Peer Effects in Limited Attention*

Nail Kashaev
Western University
nkashaev@uwo.ca

Natalia Lazzati
UC Santa Cruz
nlazzati@ucsc.edu

May 13, 2026

Abstract We develop a dynamic model of discrete choice that allows peers to affect the sets of alternatives that agents pay attention to —consideration sets. We characterize the equilibrium behavior and study the empirical content of the model. In our setup, changes in the choices of friends affect the probability of considering different sets of alternatives. We exploit this variation to recover the ranking of preferences, attention mechanisms, and network connections. These nonparametric identification results allow for unrestricted heterogeneity across people and do not rely on the variation of either covariates or the set of available options. We apply our results to an experimental dataset that has been designed to study the visual focus of attention. We find robust left-to-right bias and positive peer effects in gazing.

JEL codes: C31, C33, D83, O33

Keywords: Peer Effects, Random Consideration Sets, Continuous Time Markov Process

*This paper subsumes Kashaev and Lazzati (2019). We thank Victor Aguiar, Roy Allen, Tim Conley, Steven Durlauf, Paola Manzini, Marco Mariotti, Salvador Navarro, and Al Slivinski for their useful comments and suggestions. We would also like to thank the audience at Workshop in Experiments, Revealed Preferences, and Decisions, the Inter University Colloquium on Analytical Research in Economics, the CARE-Colloquium series at University of Kansas, NAWM ES 2021, the Econometrics of Games, Matching and Network at Toulouse, Bristol ESG, and the seminar presentations at Cornell University, UC Riverside, and at El Colegio de Mexico.

1. Introduction

In recent years, the foundational rational choice model has been revised to better reflect human cognitive limitations, leading to the development of consideration set models. These models recognize that individuals do not evaluate all available options, but rather a limited subset. Despite their growing popularity, how these sets are formed and their primary determinants remain open questions in the literature. We propose a dynamic social network model that incorporates peer effects into the formation of consideration sets. While previous studies have suggested that peers influence discovery[1], we formalize this within a structural framework. We show that all components of our model can be recovered from sequential choice data. We apply our model to an experimental dataset on visual focus of attention, confirming existing beliefs regarding left-to-right scanning bias and positive peer effects in gazing. The theoretical results and main empirical findings might guide the design of online platforms and marketing strategies.

Unlike traditional models that treat peer influence as a change in utility (perceived value), our framework thinks of peers as drivers of salience, isolating the discovery effect from the persuasion effect by focusing on limited attention. In the model, individuals are connected via a social network that defines their reference groups. When making decisions, agents do not consider all available options; instead, they choose from a subset of options, with the probability of paying attention to a specific option determined by the number of friends currently adopting it. This setup applies to various fields, such as digital music platforms where users must navigate millions of songs. Services like Spotify or Pandora often create a reference group for each user and share with her the last songs selected by the other group members. This information aims to help the user to circumscribe the subset of songs to consider. The model we offer can help the platform personalize the reference group of each user to facilitate the decision process and optimize her experience.[2]

We consider a researcher who observes a long sequence of choices made by the members of

[1] In various contexts and with different set-ups, the possibility that peers might affect the consideration sets has been incorporated into the analysis. For example, Godes and Mayzlin (2004) states that the choices of peers can help a person discover a new television show and Qiu, Shi and Whinston (2018) incorporates this idea into the problem of finding a new restaurant.

[2] A similar idea has been studied by Lee (2015) for online games.

a fixed network. Each agent is endowed with a stable-over-time consideration mechanism and strict ranking over all alternatives, which we associate with preferences.[3] From this dataset the researcher can estimate the frequency of choices of each person in the network conditional on the choices of others. We refer to these frequencies as the conditional choice probabilities (CCP). We show that all parts of the model can be recovered from the CCPs. These parts include the ranking of preferences, the attention mechanism, and the network structure. The identification strategy takes advantage of the fact that changes in the choices of friends induce stochastic variation in the consideration probabilities. We use this variation to recover the set of connections among people in the network and their rankings over alternatives. We then use this information to recover the attention mechanism of each person, i.e., the probability of paying attention to a particular option as a function of the number of friends currently choosing it.

We apply our model to an experimental dataset that has been used by Bai, Kumar, Leskovec, Metzger, Nunamaker Jr and Subrahmanian (2019) and Kumar, Bai, Subrahmanian and Leskovec (2021) to compare the effectiveness of various models that aim to predict the visual focus of attention of individuals within a group. In the experiment, a group of people were asked to play a party game in which they communicated with each other to find the deceptive players. The players were seated around a circle, and a tablet, which was placed in front of each player, recorded the direction of the player's sight every third of a second. We consider each person to be looking to the left, to the tablet, or to the right. The raw frequencies (in percent) of the direction of sight of the five players in the data are as follows.

	Player 1	Player 2	Player 3	Player 4	Player 5
left	40	35	30	29	50
tablet	20	38	5	15	16
right	40	27	65	56	34

The model we offer allows us to separate two key forces in explaining the data: directional sight rankings of players and peer effects in gazing. These effects are related to well-known observations: First, visual designers create online platforms on the premise that people scan certain areas of the screen before others. In particular, it is believed that, everything else equal, people spend more

[3] This ranking informally captures all aspects of the agent's decision process that are not accounted for by attention captured by the consideration mechanism.

time looking to the left side of the screen as compared to the right side. This phenomenon is called the left-to-right bias (Spalek and Hammad, 2004, 2005). Second, in social environments, it has been observed that people automatically redirect their visual attention by following others' gaze orientation, a phenomenon called gaze following. Although the choice set is small, participants made decisions in fractions of a second, justifying limited attention. The empirical estimates we obtain are striking: Despite considerable differences in the raw frequency of the three choices across players, as shown in the table, the estimated rankings for directional sight of all players coincide and are consistent with the left-to-right bias. Also, the estimates of peer effects in the consideration sets are positive and relevant. These findings highlight the importance of incorporating peer effects in inattentive behavior and individual rankings as different channels for decision-making.

To facilitate the exposition, we present the initial results under some simplifying restrictions. In particular, we assume that there exists one alternative (the default) that is picked if and only if nothing else is considered; we let the distribution of consideration sets be multiplicatively separable across alternatives; and we let the probability of including each option depend on the number (but not the identity) of the friends that selected that option. We then extend our model in various directions to relax these assumptions: First, we analyze a model with no default option. All the primitives of this model are identified if there are more than three options in the set of available alternatives. Second, we allow the default option not to be fully dominated by all other alternatives, and its ranking to differ across people. This extension is particularly important for applications such as the online platform we described earlier. Finally, we consider a set of extensions where consideration sets are formed arbitrarily (e.g., no multiplicative separability assumption). In these extensions, different friends may have different effects on the attention mechanism of a given person. In this model, we still can uniquely recover rankings over alternatives and the network structure. The consideration probabilities, in general, are partially identified. However, we show that different forms of symmetry between consideration probabilities are sufficient for their identification.

We finally relate our results with the existing literature. The closest paper to ours is Kashaev, Lazzati and Xiao (2026). This paper models peer effects in both consideration sets and preferences, showing that these two mechanisms have different behavioral implications in the data and can be set apart. While adding two sources of peer effects, Kashaev et al. (2026) is more restrictive in other regards. In particular, it does not allow for strict rankings over alternatives, focuses on estimation with continuous-time data, and does not extend the analysis to settings with no default or more general forms of the consideration set formation mechanisms. Our work integrates the dynamic model of social interactions developed by Blume (1993, 1995) with the single-agent model of random consideration sets proposed by Manski (1977) and Manzini and Mariotti (2014). The literature on single-agent consideration set models has mainly relied on the observable variation of the set of available options (see de Clippel and Rozen, 2024 and Strzalecki, 2025 for reviews of this literature). Other papers have relied on exogenous covariates (e.g., Abaluck and Adams, 2021, Barseghyan, Molinari and Thirkettle, 2021, Crawford, Griffith and Iaria, 2021, Goeree, 2008, Roberts and Lattin, 1991). Aguiar and Kashaev (2025), Crawford et al. (2021), and Dardanoni, Manzini, Mariotti and Tyson (2020) focus on panel datasets, but do not allow for peer effects.

The literature on identification of models of social interactions where choices of peers affect preferences but not the choice sets is quite large (see Blume, Brock, Durlauf and Ioannides, 2011, Bramoullé, Djebbari and Fortin, 2020, De Paula, 2017, and Graham, 2015 for reviews of this literature). In this paper, we depart from this literature in that in our framework the direct interdependence between choices (endogenous effects) is captured by consideration sets (choice sets), not preferences.[4] We view our work as complementing the existing results on peer effects in preferences by proposing an alternative mechanism for the social interaction effects. Kashaev et al. (2026) incorporates peer effects in both consideration and preferences and show how to recover each of them. As we mentioned earlier, similar to Kashaev et al. (2026), we can identify the network structure. In the context of linear models, Blume, Brock, Durlauf and Jayaraman (2015), Bonaldi, Hortaçsu and Kastl (2015), De Paula, Rasul and Souza (2025), Lewbel, Qu and Tang (2023), and Manresa (2013) also establish network identification. Chambers, , Masatlioglu and Turansick (2022) identifies the impact that the choices of other people have on the choices of a given person, and Chambers, Cuhadaroglu and Masatlioglu (2025) identifies the network structure in the discrete-choice model where peers do not affect consideration sets.

[4]Implicitly, preferences in our model capture the so-called contextual effects – effects of predetermined factors that are not affected by peers.

Peer effects in the formation of consideration sets are also studied in Borah and Kops (2018) and Lazzati (2020). The former considers a static framework and relies on menu variation. The latter considers a discrete-time dynamic model, with the focus on two binary options that can be acquired together and display complementarities.

The remainder of the paper is organized as follows. Section 2 presents the model, the main assumptions, and some key insights of our approach. Section 3 describes the equilibrium behavior. Section 4 studies the empirical content of the model. Section 5 extends the initial idea to two variants of the default option and more general formation processes for the consideration sets. Section 6 applies our model to an experimental dataset on visual focus of attention. Section 7 concludes, and all the proofs are collected in Appendix A. Appendix B provides some simulation evidence of the finite sample performance of our estimator. Appendix C contains additional results for our empirical application.

2. The Model

This section describes the model and the main assumptions we invoke in the paper.

2.1. Social Network, Consideration Sets, and Choices

Network and Choice Configuration There is a finite set of people connected through a social network. The network is described by a graph $(\mathcal{A}, e)$, where $\mathcal{A} = \{1, 2, \ldots, A\}$ is the finite set of nodes (or people) and e is the set of edges. Each edge identifies two connected people and the direction of the connection. For each Person $a \in \mathcal{A}$ her set of friends (or reference group) is defined as follows:

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

There is a set of alternatives $\overline{\mathcal{Y}} = \mathcal{Y} \cup \{o\}$, where $\mathcal{Y} = \{1, 2, \ldots, Y\}$ is a finite set of options, and o is a default option. Each Person a has a strict preference order $\succ_a$ over the set of options $\mathcal{Y}$.

All people agree in that the default option is the least preferred. These assumptions are standard in the literature on single-agent choice with limited consideration and allow us to connect our framework with that literature. We relax the specification of the default option in Section 5.1, and consider the case in which people differ regarding the ranking in preferences of the default option in Section 5.2. We refer to a vector $\mathbf{y} = (y_a)_{a \in \mathcal{A}} \in \overline{\mathcal{Y}}^A$ as a choice configuration.

Choice Revision Similarly to Kashaev et al. (2026), we model the revision of choices as a standard continuous-time Markov process. In particular, we assume that people are endowed with independent Poisson alarm clocks with rates $\lambda = (\lambda_a)_{a \in \mathcal{A}}$.[5] At exponentially distributed with mean $1/\lambda_a$ moments the alarm of Person a goes off. When this happens, the person selects the most preferred alternative among the ones she is paying attention to. Formally, if $\mathcal{C} \subseteq \mathcal{Y}$ is the set of alternatives that the person is paying attention to, then the choice of Person a can be represented by an indicator function

$$\mathrm{R}_a (v \mid \mathcal{C}) = \mathbb{1} (v \succ_a v' \text{ for all } v' \in \mathcal{C} \text{ and } v \in \mathcal{C})$$

that takes value 1 if v is the most preferred option in $\mathcal{C}$ according to $\succ_a$ and is 0 otherwise. If Person a does not pay attention to any alternative in $\mathcal{Y}$, then she simply selects the default option.

Peer Effects in the Formation of Consideration Sets In our model, whether Person a pays attention to a particular alternative depends on her own choice and the configuration of choices of her friends at the moment of revising her selection. Let $\mathrm{N}_a^v (\mathbf{y})$ be the number of friends of Person a who select option v in choice configuration $\mathbf{y}$. Formally,

$$\mathrm{N}_a^v (\mathbf{y}) = \sum\nolimits_{a' \in \mathcal{N}_a} \mathbb{1} (y_{a'} = v) .$$

The probability that Person a pays attention to alternative v given a choice configuration $\mathbf{y}$ is captured by $\mathrm{Q}_a (v \mid y_a, \mathrm{N}_a^v (\mathbf{y}))$. We assume that consideration sets are formed by independent

[5] See Blume (1993, 1995) for theoretical models that rely on Poisson alarm clocks and Blevins (2026) for a nice discussion of the advantages of this type of revision process from an applied perspective.

considerations. That is, the probability of facing consideration set $\mathcal{C}$ is

$$\prod_{v\in\mathcal{C}} \mathrm{Q}_a\left(v \mid y_a, \mathrm{N}_a^v\left(\mathbf{y}\right)\right) \prod_{v\notin\mathcal{C}} \left(1 - \mathrm{Q}_a\left(v \mid y_a, \mathrm{N}_a^v\left(\mathbf{y}\right)\right)\right).$$

By combining preferences and random consideration sets, the probability that Person a selects (at the moment of choosing) alternative $v \in \mathcal{Y}$ is given by

$$\mathrm{P}_a\left(v \mid \mathbf{y}\right) = \mathrm{Q}_a\left(v \mid y_a, \mathrm{N}_a^v\left(\mathbf{y}\right)\right) \prod_{v'\in\mathcal{Y}, v'\succ_a v} \left(1 - \mathrm{Q}_a\left(v' \mid y_a, \mathrm{N}_a^v\left(\mathbf{y}\right)\right)\right). \tag{1}$$

The default option is assumed to be always considered. Since it is also assumed to be the worst for every person, the default option is chosen by the person if, and only if, nothing else is considered. Thus, the probability of selecting o is just $\prod_{v\in\mathcal{Y}}\left(1 - \mathrm{Q}_a\left(v \mid y_a, \mathrm{N}_a^v\left(\mathbf{y}\right)\right)\right)$. This process of formation of consideration sets is analogous to the one studied by Manski (1977) and Manzini and Mariotti (2014). In Section 5.3 we extend our analysis to a more general setting. Note that, by construction, the peer effects in our model do not involve forward-looking behavior. This restriction is not added for technical reasons. We incorporate it to capture situations where the choices of friends simply make a person more aware of specific alternatives redirecting her attention towards that option.

Altogether, the three elements just described characterize our initial model of peer effects in random consideration sets. (As mentioned above, we consider later several extensions.) We next present three key assumptions we invoke in the paper.

2.2. Main Assumptions

Our results build on one or more of the three assumptions we discuss next. In what follows, we indicate by $|\mathcal{N}_a|$ the cardinality of $\mathcal{N}_a$.

(A1) For each $a \in \mathcal{A}$, $v \in \mathcal{Y}$, and $\mathbf{y} \in \overline{\mathcal{Y}}^A$, $1 > \mathrm{Q}_a\left(v \mid y_a, \mathrm{N}_a^v\left(\mathbf{y}\right)\right) > 0$.

(A2) For each $a \in \mathcal{A}$, $|\mathcal{N}_a| > 0$.

(A3) For each $a \in \mathcal{A}$, $v \in \mathcal{Y}$, and $y_a \in \overline{\mathcal{Y}}$, $\mathrm{Q}_a\left(v \mid y_a, 0\right) \neq \mathrm{Q}_a\left(v \mid y_a, 1\right)$.

Assumption A1 states that every alternative is considered or ignored with positive probability. This assumption implies that agents can eventually pay attention to any option for reasons that are not implicitly modeled (e.g., watching an ad on television or receiving a coupon). In line with the spirit of the limited attention mechanisms, it also allows people to eventually disregard any further consideration of a given option, including the one that she is currently adopting. Assumption A1 guarantees that each subset of options is (ex-ante) considered with nonzero probability. Assumption A2 requires each person to have at least one friend. Assumption A3 states that the attention that each person pays to each alternative changes when a friend selects that alternative. This variation is required only at zero, allowing for different levels of satiation —e.g. consideration changes till the number of friends picking the option achieves some threshold (e.g., 10 friends, 20 friends, etc).[6]

Note that Assumptions A1 and A3 allow a setting where the current selection is considered with probability close to 1, thus capturing environments with inertia. Though quite general, these modeling restrictions can be further relaxed. For instance, we could incorporate dependence of the consideration sets not just on current choices, but also on past decisions —e.g., a Markov process with memory. This extension can be achieved by enlarging the state space. Note also that we assume peer effects in consideration sets are positive as this assumption is consistent with the sort of applications that motivated our work. But our results still hold —regarding existence and identification— if peer effects were negative for some people and positive for other persons. Moreover, we could be agnostic about the signs of the peer effects and recover them (separately) for each person from the data.

3. Equilibrium Behavior

This section states equilibrium existence and offers some insights on equilibrium behavior.

The Poisson model we propose, guarantees that, at each moment of time, at most one person

[6] Assumption A3 can be relaxed. The variability of consideration probabilities does not need to hold at zero or be the same for all agents, as long as it is below $|\mathcal{N}_a|$. We use the stronger assumption strictly to simplify the exposition.

revises her selection almost surely.[7] Thus, the transition rate from choice configuration $\mathbf{y}$ to any different one $\mathbf{y}'$ takes a simple form

$$\mathrm{m}\left(\mathbf{y}' \mid \mathbf{y}\right) = \begin{cases} 0 & \text{if } \sum_{a\in\mathcal{A}} \mathbb{1}\left(y'_a \neq y_a\right) > 1 \\ \sum_{a\in\mathcal{A}} \lambda_a \, \mathrm{P}_a\left(y'_a \mid \mathbf{y}\right) \mathbb{1}\left(y'_a \neq y_a\right) & \text{if } \sum_{a\in\mathcal{A}} \mathbb{1}\left(y'_a \neq y_a\right) = 1 \end{cases}. \tag{2}$$

In the literature of continuous-time Markov processes, these transition rates are the out-of-diagonal terms of the *transition rate matrix* $\mathcal{M}$ (or the infinitesimal generator matrix). The diagonal terms are simply given by $\mathrm{m}\left(\mathbf{y} \mid \mathbf{y}\right) = -\sum_{\mathbf{y}'\in\overline{\mathcal{Y}}^A\setminus\{\mathbf{y}\}} \mathrm{m}\left(\mathbf{y}' \mid \mathbf{y}\right)$. Since the the number of choice configurations is $(Y+1)^A$, $\mathcal{M}$ is a $(Y+1)^A \times (Y+1)^A$ matrix.

An equilibrium in this model is an invariant distribution $\mu : \overline{\mathcal{Y}}^A \to [0,1]$, with $\sum_{\mathbf{y}\in\overline{\mathcal{Y}}^A} \mu\left(\mathbf{y}\right) = 1$, of the dynamic process with transition rate matrix $\mathcal{M}$. This distribution satisfies

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

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

Proposition 3.1. *If Assumption A1 is satisfied, then there exists a unique equilibrium μ. Moreover, it has full support.*

The first example describes the equilibrium behavior of a very simple specification of our model.

Example 1. The network consists of two identical people who select between two alternatives, option 1 and the default option o. The rates for their Poisson alarm clocks are 1. Assume, to simplify the setup, that the probability of paying attention to a particular option only depends on the current choice of the other person (i.e., $\mathrm{Q}_a\left(v \mid y_a, \mathrm{N}^v_a\left(\mathbf{y}\right)\right) = \mathrm{Q}_a\left(v \mid \mathrm{N}^v_a\left(\mathbf{y}\right)\right)$) and is the same for both people (i.e., $\mathrm{Q}(\cdot) = \mathrm{Q}_1(\cdot) = \mathrm{Q}_2(\cdot)$). Thus, for $a = 1, 2$, we get that

$$\mathrm{P}_a\left(1 \mid \mathbf{y}\right) = \mathrm{Q}\left(1 \mid \mathrm{N}^v_a\left(\mathbf{y}\right)\right) \text{ and } \mathrm{P}_a\left(o \mid \mathbf{y}\right) = 1 - \mathrm{Q}\left(1 \mid \mathrm{N}^v_a\left(\mathbf{y}\right)\right).$$

[7] As Blevins (2017, 2026) explain, this feature of the model translates into many zeros in known locations in the infinitesimal-transition matrix. This is an advantage over discrete time models.

The transition rate matrix $\mathcal{M}$ is as follows.

	(o,o)	$(o,1)$	$(1,o)$	$(1,1)$
(o,o)	$-2\,\mathrm{Q}\,(1 \mid 0)$	$\mathrm{Q}\,(1 \mid 0)$	$\mathrm{Q}\,(1 \mid 0)$	0
$(o,1)$	$1-\mathrm{Q}\,(1 \mid 0)$	$-1+\mathrm{Q}\,(1 \mid 0)-\mathrm{Q}\,(1 \mid 1)$	0	$\mathrm{Q}\,(1 \mid 1)$
$(1,o)$	$1-\mathrm{Q}\,(1 \mid 0)$	0	$-1+\mathrm{Q}\,(1 \mid 0)-\mathrm{Q}\,(1 \mid 1)$	$\mathrm{Q}\,(1 \mid 1)$
$(1,1)$	0	$1-\mathrm{Q}\,(1 \mid 1)$	$1-\mathrm{Q}\,(1 \mid 1)$	$-2+2\,\mathrm{Q}\,(1 \mid 1)$

.

The matrix $\mathcal{M}$ is naturally more complex when there are more alternatives or more people. However, the structure of the zeros in $\mathcal{M}$ is rather similar. In particular, there are many zeros in known locations. As we mentioned earlier, this feature of the model is particularly helpful for identification.

Solving the system of equations $\mu\mathcal{M} = \mathbf{0}$, we get that the equilibrium is

$$\mu\,(o,o) = \frac{[1-\mathrm{Q}\,(1 \mid 0)]\,[1-\mathrm{Q}\,(1 \mid 1)]}{1-\mathrm{Q}\,(1 \mid 1)+\mathrm{Q}\,(1 \mid 0)}$$
$$\mu\,(o,1) = \mu\,(1,o) = \frac{\mathrm{Q}\,(1 \mid 0)\,[1-\mathrm{Q}\,(1 \mid 1)]}{1-\mathrm{Q}\,(1 \mid 1)+\mathrm{Q}\,(1 \mid 0)}$$
$$\mu\,(1,1) = \frac{\mathrm{Q}\,(1 \mid 0)\,\mathrm{Q}\,(1 \mid 1)}{1-\mathrm{Q}\,(1 \mid 1)+\mathrm{Q}\,(1 \mid 0)}.$$

The equilibrium is a joint distribution on the pair of choice configurations. It states the average fraction of time that each pair of choices (y_1, y_2) realizes in the system. ■

4. Identification

This section provides conditions under which the researcher can uniquely recover (from a long sequence of choices) the reference group of each person $\mathcal{N} = (\mathcal{N}_a)_{a\in\mathcal{A}}$, the profile of strict preferences $\succ = (\succ_a)_{a\in\mathcal{A}}$, the attention mechanism $\mathrm{Q} = (\mathrm{Q}_a)_{a\in\mathcal{A}}$, and the rates of the Poisson alarm clocks $\lambda = (\lambda_a)_{a\in\mathcal{A}}$. In Appendix B, we propose a maximum likelihood estimator of the model parameters and conduct several Monte Carlo experiments to evaluate its finite-sample properties.

We will separate the identification analysis into two parts. Recall that $\mathrm{P} = (\mathrm{P}_a)_{a\in\mathcal{A}}$ are the CCPs. Each $\mathrm{P}_a(v \mid \mathbf{y}) : \overline{\mathcal{Y}} \times \overline{\mathcal{Y}}^A \to (0,1)$ specifies the (ex-ante) probability that Person a selects option v when the choice configuration is $\mathbf{y}$. In our setting, the observable CCPs relate to the unobservable parts of the model by

$$\mathrm{P}_a(v \mid \mathbf{y}) = \mathrm{Q}_a(v \mid \mathbf{y}) \prod\nolimits_{v'\in\mathcal{Y}, v'\succ_a v} (1 - \mathrm{Q}_a(v' \mid \mathbf{y})).$$

Also, the probability of selecting the default option o is $\prod_{v\in\mathcal{Y}} (1 - \mathrm{Q}_a(v \mid \mathbf{y}))$. First, we show that each set of CCPs P maps into a different set of connections, profile of strict preferences, and attention mechanism. Thus, knowledge of P allows us to uniquely recover all the elements of the model. Second, we show that P can be recovered from a long sequence of choices.

4.1. Identification of the Model from P

Under Assumptions A1, A2 and A3, changes in the choices of friends induce stochastic variation in the attention probabilities. This stochastic variation in choices allows us to recover the set of connections between the people in the network and the ranking of preferences of each of them. We then sequentially identify the attention mechanism of each person moving from the most preferred alternative to the least preferred one. Proposition 4.1 presents our first identification result.

Proposition 4.1. *Under Assumptions A1, A2, and A3, the reference groups $\mathcal{N}$, the profile of strict preferences $\succ$, and the attention mechanism* Q *are point identified from* P.

The next example sheds some light on the identification strategy in Proposition 4.1.

Example 4. Suppose that three people $\mathcal{A} = \{1,2,3\}$ select between two alternatives $\mathcal{Y} = \{1,2\}$ and the default option o. The researcher knows P_1, P_2, and P_3. Let us consider Person 1. Let $\mathbf{y}$ be such that $y_1 = o$. The probability that Person 1 selects the default option o (given a profile of choices $\mathbf{y}$ with $y_1 = o$) is

$$\mathrm{P}_1(o \mid \mathbf{y}) = \left(1 - \mathrm{Q}_1\left(1 \mid o, \mathrm{N}_1^1(\mathbf{y})\right)\right)\left(1 - \mathrm{Q}_1\left(2 \mid o, \mathrm{N}_1^2(\mathbf{y})\right)\right).$$

Under A3, we get that $2 \in \mathcal{N}_1$ if and only if

$$\mathrm{P}_1(o \mid o,o,o) \neq \mathrm{P}_1(o \mid o,1,o).$$

In words, if $2 \in \mathcal{N}_1$, then the probability of choosing the default option by Person 1 changes if Person 2 picks something else. Also, if $2 \notin \mathcal{N}_1$, then the probability of choosing the default option by Person 1 should be invariant to the choices of Person 2. Similarly, $3 \in \mathcal{N}_1$ if and only if $\mathrm{P}_1(o \mid o,o,o) \neq \mathrm{P}_1(o \mid o,o,1)$. Thus, we can learn the set of friends of Person 1 from observed P_1. Let us assume that $\mathcal{N}_1 = \{2\}$. To recover the preferences of Person 1 note that

$$\begin{array}{ll} \mathrm{P}_1(1 \mid \mathbf{y}) = \mathrm{Q}_1\left(1 \mid o, \mathrm{N}_1^1(\mathbf{y})\right) & \text{if} \quad 1 \succ_1 2 \\ \mathrm{P}_1(1 \mid \mathbf{y}) = \mathrm{Q}_1\left(1 \mid o, \mathrm{N}_1^1(\mathbf{y})\right)\left(1 - \mathrm{Q}_1\left(2 \mid o, \mathrm{N}_1^2(\mathbf{y})\right)\right) & \text{if} \quad 2 \succ_1 1 \end{array}.$$

Thus, $2 \succ_1 1$ if and only if

$$\mathrm{P}_1(1 \mid o,o,o) \neq \mathrm{P}_1(1 \mid o,2,o).$$

Suppose that, indeed, we get that $2 \succ_1 1$. We can finally recover the attention mechanism (for $y_1 = o$) via the next four probabilities in the data

$$\begin{array}{ll} \mathrm{P}_1(2 \mid o,o,o) = \mathrm{Q}_1(2 \mid o,0) & \mathrm{P}_1(2 \mid o,2,o) = \mathrm{Q}_1(2 \mid o,1) \\ \mathrm{P}_1(1 \mid o,o,o) = \mathrm{Q}_1(1 \mid o,0)(1 - \mathrm{Q}_1(2 \mid o,0)) & \mathrm{P}_1(1 \mid o,1,o) = \mathrm{Q}_1(1 \mid o,1)(1 - \mathrm{Q}_1(2 \mid o,0)) \end{array}.$$

By considering two other choice profiles $\mathbf{y}$ with $y_1 = 1$ and $y_1 = 2$ (instead of $y_1 = o$), respectively, we can fully recover the attention mechanism of Person 1. By a similar exercise we can recover the sets of friends, preferences, and the attention mechanisms for Persons 2 and 3. ■

4.2. Identification of P from Discrete Dataset

This section uses the identification of the CCPs, P, and the rates of the Poisson alarm clocks from a discrete dataset established in Kashaev et al. (2026). In discrete dataset the researcher observes the joint configuration of choices at fixed time intervals as in our empirical application.[8]

Let us assume the researcher observes people's choices at time intervals of length Δ and can consistently estimate $\Pr\left(\mathbf{y}^{t+\Delta} = \mathbf{y}' \mid \mathbf{y}^{t} = \mathbf{y}\right)$ for each pair $\mathbf{y}', \mathbf{y} \in \overline{\mathcal{Y}}^{A}$. We will capture these transition probabilities by a matrix $\mathcal{P}(\Delta)$. The connection between $\mathcal{P}(\Delta)$ and the transition rate matrix $\mathcal{M}$ described in Equation (2) is given by $\mathcal{P}(\Delta) = e^{(\Delta\mathcal{M})}$, where $e^{(\Delta\mathcal{M})}$ is the matrix exponential of $\Delta\mathcal{M}$. We allow the time interval Δ to be of arbitrary size.

The next proposition states that, by adding an extra restriction, it is possible to uniquely recover $\mathcal{M}$ from $\mathcal{P}(\Delta)$. In this case, the researcher needs to know the rates of the Poisson alarm clocks or normalize them in empirical work.

Proposition 4.2 (Proposition 3.8 from Kashaev et al., 2026)**.** *If Assumption A2 is satisfied and $\mathcal{M}$ has distinct eigenvalues that do not differ by an integer multiple of $2\pi i/\Delta$, where i here denotes the imaginary unit, then the conditional choice probabilities* P *are generically identified.*

5. Extensions of the Model

5.1. No Default Option

In the initial model, the default option plays a special role: it is chosen if, and only if, nothing else is considered. In some settings, such a default option does not exist.[9] This section modifies the initial assumptions to model this situation.[10]

Assume that there is no default option o, so that $\overline{\mathcal{Y}} = \mathcal{Y}$. The formation process of the consideration set is as before except that, since there is no default option, we need to specify what people do when the consideration set is empty. Given the dynamic nature of our model, we will simply assume that each person sticks to her previous choice if no alternative receives further consideration.[11] Formally, the probability that Person a selects (at the moment of choosing)

[8] Kashaev et al. (2026) also establish identification in cases when choices are observed in real-time.
[9] Also, see Horan (2019).
[10] For alternative ways to close the model see, for example, Barseghyan et al. (2021).
[11] In our empirical application we use this version of our model.

alternative $v \in \mathcal{Y}$ is given by

$$\mathrm{P}_a\left(v \mid \mathbf{y}\right) = \mathrm{Q}_a\left(v \mid \mathbf{y}\right) \prod\nolimits_{v' \in \mathcal{Y}, v' \succ_a v} \left(1 - \mathrm{Q}_a\left(v' \mid \mathbf{y}\right)\right) + 1(v = y_a) \prod\nolimits_{v' \in \mathcal{Y}} \left(1 - \mathrm{Q}_a\left(v' \mid \mathbf{y}\right)\right). \tag{3}$$

Proposition 5.1. *Suppose that Assumptions A1-A3 are satisfied and $Y \geq 3$. Then, the reference groups $\mathcal{N}$, the profile of strict preferences $\succ$, and the attention mechanism* Q *are point identified from* P.

The condition that $Y \geq 3$ is equivalent to requiring at least 2 non-default alternatives in the initial model. Thus, it is not restrictive. The identification proof for the network and the consideration probabilities is very similar to the proof of the initial model. However, since there is no default option, identification of preferences is more involved.

5.2. Non-Dominated Default Option

In some applications, it might be reasonable to think that people differ in regard to the ranking of the default option. In this section, we show that this variant of the model is also identified with an additional condition.

First note that, from a technical perspective, we can model the possibility that people differ regarding their ranking of the default option by letting the set of available alternatives for each Person a be a subset of $\mathcal{Y}$ that is not dominated by the default option. We can do so because any option outside this set will never be chosen by the person. We call this set $\mathcal{Y}_a$, to make it clear that the ranking of the default alternative varies for different people. The identification result is as follows.

Proposition 5.2. *Suppose that Assumptions A1-A3 are satisfied and for each $a \in \mathcal{A}$ and each $y \in \mathcal{Y}_a$ there exists $a' \in \mathcal{N}_a$ such that $y \in \mathcal{Y}_{a'}$. Then, the reference groups $\mathcal{N}$, the profile of strict preferences $\succ$, and the attention mechanism* Q *are point identified from* P.

The extra condition simply requires each person to have, for each alternative that is not dominated by the default option, at least one friend for whom that alternative is also not dominated by the default option. This condition ensures enough variability in the consideration probabilities to recover the ranking of preferences of each person. The identification strategy is otherwise similar to the initial one. (Note that for each Person a we can only recover the set of friends —or members of her reference group— that share with the person at least one element of the set of alternatives that are not dominated by the default.)

5.3. More General Peer Effects in Consideration Sets

In our initial model, the probability that Person a faces consideration set $\mathcal{C}$ given a choice configuration $\mathbf{y}$ takes the form of

$$\prod_{v\in\mathcal{C}} \mathrm{Q}_a\left(v \mid y_a, \mathrm{N}_a^v\left(\mathbf{y}\right)\right)\prod_{v\notin\mathcal{C}}\left(1-\mathrm{Q}_a\left(v \mid y_a, \mathrm{N}_a^v\left(\mathbf{y}\right)\right)\right).$$

This approach entails a multiplicative separable specification of alternatives in the consideration sets. It also assumes that the probability of considering an option depends only on the total number of friends who pick that option, but not on the identity of these friends (i.e., the effect of friends' choices is symmetric across friends). We next show that neither of these assumptions is essential for our approach. Indeed, all previous insights can be used here to extend the initial results.

For each Person a and configuration $\mathbf{y}$, let $\eta_a(\cdot \mid \mathbf{y})$ be an attention index function from $2^{\mathcal{Y}}$ to the positive reals. The value $\eta_a(\mathcal{C} \mid \mathbf{y})$ captures the attention that Person a pays to the set of alternatives $\mathcal{C} \in 2^{\mathcal{Y}}$ given the choice configuration $\mathbf{y}$. The attention-index measures how enticing a consideration set is (see Aguiar, Boccardi, Kashaev and Kim (2023) for further details). We define the probability of facing consideration set $\mathcal{C}$ as

$$\frac{\eta_a(\mathcal{C} \mid \mathbf{y})}{\sum_{\mathcal{D}\subseteq\mathcal{Y}}\eta_a(\mathcal{D} \mid \mathbf{y})}.$$

In this model, the consideration set probabilities are as in Brady and Rehbeck (2016). Since η_a can be identified only up to scale, we normalize the attention-index for the empty set to be 1 (i.e.,

$\eta(\emptyset \mid \mathbf{y}) = 1$ for all $\mathbf{y} \in \overline{\mathcal{Y}}^A$).[12] This more general setting covers our initial model, which is based on Manzini and Mariotti (2014), as a special case.

By combining preferences and this specification of stochastic consideration sets, the probability that Person a selects (at the moment of choosing) alternative $v \in \mathcal{Y}$ is given by

$$\mathrm{P}_a(v \mid \mathbf{y}) = \sum\nolimits_{\mathcal{C} \in 2^{\mathcal{Y}} : v \in \mathcal{C}} \mathrm{R}_a(v \mid \mathcal{C}) \frac{\eta_a(\mathcal{C} \mid \mathbf{y})}{\sum_{\mathcal{D} \subseteq \mathcal{Y}} \eta_a(\mathcal{D} \mid \mathbf{y})}.$$

The probability of selecting the default option o is just $\eta_a(\emptyset \mid \mathbf{y}) / \sum_{\mathcal{D} \subseteq \mathcal{Y}} \eta_a(\mathcal{D} \mid \mathbf{y})$. To better understand the connection between our initial model and this extension note that, for any $v \in \mathcal{Y}$ and $\mathcal{C} \subseteq \mathcal{Y}$ such that $v \notin \mathcal{C}$, we have that

$$\frac{\eta_a(\mathcal{C} \cup \{v\} \mid \mathbf{y})}{\eta_a(\mathcal{C} \mid \mathbf{y})} = \frac{\mathrm{Q}_a(v \mid \mathbf{y})}{1 - \mathrm{Q}_a(v \mid \mathbf{y})}.$$

Thus,

$$\mathrm{Q}_a(v \mid \mathbf{y}) = \frac{\eta_a(\mathcal{C} \cup \{v\} \mid \mathbf{y})}{\eta_a(\mathcal{C} \cup \{v\} \mid \mathbf{y}) + \eta_a(\mathcal{C} \mid \mathbf{y})}.$$

Based on this alternative specification, we accommodate Assumptions A1 and A3 as follows.

(A1') For each $a \in \mathcal{A}$, $v \in \mathcal{Y}$, and $\mathbf{y} \in \overline{\mathcal{Y}}^A$, there exists $\mathcal{C} \in 2^{\mathcal{Y}}$ such that $v \succ_a v'$ for all $v' \in \mathcal{C}$ and $\eta_a(\mathcal{C} \cup \{v\} \mid \mathbf{y}) > 0$.

(A3') For each $a \in \mathcal{A}$, $\mathcal{C} \in 2^{\mathcal{Y}}$, and $\mathbf{y}, \mathbf{y}^* \in \overline{\mathcal{Y}}^A$, such that $\mathbf{y}$ is different from $\mathbf{y}^*$ just in one component a^*,

(i) $a^* \notin \mathcal{N}_a$ or $y_{a^*},\ y^*_{a^*} \notin \mathcal{C} \implies \eta_a(\mathcal{C} \mid \mathbf{y}) = \eta_a(\mathcal{C} \mid \mathbf{y}^*)$;

(ii) $a^* \in \mathcal{N}_a,\ y_{a^*} \in \mathcal{C}$ and $y^*_{a^*} \notin \mathcal{C} \implies \eta_a(\mathcal{C} \mid \mathbf{y}) \geq \eta_a(\mathcal{C} \mid \mathbf{y}^*)$ with strict inequality for some $\mathcal{C}$.

Assumption A3'(i) states that the attention a person pays to a given set is invariant to the choices of those who are not connected with the person and to alternatives that do not enter the set. Assumption A3'(ii) means that switches of friends to a new alternative boost the attention for all sets that contain this new alternative. Note that Assumption A3' does not assume that different

[12] If one instead normalizes $\sum_{\mathcal{D} \subseteq \mathcal{Y}} \eta_a(\mathcal{D} \mid \mathbf{y}) = 1$, then we get the consideration rule of Aguiar (2017).

friends affect consideration probabilities symmetrically. That is, the model allows for the possibility that some friends have a bigger effect than others. Since the probability of facing consideration set $\mathcal{C}$ is proportional to the inverse of the total attention, $\sum_{\mathcal{D}\subseteq\mathcal{Y}} \eta_a(\mathcal{D} \mid \mathbf{y})$, even if peers are switching between alternatives that are not elements of $\mathcal{C}$, the probability of facing $\mathcal{C}$ may change.

The next proposition states that, with this general alternative specification, the network structure and the profile of preferences can be uniquely recovered from the conditional choice probabilities. Though the attention mechanism is just partially-identified, it is point identified under additional restrictions that reduce the dimensionality of the problem; we describe some of these restrictions in the next result.

Proposition 5.3. *Suppose that Assumptions A1', A2, and A3' are satisfied. Then, the reference groups $\mathcal{N}$ and the profile of strict preferences $\succ$ are point identified from* P*. If additionally for any $a \in \mathcal{A}$, $\mathbf{y} \in \overline{\mathcal{Y}}^A$, and $\mathcal{C}, \mathcal{D} \in 2^{\mathcal{Y}}$ one of the following holds*

(i) (Manzini and Mariotti, 2014) $\eta_a(\mathcal{C} \mid \mathbf{y}) = \prod_{v\in\mathcal{C}} \eta_a(\{v\} \mid \mathbf{y})$;

(ii) (Dardanoni et al., 2020) $|\mathcal{C}| = |\mathcal{D}| \implies \eta_a(\mathcal{C} \mid \mathbf{y}) = \eta_a(\mathcal{D} \mid \mathbf{y})$;

(iii) $\eta_a(\mathcal{C} \mid \mathbf{y}) = \sum_{v\in\mathcal{C}} \eta_a(\{v\} \mid \mathbf{y})$;

(iv) $\eta_a(\mathcal{C} \mid \mathbf{y}) = \eta_a(\{v^*\} \mid \mathbf{y})$, *where $v^* \in \mathcal{C}$ satisfies $v^* \succ_a v'$ for all $v' \in \mathcal{C}$;*

then $\{\eta_a\}_{a\in\mathcal{A}}$ is also pointidentified.

Condition (i) in Proposition 5.3 restates our main result using the consideration set formation model of Manzini and Mariotti (2014) —see also Manski (1977). Condition (ii) shows that the model analyzed in Dardanoni et al. (2020) —where the sets of equal cardinality are considered with equal probability— also imposes enough restrictions to recover consideration probabilities. Conditions (iii) and (iv) are new. Condition (iii) is similar to condition (i), but defines "aggregate" attention as a sum of singleton-attentions. Condition (iv) postulates that consideration probability of a set only depends on the best alternative in that set.

6. Empirical Application

The application we study is based on an experimental dataset that has been used to compare the effectiveness of various models that aim to predict the visual focus of attention of individuals within a group (see Bai et al., 2019 and Kumar et al., 2021).[13] In the experiment, a group of people were asked to play the Resistance game. This is a party game in which people communicate with each other in order to find the deceptive players. The game was implemented in several rounds and lasted about 30 minutes. Five participants were seated around a circle. A tablet was placed in front of each player and it recorded the direction of the sight of the player every 1/3 of a second. Each player can focus on 5 points —4 other players and the tablet. The dataset only shows information about the exact point of focus of each player (i.e., we do not know whether the person of focus was speaking or not). To simplify the exposition, we let the choice set of each person to be looking to the left, to the tablet, or to the right (i.e., left, tablet, or right). Although the choice set is rather small, players on average make at least 1 choice every 1.76 seconds in our data. We believe that such high frequency of choice forces players to not pay attention to all available options, thus justifying the use of our model. The raw frequency of choices in the data is displayed in Table 1.

Table 1 – Marginal Shares (%)

	Player 1	Player 2	Player 3	Player 4	Player 5
left	40	35	30	29	50
tablet	20	38	5	15	16
right	40	27	65	56	34

Notes: The sample size is 7875. Numbers were rounded.

Our structural approach allows us to separate two key forces in explaining the data: directional sight preferences of players and peer effects in gazing. These two estimated effects are motivated by well-known observations: First, visual designers create online platforms under the premise that people scan certain areas of the screen before others. In particular, it is believed that, everything else equal, people spend more time looking to the left side of the screen as compared to the right

[13]This dataset is publicly available at SNAP, Stanford University (https://snap.stanford.edu/index.html).

side.[14] This so-called left-to-right bias has been also documented by experimental studies.[15] Second, in social environments, it has been observed that people automatically redirect their visual attention by following the gaze orientation of other people, a phenomenon called gaze following.[16]

Before offering more details on the dataset and our estimates, let us briefly describe the main findings. Interestingly, despite considerable differences in the raw marginal shares of the three choices across players, as reflected in Table 1, the estimated preferences for directional sight of all players coincide. (Let us highlight here that in our estimation we allowed for full heterogeneity in preferences across people.) Consistent with the left-to-right bias, all players preferences agree in that

$$\text{left} \succ \text{tablet} \succ \text{right}.$$

The estimates for peer effects in consideration sets (in the data) are positive and relevant. These findings emphasize the importance of peer effects on attention and provides indirect evidence of the left-to-right bias, which is important for advertisement and marketing.

We next present more details on the dataset we use and some assumptions we make to simplify the exposition. (See Appendix C for extra information.)

Data The dataset contains 7,875 observations of the direction of each of 5 player's sight. Although the frequency of sight measure was rather high (3 measurements per second), about 26 percent of the observations have more than one player changing the direction of sight between consecutive measurements. Thus, we treat the dataset as discrete time data (i.e., Dataset 2 in Section 4).

Choice sets To capture possible preferences for direction in visual focus of attention, we aggregate the data to three options: left, tablet, and right.[17] (The original dataset has five options as each person can look at one of her four opponents or her tablet.)

[14] See, for example, https://www.nngroup.com/articles/horizontal-attention-original-research/.

[15] See Spalek and Hammad (2004, 2005) and Reutskaja, Nagel, Camerer and Rangel (2011). See also Maass and Russo (2003) and reference therein.

[16] See, for example, Gallup, Hale, Sumpter, Garnier, Kacelnik, Krebs and Couzin (2012) and https://www.nationalgeographic.com/science/article/what-are-you-looking-at-people-follow-each-others-gazes-but-without-a-tipping-point#close.21.

[17] This aggregation has several other advantages. For instance, the induced model has more players than alternatives. It also reduces the cardinality of the outcome space and the possible number of preference orders, i.e., $3^5 = 243$ vs. $5^5 = 3125$.

Consideration probabilities and choices of friends We assume that consideration probabilities on the direction of sight (i.e., left, tablet, or right) vary with the number of peers that look in the same direction. That is, each player is more likely to consider looking at a particular area if other players are doing so.[18]

Network structure Given the small number of players and the fact that participants are playing a party game with monetary prizes, we think it is natural to assume the network is complete. That is, we assume each player is connected to all other players in the group. This assumption is not necessary for our analysis and can be dropped. However, it decreases the computational burden and simplifies the exposition. Moreover, our estimation results suggest that this assumption holds. In particular, the estimated consideration probabilities increase from 3 to 4 peers, which is the maximal number of friends, for all players (see Figure 2).

Preferences and the default option Since, in this setting, there is no reason for us to treat one of the directions as the default option (i.e., always considered and the least preferred), we use the specification of the model with no default option presented in Section 5.1 for the estimates. Also, we do not impose any restrictions on preferences neither for each person nor across different people. Thus, in total, we have 3 possible preference orders for each player and $3^5 = 243$ possible combinations of preference orders for the five players.

Estimation Since we work with discrete-time data (Dataset 2), we follow the following general estimation procedure. Let $\theta = (e, \succ, \mathrm{Q})$ be an element of the space of possible parameters we want to estimate. (We normalize the intensity parameter λ_a to 1 for all $a \in \mathcal{A}$.) Note that since the number of people in the network and the number of choices is finite, the parameter space for e and $\succ$ is a finite set. For each θ we can construct the transition rate matrix $\mathcal{M}(\theta)$ using Equations (1) and (2). In turn, this information allows us to calculate the transition matrix

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

[18] Allowing for dependence of consideration probabilities on the current selection does not change the estimated preferences. See Appendix C for further details.

Given a sample of network configurations $\{\mathbf{y}_t\}_{t=0}^{T}$, we can use the latter to build the log-likelihood function $\mathrm{L}_T(\theta) = \Sigma_{t=0}^{T-1} \ln \mathcal{P}_{\iota(\mathbf{y}_t),\iota(\mathbf{y}_{t+1})}(\theta, \Delta)$, where $\iota(\mathbf{y}) \in \left\{1, 2, \ldots, \overline{\mathcal{Y}}^{A}\right\}$ is the position of $\mathbf{y}$ according the lexicographic order, and $\mathcal{P}_{k,m}(\theta, \Delta)$ is the (k, m)-th element of the matrix $\mathcal{P}(\theta, \Delta)$. Finally, the maximum likelihood estimator of the true parameter value can be defined as[19]

$$\widehat{\theta}_T = \arg\max_{\theta} \mathrm{L}_T(\theta).$$

Given that we assume a totally connected network structure, we only need to estimate $\succ$ and Q.

We estimate a few specifications of the model. We present two of these specifications next and display the other ones in Appendix C.3. In Model I(a), we assume there is no heterogeneity in consideration probabilities across options and players, i.e., $\mathrm{Q}(\cdot) = \mathrm{Q}_a(v \mid \cdot)$ for all a and v. Model I(b) adds heterogeneity in consideration probabilities across players. All the models we estimate allow for unrestricted heterogeneity in preferences regarding directional visual sight.

Figure 1 shows the estimates for the consideration probabilities of Model I(a). Note that although monotonicity of Q in the number of friends was not imposed in the estimation, the estimated probabilities are indeed monotone increasing. The estimated preferences for directional sight coincide for all players. Specifically, all players prefer looking to the left, then to tablet, and then to the right. That is, we obtain that, for each $a \in \mathcal{A}$,

$$\text{left} \succ_a \text{table} \succ_a \text{right}.$$

Thus, in the first specification of our model, all the players show the left-to-right bias.

Figure 2 shows the estimates for the consideration probabilities for Model I(b). Similar to the case with homogeneity, many of the consideration probabilities for each direction of sight are indeed increasing in the number of players looking at that direction. This monotonicity also suggests that indeed every player is connected to all other players (e.g., if Player a had only 2 friends, then the consideration probability would be the same for 2, 3, and 4 friends). The estimated preference orders coincide in Models I(a) and (b). That is, here again, all the players prefer looking to the left,

[19] We assesses the finite-sample performance of the proposed estimator in Appendix B.

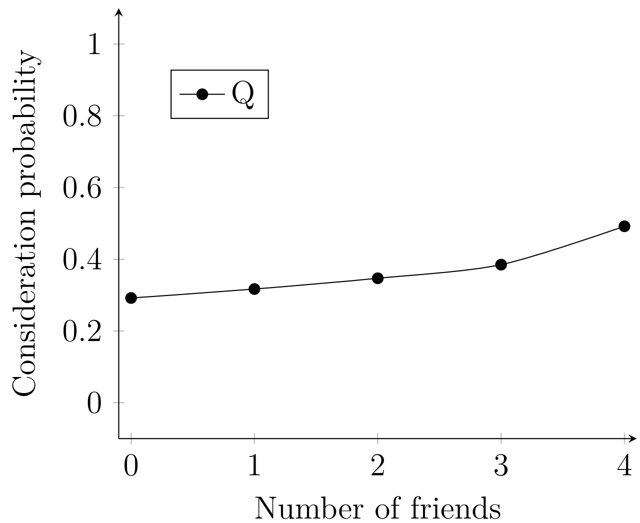


Figure 1 – Consideration probability as a function of the number of friends looking in the same direction. Model I(a).

then to tablet, and then to the right.

To sum up, despite considerable differences in raw marginal shares (see Table 1), the preferences for directional sight of all players coincide and are consistent with the left-to-right bias. As a robustness check, to address possible concerns with the dynamic of players' interactions in different stages of the game (e.g., learning), we estimated the previous two models using the last half of observations and the middle half of observations (i.e., we did not use the first and the last quarter of observations). The consideration probabilities are similar to the ones estimated using the whole sample, and the estimated preferences of all players are again left $\succ_a$ table $\succ_a$ right. (See Appendix C.2 for further details.)

7. Final Remarks

This paper adds peer effects to the consideration set models. It does so by combining the dynamic model of social interactions of Blume (1993, 1995) with the (single-agent) model of random consideration sets of Manski (1977) and Manzini and Mariotti (2014). The model we build differs from most of the social interaction models in that the choices of friends do not affect preferences

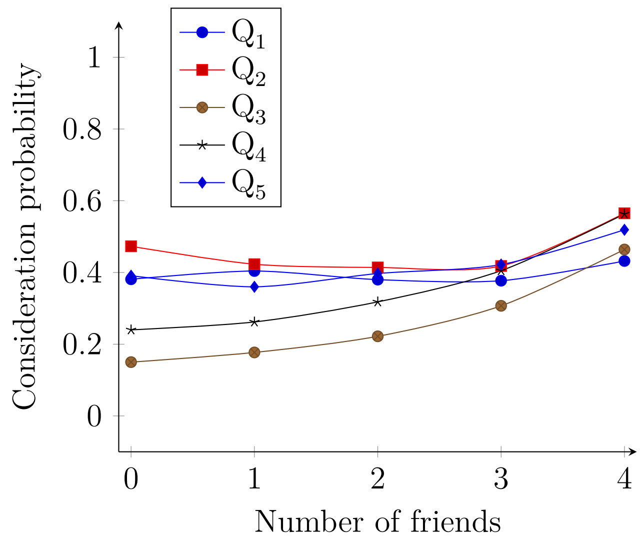


Figure 2 – Consideration probabilities for different players as functions of the number of friends looking in the same direction. Model I(b).

but the subset of options that people end up considering.

From an applied perspective, changes in the choices of friends induce stochastic variation of the considerations sets. We show that this variation can be used to recover the main parts of the model. On top of nonparametrically recovering the preference ranking of each person and the attention mechanism, we identify the set of connections between the people in the network. The identification strategy allows unrestricted heterogeneity across people. We propose a consistent estimator of model parameters and apply it to an experimental dataset. The structural approach we offer allows us to identify and estimate two main determinants on people visual focus of attention: directional sight preferences and peer effects in gazing. Our results are consistent with the documented left-to-right bias in directional sighting and positive peers effects in gazing.

We believe that our approach to peer effects in consideration sets could be incorporated in various empirical studies of technology adoption and diffusion. For instance, Aral, Muchnik and Sundararajan (2009) develop a dynamic matched sample estimation framework to distinguish influence and homophily effects in the day-by-day adoption of a mobile service application (i.e., Yahoo! Go). A similar dynamic matched sample estimation framework could be used to recover preferences over adoption and the influence of neighbor adoption rates on likelihood of considering the possibility of incorporating the mobile service application.

References

Abaluck, Jason and Abi Adams (2021) "What do consumers consider before they choose? Identification from asymmetric demand responses," *The Quarterly Journal of Economics*, 136 (3), 1611–1663.

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

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

Aguiar, Victor H and Nail Kashaev (2025) "Identification and estimation of discrete choice models with unobserved choice sets," *Journal of Business and Economic Statistics*, 43 (1), 204–215.

Aral, Sinan, Lev Muchnik, and Arun Sundararajan (2009) "Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks," *Proceedings of the National Academy of Sciences*, 106 (51), 21544–21549.

Bai, Chongyang, Srijan Kumar, Jure Leskovec, Miriam Metzger, Jay F Nunamaker Jr, and VS Subrahmanian (2019) "Predicting the visual focus of attention in multi-person discussion videos.," in *IJCAI 2019*, International Joint Conferences on Artificial Intelligence.

Barseghyan, Levon, Francesca Molinari, and Matthew Thirkettle (2021) "Discrete choice under risk with limited consideration," *American Economic Review*, 11 (6), 1972–2006.

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.

Blume, Lawrence E (1993) "The statistical mechanics of strategic interaction," *Games and economic behavior*, 5 (3), 387–424.

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

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

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

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

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

Brady, Richard L and John Rehbeck (2016) "Menu-dependent stochastic feasibility," *Econometrica*, 84 (3), 1203–1223.

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

Chambers, Christopher P, , Yusufcan Masatlioglu, and Christopher Turansick (2022) "Revealed social networks," *Journal of the European Economic Association*, 21 (1), 135–166.

Chambers, Christopher P, Tugce Cuhadaroglu, and Yusufcan Masatlioglu (2025) "Behavioral influence,"Technical report, Mimeo.

de Clippel, Geoffroy and Kareen Rozen (2024) "Bounded rationality in choice theory: A survey," *Journal of economic literature*, 62 (3), 995–1039.

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

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

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

De Paula, Áureo, Imran Rasul, and Pedro Souza (2025) "Identifying network ties from panel ata: Theory and an application to tax competition," *The Review of Economic Studies*, 92 (4), 2691–2729.

Gallup, Andrew C, Joseph J Hale, David JT Sumpter, Simon Garnier, Alex Kacelnik, John R Krebs, and Iain D Couzin (2012) "Visual attention and the acquisition of information in human crowds," *Proceedings of the National Academy of Sciences*, 109 (19), 7245–7250.

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

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

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

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

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*, Accepted.

Kumar, Srijan, Chongyang Bai, VS Subrahmanian, and Jure Leskovec (2021) "Deception Detection in Group Video Conversations using Dynamic Interaction Networks.," in *ICWSM 2021*, International AAAI Conference on Web and Social Media.

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

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

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

Maass, Anne and Aurore Russo (2003) "Directional bias in the mental representation of spatial events: Nature or culture?" *Psychological science*, 14 (4), 296–301.

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

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

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

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

Reutskaja, Elena, Rosemarie Nagel, Colin F Camerer, and Antonio Rangel (2011) "Search dynamics in consumer choice under time pressure: An eye-tracking study," *American Economic Review*, 101 (2), 900–926.

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

Spalek, Thomas M and Sherief Hammad (2004) "Supporting the attentional momentum view of IOR: Is attention biased to go right?" *Perception & Psychophysics*, 66 (2), 219–233.

——— (2005) "The left-to-right bias in inhibition of return is due to the direction of reading," *Psychological Science*, 16 (1), 15–18.

Strzalecki, Tomasz (2025) *Stochastic Choice Theory*, Econometric Society Monographs: Cambridge University Press.

A. Proofs

Proof of Proposition 3.1: For an irreducible, finite-state, continuous Markov chain the equilibrium μ exists and it is unique. Thus, we only need to prove that A1 implies that the Markov chain induced by our model is irreducible. First note that, under A1, for each $a \in \mathcal{A}$, $v \in \mathcal{Y}$, and $\mathbf{y} \in \overline{\mathcal{Y}}^A$, we have that

$$1 > \mathrm{P}_a\left(v \mid \mathbf{y}\right) = \mathrm{Q}_a\left(v \mid y_a, \mathrm{N}_a^v\left(\mathbf{y}\right)\right) \prod\nolimits_{v' \in \mathcal{Y}, v' \succ_a v} \left(1 - \mathrm{Q}_a\left(v' \mid y_a, \mathrm{N}_a^{v'}\left(\mathbf{y}\right)\right)\right) > 0.$$

To show irreducibility, let $\mathbf{y}$ and $\mathbf{y}'$ be two different choice configurations. It follows from expression (2) that we can go from one configuration to the other one in less than A steps with positive probability. ■

Proof of Proposition 4.1: By A1, $\mathrm{P}_a\left(\cdot \mid \mathbf{y}\right)$ has full support for each $\mathbf{y} \in \overline{\mathcal{Y}}^A$. Let $\mathbf{0}_{a'}^v$ be a profile of choices that differs from $\mathbf{0}$ only in that the action of agent a' changes from 0 to v. Under A3, $a' \in \mathcal{N}_a$ if, and only if,

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

Since this is true for each $a \in \mathcal{A}$, we can get $(\mathcal{A}, e)$. Also, if we add A2, from variation in $\mathrm{N}_a^{v'}\left(\mathbf{y}\right)$ for each $v' \neq v$, we can recover Person a's upper level set that corresponds to option v. That is,

$$\left\{v' \in \mathcal{Y} : v' \succ_a v\right\}.$$

By repeating the exercise with each alternative, we can recover $\succ_a$. Finally, suppose that v_a^* is the most preferred alternative for Person a. Then,

$$\mathrm{P}_a\left(v_a^* \mid \mathbf{y}\right) = \mathrm{Q}_a\left(v_a^* \mid y_a, \mathrm{N}_a^{v_a^*}\left(\mathbf{y}\right)\right).$$

It follows that we can recover $\mathrm{Q}_a\left(v_a^* \mid y_a, \mathrm{N}_a^{v_a^*}\left(\mathbf{y}\right)\right)$ for each $\mathbf{y} \in \overline{\mathcal{Y}}^A$. By proceeding in descending preference ordering we can then recover $\mathrm{Q}_a\left(v \mid y_a, \mathrm{N}_a^{v}\left(\mathbf{y}\right)\right)$ for all $v \in \mathcal{Y}$ (and each $\mathbf{y} \in \overline{\mathcal{Y}}^A$). ■

Proposition A.1 (Perfect Dataset)**.** *The researcher knows* $\lim_{\Delta \to 0} \mathcal{P}\left(\Delta\right) = e^{(\Delta \mathcal{M})}$. *Then the conditional choice probabilities* P *and the rates of the Poisson alarm clocks* λ *are identified.*

Proof of Proposition A.1: Since $\lim_{\Delta \to 0} \mathcal{P}\left(\Delta\right) = \mathcal{M}$, we can recover transition rate matrix from the data. Recall that

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

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

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

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

Proof of Proposition 4.2: 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 $\mathcal{M}$ is generically identified if, in addition to the conditions in Proposition 4.2, we have that

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

This condition is always satisfied if $A > 1$. Thus, identification of $\mathcal{M}$ follows because, by A2, $A \geq 2$. We can then uniquely recover $\left(\mathrm{P}_a\right)_{a \in \mathcal{A}}$ from $\mathcal{M}$ as in the proof of Proposition A.1 ■

Proof of Proposition 5.1: This proof is divided in three steps.

Step 1. (Identification of the Set of Connections) Take any two different people with arbitrary designations a_1 and a_2. Note that if $a_2 \notin \mathcal{N}_{a_1}$, then $\mathrm{P}_{a_1}(v \mid \mathbf{y}) = \mathrm{P}_{a_1}(v \mid \mathbf{y}')$ for any $\mathbf{y}$ and $\mathbf{y}'$ such that $y_a = y'_a$ for all $a \neq a_2$ and $y_{a_1} \neq v$. Also, let $v^*_{a_1}$ be the best preferred alternative of a_1. Then, by A1 and A3, for any $\mathbf{y}$ such that $y_{a_1} \neq v^*_{a_1}$

$$\mathrm{P}_{a_1}\left(v^*_{a_1} \mid \mathbf{y}\right) = \mathrm{Q}_{a_1}\left(v^*_{a_1} \mid \mathbf{y}\right)$$

is constant in y_{a_2} if and only if $a_2 \notin \mathcal{N}_{a_1}$. Altogether, $a_2 \notin \mathcal{N}_{a_1}$ if and only if $\mathrm{P}_{a_1}\left(v \mid \mathbf{y}\right)$ with $y_{a_1} \neq v$ is constant in y_{a_2}. As a result, we can identify whether a_2 is in the set of friends of a_1. Since the choice of a_1 and a_2 was arbitrary, we can identify the set of connections e. Note that for this result we only need $Y \geq 2$.

Step 2. (Identification of the Preferences) Fix some Person a_1. We will show the result for a set of alternatives $\mathcal{Y} = \{1, 2, 3\}$ of size 3. (The proof easily extends to the case of more alternatives. The only cost is extra notation.) Recall that for the probability of picking the best alternative of Person a_1, $v^*_{a_1}$, is

$$\mathrm{P}_{a_1}(v^*_{a_1} \mid \mathbf{y}) = \mathrm{Q}_{a_1}\left(v^*_{a_1} \mid y_{a_1}, \mathrm{N}^{v^*_{a_1}}_{a_1}(\mathbf{y})\right)$$

for any $y_{a_1} \neq v^*_{a_1}$. Hence, if any of friends of a_1 switches from the second best $(v^{**}_{a_1})$ to the third best $(v^{***}_{a_1})$ option or from $v^{***}_{a_1}$ to $v^{**}_{a_1}$, then $\mathrm{P}_{a_1}(v^*_{a_1} \mid \mathbf{y})$ will not change.

Similarly, the probability of picking the second best is

$$\mathrm{P}_{a_1}(v^{**}_{a_1} \mid \mathbf{y}) = \mathrm{Q}_{a_1}\left(v^{**}_{a_1} \mid y_{a_1}, \mathrm{N}^{v^{**}_{a_1}}_{a_1}(\mathbf{y})\right)\left(1 - \mathrm{Q}_{a_1}\left(v^*_{a_1} \mid y_{a_1}, \mathrm{N}^{v^*_{a_1}}_{a_1}(\mathbf{y})\right)\right)$$

for any $\mathbf{y}$ such that $y_{a_1} \neq v^{**}_{a_1}$ and $\mathrm{N}^{v^*_{a_1}}_{a_1}(\mathbf{y}) \in \{0, 1\}$. Thus, the friend changing from $v^*_{a_1}$ $(v^{***}_{a_1})$ to $v^{***}_{a_1}$ $(v^*_{a_1})$ will increase (decrease) $\mathrm{P}_{a_1}(v^{**}_{a_1} \mid \mathbf{y})$.

Finally, the probability of picking the last option is

$$\mathrm{P}_{a_1}(v^{***}_{a_1} \mid \mathbf{y}) = \mathrm{Q}_{a_1}\left(v^{***}_{a_1} \mid y_{a_1}, \mathrm{N}^{v^{***}_{a_1}}_{a_1}(\mathbf{y})\right)\left(1 - \mathrm{Q}_{a_1}\left(v^*_{a_1} \mid y_{a_1}, \mathrm{N}^{v^*_{a_1}}_{a_1}(\mathbf{y})\right)\right)\left(1 - \mathrm{Q}_{a_1}\left(v^{**}_{a_1} \mid y_{a_1}, \mathrm{N}^{v^{**}_{a_1}}_{a_1}(\mathbf{y})\right)\right).$$

The changes in this probability because of switches of a friend from $v^*_{a_1}$ to $v^{**}_{a_1}$ or from $v^{**}_{a_1}$ to $v^*_{a_1}$ are ambiguous: $\mathrm{P}_{a_1}(v^{***}_{a_1} \mid \mathbf{y})$ may increase, decrease, or stay constant.

Next take option 1 and consider changes in the probability of picking it by Person 1 when her friend of switches from option 2 to option 3 and from option 3 to option 2. Then repeat the same exercise for option 2 (switches from 1 to 3 and from 3 to 1) and option 3 (switches from 1 to 3 and from 3 to 1). Note that one of these probabilities, say for option 1, will be invariant to switches, another one will increase and decrease (say option 2). The third probability (for option 3) may behave differently.

Case 1. The third probability is increasing with one switch and decreasing with the opposite one. Then we can conclude with certainty that option 1 is the first best option. Hence we can identify

$$\mathrm{Q}_{a_1}\left(v^*_{a_1} \mid y_{a_1}, \mathrm{N}^{v^*_{a_1}}_{a_1}(\mathbf{y})\right) = \mathrm{P}_{a_1}(v^*_{a_1} \mid \mathbf{y})$$

for all $y_{a_1} \neq v^*_{a_1} = 1$. Next note that

$$\frac{\mathrm{P}_{a_1}(2 \mid \mathbf{y})}{1 - \mathrm{P}_{a_1}(1 \mid \mathbf{y})} = \begin{cases} \mathrm{Q}_a(2 \mid y_{a_1}, \mathrm{N}^2_{a_1}(\mathbf{y})), & \text{if } 2 \succ_{a_1} 3 \\ \mathrm{Q}_a(2 \mid y_{a_1}, \mathrm{N}^2_{a_1}(\mathbf{y}))(1 - \mathrm{Q}_a(3 \mid y_{a_1}, \mathrm{N}^3_{a_1}(\mathbf{y}))), & \text{if } 3 \succ_{a_1} 2 \end{cases}$$

for any $y_{a_1} \notin \{1, 2\}$. Thus a switch from option 1 to option 3 (or from option 3 to option 1) should identify the second best and the third best options.

Case 2. The third probability is invariant to changes. Then we can conclude with certainty that option 2 is the second best option. We also identify the fist best the the third best options since $\mathrm{P}_{a_1}(v^{**}_{a_1} \mid \mathbf{y})$ increases when a friend switches from the first best to the third best option and increases when the switch is opposite.

Case 3. Non of the above two cases. Then we can conclude with certainty that option 1 is the first best, option 2 is the second best, and option 3 is the third best option.

Thus we can identify preferences of a_1. Since the choice of a_1 was arbitrary we can identify $\succ = (\succ_a)_{a \in \mathcal{A}}$.

Step 3. (Identification of the Attention Mechanism) Fix some a_1 and let $y^*_{a_1}$ be the most preferred alternative of Person a_1. Then we can identify $\mathrm{Q}_{a_1}(y^*_{a_1} \mid y_{a_1}, \mathrm{N}^{y^*_{a_1}}_{a_1}(\mathbf{y}))$ for any $\mathbf{y}$ such that $y^*_{a_1} \neq y_{a_1}$ since

$$\mathrm{Q}_{a_1}\left(y^*_{a_1} \mid y_{a_1}, \mathrm{N}^{y^*_{a_1}}_{a_1}(\mathbf{y})\right) = \mathrm{P}_{a_1}(y^*_{a_1} \mid \mathbf{y}).$$

By proceeding in decreasing preference order we can recover $\mathrm{Q}_{a_1}(y'_{a_1} \mid y_{a_1}, N^{y'_{a_1}}_{a_1}(\mathbf{y}))$ for any y'_{a_1} and $\mathbf{y}$ such that $y'_{a_1} \neq y_{a_1}$. Moreover, we can identify

$$\prod\nolimits_{v' \neq y_{a_1}} \left(1 - \mathrm{Q}_a\left(v' \mid y_{a_1}, \mathrm{N}^{v'}_{a_1}(\mathbf{y})\right)\right)$$

Next note that for any $\mathbf{y}$ such that $y^*_{a_1} = y_{a_1}$

$$\mathrm{Q}_{a_1}(y^*_{a_1} \mid y_{a_1}, \mathrm{N}^{y^*_{a_1}}_{a_1}(\mathbf{y})) = \frac{\mathrm{P}_{a_1}(y^*_{a_1} \mid \mathbf{y}) - \prod_{v' \neq y^*_{a_1}}\left(1 - \mathrm{Q}_a\left(v' \mid y^*_{a_1}, \mathrm{N}^{v'}_{a_1}(\mathbf{y})\right)\right)}{1 - \prod_{v' \neq y^*_{a_1}}\left(1 - \mathrm{Q}_a\left(v' \mid y^*_{a_1}, \mathrm{N}^{v'}_{a_1}(\mathbf{y})\right)\right)}.$$

Hence, we can identify $\mathrm{Q}_{a_1}(y^*_{a_1} \mid y_{a_1}, \mathrm{N}^{y^*_{a_1}}_{a_1}(\mathbf{y}))$ for all $\mathbf{y}$. Let $y^{**}_{a_1}$ be the second best alternative of Person a_1, then for any $\mathbf{y}$ such that $y^{**}_{a_1} = y_{a_1}$ similarly to the case with $y^*_{a_1}$ we can identify $\mathrm{Q}_{a_1}(y^{**}_{a_1} \mid y_{a_1}, \mathrm{N}^{y^{**}_{a_1}}_{a_1}(\mathbf{y}))$ since

$$\mathrm{Q}_{a_1}(y^{**}_{a_1} \mid y_{a_1}, N^{y^{**}_{a_1}}_{a_1}(\mathbf{y})) = \frac{\mathrm{P}_{a_1}(y^{**}_{a_1} \mid \mathbf{y}) - \prod_{v' \neq y^{**}_{a_1}}\left(1 - \mathrm{Q}_{a_1}\left(v' \mid y^{**}_{a_1}, \mathrm{N}^{v'}_{a_1}(\mathbf{y})\right)\right)}{1 - \mathrm{Q}_{a_1}(y^*_{a_1} \mid y_{a_1}, \mathrm{N}^{v'}_{a_1}(\mathbf{y})) - \prod_{v' \neq y^{**}_{a_1}}\left(1 - \mathrm{Q}_{a_1}\left(v' \mid y^{**}_{a_1}, \mathrm{N}^{v'}_{a_1}(\mathbf{y})\right)\right)},$$

and thus we recover $\mathrm{Q}_{a_1}(y^{**}_{a_1} \mid y_{a_1}, N^{y^{**}_{a_1}}_{a_1}(\mathbf{y}))$ for all $\mathbf{y}$. By proceeding in decreasing preference order we can recover $\mathrm{Q}_{a_1}(y'_{a_1} \mid y_{a_1}, N^{y'_{a_1}}_{a_1}(\mathbf{y}))$ for any y'_{a_1} and $\mathbf{y}$. Since the choice of a_1 was arbitrary we can identify $(\mathrm{Q}_a)_{a \in \mathcal{A}}$. ■

Proof of Proposition 5.2: By A1, $\mathrm{P}_a(\cdot \mid \mathbf{y})$ has full support on $\overline{\mathcal{Y}}_a = \mathcal{Y}_a \cup \{o\}$ for each $\mathbf{y} \in \times_{a \in \mathcal{A}} \overline{\mathcal{Y}}_a$ and for every a. By A2 and A3, $\mathrm{P}_a(v \mid \mathbf{y})$ is weakly decreasing in $\mathrm{N}^{v'}_a(\mathbf{y})$ for each $v' \succ_a v$. Since $\mathrm{N}^{v'}_a(\mathbf{y})$ can take at lest two values (every option is in the choices set of at least one friend), we can recover $\mathcal{N}_a$. Since this is true for each $a \in \mathcal{A}$, we can get e. Also, from variation in $\mathrm{N}^{v'}_a(\mathbf{y})$ for each $v' \neq v$, we can recover

$$\{v' \in \mathcal{Y}_a : v' \succ_a v\}$$

By repeating the exercise with each alternative, we can recover $\succ_a$ over $\mathcal{Y}_a$. Finally, suppose that y_a^* is the most preferred alternative for Person a. Then,

$$\mathrm{P}_a\left(y_a^* \mid \mathbf{y}\right) = \mathrm{Q}_a\left(y_a^* \mid y_a, \mathrm{N}_a^{y_a^*}(\mathbf{y})\right).$$

It follows that we can recover $\mathrm{Q}_a\left(y_a^* \mid y_a, \mathrm{N}_a^{y_a^*}(\mathbf{y})\right)$ for each $\mathbf{y} \in \times_{a\in\mathcal{A}}\overline{\mathcal{Y}}_a$. By proceeding in descending preference ordering we can then recover $\mathrm{Q}_a\left(v \mid y_a, \mathrm{N}_a^{v}(\mathbf{y})\right)$ for all $v \in \mathcal{Y}_a$ (and each $\mathbf{y} \in \times_{a\in\mathcal{A}}\overline{\mathcal{Y}}_a$). ■

Proof of Proposition 5.3: Step 1. (Identification of the Set of Connections) Take any two different agents a_1 and a_2. Note that if $a_2 \notin \mathcal{N}_{a_1}$, then $\mathrm{P}_{a_1}(v \mid \mathbf{y}) = \mathrm{P}_{a_1}(v \mid \mathbf{y}')$ for any $\mathbf{y}$ and $\mathbf{y}'$ such that $y_a = y_a'$ for all $a \neq a_2$ and $y_{a_1} \neq v$. Also, if $v_{a_1}^*$ is the best preferred alternative of a_1, then by Assumptions A1' and A3', for any $\mathbf{y}$

$$\frac{\mathrm{P}_{a_1}\left(v_{a_1}^* \mid \mathbf{y}\right)}{\mathrm{P}_{a_1}\left(o \mid \mathbf{y}\right)} = \sum\nolimits_{\mathcal{C}\in 2^{\mathcal{Y}}:\, v_{a_1}^* \in \mathcal{C}} \eta_{a_1}\left(\mathcal{C} \mid \mathbf{y}\right)$$

is constant in choices of Person a_2 if and only if $a_2 \notin \mathcal{N}_{a_1}$. Hence, $a_2 \notin \mathcal{N}_{a_1}$ if and only if $\mathrm{P}_{a_1}\left(v \mid \mathbf{y}\right) / \mathrm{P}_{a_1}\left(o \mid \mathbf{y}\right)$ is constant in y_{a_2}. As a result, we can identify whether a_2 is a friend of a_1. Since the choice of a_1 and a_2 was arbitrary we can identify the whole e. Note that for this result to hold we only need $|\mathcal{Y}| \geq 2$.

Step 2. (Identification of the Preferences) Fix Person a_1 and $\mathbf{y}^* = (o, o, \ldots, o)\prime$. Note that

$$\frac{\mathrm{P}_{a_1}\left(v_{a_1}^* \mid \mathbf{y}^*\right)}{\mathrm{P}_{a_1}\left(o \mid \mathbf{y}^*\right)} = \sum\nolimits_{\mathcal{C}\in 2^{\mathcal{Y}}:\, v_{a_1}^* \in \mathcal{C}} \eta_{a_1}\left(\mathcal{C} \mid \mathbf{y}^*\right)$$

will increase if any of friends of a_1 switches to anything else. Let $v_{a_1}^{**}$ be the second best preferred alternative of a_1. Then

$$\frac{\mathrm{P}_{a_1}\left(v_{a_1}^{**} \mid \mathbf{y}^*\right)}{\mathrm{P}_{a_1}\left(o \mid \mathbf{y}^*\right)} = \sum\nolimits_{\mathcal{C}\in 2^{\mathcal{Y}}:\, v_{a_1}^* \notin \mathcal{C},\, v_{a_1}^{**} \in \mathcal{C}} \eta_{a_1}\left(\mathcal{C} \mid \mathbf{y}^*\right)$$

will increase if any of friends of a_1 switches to an element of $\mathcal{C}$ that corresponds to a strict inequality that and is different from $v^*_{a_1}$. Moreover, this probability will not change if any of friends of a_1 switches to $v^*_{a_1}$. Hence, we can identify $v^*_{a_1}$. Applying the above step in decreasing order, we can identify the whole preference order of Person a_1. Since the choice of a_1 was arbitrary we identify preferences of all persons.

Step 3. (Identification of the Attention Mechanism) Take any Person a_1 and configuration $\mathbf{y}$. Since preferences are identified, assume without loss of generality that $Y \succ_{a_1} Y-1 \succ_{a_1} Y-2 \succ_{a_1} \cdots \succ_{a_1} 2 \succ_{a_1} 1$. Note that we have the following system of Y equations

$$\frac{\mathrm{P}_{a_1}(k \mid \mathbf{y})}{\mathrm{P}_{a_1}(o \mid \mathbf{y})} = \begin{cases} \eta_{a_1}(\{1\} \mid \mathbf{y}), & k = 1, \\ \sum_{\mathcal{C} \subseteq \{1,\ldots,k-1\}} \eta_{a_1}(\mathcal{C} \cup \{k\} \mid \mathbf{y}), & k = 2, \ldots, Y. \end{cases}$$

Note that η_{a_1} could have generated the data if and only if it solves the above system of equations. Since, there are $2^Y - 1$ unknown parameters (recall that attention to the empty set is normalized to be 1) and Y equations, and there is no single attention parameter that enters more than one equation, η_{a_1} can not be identified without more restrictions. Suppose η_{a_1} is multiplicative. So we only need to identify $\eta_{a_1}(\{k\} \mid \mathbf{y})$ for all $k \in \mathcal{Y}$. Then, first, we can identify $\eta_{a_1}(\{1\} \mid \mathbf{y})$ since

$$\frac{\mathrm{P}_{a_1}(1 \mid \mathbf{y})}{\mathrm{P}_{a_1}(o \mid \mathbf{y})} = \eta_{a_1}(\{1\} \mid \mathbf{y}).$$

Next,

$$\frac{\mathrm{P}_{a_1}(2 \mid \mathbf{y})}{\mathrm{P}_{a_1}(o \mid \mathbf{y})} = \eta_{a_1}(\{2\} \mid \mathbf{y}) + \eta_{a_1}(\{1,2\} \mid \mathbf{y}) = \eta_{a_1}(\{2\} \mid \mathbf{y}) + \eta_{a_1}(\{1\} \mid \mathbf{y})\eta_{a_1}(\{2\} \mid \mathbf{y}).$$

Hence, we identify $\eta_{a_1}(\{2\} \mid \mathbf{y})$. Repeating the above steps recursively we can identify η_{a_1} since

$$\eta_{a_1}(\{k\} \mid \mathbf{y}) = \frac{\mathrm{P}_{a_1}(k \mid \mathbf{y})}{\mathrm{P}_{a_1}(o \mid \mathbf{y})} \cdot \frac{1}{\sum_{\mathcal{C} \subseteq \{1,\ldots,k-1\}} \eta_a(\mathcal{C} \mid \mathbf{y})}.$$

The same argument can be applied if η_{a_1} is additive. This approach can be generalized if one models the attention index of a set as a strictly increasing transformation of attention indexes of elements of that set (i.e., $\eta_{a_1}(\{1,2\} \mid \mathbf{y}) = \phi_{\{1,2\}}(\eta_{a_1}(\{1\} \mid \mathbf{y}), \eta_{a_1}(\{2\} \mid \mathbf{y}))$, where $\phi_{1,2}$ is strictly increasing in both arguments).

Suppose that η_{a_1} is the same for sets of the same cardinality. Since we know $\eta_{a_1}(\{1\})$ from the first equation we identify attention indexes for all singleton sets including $\eta_{a_1}(\{2\} \mid \mathbf{y})$. Hence, using the second equation we identify $\eta_{a_1}(\{1,2\} \mid \mathbf{y})$ and, thus, attention indexes for all sets of cardinality 2. Repeating the above arguments recursively we can identify η_{a_1}.

It is left to show that if η_{a_1} is the same for sets that have the same best option, then it is also identified. Again, $\eta_{a_1}(\{1\})$is identified from the first equation. Since $\eta_{a_1}(\{2\} \mid \mathbf{y}) = \eta_{a_1}(\{1,2\} \mid \mathbf{y})$, from the second equation we identify

$$\eta_{a_1}(\{2\} \mid \mathbf{y}) = \eta_{a_1}(\{1,2\} \mid \mathbf{y}) = \frac{\mathrm{P}_{a_1}(2 \mid \mathbf{y})}{2\,\mathrm{P}_{a_1}(o \mid \mathbf{y})}.$$

Repeating the above arguments for every equation we get that

$$\eta_{a_1}(\mathcal{C} \mid \mathbf{y}) = \frac{\mathrm{P}_{a_1}\left(v^*_{a_1,\mathcal{C}} \mid \mathbf{y}\right)}{v^*_{a_1,\mathcal{C}}\,\mathrm{P}_{a_1}(o \mid \mathbf{y})},$$

where $v^*_{a_1,\mathcal{C}}$ is the best alternative in $\mathcal{C}$ according to $\succ_{a_1}$. Since the choice of a_1 and $\mathbf{y}$ was arbitrary, we identify η_a for all $a \in \mathcal{A}$.

We conclude the proof by showing that restrictions (i) and (ii) correspond to the models of consideration sets formation in Manzini and Mariotti (2014) and Dardanoni et al. (2020), respectively. Equivalence between (ii) and the model in Dardanoni et al. (2020) is straightforward since

$$\frac{\eta_a(\mathcal{C} \mid \mathbf{y})}{\sum_{\mathcal{B}\subseteq\mathcal{Y}} \eta_a(\mathcal{B} \mid \mathbf{y})} = \frac{\eta_a(\mathcal{D} \mid \mathbf{y})}{\sum_{\mathcal{B}\subseteq\mathcal{Y}} \eta_a(\mathcal{B} \mid \mathbf{y})} \iff \eta_a(\mathcal{C} \mid \mathbf{y}) = \eta_a(\mathcal{D} \mid \mathbf{y}).$$

To show equivalence between (i) and the model in Manzini and Mariotti (2014) assume first that $\eta_a(\cdot \mid \mathbf{y})$ is multiplicative (i.e., satisfies condition (i)). Then for every $v \in \mathcal{Y}$ define

$$\mathrm{Q}_a(v \mid \mathbf{y}) = \frac{\eta_a(\{y\} \mid \mathbf{y})}{1 + \eta_a(\{v\} \mid \mathbf{y})}.$$

Since $\eta_a(\{v\} \mid \mathbf{y}) > 0$ (otherwise multiplicativity of η_a would imply that sets that contain v are never considered, which would contradict Assumption A1') and finite, we get that $\mathrm{Q}_a(v \mid \mathbf{y})) \in (0,1)$.

Take any A. By multiplicativity of η_a we get that

$$\eta_a(A \mid \mathbf{y}) = \prod_{y \in A} \eta_a(\{y\} \mid \mathbf{y}).$$

Note that for singleton $\mathcal{Y}$ we have that $\sum_{C \subseteq \mathcal{Y}} \eta_a(C \mid \mathbf{y}) = 1 + \eta_a(\mathcal{Y} \mid \mathbf{y})$. Then using multiplicativity we get that for $\mathcal{Y} \cup \{y^*\}$ such that $y^* \notin \mathcal{Y}$

$$\sum_{C \subseteq \mathcal{Y} \cup \{y^*\}} \eta_a(C \mid \mathbf{y}) = \sum_{C \subseteq \mathcal{Y}} \eta_a(C \mid \mathbf{y}) + \sum_{C \subseteq \mathcal{Y}} \eta_a(C \cup \{y^*\} \mid \mathbf{y}) = (1 + \eta_a(\{y^*\} \mid \mathbf{y})) \sum_{C \subseteq \mathcal{Y}} \eta_a(C \mid \mathbf{y}).$$

Hence, by induction

$$\sum_{C \subseteq \mathcal{Y}} \eta_a(C \mid \mathbf{y}) = \prod_{y \in \mathcal{Y}} (1 + \eta_a(\{y\} \mid \mathbf{y})).$$

As a result,

$$\begin{aligned} \frac{\eta_a(A \mid \mathbf{y})}{\sum_{C \subseteq \mathcal{Y}} \eta_a(C \mid \mathbf{y})} &= \prod_{y \in A} \frac{\eta_a(\{y\} \mid \mathbf{y})}{1 + \eta_a(\{y\} \mid \mathbf{y})} \prod_{y' \in \mathcal{Y} \setminus A} \left(1 - \frac{\eta_a(\{y'\} \mid \mathbf{y})}{1 + \eta_a(\{y'\} \mid \mathbf{y})}\right) \\ &= \prod_{y \in A} \mathrm{Q}_a(y \mid \mathbf{y}) \prod_{y' \in \mathcal{Y} \setminus A} (1 - \mathrm{Q}_a(y' \mid \mathbf{y})). \end{aligned}$$

Thus, condition (i) implies the model in Manzini and Mariotti (2014).

To show the opposite assume that the probability of considering a set $\mathcal{A}$ is

$$\prod_{y \in A} \mathrm{Q}_a(y \mid \mathbf{y}) \prod_{y' \in \mathcal{Y} \setminus A} (1 - \mathrm{Q}_a(y' \mid \mathbf{y}))$$

for some Q_a. For singleton sets define $\eta_a(\{y\} \mid \mathbf{y}) = \dfrac{\mathrm{Q}_a(y \mid \mathbf{y})}{1 - \mathrm{Q}_a(y \mid \mathbf{y})}$. For sets with cardinality bigger than one define $\eta_a(A \mid \mathbf{y}) = \prod_{y \in A} \eta_a(\{y\} \mid \mathbf{y})$. By construction, the constructed η_a satisfies multiplicativity and it is easy to verify that

$$\prod_{y \in A} \mathrm{Q}_a(y \mid \mathbf{y}) \prod_{y' \in \mathcal{Y} \setminus A} (1 - \mathrm{Q}_a(y' \mid \mathbf{y})) = \frac{\eta_a(A \mid \mathbf{y})}{\sum_{C \subseteq \mathcal{Y}} \eta_a(C \mid \mathbf{y})}.$$

■

B. Finite Sample Performance of the Estimator

In this appendix we evaluate the performance of our estimator by means of simulations.

Recall that $\theta = (e, \succ, \mathrm{Q})$ is an element of the space of possible parameters we want to estimate. Since the number of people in the network and the number of choices is finite, the parameter space for e and $\succ$ is a finite set. For each θ we construct the transition rate matrix $\mathcal{M}(\theta)$ using Equations (1) and (2). The transition matrix then is

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

The log-likelihood function is $\mathrm{L}_T(\theta) = \Sigma_{t=0}^{T-1} \ln \mathcal{P}_{\iota(\mathbf{y}_t),\iota(\mathbf{y}_{t+1})}(\theta, \Delta)$, where $\iota(\mathbf{y}) \in \left\{1, 2, \ldots, \overline{\mathcal{Y}}^A\right\}$ is the position of $\mathbf{y}$ according the lexicographic order, and $\mathcal{P}_{k,m}(\theta, \Delta)$ is the (k, m)-th element of the matrix $\mathcal{P}(\theta, \Delta)$. Finally, the maximum likelihood estimator of the true parameter value can be defined as

$$\widehat{\theta}_T = \arg\max\nolimits_\theta \mathrm{L}_T(\theta).$$

To evaluate the finite-sample properties of our estimator we conducted several Monte Carlo experiments. In all experiments we simulated discrete-time data from the specification presented in Example 2 in Section 3. In particular, we assumed that

$$2 \succ_1 1, \quad 1 \succ_2 2, \quad 2 \succ_3 1, \quad 1 \succ_4 2, \quad 1 \succ_5 2,$$

$$\mathrm{Q}(v \mid 0) = \frac{1}{4}, \quad \mathrm{Q}(v \mid 1) = \frac{3}{4}, \quad \mathrm{Q}(v \mid 2) = \frac{7}{8},$$

and

$$\mathcal{N}_1 = \{2, 3\}, \quad \mathcal{N}_2 = \{1, 3\}, \quad \mathcal{N}_3 = \{1, 2\}, \quad \mathcal{N}_4 = \{5\}, \quad \mathcal{N}_5 = \{4\}.$$

The data was simulated according to the following algorithm Let $\lambda = \sum_{a \in \mathcal{A}} \lambda_a$. We generate the data according to an iterative procedure for a fixed time period $\mathcal{T}$. The k-th iteration of the procedure is as follows:

Table 2 – Bias and Root Mean Squared Error (RMSE) ($\times 10^{-3}$)

		Sample Size					
Attention Probabilities		10	50	100	500	1000	5000
$Q(v \mid 0)$	Bias	6.8	2.2	1.5	0.0	-0.1	0.0
	RMSE	87.8	38.5	28.2	12.5	8.8	3.7
$Q(v \mid 1)$	Bias	-5.1	-3.7	-1.1	0.0	0.3	0.0
	RMSE	105.5	48.4	34.2	15.7	10.6	4.7
$Q(v \mid 2)$	Bias	-14.6	-2.5	-1.2	0.1	0.3	0.1
	RMSE	119.9	44.7	30.0	13.5	9.4	4.1

Notes: The sample sizes of 10, 50, 100, 500, 1000, and 5000 correspond to Δ equal to 2500, 500, 250, 50, 25, and 5. The number of replications is 1000.

(i) Given $\mathbf{y}_{k-1}$ set $\mathbf{y}_k = \mathbf{y}_{k-1}$;

(ii) Generate a draw from the exponential distribution with mean $1/\lambda$ and call it x_k;

(iii) Randomly sample an agent from the set $\mathcal{A}$, such that the probability that a is picked is λ_a/λ;

(iv) Given the agent selected in the previous step and the current choice configuration $\mathbf{y}_k$ construct a consideration set using Q_a;

(v) If the consideration set is empty, then set $y_{a,k} = o$. Otherwise pick the best alternative according to the preference order of agent a from the consideration set and assign it to $y_{a,k}$.

Given the initial configuration of choices $\mathbf{y}_0$ we applied the above algorithm till we reached $\sum_k x_k > \mathcal{T}$ (On average the length of the sequence is $\lambda\mathcal{T}$). Define $z_k = \sum_{l \leq k} x_l$. The continuous time data is $\{(y_k, z_k)\}$. The discrete time data is obtained from the continuous time data by splitting the interval $[0, \mathcal{T}]$ into $T = [\mathcal{T}/\Delta]$ intervals and recording the configuration of the network at every time period $t = i\Delta$, $i = 0, 1, \ldots, [\mathcal{T}/\Delta]$.

First, we estimate Q under the assumption that the network structure and preferences are known. The experiment was replicated 1000 times for 6 different sample sizes. The results of these simulations are presented in Table 2. The estimator of Q performs well in terms of the mean bias and the root mean squared error. As expected, the bias and the root mean squared error decrease with the sample size.

Next we estimate the whole parameter vector since our method allows to consistently estimate the network structure and the preference orders of individuals as well. Since in our example, without any restrictions, there are $2^{A(A-1)} = 1,048,576$ possible networks and $(Y!)^A = 32$ strict preference orders, to make the problem computationally tractable we restricted the parameter space for e by making the following assumptions: (i) each person has at most two friends; (ii) the attention mechanism is invariant across people and alternatives; and (iii) the network is undirected. As a result, the number of possible networks becomes 112 (the number of possible preference orders is still 32). The experiment was conducted 500 times for different sample sizes. Table 3 presents the results of these simulations. With just 50 observations the network structure is correctly estimated 94.4 percent times. For the sample size of 500 the network structure and the preferences are correctly estimated in all simulations.

Table 3 – Correctly Estimated Network & Preferences

Sample Size	10	50	100	500
Network	26.8%	83.2%	97.4%	100%
Preferences	37.8%	91.8%	99.0%	100%
Network & Preferences	13.6%	76.6%	96.4%	100%

Notes: The sample sizes of 10, 50, 100, and 500 correspond to Δ equal to 2500, 500, 250, and 50. The number of replications is 500.

C. Additional Estimation Results

C.1. Additional Experiment

The dataset in Bai et al. (2019) contains two experiments with five players. We presented the results for one of them. We next state that the estimates for the other one are quire similar.[20] This experiment contains 4248 observations. Similar to the experiment in the main text (the Main Experiment), we present marginal shares of different alternatives in the Additional Experiment in

[20]The full dataset contains five experiments of up to eight people. For our study, we use the data coming from the two experiments that have five players.

Table 4.

Table 4 – Marginal Shares in the Additional Experiment (%)

	Player 1	Player 2	Player 3	Player 4	Player 5
left	24.58	50.66	56.57	78.32	55.25
tablet	6.97	2.71	0.05	2.68	0.94
right	68.45	46.63	43.38	19	43.81

Notes: The sample size is 4248.

Looking at Table 4 we see that, in the Additional Experiment, all the players spend very little time looking at the tablet. In addition, all players but Player 1 look to the left more often than to the right. Recall, that the Main Experiment has more balanced shares. In other words, based on these initial results, one may expect considerable differences in the two experiments regarding preferences for directional sight. Interestingly, we get that this is not the case.

We estimated the same specification. Recall that in Model I(a), we assume there is no heterogeneity in consideration probabilities across options and players, i.e., $\mathrm{Q}(\cdot) = \mathrm{Q}_a (v \mid \cdot)$ for all a and v. Model I(b) adds heterogeneity in consideration probabilities across players. All the models we estimate allow for unrestricted heterogeneity in preferences regarding directional visual sight.

Figure 3 shows the estimates for the consideration probabilities of Model I(a) for the Additional Experiment. Identically to the Main Experiment, the estimated preferences for directional sight coincide for all players. Specifically, all players prefer looking to the left, then to tablet, and then to the right. Thus, in the first specification of our model, all the players, as in the Main Experiment, show the left-to-right bias.

Figure 4 shows the estimates for the consideration probabilities for Model I(b) in Experiments 1 and 2. Similar to the case with homogeneity, many of the consideration probabilities for each direction of sight are indeed increasing in the number of players looking at that direction. The estimated preference orders coincide in Models I(a) and (b), for the two experiments. That is, here again, all the players prefer looking to the left, then to tablet, and then to the right.

To sum up, despite considerable differences in raw marginal shares across players and across experiments (see Tables 4 and 1), the preferences for directional sight of all players in both experiments coincide.

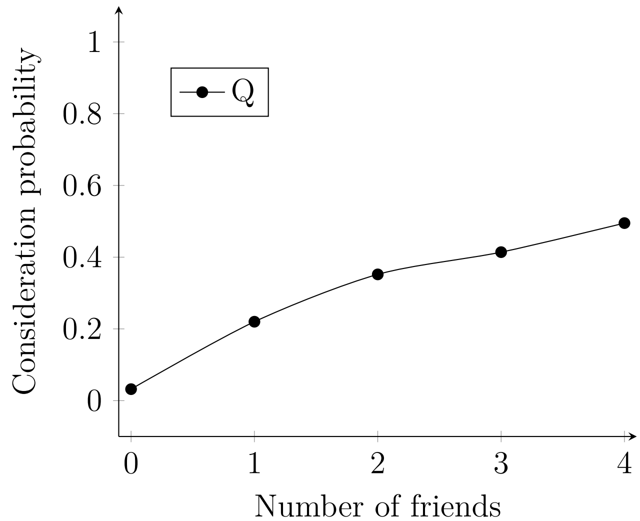


Figure 3 – Consideration probability as a function of the number of friends looking in the same direction. Model I(a). Additional Experiment.

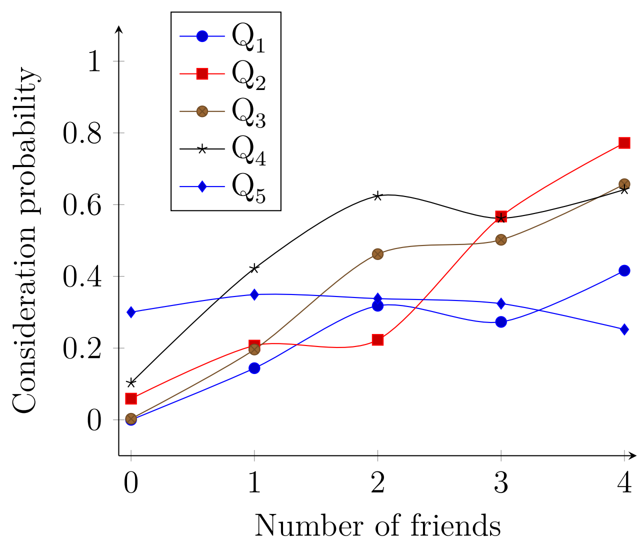
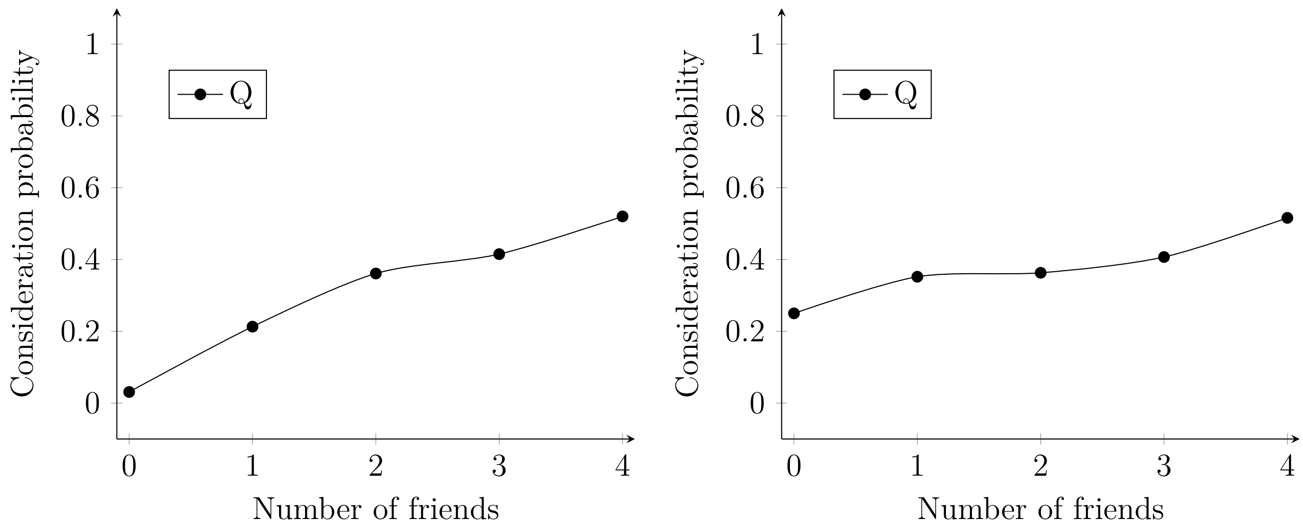


Figure 4 – Consideration probabilities for different players as functions of the number of friends looking in the same direction. Model I(b). Additional Experiment.



(a) Additional Experiment **(b)** Main Experiment

Figure 5 – Consideration probability. Model Ia. Middle half of the samples.

C.2. Different Parts of the Sample

Given that the Resistance is a dynamic party game, in order to analyse the behavior of consideration probabilities and preferences over time we estimated Models I(a) and I(b) presented in Section 6 using two different subsamples. The first subsample consists of the observations coming from the middle half (i.e., we trowed away the first and the last quarter of observations). The second one contains observations from the second half.

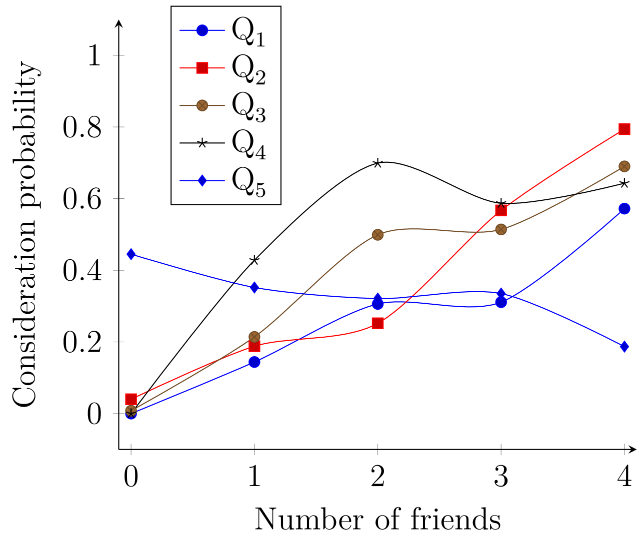
The estimated consideration set probabilities for the first subsample are presented in Figures 5-8. The estimated preference orders are presented in Tables 5-6.

Table 5 – Preferences. Model I(a).

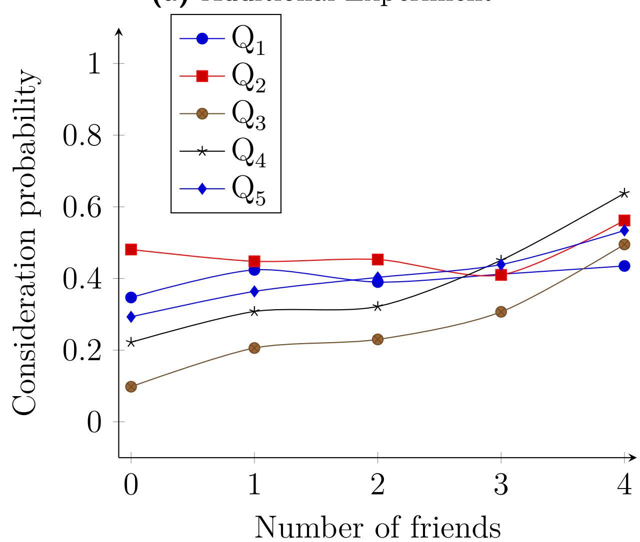
		Player 1	Player 2	Player 3	Player 4	Player 5
Middle half	Additional Experiment	L,T,R	L,T,R	L,T,R	L,T,R	L,T,R
	Main Experiment	L,T,R	L,T,R	L,T,R	L,T,R	L,T,R
2nd half	Additional Experiment	L,T,R	L,T,R	L,T,R	L,T,R	L,T,R
	Main Experiment	L,T,R	L,T,R	L,T,R	L,T,R	L,T,R

C.3. Heterogeneity in Consideration Probabilities Across Choices

In this section we present results of estimation of the following specifications of the model:

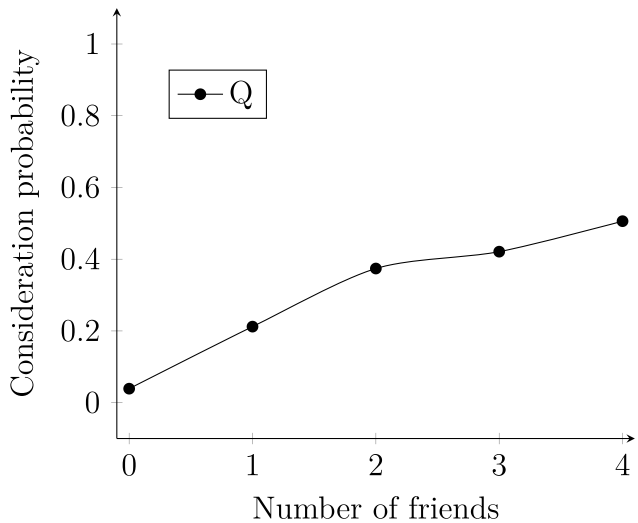


(a) Additional Experiment

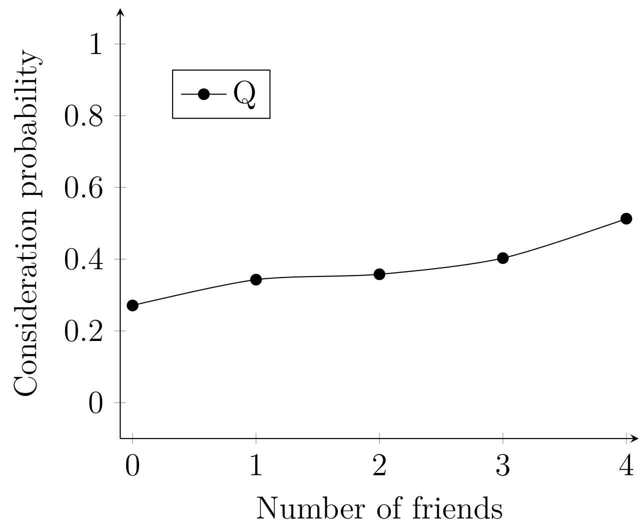


(b) Main Experiment

Figure 6 – Consideration probabilities. Model I(b). Middle half of the samples.

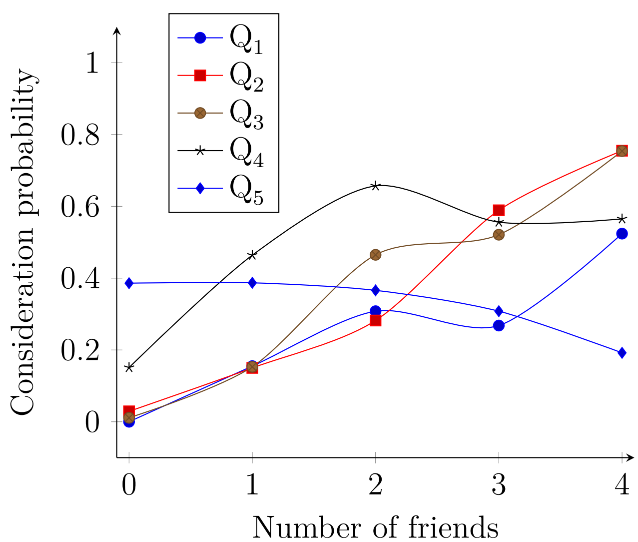


(a) Additional Experiment

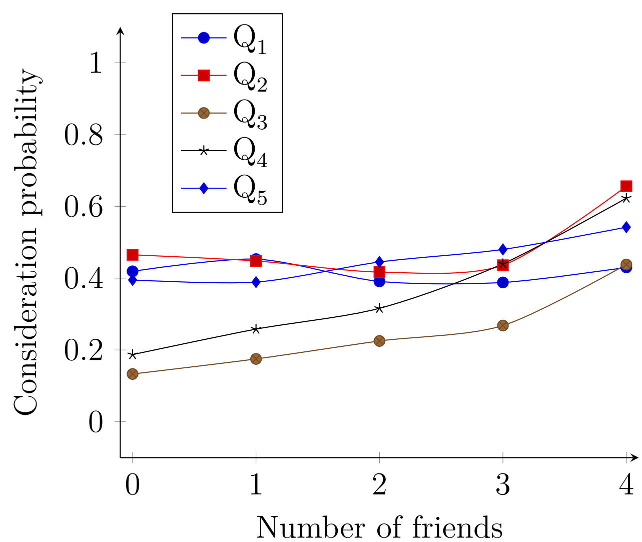


(b) Main Experiment

Figure 7 – Consideration probability. Model I(a). Second half of the samples.



(a) Additional Experiment

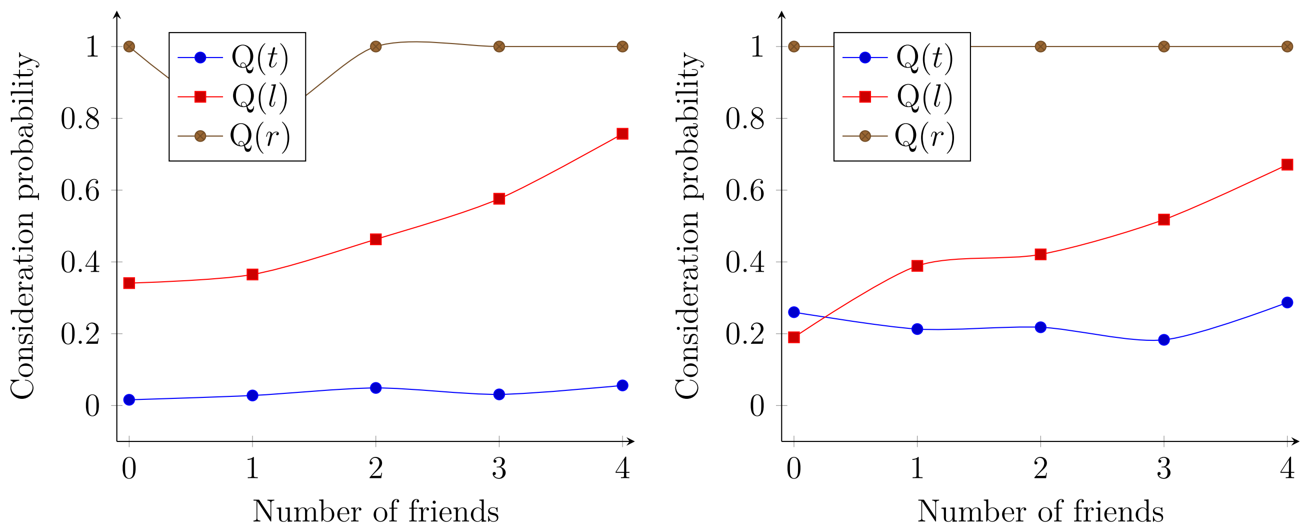


(b) Main Experiment

Figure 8 – Consideration probabilities. Model I(b). Second half of the samples.

Table 6 – Preferences. Model I(b).

	Player 1	Player 2	Player 3	Player 4	Player 5	
Middle half	Additional Experiment	L,T,R	L,T,R	L,T,R	L,T,R	L,T,R
	Main Experiment	L,T,R	L,T,R	L,T,R	L,T,R	L,T,R
2nd half	Additional Experiment	L,T,R	L,T,R	L,T,R	L,T,R	L,T,R
	Main Experiment	L,T,R	L,T,R	L,T,R	L,T,R	L,T,R



(a) Additional Experiment

(b) Main Experiment

Figure 9 – Consideration probabilities. Model Ic.

(i) Model Ic: $\mathrm{Q}_a\left(v \mid y_a, \mathrm{N}_a^v(\mathbf{y})\right) = \mathrm{Q}\left(v \mid \mathrm{N}_a^v(\mathbf{y})\right)$ for all $a, v, \mathbf{y}$;

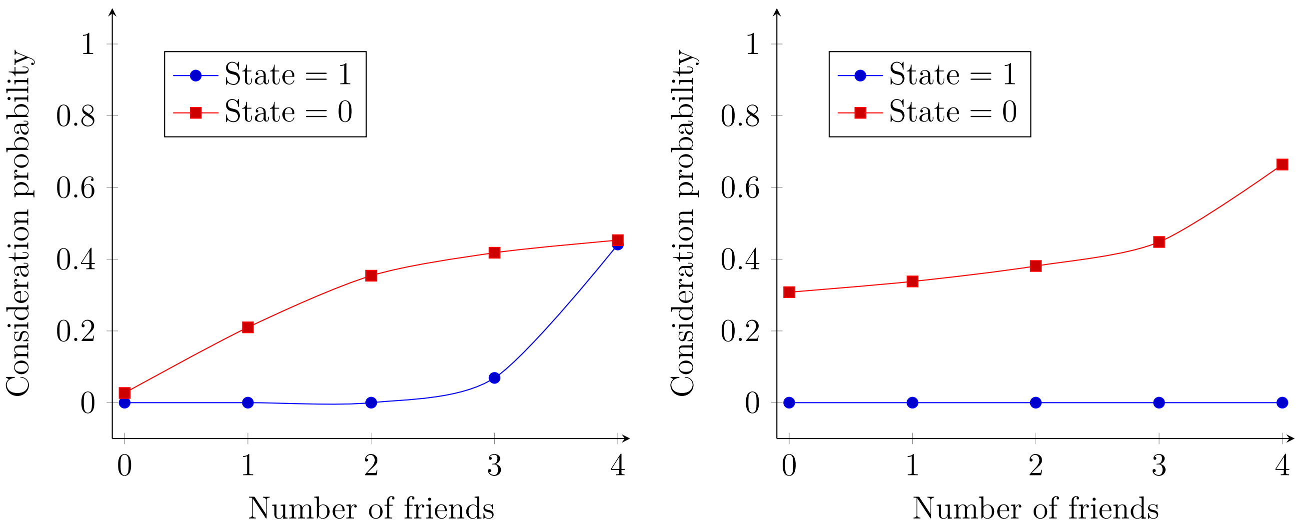
(ii) Model IIa: $\mathrm{Q}_a\left(v \mid y_a, \mathrm{N}_a^v(\mathbf{y})\right) = \mathrm{Q}\left(\mathrm{State}, \mathrm{N}_a^v(\mathbf{y})\right)$ for all $a, v, \mathbf{y}$. State equals to 1 if $v = y_a$, and 0 otherwise.

(iii) Model IIb: $\mathrm{Q}_a\left(v \mid y_a, \mathrm{N}_a^v(\mathbf{y})\right) = \mathrm{Q}_a\left(\mathrm{State}, \mathrm{N}_a^v(\mathbf{y})\right)$ for all $a, v, \mathbf{y}$. State equals to 1 if $v = y_a$, and 0 otherwise.

The estimated consideration probabilities are presented in Figures 9- 12 and Tables 7- 9.

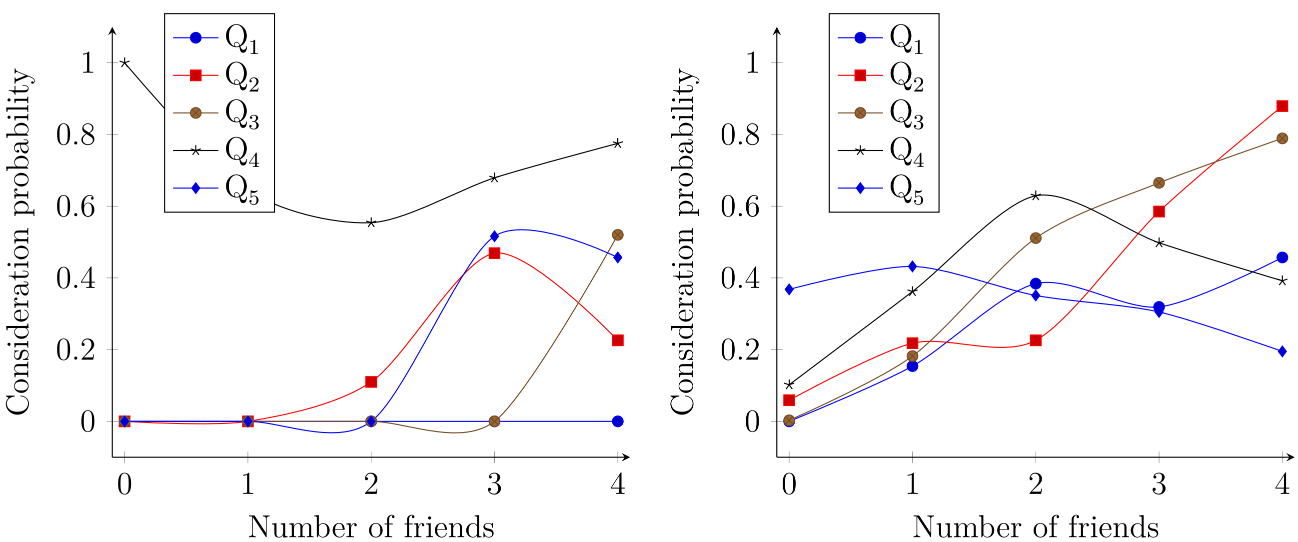
Table 7 – Preferences. Model Ic

	Player 1	Player 2	Player 3	Player 4	Player 5
Additional Experiment	T,L,R	T,L,R	L,T,R	T,L,R	L,T,R
Main Experiment	T,L,R	T,L,R	L,T,R	T,L,R	T,L,R



(a) Additional Experiment **(b)** Main Experiment

Figure 10 – Consideration probabilities. Model IIa.

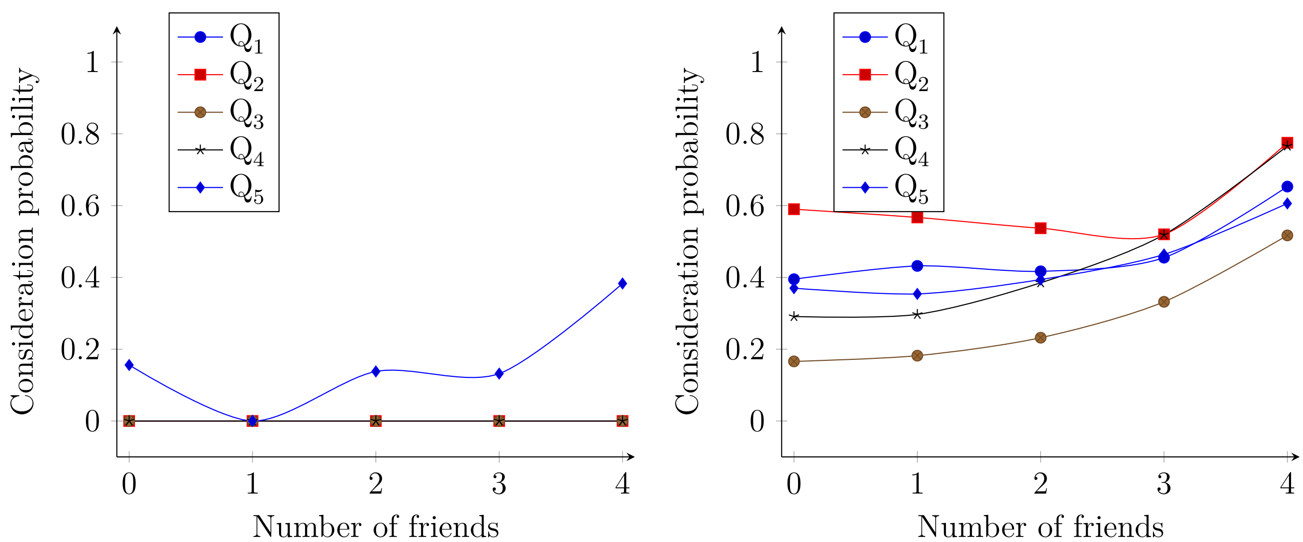


(a) State= 1 **(b)** State= 0

Figure 11 – Consideration probabilities. Model IIb. Additional Experiment.

Table 8 – Preferences. Model IIa

	Player 1	Player 2	Player 3	Player 4	Player 5
Additional Experiment	L,T,R	L,T,R	L,T,R	L,T,R	L,T,R
Main Experiment	L,T,R	L,T,R	L,T,R	L,T,R	L,T,R



(a) State= 1 **(b)** State= 0

Figure 12 – Consideration probabilities. Model IIb. Main Experiment.

Table 9 – Preferences. Model IIb

	Player 1	Player 2	Player 3	Player 4	Player 5
Additional Experiment	L,T,R	L,T,R	L,T,R	L,T,R	L,T,R
Main Experiment	L,T,R	L,T,R	L,T,R	L,T,R	L,T,R