, if $a'' \in \mathcal{NCR}_a$, then let $i = 3$. Take $v = v_i$, $w = w_i$, and $\mathbf{y}$ such that $y_{a'} = y_{a''} = 0$, $\mathrm{NR}^{\mathcal{Y}}_a(\mathbf{y}) = \mathbf{nr}_i$, and $\mathrm{NC}^{\mathcal{Y}}_a(\mathbf{y}) = \mathbf{nc}_i$ from Assumption 5(ii). Then

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

where the first equality follows from the exchangeability of the difference operator, the second equality follows from the definition of $\bar{\mathrm{P}}_a$, and the last inequality follows from Assumption 5(ii). So in all possible cases, Assumption 5(ii) implies that if $a' \in \mathcal{NR}_a$, then there exist v, w, and $\mathbf{y}$ with

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

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

A.5. Proof of Proposition 3.3

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

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

Thus, for any $a'' \in \mathcal{NR}_a$, by Assumption 2,

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

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

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

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

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

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

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

A.6. Proof of Proposition 3.4

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

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

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

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

Repeating this procedure, we identify

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

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

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

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

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

0. As a result,

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

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

A.7. Proof of Proposition 3.5

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

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

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

Moreover,

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

where $\mathrm{R}^*_a(v|v')$ denotes the probabilities that Agent a picks v conditional on considering v'. Since, $\mathrm{Q}_a(v'|0)t_{v'} = \mathrm{Q}_a(v'|1)$, we obtain from the above two equations that

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

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

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

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

A.8. Proof of Proposition 3.6

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

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

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

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

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

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

A.9. Proof of Proposition 3.7

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

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

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

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

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

A.10. Proof of Proposition 3.8

This proof builds on Theorem 1 of Blevins (2017) and Theorem 3 of Blevins (2026). For the present case, it follows from these two theorems, that the transition rate matrix $\mathcal{W}$ is generically identified if, in addition to the conditions in Proposition 3.8, we have that

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

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

B. Simulation and Estimation of the Running Example

This section offers Monte Carlo simulation and estimation results for Example 1 in the paper. This exercise aims to show that with a rich dataset the primitives of the model can be estimated by following our main identification strategy step-by-step. Specifically, if we have enough observations, we can first reliably and nonparametrically estimate the CCPs for each agent using a frequency estimator. Then, we can use these estimates to recover which peers are in the consideration and preference group of each agent —by following Propositions 3.1- 3.3. With the network structure being estimated for each agent separately, we can then estimate the choice probability and the consideration mechanism following Propositions 3.4- 3.6. We show later that when the dataset is not long enough one could implement a parametric version of this approach.

Simulation Design Recall that there are four agents and three alternatives (i.e., $\mathcal{A} = \{1, 2, 3, 4\}$ and $\mathcal{Y} = \{0, 1, 2\}$). The reference groups for consideration and preferences are as follows

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

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

We specify the preferences of the agents and their consideration mechanisms as follows. For all a and $v \in \{1, 2\}$,

$$\mathrm{Q}_a(v \mid n) = \begin{cases} 1/4 & \text{if } n = 0 \\ 3/4 & \text{if } n = 1 \\ 1 & \text{if } n = 2. \end{cases}$$

As in the paper, the default 0 is always considered.

The mean utility for all a, $v \in \{1, 2\}$, and $\mathcal{C}$ is

$$\bar{u}_{a,v,\mathcal{C}}(n) = \begin{cases} 3 & \text{if } n = 0 \\ 9/2 & \text{if } n = 1 \\ 5 & \text{if } n = 2. \end{cases}$$

The mean utility from the default is normalized to be 0 regardless of how many peers choose the default. We assume the payoff shocks are generated by a Type I extreme value distribution, so the choice probability has the logit form. Note that the agent's previous choice does not affect her consideration probabilities or preferences.

We can calculate the implied (population) CCPs of Agent 1 as a function of choices of Agents 2 and 3. To simplify the notation, we ignore the previous choice of Agent 4 in the choice configuration $\mathbf{y}$ as it does not affect the CCPs of Agent 1. For example, when the previous choices of Agents 2 and 3 are the default, the probability that Agent 1 selects option 1 is

$$
\begin{aligned}
\mathrm{P}_1\left(v=1 \mid (0,0)\right) &= \underbrace{\mathrm{Q}_1(1 \mid (0,0))(1-\mathrm{Q}_2(1 \mid (0,0)))}_{\text{prob of considering}\{0,1\}} \qquad \underbrace{\frac{\exp\left(3\right)}{1+\exp\left(3\right)}}_{\text{prob of choosing 1 when considers }\{0,1\}} \\
&+ \underbrace{\mathrm{Q}_1(1 \mid (0,0))\,\mathrm{Q}_2(1 \mid (0,0))}_{\text{prob of considering}\{0,1,2\}} \quad \underbrace{\frac{\exp\left(3\right)}{1+\exp\left(3\right)+\exp\left(3\right)}}_{\text{prob of choosing 1 when considers }\{0,1,2\}} \\
&\approx 1/4 \times (1-1/4) \times 0.9526 + 1/4 \times 1/4 \times 0.4879 = 0.209.
\end{aligned}
$$

Similarly, we can calculate the CCPs for other configurations of $\mathbf{y}$ and alternative 2.

Simulation Procedure We simulate the data by the following procedure. First, we simulate four different Poisson alarms with an arrival rate of 1 and record the specific time at which the alarm goes off for each of the four agents. We start by assuming that all of them have selected the default. We then simulate the choices of each agent based on the order of the Poisson alarms. We start the simulation for $t=0$ and continue until the time reaches T from which we can collect the profile of actions at different times that we indicate by $\{y_{1t}, y_{2t}, y_{3t}, y_{4t}\}_{t\leq T}$. We assume that we can observe when the agents select each alternative, including the default —Dataset 1 in the paper.

Estimation of CCPs With the simulated data, we can first estimate the CCPs of each agent by using a simple frequency estimator. We use Agent 1 to illustrate the ideas. With a slight abuse of notation, we indicate by t the time at which the alarm of Agent 1 is off and reserve t', t'', and t'''' to denote the previous time at which the alarms of Agents 2, 3, and 4 were off, respectively. Then the estimator of P_1 is

$$
\hat{\mathrm{P}}_1(v|\mathbf{y}) = \frac{\#_{\{t:t'<t,t''<t,t'''<t\}}\{a_{1t}=v, (y_{2t'}, y_{3t''}, y_{4t'''})=\mathbf{y}\}}{\#_{\{t:t'<t,t''<t,t'''<t\}}\{(y_{2t'}, y_{3t''}, y_{4t'''})=\mathbf{y}\}}.
$$

Table 1 displays the average of the estimated CCPs of Agent 1 for 1000 replications and $T=800$ (about 800 choices per agent). We also present the true (population) CCPs calculated by using the primitives of the model.

The nonparametric estimator performs reasonably well for $T=800$ observations per agent. This simple framework has four agents and three alternatives. For each agent, the configuration of the choices of the other agents can take $3^3=27$ values. Thus, one needs to estimate 27 probabilities per agent. This requirement increases exponentially with the number of agents in the model.

Estimation of the Network We estimate the network using the nonparametric estimators

Table 1 – Agent 1's CCPs

	population	estimates	estimates for different y_4 $y_4 = 0$	$y_4 = 1$	$y_4 = 2$
$\mathrm{P}_1(1\|00)$	0.209	0.208	0.207	0.207	0.212
$\mathrm{P}_1(1\|01)$	0.708	0.710	0.710	0.712	0.713
$\mathrm{P}_1(1\|02)$	0.093	0.095	0.094	0.096	0.095
$\mathrm{P}_1(1\|10)$	0.627	0.629	0.628	0.629	0.630
$\mathrm{P}_1(1\|11)$	0.944	0.944	0.944	0.945	0.943
$\mathrm{P}_1(1\|12)$	0.280	0.280	0.279	0.280	0.285
$\mathrm{P}_1(1\|20)$	0.151	0.151	0.153	0.151	0.148
$\mathrm{P}_1(1\|21)$	0.641	0.643	0.641	0.646	0.641
$\mathrm{P}_1(1\|22)$	0.045	0.045	0.046	0.044	0.045
$\mathrm{P}_1(2\|00)$	0.209	0.209	0.208	0.210	0.211
$\mathrm{P}_1(2\|01)$	0.093	0.094	0.094	0.099	0.087
$\mathrm{P}_1(2\|02)$	0.708	0.709	0.709	0.704	0.711
$\mathrm{P}_1(2\|10)$	0.151	0.149	0.152	0.145	0.146
$\mathrm{P}_1(2\|11)$	0.045	0.045	0.046	0.044	0.046
$\mathrm{P}_1(2\|12)$	0.641	0.642	0.642	0.648	0.636
$\mathrm{P}_1(2\|20)$	0.627	0.627	0.626	0.624	0.628
$\mathrm{P}_1(2\|21)$	0.280	0.279	0.282	0.279	0.278
$\mathrm{P}_1(2\|22)$	0.944	0.944	0.943	0.945	0.944

of CCPs we just described. We first recover the reference group of Agent 1 by following the identification strategy in Proposition 3.1. Specifically, we compute the difference of the ln of the probability that Agent 1 selects alternative 1 when all other agents initially select the default and the ones we obtain when we change, one by one, each other agent to alternative 1: $\Delta_a^1 \ln \mathrm{P}_1(1 \mid \mathbf{0})$, $a = 2, 3, 4$.

Note that because we work with a finite sample, we cannot directly conclude that there is a link when $\Delta_a^1 \ln \hat{\mathrm{P}}_1(1 \mid \mathbf{0}) \neq 0$. To decide whether there is a link or not, we implement a t-statistic test with a sample-size-driven critical value. Specifically, we say that a is a peer of Agent 1 if and only if $\left|\Delta_a^1 \ln \hat{\mathrm{P}}_1(1 \mid \mathbf{0})/\mathrm{std}(\Delta_a^1 \ln \hat{\mathrm{P}}_1(1 \mid \mathbf{0}))\right| > \kappa_T$, where $\kappa_T = 0.15 \ln T$ and $\mathrm{std}(A)$ is the estimated standard error of A.[1] [2] By following this criterion we calculate the frequency of correct estimates of connections for for each agent in 1000 replications. We present the results in Table 2 for different sample sizes. The results improve with the sample size, and the estimation performs reasonably well when the sample size is about 2000.

To state whether Agents 2 or 3 are consideration-only peers we implement double differences —

[1]We use the estimated CCPs to bootstrap standard errors (Kline and Tamer, 2016).

[2]Any sequence that converges to 0 slower than $1/\sqrt{T}$ will work asymptotically.

Table 2 – Percentage of correctly estimated peers of Agent 1

Agent	Peer	T=800	T=2000	T=4000	T=8000	T=10000	T=100000
2	Yes	100.0	100.0	100.0	100.0	100.0	100.0
3	Yes	100.0	100.0	100.0	100.0	100.0	100.0
4	No	66.2	75.1	79.5	81.3	82.8	91.6

as we do in the proof of Proposition 3.2. We use the population CCPs to illustrate the identification strategy in Proposition 3.2. Specifically, to determine whether Agent 2 is a consideration-only peer of Agent 1, we check the following double difference:

$$\begin{gathered}[\ln \mathrm{P}_1(1|12) - \ln \mathrm{P}_1(1|10)] - [\ln \mathrm{P}_1(1|02) - \ln \mathrm{P}_1(1|00)] \\ = \ln(0.280) - \ln(0.627) - \ln(0.093) + \ln(0.209) = 0.\end{gathered}$$

Since the difference is 0 we could conclude that Agent 2 is a consideration-only peer of Agent 1. We conduct the same analysis for Agent 3,

$$\begin{gathered}[\ln \mathrm{P}_1(1|21) - \ln \mathrm{P}_1(1|01)] - [\ln \mathrm{P}_1(1|20) - \ln \mathrm{P}_1(1|00)] \\ = \ln(0.641) - \ln(0.708) - \ln(0.151) + \ln(0.209) = 0.2256 \neq 0,\end{gathered}$$

and conclude that Agent 3 affects the preferences and maybe consideration of Agent 1.

We apply the above double differences to the estimated CCPs with the same threshold rule to address whether the peer is a consideration-only peer or not. The results of the simulation are presented in Table 3. We can see that the percentage of correct estimates of the network increases with the sample size.

Table 3 – Percentage of correctly estimated consideration-only peers of Agent 1

Agent	Consideration-only Peers	T=800	T=2000	T=4000	T=8000	T=10000	T=100000
2	Yes	67.1	72.3	78.7	79.5	81.3	92.1
3	No	37.4	49.4	56.9	74.5	78.7	100.0

We finally identify whether Agent 3 affects both preferences and consideration. We follow the identification strategy in Proposition 3.3. Instead of switching Agent 2's choice from default to alternative 2, we check the changes in Agent 1's choice when switching the choice of Agent 2, the consideration-only peer, from the default to alternative 1 in the following double difference:

$$\begin{aligned} &[\ln \mathrm{P}_1(1|11) - \ln \mathrm{P}_1(1|01)] - [\ln \mathrm{P}_1(1|10) - \ln \mathrm{P}_1(1|00)] \\ =& \ln(0.944) - \ln(0.627) - \ln(0.708) + \ln(0.209) = -0.8109 \neq 0. \end{aligned}$$

Since the result differs from 0, we can correctly conclude that Agent 3 is a consideration-preference peer of Agent 1.

The results of applying the same logic to estimated CCPs (with the same threshold rule) are presented in Table 4. We can see that the percentage of correct network estimates increases with the sample size. Moreover, even with a sample size of 800, we can correctly estimate the link in 99.8% of 1000 replicated samples.

Table 4 – Percentage of correctly estimated preference-only peers of Agent 1

Agent	Preference-only Peer	T=800	T=2000	T=4000	T=8000	T=10000	T=100000
3	No	99.8	100.0	100.0	100.0	100.0	100.0

Estimation of Consideration Mechanism Once the network structure is recovered, we can proceed to identify and estimate the consideration mechanism. First, given that Agent 2 is a consideration-only peer of Agent 1, we can switch Agent 2 first and then Agent 3 from default to alternative 1 to identify ratios of consideration probabilities of alternative 1 for different numbers of peers selecting it —following the ideas in the proof of Proposition 3.4. Specifically, we get that

$$\begin{aligned} \frac{\mathrm{Q}_1(1|1)}{\mathrm{Q}_1(1|0)} &= \frac{\mathrm{P}_1(1|10)}{\mathrm{P}_1(1|00)} = \frac{0.627}{0.209} = 3 = \underbrace{\frac{3/4}{1/4}}_{\text{the ratio from the model}} \\ \frac{\mathrm{Q}_1(1|2)}{\mathrm{Q}_1(1|1)} &= \frac{\mathrm{P}_1(1|11)}{\mathrm{P}_1(1|01)} = \frac{0.944}{0.708} = 1.3333 = \underbrace{\frac{1}{3/4}}_{\text{the ratio from the model}} . \end{aligned}$$

Assuming that when all consideration peers are picking the alternative, the alternative is considered with probability 1, i.e., $\mathrm{Q}_1(1|2) = 1$, we can fully identify and estimate the rest of the consideration probabilities. In particular,

$$\begin{aligned} \mathrm{Q}_1(1|1) =& \frac{\mathrm{P}_1(1|01)}{\mathrm{P}_1(1|11)}, \\ \mathrm{Q}_1(1|0) =& \frac{\mathrm{P}_1(1|01)}{\mathrm{P}_1(1|11)} \frac{\mathrm{P}_1(1|00)}{\mathrm{P}_1(1|10)}. \end{aligned}$$

We estimate those consideration probabilities and present the bias and the root mean squared error (RMSE) in Table 5 for different sample sizes with 1000 replications.

Table 5 – Consideration Mechanism

Consideration Probability	Performance	T=800	T=2000	T=4000
$Q_1(1\|1)$	Bias	0.0022	-0.0001	0.0002
	RMSE	0.0723	0.0292	0.0151
$Q_1(1\|0)$	Bias	0.0006	0.0037	0.0007
	RMSE	0.0674	0.0255	0.0130

Estimation of Counterfactual CCPs and Choice Rule We now follow the proof of Proposition 3.5 to identify and estimate the counterfactual CCPs. For instance, we can identify and estimate the counterfactual CCPs if we shrink the menu from $\{0,1,2\}$ to $\{0,1\}$. We denote these counterfactual CCPs as $\mathrm{P}_1^*(1|00,\{0,1\})$, where 00 denotes the previous choices of Agents 2 and 3, respectively, and $\{0,1\}$ indicates the counterfactual menu. Given the model primitives, this counterfactual choice probability is given by

$$\mathrm{P}_1^*(1|00,\{0,1\}) = \mathrm{Q}_1(1|0)\,\mathrm{R}_1(1|0,\{0,1\}) = 1/4 \times \underbrace{\frac{\exp(3)}{1+\exp(3)}}_{\text{choice prob}} \approx 0.2381.$$

The proof of Proposition 3.5 implies that

$$\mathrm{P}_1^*(1|00,\{0,1\}) = \frac{\mathrm{P}_1(1|20) - \frac{\mathrm{Q}_1(2|1)}{\mathrm{Q}_1(2|0)}\,\mathrm{P}_1(1|00)}{1 - \frac{\mathrm{Q}_1(2|1)}{\mathrm{Q}_1(2|0)}} \approx \frac{0.151 - 3\times 0.209}{1-3} = 0.2380.$$

Hence, up to a numerical error, the results agree. Given the identified and estimated $\mathrm{Q}_1(1|0)$ we can recover

$$\mathrm{R}_1(1|0,\{0,1\}) = \frac{\mathrm{P}_1(1|20) - \frac{\mathrm{P}_1(2|20)}{\mathrm{P}_1(2|00)}\,\mathrm{P}_1(1|00)}{\left(1 - \frac{\mathrm{P}_1(2|20)}{\mathrm{P}_1(2|00)}\right)\frac{\mathrm{P}_1(1|00)}{\mathrm{P}_1(1|11)}}.$$

A similar formula can be used to identify the rest of the counterfactual CCPs and the values of the choice rules for each consideration set and agent. We estimate the counterfactual CCP and the choice rules and present the bias and RMSE in Table 6 for different sample sizes in 1000 replications.

Table 6 – Counterfactual CCP and Choice Rule

Probability	Performance	T=800	T=2000	T=4000
$P_1^*(1\|00,\{0,1\})$	Bias	-0.0041	0.0029	0.0007
	RMSE	0.0713	0.0278	0.0134
$R_1(1\|0,\{0,1\})$	Bias	-0.0153	-0.0004	0.0022
	RMSE	0.1949	0.0732	0.0368

C. The Empirical Application

This appendix offers supplemental material for the empirical application. First, we provide details for the data we use for estimation. Second, we provide extra details about the network estimation we use in the main text. Lastly, we conduct a robustness check of our empirical findings where we allow own stores in nearby markets to affect payoffs.

C.1. Data

We purchased the tea chain expansion data from CnOpenData, a data marketing company that scraps all the registration data from the National Enterprise Credit Information Publicity system. This system, which is an information-query platform for all types of enterprises (market entities) in the People's Republic of China, was launched by the State Administration for Industry and Commerce of the People's Republic of China in February 2014. Users can employ it to search for enterprise registration and filing details, license approvals, administrative penalties, records of abnormal business operations, and other related information.

Enterprises are required to register at the local (city/district) Administration for Industry and Commerce. For each new store, the enterprise has to register and provide the required information to obtain the approval for operation —the required information includes the specific street location of the store. The entry dates are the registration dates. If a store is closed, the enterprise is required by law to update this information and the date of the change of status is recorded. The overall framework, document standards, and processing time frames are unified nationwide by law, but individual service windows may apply slightly different procedures.

The market characteristics, including population, Gross Regional Product (GDP) of the city and area, were mostly collected from the China City Statistics Yearbook with two exceptions. First, the yearbook of 2016 - 2020 reported the registered population in the city in 2015-2019, but China conducted its seventh national population census in 2020, so the population reported in the Yearbook of 2021 is the resident population instead of the registered population in 2020. Second, the Yearbook of 2018 only reported GDP in the Districts under City in 2017, excluding the suburban area. We supplemented the missing of the total city GDP in 2017 and the registered population in 2020 through the China Economic and Social Big Data Research Platform, which is a large-scale, integrated statistical database that aggregates China's official economic and social development data from 1949 to the present. It brings together all central-, provincial-, and major municipal-level statistical yearbooks, as well as census reports, survey results and historical statistical compilations, covering 32 sectors and industries of the Chinese economy and society (`https://data.oversea.cnki.net/`). We also used the Consumer Price Index from the National Bureau of Statistics of China to convert the GDP into real terms (`https://data.stats.gov.cn/english/easyquery.htm?cn=C01`).

C.2. Estimation Details for the Network Structure

The vector of parameters θ consists of two parts: the parameters of the network structure (i.e., $\mathcal{NC}_a$, $a \in \mathcal{A}$), and the parameters of the attention index and marginal profits. To maximize the likelihood value by searching θ in its parameter space, we can proceed as follows: In the inner loop, fixing the network structure, we maximize the likelihood function over the consideration and payoff parameters by using the profiled likelihood estimation. The outer loop then searches for the network structure that leads to the highest likelihood.

The set of parameters in the inner loop —attention and payoff set of parameters— is rather standard and does not pose any particular challenge. Unfortunately, checking all possible network structures is often computationally prohibitive without some extra restrictions. In our application, the parameter space for $(\mathcal{NC}_a)_{a\in\mathcal{A}}$, consists of $2^{2\times 71\times(71-1)} = 2^{9940} > 10^{2400}$ possible network structures (i.e., there are $2 \times 71 \times (71-1)$ binary variables). Even if we impose full symmetry, then the size of the parameter space drops to $2^{71\times(71-1)/2} = 2^{2485}$. Instead, to simplify the estimation, we use spatial information about markets. In particular, we assume that if market m' is in the neighborhood of market m, then at least one of the following three conditions holds: m' and m are in the same province; the prefectures where m' and m are located share a border; and/or m' is at least the 5-th closest (in terms of geographical distance) market to market m.

With these additional constraints, the number of binary parameters describing the network structure is 563. Searching through all possible network structures still involves 2^{563} possibilities. To further facilitate the empirical analysis, instead of searching every possible network, we start the search from the initial/largest possible network and then shut down one link at a time to find the best improvement of the likelihood. We repeat this procedure until no link shutdown leads to any improvement.[3] Though this method is only guaranteed to converge to a local optimum, we believe that, in our application, it provides an informative approximation of the solution.

C.3. Robustness Checks: Own Stores in Nearby Markets Affecting Profits

In our benchmark analysis, we assume that only own and rival stores in the focal market can affect the payoff of the firms in the focal market. This rules out spillover effects across different markets due to various channels. A potential channel could be transportation cost saving, though, as we argue in the paper, transportation costs seem minimal in the tea industry. Another potential spillover channel could be due to information aggregation. To incorporate these potential spillover effects, we relax our benchmark assumption and allow own stores in the nearby markets to affect the firm's profitability in the focal market. We still sustain that rival's stores in the nearby markets only affect consideration. That is, the number of rival stores in nearby markets still acts as the excluded variable. We, moreover, assume that the consideration and preference networks for each market coincide. As a result, the marginal profit and the attention index, respectively, are represented by

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

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

We present the estimated consideration and expansion probabilities in Figures 1 and 2. We find that the directions of the effects are similar to the ones in the main paper. Both firms displayed limited consideration initially but became almost full consideration at the end of the measurement period because of the increased number of stores in different markets. As in the main paper, the expansion probabilities estimated under full consideration substantially underestimate the profitability of markets.

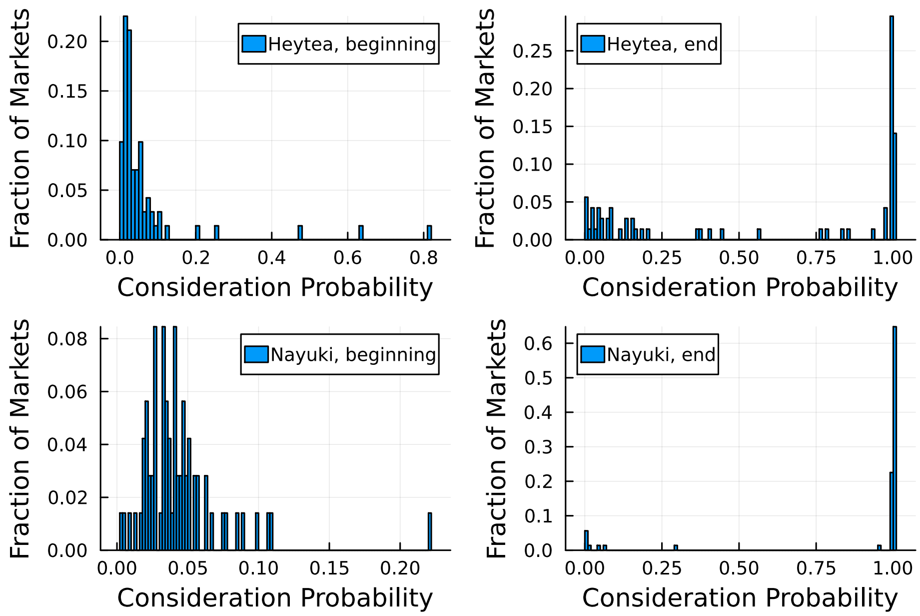

Figure 1 – Normalized histogram of consideration probabilities for both firms at the data's beginning and end.

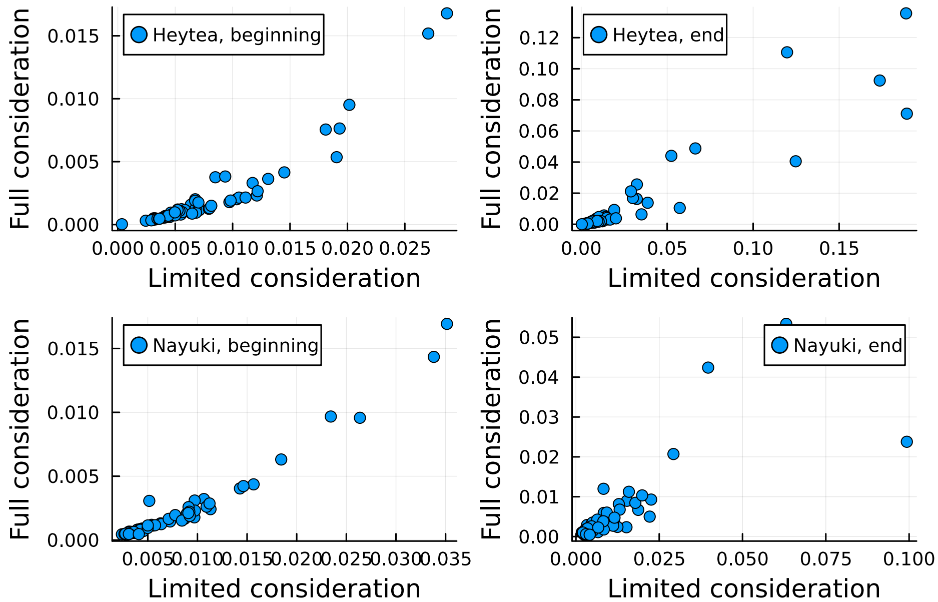

Figure 2 – Limited consideration vs. full consideration expansion probabilities.

References

Blevins, J. 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*, Forthcoming.

Kitagawa, T. and G. Wang (2023) "Individualized treatment allocation in sequential network games."

Kline, B. and E. Tamer (2016) "Bayesian inference in a class of partially identified models," *Quantitative Economics*, 7 (2), 329–366.