DISCUSSION PAPER SERIES

DP17740

IDENTIFICATION IN SEARCH MODELS WITH SOCIAL INFORMATION

Niccolò Lomys and Emanuele Tarantino

INDUSTRIAL ORGANIZATION

CEPR

ISSN 0265-8003

IDENTIFICATION IN SEARCH MODELS WITH SOCIAL INFORMATION

Niccolò Lomys and Emanuele Tarantino

Discussion Paper DP17740
Published 11 December 2022
Submitted 24 November 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

IDENTIFICATION IN SEARCH MODELS WITH SOCIAL INFORMATION

Abstract

We theoretically study how social information affects agents' search behavior and the resulting observable outcomes that identify search models. We generalize canonical empirical search models by allowing a share of agents in the population to observe some peers' choices. Social information changes optimal search. First, we show that neglecting social information leads to non-identification and inconsistent estimation of search cost distributions under various standard datasets. Whether search costs are under or overestimated depends on the dataset. Second, we propose several remedies—such as data requirements, offline estimation techniques, exogenous variations, and partial identification approaches—that restore identification and consistent estimation.

JEL Classification: C1, C5, C8, D1, D6, D8

Keywords: Search and Learning, Social Information, Identification, Networks

Niccolò Lomys - niccolomys@gmail.com
Università degli studi di Napoli Federico II, CSEF

Emanuele Tarantino - etaranti@gmail.com
LUISS Guido Carli university

Acknowledgements
We thank Michele Bisceglia, Federico Carlini, Daniel Ershov, Daniel Garcia, Cristina Gualdani, Philipp Kircher, Lorenzo Magnolfi, José Luis Moraga-González, Philipp Peitler, Jacopo Perego, and Yutec Sun for helpful conversations and discussions. We are grateful to various seminar and conference audiences for helpful comments. We thank Xingwei Gan for excellent research assistance. Niccolò Lomys acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 714693) and from the ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program).

Identification in Search Models with Social Information*

Niccolò Lomys† Emanuele Tarantino‡

December 11, 2022

Abstract

We theoretically study how social information affects agents' search behavior and the resulting observable outcomes that identify search models. We generalize canonical empirical search models by allowing a share of agents in the population to observe some peers' choices. Social information changes optimal search. First, we show that neglecting social information leads to non-identification of search cost distributions under various standard datasets. Whether search costs are under or overestimated depends on the dataset. Next, we propose remedies—such as data requirements, offline estimation techniques, exogenous variations, and partial identification approaches—that restore identification and consistent estimation.

Keywords: Search and Learning; Social Information; Identification; Networks; Robustness.

JEL Classification: C1; C5; C8; D1; D6; D8.

*We thank Michele Bisceglia, Federico Carlini, Daniel Ershov, Daniel Garcia, Cristina Gualdani, Philipp Kircher, Lorenzo Magnolfi, José Luis Moraga-González, Philipp Peitler, Jacopo Perego, and Yutec Sun for helpful conversations and discussions. We are grateful to various seminar and conference audiences for helpful comments. We thank Xingwei Gan for excellent research assistance. Niccolò Lomys acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreements № 714147 and № 714693).

†Simons Institute for the Theory of Computing at UC Berkeley, CSEF, and Università degli Studi di Napoli Federico II; niccolomys@gmail.com.

‡Luiss University, EIEF, and CEPR; etarantino@luiss.it.

1 Introduction

Agents seldom search in isolation: others' choices and experiences are readily available via observation, communication, and social networks. The theoretical and empirical literature on social learning (see, e.g., Mobius and Rosenblat, 2014; Golub and Sadler, 2016; Bikhchandani, Hirshleifer, Tamuz and Welch, 2022, for surveys) documents how agents heavily rely on the information of others in shaping their behavior. Survey evidence documents that referrals on social media influence agents' purchases (Forbes, 2012, 2022). However, empirical search models typically assume that agents search in isolation, ignoring their social information (i.e., the information in their peers' decisions).

We theoretically investigate how social information affects agents' search behavior and the resulting observable outcomes that identify search models. We generalize the canonical empirical search model by allowing a share of agents in the population to observe the choice of some peers. First, we show that neglecting social information leads to non-identification of search cost distributions. Next, we propose approaches that restore identification.

We consider a stylized version of Weitzman (1979) sequential search model. An agent must choose between two alternatives whose high or low utilities are i.i.d. draws. Before searching, the agent knows the utility distribution but not the realized utilities. Searching an alternative reveals its utility to the agent. After searching the first alternative for free, the agent decides whether to search the second alternative at a cost, high or low. Before searching, with some probability, the agent observes the choice of one of her peers, but neither the peer's search behavior nor the peer's search cost. The peer acts in isolation and faces the same utility realizations as the agent. Thus, the agent draws inferences about realized utilities from the peer's choice.

The optimal decisions of an agent with social information differ from those of an isolated agent in two ways. First, an agent with social information is not indifferent about which alternative to search first. Since the peer searches both alternatives with positive probability, the utility distribution of the alternative chosen by the peer first-order stochastically dominates that of the other alternative. Thus, an agent with social information searches first the alternative chosen by her peer. Second, the expected gain from the second search for an agent with social information is lower than that of an isolated agent. The reason is that a Bayesian agent discounts the value of the second search by the probability that the peer searched only once, as only in this case the second search is valuable to the agent.

Thus, social information reduces the incentive to search.

These differences imply that an agent with social information obtains a higher utility and searches less than an isolated agent. Since social information changes the distributions of observable outcomes that identify search models, neglecting social information leads to non-identification of search cost distributions. Whether search costs are under or overestimated depends on the dataset. Below, we illustrate these insights with three datasets commonly available to researchers (see Honka, Hortaçsu and Wildenbeest, 2019, for a survey of the empirical search literature). Each dataset is also collectible by tracking agents' online (Ursu, Seiler and Honka, 2022).

Example 1 (Data on Choice). Suppose agents search online bookstores for the best price to buy a certain book. The researcher observes the price at which each agent buys the book at various bookstores. Some agents pay a low price, whereas others pay a high price for the book. If the researcher assumes agents search in isolation, she infers low search costs for all agents paying a low price. Before searching, however, some agents observe the bookstore where one of their peers has purchased the book. Since the peer has likely searched other bookstores, they infer that that specific bookstore offers a good bargain. Therefore, they buy at a low price because they exploit social information and not because of low search costs. Hence, search costs are underestimated.

Example 2 (Data on Optimal Stopping). Suppose agents search for the best price to buy a certain camera. The researcher observes the search history of each agent and the price they pay at the store. Some agents conduct a single search and pay a high price for the camera. If the researcher assumes agents search in isolation, she infers high search costs for all agents who discontinue search after searching only one store selling the camera at a high price. Before searching, however, some agents observe one of their peers buying the camera from a specific store. Since the peer has possibly searched both stores, they infer that the camera is not available at the other store at a lower price. Therefore, they discontinue search despite the high price because they exploit social information and not because of high search costs. Hence, search costs are overestimated.

Example 3 (Data on the Number of Searches). Suppose agents search for restaurants online. The researcher observes the number of searches each agent conducts before choosing a restaurant. Some agents search twice, whereas others search only once. If the researcher assumes agents search in isolation, she infers high search costs for all agents searching only once. Before searching, however, some agents observe an online referral of a specific restaurant by one of their peers. Since the peer has possibly searched both restaurants before posting the referral, they infer that the referred restaurant offers high-quality meals. Therefore, they search only once because they exploit social information and not because of high search costs. Hence, search costs are overestimated.

Search costs are a key determinant of agents' choice, pricing behavior, and market outcomes. Quantifying search frictions is crucial for computing price elasticities, assessing market competitiveness, and performing counterfactuals in regulated markets. To help fix misguided conclusions that may arise when neglecting social information, we present remedies that restore identification.

We begin with three methods that restore point identification of the search cost distribution. First, identification obtains by using agent-level data on social information, distinguishing isolated agents from agents with social information. Such data, however, are hardly available. Second, identification can obtain by estimating offline the share of agents with social information. For instance, the researcher can acquire detailed network data on the agents with access to social media or survey evidence on agents' reliance on peers' choices to make their purchases. Third, identification can obtain by exploiting variations in observables, like the utility distribution. For instance, the researcher can exploit the changes in observed search decisions due to a product upgrade shifting the utility distribution.

The three remedies above require additional information or exogenous variations and may be sensitive to the details of the search model. We next present partial identification approaches that allow the researcher to recover bounds on search cost distributions under weak assumptions on the environment. First, we show how to construct bounds with no information on the share of agents who observe a peer's choice. Second, we consider a general structure for social information, allowing the social information available to agents to vary and go beyond the simple observation of one isolated peer. This approach does not require specifying: the agents' amount of social information (e.g., how many peers they interact with); the agents' type of social information (e.g., whether it comes from observational or communication learning, advertising, ratings, or reviews); the search procedure (sequential or simultaneous). As a result, the researcher can recover robust bounds on search cost distributions when agents' amount and type of social information are unobserved, and so is the search procedure. Such bounds can be constructed by estimating the search cost distribution under two opposite assumptions: each agent acts in isolation; each agent's social information is such that the agent chooses the alternative with the highest utility at the first search. How far apart these bounds are is a measure of how misguided conclusions can be when neglecting social information.

We conclude by generalizing the model to provide additional guidance on how to account for social information in alternative empirically relevant contexts focusing, in particular, on continuous search cost distributions and Stigler (1961) simultaneous search.

Related Literature. Recovering economic primitives from observable behavior has a longstanding tradition (see, e.g., the revealed preference literature, Chambers and Echenique, 2016). Our approach draws inspiration from recent theoretical work on identification, such as: Heidhues and Strack (2021) for the identification of present bias from the timing of choices; Bergemann, Brooks and Morris (2022) for counterfactual predictions with latent information; Liu and Netzer (2021) for the identification of happiness measures from ordered response data; Libgober (2021) for the identification of information structures from posterior beliefs; Shmaya and Yariv (2016); Deb and Renou (2021); De Oliveira and Lamba (2022) for the testable implications of learning on observed choices; Heumann (2019) for informationally robust comparative statics. None of these papers studies identification in search models or the role of social information.

Our approach to partial identification and general social information is close in spirit to the theoretical literature on robust predictions in incomplete information or extensive form games (Bergemann and Morris, 2016; Doval and Ely, 2020), and to Barseghyan, Coughlin, Molinari and Teitelbaum (2021), who propose a robust method of discrete choice to partially identify preferences under heterogeneous choice sets. Recent work uses the Bayes correlated equilibrium notion of Bergemann and Morris (2016) to develop informationally robust identification and estimation strategies. Examples are: Magnolfi and Roncoroni (2022) for entry games; Syrgkanis, Tamer and Ziani (2021) for auctions; Gualdani and Sinha (2020) for single-agent models of voting; Canen and Song (2022) for counterfactual analyses.

Recent work analyzes how social learning affects individual search behavior (see, e.g., Kircher and Postlewaite, 2008; Galeotti, 2010; Hendricks, Sorensen and Wiseman, 2012; Mueller-Frank and Pai, 2016; Garcia and Shelegia, 2018; Lomys, 2020). We use the results in this literature to motivate the importance of social information in shaping the search process. However, our goals are distinct, as none of these papers studies identification in search models.

2 Model

Basic Setting. Consider (a stylized version of) the canonical sequential search model by Weitzman (1979) as developed by the empirical search literature (see Honka et al., 2019). There are countably many search problems, $n = 1, 2, \ldots$. In search problem n, the Bayesian agent n must select an alternative from the set $X \coloneqq \{0, 1\}$. We denote by x an alternative in X and by $\neg x$ the alternative in X other than x. Let $u_n^x \in \{\underline{u}, \overline{u}\}$, where $0 \leq \underline{u} < \overline{u}$, denote agent n's (indirect) utility from alternative x. Utilities are i.i.d. across alternatives within search problems and across search problems. Let $\alpha \coloneqq \mathbb{P}(u_n^x = \overline{u}) \in (0, 1)$ and $\Delta u \coloneqq \overline{u} - \underline{u}$.

Agent n knows the utility distribution, but not the realized utilities, about which she collects information via costly sequential search with recall:

1. Agent n decides which alternative to search first, $s_n^1 \in \{0, 1\}$. By searching alternative s_n^1, agent n perfectly learns its realized utility $u_n^{s_n^1}$.
2. Agent n decides whether to search the remaining alternative, $s_n^2 = \neg s_n^1$, and perfectly learn its realized utility $u_n^{\neg s_n^1}$, or to discontinue search, $s_n^2 = d$.
3. Agent n chooses an alternative $a_n \in S_n$, where S_n is the set of alternatives agent n has searched.

The first search is free. The second search costs $c_n \in \{\underline{c}, \overline{c}\}$. Search cost c_n is known to agent n. We assume $0 < \underline{c} < \alpha \Delta u < \overline{c}$. Absent this assumption, a search problem of type I (see below) is trivial: an agent would always search either both or only one alternative, irrespective of her search cost. Search costs are i.i.d. across agents. Let $\beta \coloneqq \mathbb{P}(c_n = \overline{c}) \in (0, 1)$.

Agent n maximizes the difference between the utility of the alternative she chooses and the search cost she incurs: $u_n^{a_n} - c_n(|S_n| - 1)$.

Social Information. Let $\theta_n \in \{\mathrm{I}, \mathrm{S}\}$ be the type of search problem n. Types θ_n are i.i.d. across search problems. Let $\gamma \coloneqq \mathbb{P}(\theta_n = \mathrm{I}) \in (0, 1]$.

- If $\theta_n = \mathrm{I}$, agent n is isolated. Her search problem is as described above.
- If $\theta_n = \mathrm{S}$, agent n has social information. Before engaging in sequential search (as described above), agent n observes the alternative a_{n_0} chosen by a fictitious Bayesian agent n_0 who: (i) faces a sequential search problem of type I, $\mathbb{P}(\theta_{n_0} = \mathrm{I}) = 1$; (ii) has the same realized utilities as agent n, $(u_{n_0}^0, u_{n_0}^1) = (u_n^0, u_n^1)$; (iii) has an idiosyncratic search cost c_{n_0} drawn independently of c_n, but from the same distribution. Agent n, however, observes neither agent n_0's search cost nor agent n_0's search decisions.

Examples. If $u_n^x := U - p_n^x$, we have a price search model for homogeneous goods with identical agents and ex-ante identical firms. If $u_n^x := \varepsilon_n^x$, we have a match-value search model with ex-ante identical agents and firms. Other search models can be accommodated similarly.

3 Optimal Decisions

3.1 Search Problem of Type I

First Search. Since the utilities of the two alternatives are i.i.d., agent n decides which alternative to search first uniformly at random: $s_n^1 = \frac{1}{2} \circ 0 + \frac{1}{2} \circ 1$, where $\sum_x \xi(x) \circ x$ denotes the mixture assigning probability $\xi(x)$ to alternative x.[1]

Since utilities are i.i.d.,

$$u_n^{s_n^1} = \begin{cases} \overline{u} & \text{with probability } \alpha \\ \underline{u} & \text{with probability } 1-\alpha \end{cases}.$$

Second Search. Agent n searches the second alternative if and only if the expected gain from doing so is no less than her search cost. Given the utility of the first alternative searched, $u_n^{s_n^1}$, agent n's expected gain from the second search is

$$V_{\mathrm{I}}\Big(u_n^{s_n^1}\Big) := \mathbb{E}\Big[\max\Big\{u - u_n^{s_n^1}, 0\Big\}\Big] = \begin{cases} 0 & \text{if } u_n^{s_n^1} = \overline{u} \\ \alpha\Delta u & \text{if } u_n^{s_n^1} = \underline{u} \end{cases}. \tag{1}$$

Since $0 < \underline{c} < \alpha\Delta u < \overline{c}$,

$$s_n^2 = \begin{cases} d & \text{if } u_n^{s_n^1} = \overline{u} \\ d & \text{if } u_n^{s_n^1} = \underline{u} \text{ and } c_n = \overline{c} \\ \neg s_n^1 & \text{if } u_n^{s_n^1} = \underline{u} \text{ and } c_n = \underline{c} \end{cases}.$$

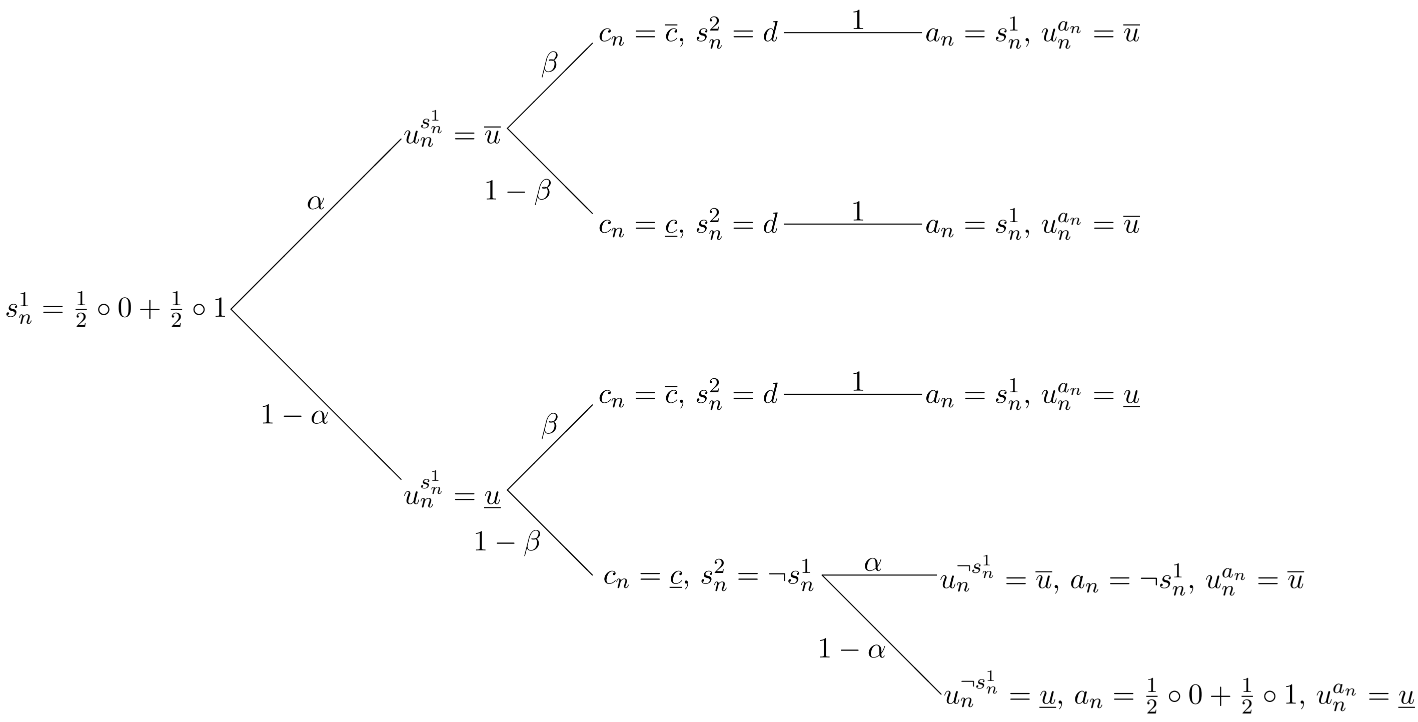
Choice. Agent n chooses the best alternative among those she sampled, randomizing uniformly if indifferent:

$$a_n = \begin{cases} s_n^1 & \text{if } s_n^2 = d \\ \neg s_n^1 & \text{if } s_n^2 = \neg s_n^1 \text{ and } u_n^{\neg s_n^1} = \overline{u} \\ \frac{1}{2} \circ 0 + \frac{1}{2} \circ 1 & \text{if } s_n^2 = \neg s_n^1 \text{ and } u_n^{\neg s_n^1} = \underline{u} \end{cases}. \tag{2}$$

[1] Breaking indifferences uniformly at random captures that labels do not convey information about alternatives' utilities or agents' behavior.

Decision Tree. Agent n's decisions in a search problem of type I are in Figure 1.

Figure 1: Decision Tree for a Search Problem of Type I.



3.2 Search Problem of Type S

First Search. In a search problem of type S, agent n's belief about the utilities of the two alternatives depends on agent n_0's optimal choice in a search problem of type I. There are two possibilities, each having positive probability:

1. Agent n_0 did not search alternative $\neg a_{n_0}$. If so, agent n_0's choice is uninformative about the utility of alternative $\neg a_{n_0}$.

2. Agent n_0 searched alternative $\neg a_{n_0}$. If so, since agent n_0 chose alternative a_{n_0}, it must be that $u_n^{a_{n_0}} \geq u_n^{\neg a_{n_0}}$ and, with positive probability, $u_n^{a_{n_0}} > u_n^{\neg a_{n_0}}$.

Agent n's belief about the utility of alternative a_{n_0} strictly first-order stochastically dominates her belief about the utility of alternative $\neg a_{n_0}$. Hence, by Weitzman (1979)'s optimal search rule, agent n searches alternative a_{n_0} first: $s_n^1 = a_{n_0}$. This is the first difference between search problems of types I and S.

Therefore,

$$u_n^{s_n^1} = u_{n_0}^{a_{n_0}} = \begin{cases} \overline{u} & \text{with probability } \alpha + \alpha(1-\alpha)(1-\beta) \\ \underline{u} & \text{with probability } \alpha\beta(1-\alpha) + (1-\alpha)^2 \end{cases},$$

where the probabilities are calculated from Figure 1.

Second Search. Agent n searches the second alternative if and only if the expected gain from doing so is no less than her search cost. The expected gain from the second search depends on the probability that agent n_0 did not search alternative $\neg s_n^1$ given that an alternative with utility $u_n^{s_n^1}$ was chosen, denoted by $P\left(u_n^{s_n^1}\right)$. With remaining probability, agent n_0 searched alternative $\neg s_n^1$ but chose alternative s_n^1, in which case alternative s_n^1 is non-inferior by revealed preference. Thus, agent n's expected gain from the second search is

$$V_{\mathrm{S}}\left(u_n^{s_n^1}\right) := P\left(u_n^{s_n^1}\right)\mathbb{E}\left[\max\{u - u_n^{s_n^1}, 0\}\right] = P\left(u_n^{s_n^1}\right)V_{\mathrm{I}}\left(u_n^{s_n^1}\right). \tag{3}$$

From Bayes rule and Figure 1,

$$P(\underline{u}) := \mathbb{P}\left(s_{n_0}^2 = d \mid u_{n_0}^{a_{n_0}} = \underline{u}\right) = \frac{\beta}{\beta + (1-\alpha)(1-\beta)},$$

and so

$$V_{\mathrm{S}}\left(u_n^{s_n^1}\right) = \begin{cases} 0 & \text{if } u_n^{s_n^1} = \overline{u} \\ \frac{\alpha\beta\Delta u}{\beta+(1-\alpha)(1-\beta)} & \text{if } u_n^{s_n^1} = \underline{u} \end{cases}. \tag{4}$$

Thus:[2]

- If $\underline{c} < V_{\mathrm{S}}(\underline{u})$,

$$s_n^2 = \begin{cases} d & \text{if } u_n^{s_n^1} = \overline{u} \\ d & \text{if } u_n^{s_n^1} = \underline{u} \text{ and } c_n = \overline{c} \\ \neg s_n^1 & \text{if } u_n^{s_n^1} = \underline{u} \text{ and } c_n = \underline{c} \end{cases}.$$

- If $\underline{c} > V_{\mathrm{S}}(\underline{u})$,

$$s_n^2 = d.$$

Conditional on the first searched alternative having utility $\underline{u}$, the expected gain from the second search for an agent with social information is lower than that for an isolated agent (compare equations (1) and (4) for $u_n^{s_n^1} = \underline{u}$). This is the second difference between search problems of types I and S.

Choice. Optimal choice is as in a search problem of type I (see equation (2)).

[2] Since non-generic in the parameter space, we ignore $\underline{c} = V_{\mathrm{S}}(\underline{u})$.

Decision Tree. Agent n's decisions in a search problem of type S are in Figure 2: Panel A if $\underline{c} < V_{\mathrm{S}}(\underline{u})$ and Panel B if $\underline{c} > V_{\mathrm{S}}(\underline{u})$.

3.3 Comparison of Optimal Decisions

The two differences between types of search problems have the following implications, which are key to understanding the identification results we present next.

Choice. The probability with which an agent with social information chooses an alternative with utility $\underline{u}$ is smaller than that with which an isolated agent does so if $\underline{c} < V_{\mathrm{S}}(\underline{u})$,

$$\underline{c} < V_{\mathrm{S}}(\underline{u}) \Longrightarrow \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S}) < \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I}), \tag{5}$$

and the same as that one if $\underline{c} > V_{\mathrm{S}}(\underline{u})$,

$$\underline{c} > V_{\mathrm{S}}(\underline{u}) \Longrightarrow \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S}) = \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I}). \tag{6}$$

Suppose $\theta_n = \mathrm{S}$. Then, agent n searches alternative a_{n_0} first: $s_n^1 = a_{n_0}$.

- If $\underline{c} < V_{\mathrm{S}}(\underline{u})$, agent n searches alternative $\neg s_n^1$, $s_n^2 = \neg s_n^1$, if $u_n^{s_n^1} = \underline{u}$ and $c_n = \underline{c}$, which occurs with positive probability. Moreover, with positive probability, $u_n^{\neg s_n^1} = \overline{u}$, in which case $a_n = \neg s_n^1$, and so $u_n^{a_n} = \overline{u}$. Thus, $\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S}) < \mathbb{P}(u_n^{a_{n_0}} = \underline{u}) = \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I})$. Implication (5) follows.
- If $\underline{c} > V_{\mathrm{S}}(\underline{u})$, agent n always discontinues search, $s_n^2 = d$, and so $a_n = s_n^1 = a_{n_0}$. Thus, $\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S}) = \mathbb{P}(u_n^{a_{n_0}} = \underline{u}) = \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I})$. Implication (6) follows.

Optimal Stopping. The probability with which an agent with social information discontinues search after searching first an alternative with utility $\underline{u}$ is the same as that with which an isolated agent does so if $\underline{c} < V_{\mathrm{S}}(\underline{u})$,

$$\underline{c} < V_{\mathrm{S}}(\underline{u}) \Longrightarrow \mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}, \theta_n = \mathrm{S}\Big) = \mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}, \theta_n = \mathrm{I}\Big), \tag{7}$$

and larger than that one if $\underline{c} > V_{\mathrm{S}}(\underline{u})$,

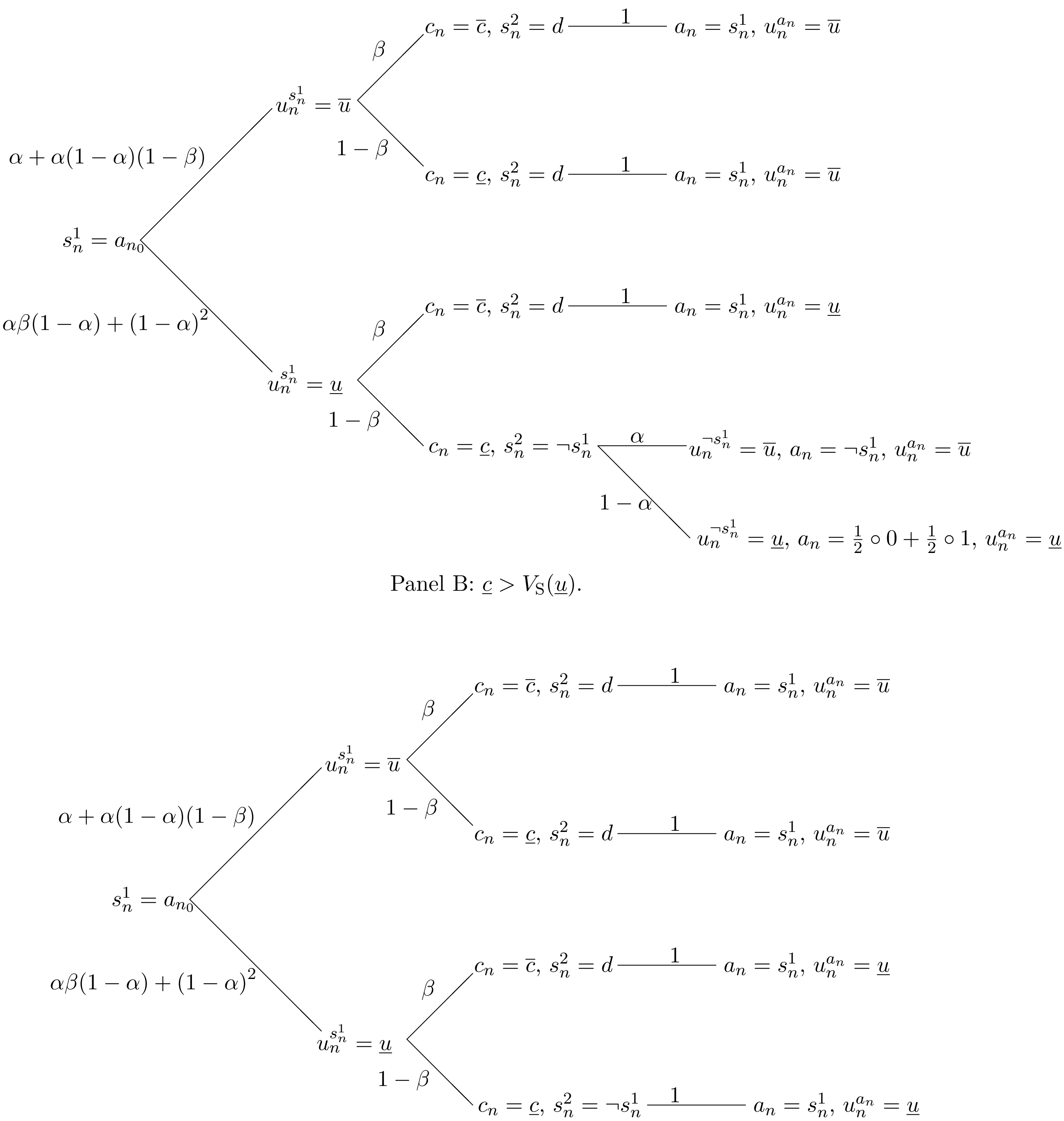
$$\underline{c} > V_{\mathrm{S}}(\underline{u}) \Longrightarrow \mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}, \theta_n = \mathrm{S}\Big) > \mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}, \theta_n = \mathrm{I}\Big). \tag{8}$$

Suppose the first searched alternative has utility $\underline{u}$: $u_n^{s_n^1} = \underline{u}$.

- If $\underline{c} < V_{\mathrm{S}}(\underline{u})$, agent n discontinues search if and only if $c_n = \overline{c}$ independently of whether $\theta_n = \mathrm{S}$ or $\theta_n = \mathrm{I}$. Implication (7) follows.
- If $\underline{c} > V_{\mathrm{S}}(\underline{u})$, agent n always discontinues search when $\theta_n = \mathrm{S}$, but discontinues search if and only if $c_n = \overline{c}$ when $\theta_n = \mathrm{I}$. Implication (8) follows.

Figure 2: Decision Trees for a Search Problem of Type S.

Panel A: $\underline{c} < V_{\mathrm{S}}(\underline{u})$.



Number of Searches. The probability with which an agent with social information discontinues search is larger than that with which an isolated agent does so:

$$\mathbb{P}\Big(s_n^2 = d \mid \theta_n = \mathrm{S}\Big) > \mathbb{P}\Big(s_n^2 = d \mid \theta_n = \mathrm{I}\Big). \tag{9}$$

Suppose $\theta_n = \mathrm{S}$. Then, agent n searches alternative a_{n_0} first, $s_n^1 = a_{n_0}$, and so $\mathbb{P}\Big(u_n^{s_n^1} = \overline{u} \mid \theta_n = \mathrm{S}\Big) = \mathbb{P}\Big(u_{n0}^{a_{n0}} = \overline{u}\Big)$. Since the utility of the two alternatives is different with positive probability and agent n_0 searches both alternatives with positive probability, $\mathbb{P}\Big(u_{n0}^{a_{n0}} = \overline{u}\Big) > \mathbb{P}\Big(u_{n0}^{s_{n0}^1} = \overline{u}\Big) = \mathbb{P}\Big(u_n^{s_n^1} = \overline{u} \mid \theta_n = \mathrm{I}\Big)$. Thus, $\mathbb{P}\Big(u_n^{s_n^1} = \overline{u} \mid \theta_n = \mathrm{S}\Big) > \mathbb{P}\Big(u_n^{s_n^1} = \overline{u} \mid \theta_n = \mathrm{I}\Big)$. Moreover: agent n discontinues search if $u_n^{s_n^1} = \overline{u}$ independently of whether $\theta_n = \mathrm{S}$ or $\theta_n = \mathrm{I}$; the probability with which agent n discontinues search if $u_n^{s_n^1} = \underline{u}$ when $\theta_n = \mathrm{S}$ is at least as large that with which agent n does so when $\theta_n = \mathrm{I}$ (see implications (7) and (8)). Inequality (9) follows.

4 Identification and Social Information

Consider a researcher who knows (or can consistently estimate) the utility distribution and the support of the search cost distribution. The researcher wants to identify the search cost distribution, i.e., $\beta := \mathbb{P}(c_n = \overline{c})$. We consider three standard datasets available to the researcher: choice, optimal stopping, and the number of searches. Observations are i.i.d., and N denotes the sample size.[3]

For each dataset, we show how to identify the parameter β when all agents are isolated. Next, we explain why identification fails if a positive share of agents has social information. Depending on the dataset, neglecting social information may lead to under or overestimation of search cost distributions.

4.1 Data on Choice

The researcher observes $\underline{u}_N^a$, the share of agents who choose an alternative with utility $\underline{u}$:

$$\underline{u}_N^a := \frac{\sum_{n=1}^N \mathbb{1}_{\{u_n^{a_n} = \underline{u}\}}}{N}. \tag{10}$$

Data on choice are readily available for price search models: the researcher needs the shares of transactions occurring at different prices. Data on choice, instead, may be harder to obtain in match-value search models as match values may not be observable.

[3]Given that types of search problems are i.i.d. across agents and the other model's assumptions, agents' decisions in each dataset are i.i.d.

Preliminary Observations for Identification. We characterize the probability with which an agent chooses an alternative with utility $\underline{u}$ in the data generating process. By the law of total probability,

$$\mathbb{P}(u_n^{a_n} = \underline{u}) = \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I})\gamma + \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S})(1-\gamma). \tag{11}$$

By Figures 1 and 2,

$$\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I}) = \alpha\beta(1-\alpha) + (1-\alpha)^2, \tag{12}$$

and

$$\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S}) = \begin{cases} \left[\alpha\beta(1-\alpha) + (1-\alpha)^2\right][1-\alpha(1-\beta)] & \text{if } \underline{c} < V_{\mathrm{S}}(\underline{u}) \\ \alpha\beta(1-\alpha) + (1-\alpha)^2 & \text{if } \underline{c} > V_{\mathrm{S}}(\underline{u}) \end{cases}. \tag{13}$$

By equations (11)–(13),

$$\mathbb{P}(u_n^{a_n} = \underline{u}) = \begin{cases} \left[\alpha\beta(1-\alpha) + (1-\alpha)^2\right][1-(1-\gamma)\alpha(1-\beta)] & \text{if } \underline{c} < V_{\mathrm{S}}(\underline{u}) \\ \alpha\beta(1-\alpha) + (1-\alpha)^2 & \text{if } \underline{c} > V_{\mathrm{S}}(\underline{u}) \end{cases}. \tag{14}$$

By the strong law of large numbers,

$$\underline{u}_N^a \xrightarrow{a.s.} \mathbb{E}[\underline{u}_N^a] = \mathbb{P}(u_n^{a_n} = \underline{u}). \tag{15}$$

All Agents Are Isolated. Suppose $\gamma = 1$. By equations (11) and (12), β is identified by

$$\mathbb{P}(u_n^{a_n} = \underline{u}) = \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I}) = \alpha\beta(1-\alpha) + (1-\alpha)^2. \tag{16}$$

Replacing $\mathbb{P}(u_n^{a_n} = \underline{u})$ with its sample analog $\underline{u}_N^a$ in equation (16), we obtain

$$\hat{\beta}_N^1 = \frac{\underline{u}_N^a}{\alpha(1-\alpha)} - \frac{1-\alpha}{\alpha}, \tag{17}$$

which, by the convergence in (15), is a consistent estimator of β.

Social Information. The next proposition summarizes the identification of β with data on choice when $\gamma < 1$, but the researcher assumes $\gamma = 1$.

Proposition 1. *Let $\gamma < 1$. Suppose the researcher observes data on choice and assumes $\gamma = 1$. Then:*

- *(i) If $\underline{c} < V_S(\underline{u})$, the parameter β is not identified by equation (16), the estimator $\hat{\beta}_N^1$ in equation (17) is inconsistent, and search costs are underestimated.*
- *(ii) If $\underline{c} > V_S(\underline{u})$, the parameter β is identified by equation (16), and the estimator $\hat{\beta}_N^1$ in equation (17) is consistent.*

Proof. [Part (i)] If $\underline{c} < V_{\mathrm{S}}(\underline{u})$, by equations (5) and (11), $\mathbb{P}(u_n^{a_n} = \underline{u}) < \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I})$. Therefore, β is not identified by equation (16). To see that $\hat{\beta}_N^1$ is inconsistent, and search costs are underestimated, note that

$$\hat{\beta}_N^1 \xrightarrow{a.s.} \mathbb{E}\left[\hat{\beta}_N^1\right] = \beta - (1-\gamma)(1-\beta)[1-\alpha(1-\beta)] < \beta, \tag{18}$$

where the equality holds by equation (14) for $\underline{c} < V_{\mathrm{S}}(\underline{u})$.

[Part (ii)] If $\underline{c} > V_{\mathrm{S}}(\underline{u})$, by equations (6) and (11), $\mathbb{P}(u_n^{a_n} = \underline{u}) = \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I})$. ■

If $\underline{c} < V_{\mathrm{S}}(\underline{u})$ and social information is neglected, β is not identified. The intuition is as in Example 1 in the Introduction. By assuming all agents are isolated, the researcher infers low search costs for all agents obtaining a high utility. Some of these agents, however, obtain a high utility because they exploit social information. Thus, search costs are underestimated.

4.2 Data on Optimal Stopping

The researcher observes $d_N^{\underline{u}}$, the share of agents who discontinue search after searching first an alternative with utility $\underline{u}$:

$$d_N^{\underline{u}} := \frac{\sum_{n=1}^N \mathbb{1}_{\{s_n^2 = d\}} \mathbb{1}_{\left\{u_n^{s_n^1} = \underline{u}\right\}}}{\sum_{n=1}^N \mathbb{1}_{\left\{u_n^{s_n^1} = \underline{u}\right\}}}.$$

Data on optimal stopping are readily available for price search models. The researcher needs the shares of agents who discontinue search when the price of the first searched alternative is high. Data on optimal stopping, instead, may be harder to obtain in match-value search models.

Preliminary Observations for Identification. We characterize the probability with which an agent discontinues search after searching first an alternative with utility $\underline{u}$ in the data generating process. By the law of total probability,

$$\begin{aligned} \mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}\Big) = \mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}, \theta_n = \mathrm{I}\Big)\mathbb{P}\Big(\theta_n = \mathrm{I} \mid u_n^{s_n^1} = \underline{u}\Big) \\ + \mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}, \theta_n = \mathrm{S}\Big)\mathbb{P}\Big(\theta_n = \mathrm{S} \mid u_n^{s_n^1} = \underline{u}\Big). \end{aligned} \tag{19}$$

By Figures 1 and 2,

$$\mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}, \theta_n = \mathrm{I}\Big) = \beta, \tag{20}$$

$$\mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}, \theta_n = \mathrm{S}\Big) = \begin{cases} \beta & \text{if } \underline{c} < V_{\mathrm{S}}(\underline{u}) \\ 1 & \text{if } \underline{c} > V_{\mathrm{S}}(\underline{u}) \end{cases}, \tag{21}$$

$$\mathbb{P}\Big(\theta_n = \mathrm{I} \mid u_n^{s_n^1} = \underline{u}\Big) = \frac{\gamma}{1-\alpha(1-\beta)(1-\gamma)}, \tag{22}$$

$$\mathbb{P}\Big(\theta_n = \mathrm{S} \mid u_n^{s_n^1} = \underline{u}\Big) = \frac{(1-\gamma)[1-\alpha(1-\beta)]}{1-\alpha(1-\beta)(1-\gamma)}. \tag{23}$$

By equations (19)–(23),

$$\mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}\Big) = \begin{cases} \beta & \text{if } \underline{c} < V_{\mathrm{S}}(\underline{u}) \\ \beta + \frac{(1-\beta)(1-\gamma)[1-\alpha(1-\beta)]}{1-\alpha(1-\beta)(1-\gamma)} & \text{if } \underline{c} > V_{\mathrm{S}}(\underline{u}) \end{cases}. \tag{24}$$

By the strong law of large numbers,

$$d_N^{\underline{u}} \xrightarrow{a.s.} \mathbb{E}\Big[d_N^{\underline{u}}\Big] = \mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}\Big). \tag{25}$$

All Agents Are Isolated. Suppose $\gamma = 1$. By equations (19), (20), and (23), β is identified by

$$\mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}\Big) = \mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}, \theta_n = \mathrm{I}\Big) = \beta. \tag{26}$$

Replacing $\mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}\Big)$ with its sample analog $d_N^{\underline{u}}$ in equation (26), we obtain

$$\hat{\beta}_N^2 = d_N^{\underline{u}}, \tag{27}$$

which, by the convergence in (25), is a consistent estimator of β.

Social Information. The next proposition summarizes the identification of β with data on optimal stopping when $\gamma < 1$, but the researcher assumes $\gamma = 1$.

Proposition 2. *Let $\gamma < 1$. Suppose the researcher observes data on optimal stopping and assumes $\gamma = 1$. Then:*

(i) If $\underline{c} < V_S(\underline{u})$, the parameter β is identified by equation (26), and the estimator $\hat{\beta}_2$ in equation (27) is consistent.

(ii) If $\underline{c} > V_S(\underline{u})$, the parameter β is not identified by equation (26), the estimator $\hat{\beta}_2$ in equation (27) is consistent, and search costs are overestimated.

Proof. [Part (i)] If $\underline{c} < V_{\mathrm{S}}(\underline{u})$, by equations (7) and (19), $\mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}\Big) = \mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}, \theta_n = \mathrm{I}\Big)$.

[Part (ii)] If $\underline{c} > V_{\mathrm{S}}(\underline{u})$, by equations (8), (19), (22), and (23), $\mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}\Big) > \mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}, \theta_n = \mathrm{I}\Big)$. Therefore, β is not identified by equation (26). To see that $\hat{\beta}_N^2$ is inconsistent, and search costs are overestimated, note that

$$\hat{\beta}_N^2 \xrightarrow{a.s.} \mathbb{E}\Big[\hat{\beta}_N^2\Big] = \beta + \frac{(1-\beta)(1-\gamma)[1-\alpha(1-\beta)]}{1-\alpha(1-\beta)(1-\gamma)} > \beta, \tag{28}$$

where the equality holds by equation (24) for $\underline{c} > V_{\text{S}}(\underline{u})$. ■

If $\underline{c} > V_{\text{S}}(\underline{u})$ and social information is neglected, β is not identified. The intuition is as in Example 2 in the Introduction. By assuming all agents are isolated, the researcher infers high search costs for all agents who discontinue search after searching first an alternative with low utility. Some of these agents, however, stop searching because they exploit social information. Thus, search costs are overestimated.

4.3 Data on the Number of Searches

The researcher observes d_N, the share of agents who conducted only one search:

$$d_N := \frac{\sum_{n=1}^{N} \mathbb{1}_{\{s_n^2 = d\}}}{N}.$$

Data on the number of searches are readily available for price search and match-value search models. Information about prices or match values is not necessary.

Preliminary Observations for Identification. We characterize the probability with which an agent discontinues search in the data generating process. By the law of total probability,

$$\mathbb{P}\Big(s_n^2 = d\Big) = \mathbb{P}\Big(s_n^2 = d \mid \theta_n = \text{I}\Big)\gamma + \mathbb{P}\Big(s_n^2 = d \mid \theta_n = \text{S}\Big)(1-\gamma). \tag{29}$$

By Figures 1 and 2,

$$\mathbb{P}\Big(s_n^2 = d \mid \theta_n = \text{I}\Big) = \alpha + (1-\alpha)\beta \tag{30}$$

and

$$\mathbb{P}\Big(s_n^2 = d \mid \theta_n = \text{S}\Big) = \begin{cases} \alpha + (1-\alpha)\beta + \alpha(1-\alpha)(1-\beta)^2 & \text{if } \underline{c} < V_{\text{S}}(\underline{u}) \\ 1 & \text{if } \underline{c} > V_{\text{S}}(\underline{u}) \end{cases}. \tag{31}$$

By equations (29)–(31),

$$\mathbb{P}\Big(s_n^2 = d\Big) = \begin{cases} \alpha + (1-\alpha)\beta + (1-\gamma)\alpha(1-\alpha)(1-\beta)^2 & \text{if } \underline{c} < V_{\text{S}}(\underline{u}) \\ \gamma[\alpha + (1-\alpha)\beta] + (1-\gamma) & \text{if } \underline{c} > V_{\text{S}}(\underline{u}) \end{cases}. \tag{32}$$

By the strong law of large numbers,

$$d_N \xrightarrow{a.s.} \mathbb{E}[d_N] = \mathbb{P}\Big(s_n^2 = d\Big). \tag{33}$$

All Agents Are Isolated. Suppose $\gamma = 1$. By equations (29) and (30), β is identified by

$$\mathbb{P}\Big(s_n^2 = d\Big) = \mathbb{P}\Big(s_n^2 = d \mid \theta_n = \text{I}\Big) = \alpha + (1-\alpha)\beta. \tag{34}$$

Replacing $\mathbb{P}\left(s_n^2 = d\right)$ with its sample analog d_N in equation (34), we obtain

$$\hat{\beta}_N^3 = \frac{d_N}{1-\alpha} - \frac{\alpha}{1-\alpha}, \tag{35}$$

which, by the convergence in (33), is a consistent estimator of β.

Social Information. The next proposition summarizes the identification of β with data on the number of searches when $\gamma < 1$, but the researcher assumes $\gamma = 1$.

Proposition 3. *Let $\gamma < 1$. Suppose the researcher observes data on the number of searches and assumes $\gamma = 1$. Then, the parameter β is not identified by equation (34), the estimator $\hat{\beta}_3$ in (35) is inconsistent, and search costs are overestimated.*

Proof. By equations (9) and (29), $\mathbb{P}\left(s_n^2 = d\right) > \mathbb{P}\left(s_n^2 = d \mid \theta_n = \mathrm{I}\right) = \alpha + (1-\alpha)\beta$. Therefore, β is not identified by equation (34). To see that $\hat{\beta}_N^3$ is inconsistent, and search costs are overestimated, note that

$$\hat{\beta}_N^3 \xrightarrow{a.s.} \mathbb{E}\left[\hat{\beta}_N^3\right] = \begin{cases} \beta + (1-\gamma)\alpha(1-\beta)^2 > \beta & \text{if } \underline{c} < V_{\mathrm{S}}(\underline{u}) \\ \gamma\beta + (1-\gamma) > \beta & \text{if } \underline{c} > V_{\mathrm{S}}(\underline{u}) \end{cases}, \tag{36}$$

where the equality holds by equation (32). ■

If social information is neglected, β is not identified. The intuition is as in Example 3 in the Introduction. By assuming all agents are isolated, the researcher infers high search costs for all agents searching only once. Some of these agents, however, search once because they exploit social information. Thus, search costs are overestimated.

5 Potential Remedies

Table 1 illustrates when neglecting social information leads to non-identification of search cost distributions, whether search costs are under or overestimated, and how this depends on the dataset.

Given prior knowledge of the environment, the researcher may make assumptions that restore identification. Many (theoretical) search models (e.g., Varian, 1980; Stahl, 1989; Ellison and Wolitzky, 2012) assume that a share of agents has negligible search cost and knows the utilities of all alternatives, i.e., in our model, $\underline{c} < V_{\mathrm{S}}(\underline{u})$. Under this assumption, identification of β obtains, even neglecting social information, with data on optimal stopping. Similarly, the researcher may assume search costs are "high" for all agents: $\underline{c} > V_{\mathrm{S}}(\underline{u})$. Under this assumption, identification of β obtains, even neglecting social information, with data on choice.

Table 1: Identification and Estimation.

	Under- (<0) or Over-Estimation (>0)		
	Choice	Optimal Stopping	Number of Searches
$\underline{c} < V_{\mathrm{S}}(\underline{u})$	<0	NO	>0
$\underline{c} > V_{\mathrm{S}}(\underline{u})$	NO	>0	>0

In most circumstances, however, it may be implausible to assume $\underline{c} < V_{\mathrm{S}}(\underline{u})$ or $\underline{c} > V_{\mathrm{S}}(\underline{u})$, or the required data may not be available. In this section, we present alternative approaches that restore the identification of β. To keep the exposition concise, we focus on data on choice.

5.1 Agent-Level Data on Social Information

Suppose the researcher can distinguish isolated agents from agents with social information. Formally, the researcher observes

$$\underline{u}_N^a(\mathrm{I}) := \frac{\sum_{n=1}^N \mathbb{1}_{\{u_n^{a_n}=\underline{u}\}} \mathbb{1}_{\{\theta_n=\mathrm{I}\}}}{\sum_{n=1}^N \mathbb{1}_{\{\theta_n=\mathrm{I}\}}}.$$

In this case, identification and consistent estimation of β easily obtain.

Proposition 4. *Suppose the researcher observes data on choice and agent-level data on social information. Then, the parameter β is identified by*

$$\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I}) = \alpha\beta(1-\alpha) + (1-\alpha)^2, \tag{37}$$

and

$$\hat{\beta}_N = \frac{\underline{u}_N^a(\mathrm{I})}{\alpha(1-\alpha)} - \frac{1-\alpha}{\alpha} \tag{38}$$

is a consistent estimator of β.

Proof. By equation (12), β is identified by equation (37). Replacing $\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I})$ with its sample analog $\underline{u}_N^a(\mathrm{I})$ in equation (37), we obtain the estimator $\hat{\beta}_N$ in equation (38). Consistency of $\hat{\beta}_N$ follows because, by the strong law of large numbers, $\underline{u}_N^a(\mathrm{I}) \xrightarrow{a.s.} \mathbb{E}[\underline{u}_N^a(\mathrm{I})] = \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I})$. ■

Agent-level data on social information, however, are hardly available. Thus, we consider the approaches below.

5.2 Estimating γ Offline

Suppose the researcher can identify and consistently estimate γ offline, e.g., by using detailed network data. By equation (14), if $\underline{c} > V_{\mathrm{S}}(\underline{u})$, β is identified by

$$\mathbb{P}(u_n^{a_n} = \underline{u}) = \alpha\beta(1-\alpha) + (1-\alpha)^2, \tag{39}$$

and, if $\underline{c} < V_{\mathrm{S}}(\underline{u})$, β is identified by

$$\mathbb{P}(u_n^{a_n} = \underline{u}) = \left[\alpha\beta(1-\alpha) + (1-\alpha)^2\right][1-(1-\gamma)\alpha(1-\beta)]. \tag{40}$$

Since $V_{\mathrm{S}}(\underline{u})$ is not observable, the researcher must adopt a data-driven approach to determine which equation identifies β. Let $\underline{\beta}$ be the solution for β to equation (39) and $\overline{\beta}$ be the real solution for β in $(0,1)$ to equation (40). Let $\underline{\beta}_N$ and $\overline{\beta}_N$ be the sample analogs of $\underline{\beta}$ and $\overline{\beta}$, that is, the solutions for β to equations (39) and (40) where $\mathbb{P}(u_n^{a_n} = \underline{u})$ is replaced by $\underline{u}_N^a$. Finally, let $V_{\mathrm{S}}\left(\underline{u};\underline{\beta}_N\right)$ (resp., $V_{\mathrm{S}}\left(\underline{u};\overline{\beta}_N\right)$) be the value of $V_{\mathrm{S}}(\underline{u})$ in equation (4) evaluated at $\underline{\beta}_N$ (resp., $\overline{\beta}_N$). Then, we have the following.

Proposition 5. *Suppose the researcher observes data on choice and estimates γ offline. Then:*

- *If $\underline{c} > V_S\left(\underline{u};\overline{\beta}_N\right)$ as $N \to \infty$, the parameter β is identified by equation (39), and $\underline{\beta}_N$ is a consistent estimator of β.*
- *If $\underline{c} < V_S\left(\underline{u};\underline{\beta}_N\right)$ as $N \to \infty$, the parameter β is identified by equation (40), and $\overline{\beta}_N$ is a consistent estimator of β.*

Proof. By the analysis in Section 4.1, $\underline{\beta} < \overline{\beta}$. Since $V_{\mathrm{S}}(\underline{u})$ is increasing in β, $V_{\mathrm{S}}\left(\underline{u};\underline{\beta}\right) < V_{\mathrm{S}}\left(\underline{u};\overline{\beta}\right)$. Moreover, since β is identified by either equation (39) or equation (40), either $V_{\mathrm{S}}(\underline{u}) = V_{\mathrm{S}}\left(\underline{u};\underline{\beta}\right) < V_{\mathrm{S}}\left(\underline{u};\overline{\beta}\right)$ or $V_{\mathrm{S}}\left(\underline{u};\underline{\beta}\right) < V_{\mathrm{S}}\left(\underline{u};\overline{\beta}\right) = V_{\mathrm{S}}(\underline{u})$. Thus: if $\underline{c} > V_{\mathrm{S}}\left(\underline{u};\overline{\beta}\right)$, then $\underline{c} > V_{\mathrm{S}}(\underline{u})$, and so β is identified by equation (39); if $\underline{c} < V_{\mathrm{S}}\left(\underline{u};\underline{\beta}\right)$, then $\underline{c} < V_{\mathrm{S}}(\underline{u})$, and so β is identified by equation (40). By the convergence in (15), $\underline{\beta}_N$ (resp., $\overline{\beta}_N$) is a consistent estimator of $\underline{\beta}$ (resp., $\overline{\beta}$). Therefore, $V_{\mathrm{S}}\left(\underline{u};\underline{\beta}_N\right)$ (resp., $V_{\mathrm{S}}\left(\underline{u};\overline{\beta}_N\right)$) is a consistent estimator of $V_{\mathrm{S}}\left(\underline{u};\underline{\beta}\right)$ (resp., $V_{\mathrm{S}}\left(\underline{u};\overline{\beta}\right)$). The desired result follows. ■

If $V_{\mathrm{S}}\left(\underline{u};\underline{\beta}_N\right) < \underline{c} < V_{\mathrm{S}}\left(\underline{u};\overline{\beta}_N\right)$, the available data do not suffice to detect whether β is identified by equation (39) or (40). Thus, another approach is required.

Example. Suppose $\alpha = \frac{1}{2}$, $\Delta u = 5$, $\gamma = \frac{1}{2}$, and $\underline{u}_N^a = \frac{5}{16}$. Then, $\underline{\beta}_N = \frac{1}{4}$, $\overline{\beta}_N = \sqrt{6} - 2$, $V_{\mathrm{S}}\left(\underline{u};\underline{\beta}_N\right) = 1$, and $V_{\mathrm{S}}\left(\underline{u};\overline{\beta}_N\right) = 4 - \sqrt{6}$. If $\underline{c} < 1$, the researcher uses $\overline{\beta}_N = \sqrt{6} - 2$ as an estimate of β. If $\underline{c} > 4 - \sqrt{6}$, the researcher uses $\underline{\beta}_N = \frac{1}{4}$ as an estimate of β.

5.3 Shifts to the Utility Distribution

Suppose the researcher observes a shift in the value of the high utility from $\overline{u}$ to $\overline{u}' > \overline{u}$. For shifts in the value of low utility from $\underline{u}$ to $\underline{u}' < \underline{u}$, analogous reasoning applies.

Let $V_{\mathrm{S}}'(\underline{u})$ be the value of $V_{\mathrm{S}}(\underline{u})$ in equation (4) evaluated at $\overline{u}'$. Since $V_{\mathrm{S}}(\underline{u})$ is increasing in $\overline{u}$, $V_{\mathrm{S}}(\underline{u}) < V_{\mathrm{S}}'(\underline{u})$. Moreover, let $\mathbb{P}_{\overline{u}}(u_n^{a_n} = \underline{u})$ (resp., $\mathbb{P}_{\overline{u}'}(u_n^{a_n} = \underline{u})$) be the population moment defined by equation (14) when the high utility is $\overline{u}$ (resp., $\overline{u}'$).

Suppose $V_{\mathrm{S}}(\underline{u}) < \underline{c} < V_{\mathrm{S}}'(\underline{u})$. That is, before (resp., after) the shift, an agent with social information and search cost $\underline{c}$ finds it optimal to discontinue search (resp., search the second alternative) if the utility of the first searched alternative is $\underline{u}$. If so, $\mathbb{P}_{\overline{u}}(u_n^{a_n} = \underline{u}) > \mathbb{P}_{\overline{u}'}(u_n^{a_n} = \underline{u})$ and, by equation (14), β is identified by

$$\mathbb{P}_{\overline{u}}(u_n^{a_n} = \underline{u}) = \alpha\beta(1-\alpha) + (1-\alpha)^2. \tag{41}$$

Since $V_{\mathrm{S}}(\underline{u})$ and $V_{\mathrm{S}}'(\underline{u})$ are not observable, the researcher must adopt a data-driven approach to determine whether the utility shift is enough to identify β. Let $\underline{u}_N^a$ (resp., $\underline{u}_N^a{}'$) be the sample moment in definition (10) calculated using the dataset with $\overline{u}$ (resp., $\overline{u}'$). Then, we have the following.

Proposition 6. *Suppose the researcher observes data on choice. Then, if $\underline{u}_N^a > \underline{u}_N^a{}'$ as $N \to \infty$, the parameter β is identified by equation (41), and*

$$\hat{\beta}_N = \frac{\underline{u}_N^a}{\alpha(1-\alpha)} - \frac{1-\alpha}{\alpha}, \tag{42}$$

is a consistent estimator of β.

Proof. By equation (14), $V_{\mathrm{S}}(\underline{u}) < \underline{c} < V_{\mathrm{S}}'(\underline{u}) \iff \mathbb{P}_{\overline{u}}(u_n^{a_n} = \underline{u}) > \mathbb{P}_{\overline{u}'}(u_n^{a_n} = \underline{u})$. By the strong law of large numbers,

$$\underline{u}_N^a \xrightarrow{a.s.} \mathbb{P}_{\overline{u}}(u_n^{a_n} = \underline{u}) \quad \text{and} \quad \underline{u}_N^a{}' \xrightarrow{a.s.} \mathbb{P}_{\overline{u}'}(u_n^{a_n} = \underline{u}). \tag{43}$$

Thus, $V_{\mathrm{S}}(\underline{u}) < \underline{c} < V_{\mathrm{S}}'(\underline{u}) \iff \mathbb{P}_{\overline{u}}(u_n^{a_n} = \underline{u}) > \mathbb{P}_{\overline{u}'}(u_n^{a_n} = \underline{u}) \iff \underline{u}_N^a > \underline{u}_N^a{}'$ a.s. as $N \to \infty$. Identification follows. Replacing $\mathbb{P}_{\overline{u}}(u_n^{a_n} = \underline{u})$ with its sample analog $\underline{u}_N^a$ in equation (41), we obtain the estimator $\hat{\beta}_N$ in equation (42), which is consistent by the convergence in (43). ■

If $\underline{c} < V_{\mathrm{S}}(\underline{u}) < V_{\mathrm{S}}'(\underline{u})$ or $V_{\mathrm{S}}(\underline{u}) < V_{\mathrm{S}}'(\underline{u}) < \underline{c}$, the population moments are the same before and after the shift: $\mathbb{P}_{\overline{u}}(u_n^{a_n} = \underline{u}) = \mathbb{P}_{\overline{u}'}(u_n^{a_n} = \underline{u})$. Thus, the available data do not suffice to detect whether β is identified by equation (41), and another approach is required.

Example. Suppose $\alpha = \frac{1}{2}$, $\underline{u} = 2$, $\overline{u} = 8$, $\overline{u}' = 14$, $\underline{u}_N^a = \frac{3}{8}$, and $\underline{u}_N^a{}' = \frac{9}{64}$. Since

$\underline{u}_N^a - \underline{u}_N^a{}' = \frac{15}{64} > 0$, by equation (42) evaluated at $\alpha = \frac{1}{2}$ and $\underline{u}_N^a = \frac{3}{8}$, the researcher obtains $\hat{\beta}_1 = \frac{1}{2}$ as an estimate of β.

5.4 Partial Identification

The previous remedies may not work either because of data limitations or because the conditions for their implementation are not satisfied. If so, the researcher can adopt a partial identification approach.

Suppose the researcher knows nothing about the level of social information beyond that $\gamma \in (0, 1]$. Let $p\colon (0, 1) \times (0, 1) \times (0, 1] \to [0, 1]$ be defined pointwise as

$$p(\alpha, \beta, \gamma) := \left[\alpha\beta(1-\alpha) + (1-\alpha)^2\right][1 - (1-\gamma)\alpha(1-\beta)].$$

The identified set for β, denoted by B^β, consists of all $\tilde{\beta} \in (0, 1)$ compatible with $\mathbb{P}(u_n^{a_n} = \underline{u})$ as a model prediction for some $\gamma \in (0, 1]$.

Definition 1. *With data on choice and no assumptions on γ, the identified set for β is*

$$B^\beta := \left\{\tilde{\beta} \in (0, 1) : p\left(\alpha, \tilde{\beta}, \gamma\right) = \mathbb{P}(u_n^{a_n} = \underline{u}) \text{ for some } \gamma \in (0, 1]\right\}. \tag{44}$$

The joint identified set for (β, γ), denoted by $B^{\beta,\gamma}$, consists of all $\left(\tilde{\beta}, \tilde{\gamma}\right) \in (0, 1) \times (0, 1]$ compatible with $\mathbb{P}(u_n^{a_n} = \underline{u})$ as a model prediction.

Definition 2. *With data on choice, the joint identified set for (β, γ) is*

$$B^{\beta,\gamma} := \left\{\left(\tilde{\beta}, \tilde{\gamma}\right) \in (0, 1) \times (0, 1] : p\left(\alpha, \tilde{\beta}, \tilde{\gamma}\right) = \mathbb{P}(u_n^{a_n} = \underline{u})\right\}. \tag{45}$$

Replacing $\mathbb{P}(u_n^{a_n} = \underline{u})$ with its sample analog $\underline{u}_N^a$ in definitions (44) and (45), we obtain the set estimators $\widehat{B}_N^\beta$ and $\widehat{B}_N^{\beta,\gamma}$. The next proposition follows from the convergence in (15).

Proposition 7. *As $N \to \infty$, $\beta \in \widehat{B}_N^\beta$ a.s. and $(\beta, \gamma) \in \widehat{B}_N^{\beta,\gamma}$ a.s.*

Example. Suppose $\alpha = \frac{1}{2}$ and $\underline{u}_N^a = \frac{5}{16}$. Then, $\widehat{B}_N^\beta = \left[\frac{1}{4}, \frac{1}{2}\left(\sqrt{10} - 2\right)\right)$, and $\widehat{B}_N^{\beta,\gamma} = \left\{\left(\tilde{\beta}, \tilde{\gamma}\right) \in (0, 1) \times (0, 1] : \tilde{\beta} \in \left[\frac{1}{4}, \frac{1}{2}\left(\sqrt{10} - 2\right)\right) \text{ and } \tilde{\gamma} = \frac{3 - 2\tilde{\beta}(2+\tilde{\beta})}{2(1-\tilde{\beta}^2)}\right\}$.

5.5 General Social Information

So far, we assumed agents are either isolated or observe an isolated agent. Real social networks are complex, and so are communication channels and informational externalities among agents. Hence, it is hard to specify the content and type of agents' social information. If so, the researcher may want to develop a robust approach to identify search cost distributions under weak assumptions on agents' information.

Let θ_n now denote the "amount" of agent n's social information. Such an amount can be bounded by the minimal and the maximal social information.

Agent n's social information is minimal, $\theta_n = \underline{\theta}$, if agent n is isolated. By equation (12),

$$\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \underline{\theta}) = \alpha\beta(1-\alpha) + (1-\alpha)^2.$$

Agent n's social information is maximal, $\theta_n = \overline{\theta}$, if agent n chooses the alternative with the highest realized utility at the first search by exploiting her social information. If so, agent n chooses an alternative with utility $\underline{u}$ if and only if $u_n^0 = u_n^1 = \underline{u}$, which occurs with probability $(1-\alpha)^2$:

$$\mathbb{P}\Big(u_n^{a_n} = \underline{u} \mid \theta_n = \overline{\theta}\Big) = (1-\alpha)^2.$$

Without assumptions on social information, the identified set for β, denoted by G^β, consists of all $\tilde{\beta} \in (0,1)$ compatible with $\mathbb{P}(u_n^{a_n} = \underline{u})$ as a model prediction for some level of social information between the minimal and the maximal ones.

Definition 3. *With data on choice and general social information, the identified set for β is*

$$G^\beta \coloneqq \Big\{\tilde{\beta} \in (0,1) : (1-\alpha)^2 \leq \mathbb{P}(u_n^{a_n} = \underline{u}) \leq \alpha\beta(1-\alpha) + (1-\alpha)^2\Big\}. \tag{46}$$

Replacing $\mathbb{P}(u_n^{a_n} = \underline{u})$ with its sample analog $\underline{u}_N^a$ in definition (46), we obtain the set estimator $\widehat{G}_N^\beta$. The next proposition follows from the convergence in (15).

Proposition 8. *As $N \to \infty$, $\beta \in \widehat{G}_N^\beta$ a.s.*

As discussed in the introduction, this approach requires specifying neither the agents' amount and type of social information nor the search procedure. Therefore, estimating β under the two opposite assumptions—all agents' social information is minimal and all agents' social information is maximal—allows the construction of robust bounds on β when agents' amount and type of social information are unobserved, and so is the search procedure.

Example. Suppose $\alpha = \frac{1}{2}$ and $\underline{u}_N^a = \frac{5}{16}$. Then, $\widehat{G}_N^\beta = \Big[\frac{1}{4}, 1\Big)$.

6 Generalizations

Throughout the paper, we presented our results with a portable model illustrating our insights in a coherent framework. In the appendices, we generalize the framework to allow for continuous distributions (Appendix A) and simultaneous search (Appendix B), the other workhorse search model. All qualitative insights remain valid. However, the analysis provides additional guidance on how to account for social information in other empirically relevant settings, study its impact on identification, and tailor remedies to the model specification.

Similar insights hold for other generalizations: more than two alternatives, non-binary utility distributions, non-perfectly correlated utilities across agents n and n_0, ex-ante differentiated alternatives, alternative-specific search costs, correlated utilities and search costs of agents n and n_0. Depending on the specification, the analysis may become algebraically, but not conceptually, more complex. However, our findings remain valid.

Our insights also apply to the identification of utility distributions. Consider a researcher who knows the search cost distribution and the support of the utility distribution, and wants to identify $\alpha \coloneqq \mathbb{P}(u_n^x = \overline{u})$. A similar analysis to that in Section 4 shows that neglecting social information leads to the non-identification of α. Similar observations apply to the joint identification of (α, β).

References

Barseghyan, Levon, Maura Coughlin, Francesca Molinari, and Joshua C. Teitelbaum (2021), “Heterogeneous Choice Sets and Preferences.” *Econometrica*, 89, 2015–2048.

Bergemann, Dirk, Benjamin Brooks, and Stephen Morris (2022), “Counterfactuals with Latent Information.” *The American Economic Review*, 112, 343–368.

Bergemann, Dirk and Stephen Morris (2016), “Bayes Correlated Equilibrium and the Comparison of Information Structures in Games.” *Theoretical Economics*, 11, 487–522.

Bikhchandani, Sushil, David Hirshleifer, Omer Tamuz, and Ivo Welch (2022), “Information Cascades and Social Learning.” *Working Paper*.

Canen, Nathan and Kyungchul Song (2022), “A Decomposition Approach to Counterfactual Analysis in Game-Theoretic Models.” *Working Paper*.

Chambers, Christopher P. and Federico Echenique (2016), *Revealed Preference Theory*. Cambridge University Press.

De Oliveira, Henrique and Rohit Lamba (2022), “Rationalizing Dynamic Choices.” *Working Paper*.

Deb, Rahul and Ludovic Renou (2021), “Dynamic Choices and Common Learning.” *Working Paper*.

Doval, Laura and Jeffrey C. Ely (2020), “Sequential Information Design.” *Econometrica*, 88, 2575–2608.

Ellison, Glenn and Alexander Wolitzky (2012), “A Search cost Model of Obfuscation.” *The RAND Journal of Economics*, 43, 417–441.

Forbes (2012), “Are Brands Wielding More Influence In Social Media Than We Thought?.” URL: https://www.forbes.com/sites/marketshare/2012/05/07/are-brands-wielding-more-influence-in-social-media-than-we-thought/.

——— (2022), “Studies Predict Social Media Will Be Biggest Influence On Shopping For Decades.” URL: https://www.forbes.com/sites/catherineerdly/2022/01/12/studies-predict-social-media-will-be-biggest-influence-on-shopping-for-decades/.

Galeotti, Andrea (2010), “Talking, Searching, and Pricing.” *International Economic Review*, 51, 1159–1174.

Garcia, Daniel and Sandro Shelegia (2018), “Consumer Search with Observational Learning.” *The RAND Journal of Economics*, 49, 224–253.

Golub, Benjamin and Evan Sadler (2016), “Learning in Social Networks.” Yann Bramoullé, Andrea Galeotti, and Brian Rogers eds., *The Oxford Handbook of the Economics of Network*. Oxford University Press.

Gualdani, Cristina and Shruti Sinha (2020), “Identification and Inference in Discrete Choice Models with Imperfect Information.” *Working Paper*.

Heidhues, Paul and Philipp Strack (2021), "Identifying Present Bias From the Timing of Choices." *The American Economic Review*, 111, 2594–2622.

Hendricks, Kenneth, Alan Sorensen, and Thomas Wiseman (2012), "Observational Learning and Demand for Search Goods." *American Economic Journal: Microeconomics*, 4, 1–31.

Heumann, Tibor (2019), "Informationally Robust Comparative Statics in Incomplete Information Games." *Working Paper*.

Honka, Elisabeth, Ali Hortaçsu, and Matthijs Wildenbeest (2019), "Empirical Search and Consideration Sets." *Handbook of the Economics of Marketing*, 1. Elsevier, 193–257.

Kircher, Philipp and Andrew Postlewaite (2008), "Strategic Firms and Endogenous Consumer Emulation." *The Quarterly Journal of Economics*, 123, 621–661.

Libgober, Jonathan (2021), "Hypothetical Beliefs Identify Information." *Working Paper*.

Liu, Shuo and Nick Netzer (2021), "Happy Times: Identification from Ordered Response Data." *Working Paper*.

Lomys, Niccolò (2020), "Collective Search in Networks." *Working Paper*.

Magnolfi, Lorenzo and Camilla Roncoroni (2022), "Estimation of Discrete Games with Weak Assumptions on Information." *The Review of Economic Studies (forthcoming)*.

Mobius, Markus and Tanya Rosenblat (2014), "Social Learning in Economics." *Annual Review of Economics*, 6, 827–847.

Mueller-Frank, Manuel and Mallesh M. Pai (2016), "Social Learning with Costly Search." *American Economic Journal: Microeconomics*, 8, 83–109.

Shmaya, Eran and Leeat Yariv (2016), "Experiments on Decisions under Uncertainty: A Theoretical Framework." *The American Economic Review*, 106, 1775–1801.

Stahl, Dale O. (1989), "Oligopolistic Pricing with Sequential Consumer Search." *The American Economic Review*, 700–712.

Stigler, George J. (1961), "The Economics of Information." *Journal of Political Economy*, 69, 213–225.

Syrgkanis, Vasilis, Elie Tamer, and Juba Ziani (2021), "Inference on Auctions with Weak Assumptions on Information." *Working Paper*.

Ursu, Raluca, Stephan Seiler, and Elisabeth Honka (2022), "The Sequential Search Model: A Framework for Empirical Research." *Working Paper*.

Varian, Hal R. (1980), "A Model of Sales." *The American Economic Review*, 70, 651–659.

Weitzman, Martin L. (1979), "Optimal Search for the Best Alternative." *Econometrica*, 43, 641–654.

A Online Appendix: Continuous Search Cost Distribution

In this appendix, we assume that agents' search costs are i.i.d. draws from a continuous distribution with full support, $c_n \sim F([0, \overline{v}])$. Otherwise, the model is as in Section 2. We summarize our findings below. We refer to the rest of the appendix for the formal analysis, which we keep short when the formalities mimic those in Sections 3–5.

Summary of Results. With a binary utility distribution, if all agents are isolated, the researcher can identify and estimate $F(V_{\text{I}}(\underline{u}))$.[4] Table C.1 summarizes our findings. For each dataset, the table illustrates: when neglecting social information leads to non-identification of $F(V_{\text{I}}(\underline{u}))$; whether search costs are under or overestimated. In short, if $c_n \sim F([0, \overline{v}])$, neglecting social information leads to non-identification of $F(V_{\text{I}}(\underline{u}))$ independently of the dataset. Whether search costs are under or overestimated depends on the dataset and is consistent with our previous analysis.

Table C.1: Identification and Estimation if $c_n \sim F([0, \overline{v}])$.

	Under- (< 0) or Over-Estimation (> 0)		
	Choice	Optimal Stopping	Number of Searches
$c_n \sim F([0, \overline{v}])$	< 0	> 0	> 0

We also consider potential remedies. Whereas shifts to the utility distribution become mute for identification purposes in this setting, the other remedies remain effective. With agent-level data on social information, it is possible to point identify both $F(V_{\text{I}}(\underline{u}))$ and $F(V_{\text{S}}(\underline{u}))$. If γ is estimated offline, point identification of $F(V_{\text{I}}(\underline{u}))$ is lost, but it is possible to partially identify $F(V_{\text{I}}(\underline{u}))$ and/or $F(V_{\text{S}}(\underline{u}))$, and so is the case without knowing γ.

A.1 Optimal Decisions

A.1.1 Search Problem of Type I

First Search. The optimal first search decision is as in Section 3.1.

[4] If $U := \{u_1, u_2, \ldots, u_K\}$, where $K > 2$, the researcher can identify and estimate $F(V_{\text{I}}(u_k))$ for all $1 \leq k \leq K$. To keep the exposition concise, we continue to assume that $|U| = 2$.

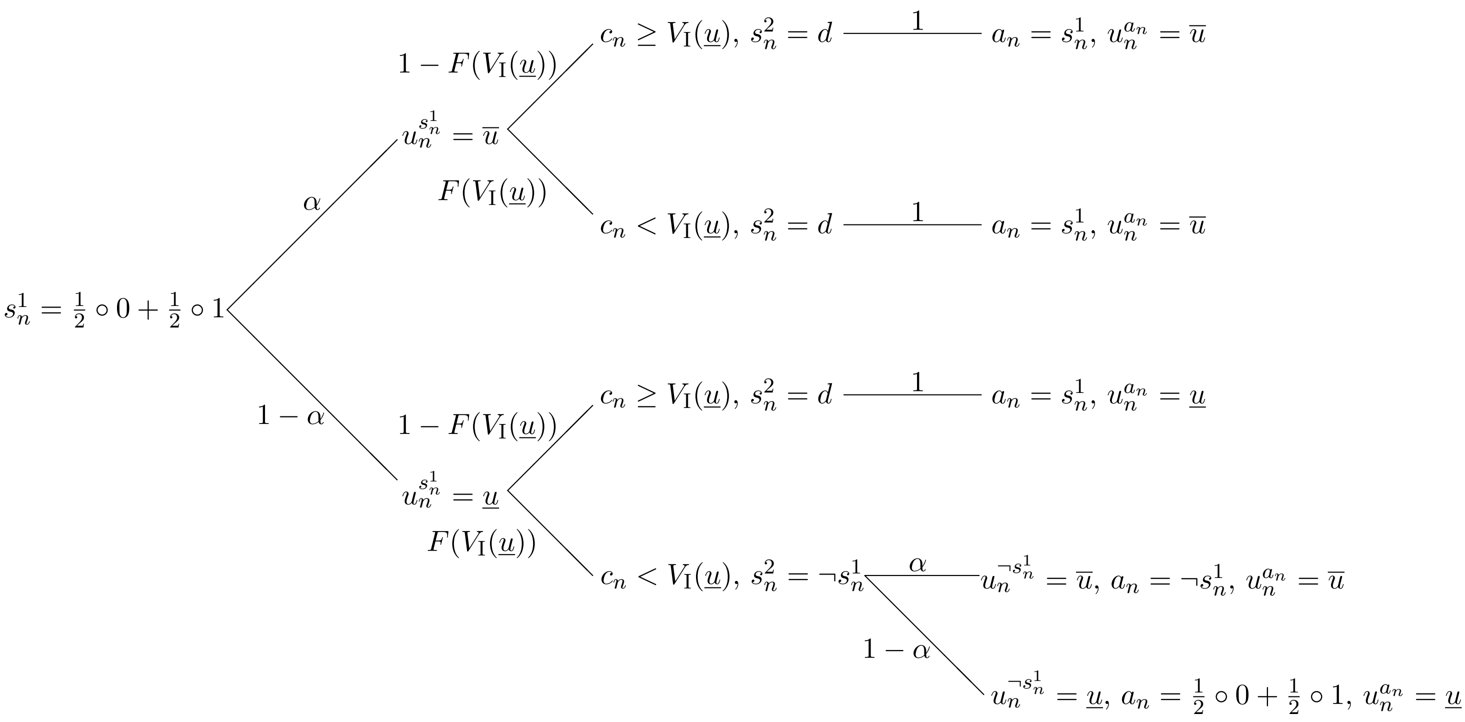
Second Search. Agent n's expected gain from the second search is given by equation (1). Thus,

$$s_n^2 = \begin{cases} d & \text{if } u_n^{s_n^1} = \overline{u} \\ d & \text{if } u_n^{s_n^1} = \underline{u} \text{ and } c_n \geq V_{\mathrm{I}}(\underline{u}) \\ \neg s_n^1 & \text{if } u_n^{s_n^1} = \underline{u} \text{ and } c_n < V_{\mathrm{I}}(\underline{u}) \end{cases}.$$

Choice. Optimal choice is as in Section 3.1.

Decision Tree. Agent n's decisions in a search problem of type I are in Figure C.1.

Figure C.1: Decision Tree for a Search Problem of Type I if $c_n \sim F([0, \overline{v}])$.



A.1.2 Search Problem of Type S

First Search. The optimal first search decision is as in Section 3.2. Therefore,

$$u_n^{s_n^1} = u_{n_0}^{a_{n_0}} = \begin{cases} \overline{u} & \text{with probability } \alpha + \alpha(1-\alpha)F(V_{\mathrm{I}}(\underline{u})) \\ \underline{u} & \text{with probability } 1 - \alpha - \alpha(1-\alpha)F(V_{\mathrm{I}}(\underline{u})) \end{cases},$$

where the probabilities are calculated from Figure C.1.

Second Search. Agent n's expected gain from the second search is given by equation (3). In this setting, from Figure C.1,

$$P(\underline{u}) \coloneqq \mathbb{P}\Big(s^2_{n_0} = d \mid u^{a_{n_0}}_{n_0} = \underline{u}\Big) = \frac{1 - F(V_\mathrm{I}(\underline{u}))}{1 - F(V_\mathrm{I}(\underline{u})) + (1-\alpha)F(V_\mathrm{I}(\underline{u}))},$$

and so

$$V_\mathrm{S}\Big(u^{s^1_n}_n\Big) = \begin{cases} 0 & \text{if } u^{s^1_n}_n = \overline{u} \\ \frac{(1-F(V_\mathrm{I}(\underline{u})))\alpha\Delta u}{1-F(V_\mathrm{I}(\underline{u}))+(1-\alpha)F(V_\mathrm{I}(\underline{u}))} & \text{if } u^{s^1_n}_n = \underline{u} \end{cases}.$$

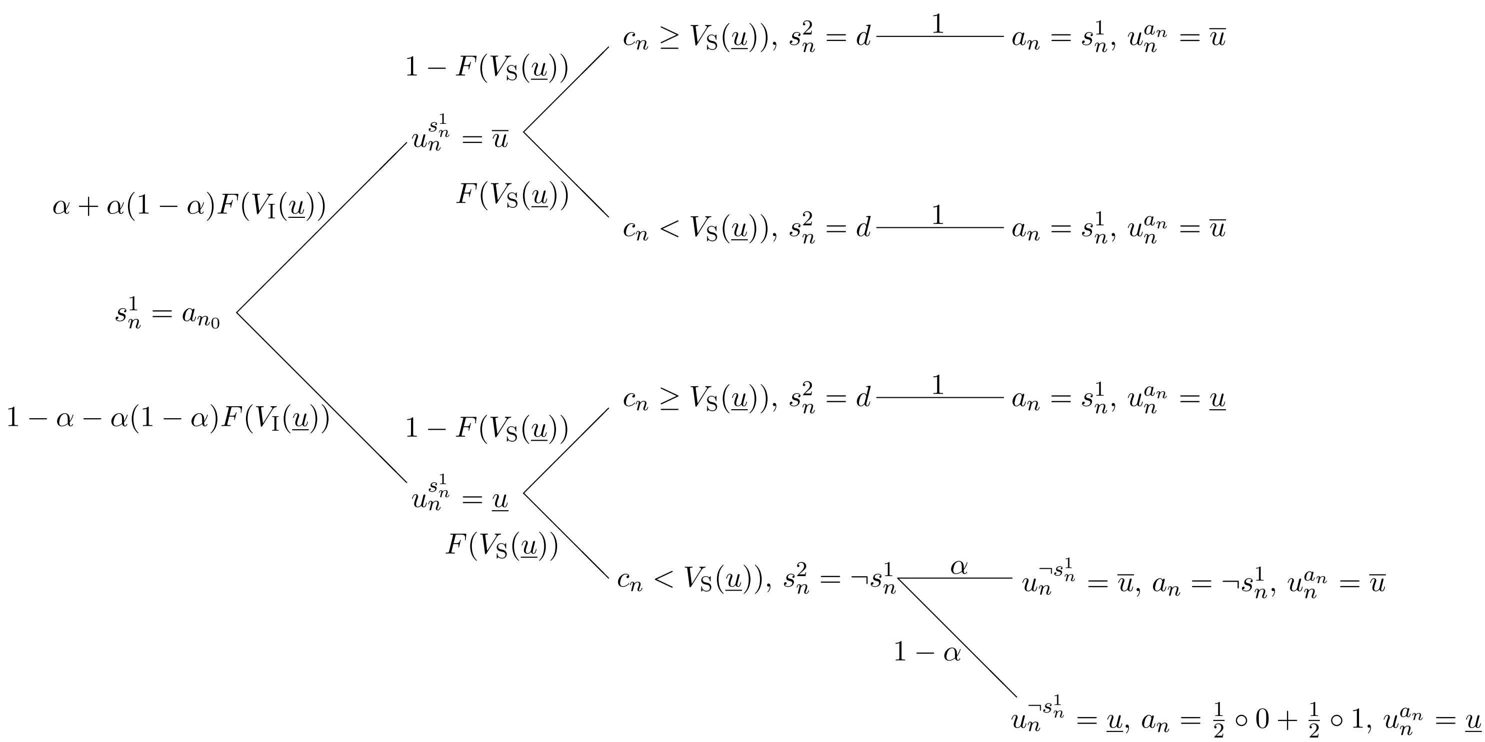
Thus,

$$s^2_n = \begin{cases} d & \text{if } u^{s^1_n}_n = \overline{u} \\ d & \text{if } u^{s^1_n}_n = \underline{u} \text{ and } c_n \geq V_\mathrm{S}(\underline{u}) \\ \neg s^1_n & \text{if } u^{s^1_n}_n = \underline{u} \text{ and } c_n < V_\mathrm{S}(\underline{u}) \end{cases}.$$

Choice. Optimal choice is as in Section 3.2.

Decision Tree. Agent n's decisions in a search problem of type S are in Figure C.2.

Figure C.2: Decision Tree for a Search Problem of Type S if $c_n \sim F([0, \overline{v}])$.



A.1.3 Comparison of Optimal Decisions

The differences between types of search problems have the following implications.

Choice. The probability with which an agent with social information chooses an alternative with utility $\underline{u}$ is smaller than that with which an isolated agent does so:

$$\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S}) < \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I}). \tag{C.1}$$

Suppose $\theta_n = \mathrm{S}$. Then, agent n searches alternative a_{n_0} first: $s_n^1 = a_{n_0}$. Since $F(V_\mathrm{S}(\underline{u})) > 0$, if $u_n^{s_n^1} = \underline{u}$, agent n searches alternative $\neg s_n^1$ with positive probability. Moreover, with positive probability, $u_n^{\neg s_n^1} = \overline{u}$, in which case $a_n = \neg s_n^1$, and so $u_n^{a_n} = \overline{u}$. Thus, $\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S}) < \mathbb{P}(u_n^{a_{n_0}} = \underline{u}) = \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I})$. Inequality (C.1) follows.

Optimal Stopping. The probability with which an agent with social information discontinues search after searching first an alternative with utility $\underline{u}$ is larger than that with which an isolated agent does so:

$$\mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}, \theta_n = \mathrm{S}\Big) > \mathbb{P}\Big(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}, \theta_n = \mathrm{I}\Big). \tag{C.2}$$

Suppose the first searched alternative has utility $\underline{u}$: $u_n^{s_n^1} = \underline{u}$. If $\theta_n = \mathrm{S}$ (resp., $\theta_n = \mathrm{I}$), agent n discontinues search if and only if $c_n \geq V_\mathrm{S}(\underline{u})$ (resp., $c_n \geq V_\mathrm{I}(\underline{u})$), which occurs with probability $1 - F(V_\mathrm{S}(\underline{u}))$ (resp., $1 - F(V_\mathrm{I}(\underline{u}))$). Since $V_\mathrm{S}(\underline{u}) < V_\mathrm{I}(\underline{u})$ and $F(\cdot)$ is increasing, we have $F(V_\mathrm{S}(\underline{u})) < F(V_\mathrm{I}(\underline{u}))$. Therefore, $1 - F(V_\mathrm{S}(\underline{u})) > 1 - F(V_\mathrm{I}(\underline{u}))$. Inequality (C.2) follows.

Number of Searches. The probability with which an agent with social information discontinues search is larger than that with which an isolated agent does so:

$$\mathbb{P}\Big(s_n^2 = d \mid \theta_n = \mathrm{S}\Big) > \mathbb{P}\Big(s_n^2 = d \mid \theta_n = \mathrm{I}\Big). \tag{C.3}$$

Suppose $\theta_n = \mathrm{S}$. Then, agent n searches alternative a_{n_0} first, $s_n^1 = a_{n_0}$, and so $\mathbb{P}\Big(u_n^{s_n^1} = \overline{u} \mid \theta_n = \mathrm{S}\Big) = \mathbb{P}\Big(u_{n_0}^{a_{n_0}} = \overline{u}\Big)$. Since the utility of the two alternatives is different with positive probability and agent n_0 searches both alternatives with positive probability, $\mathbb{P}\Big(u_{n_0}^{a_{n_0}} = \overline{u}\Big) > \mathbb{P}\Big(u_{n_0}^{s_{n_0}^1} = \overline{u}\Big) = \mathbb{P}\Big(u_n^{s_n^1} = \overline{u} \mid \theta_n = \mathrm{I}\Big)$. Thus, $\mathbb{P}\Big(u_n^{s_n^1} = \overline{u} \mid \theta_n = \mathrm{S}\Big) > \mathbb{P}\Big(u_n^{s_n^1} = \overline{u} \mid \theta_n = \mathrm{I}\Big)$. Moreover, note that: agent n discontinues search if $u_n^{s_n^1} = \overline{u}$ independently of whether $\theta_n = \mathrm{S}$ or $\theta_n = \mathrm{I}$; the probability with which agent n discontinues search if $u_n^{s_n^1} = \underline{u}$ when $\theta_n = \mathrm{S}$ is at least as large that with which agent n does so when $\theta_n = \mathrm{I}$ (see inequality (C.2)). Inequality (C.3) follows.

A.2 Identification and Social Information

Consider a researcher who wants to identify $F(V_{\mathrm{I}}(\underline{u}))$. By inequalities (C.1)–(C.3) and the same ideas as in Section 4, neglecting social information leads to non-identification of $F(V_{\mathrm{I}}(\underline{u}))$, independently of the dataset. We formalize below the argument for data on choice. With different datasets, analogous reasoning applies.

Preliminary Observations for Identification. We characterize the probability with which an agent chooses an alternative with utility $\underline{u}$ in the data generating process. By Figures C.1 and C.2,

$$\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I}) = 1 - \alpha - \alpha(1-\alpha)F(V_{\mathrm{I}}(\underline{u})) \tag{C.4}$$

and

$$\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S}) = [1 - \alpha - \alpha(1-\alpha)F(V_{\mathrm{I}}(\underline{u}))][1 - \alpha F(V_{\mathrm{S}}(\underline{u}))]. \tag{C.5}$$

By equations (11), (C.4), and (C.5),

$$\mathbb{P}(u_n^{a_n} = \underline{u}) = [1 - \alpha - \alpha(1-\alpha)F(V_{\mathrm{I}}(\underline{u}))][1 - (1-\gamma)\alpha F(V_{\mathrm{S}}(\underline{u}))]. \tag{C.6}$$

All Agents Are Isolated. Suppose $\gamma = 1$. By equations (11) and (C.4), $F(V_{\mathrm{I}}(\underline{u}))$ is identified by

$$\mathbb{P}(u_n^{a_n} = \underline{u}) = \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I}) = 1 - \alpha - \alpha(1-\alpha)F(V_{\mathrm{I}}(\underline{u})). \tag{C.7}$$

Replacing $\mathbb{P}(u_n^{a_n} = \underline{u})$ with its sample analog $\underline{u}_N^a$ in equation (C.7), we obtain

$$\widehat{F(V_{\mathrm{I}}(\underline{u}))}_N = \frac{1}{\alpha} - \frac{\underline{u}_N^a}{\alpha(1-\alpha)}, \tag{C.8}$$

which, by the convergence in (15), is a consistent estimator of $F(V_{\mathrm{I}}(\underline{u}))$.

Social Information. The next proposition summarizes the identification of $F(V_{\mathrm{I}}(\underline{u}))$ with data on choice when $\gamma < 1$, but the researcher assumes $\gamma = 1$.

Proposition C.1. *Let $\gamma < 1$. Suppose the researcher observes data on choice and assumes $\gamma = 1$. Then, $F(V_I(\underline{u}))$ is not identified by equation (C.7), the estimator $\widehat{F(V_I(\underline{u}))}_N$ in equation (C.8) is inconsistent, and search costs are underestimated.*

Proof. By equations (11) and (C.1), $\mathbb{P}(u_n^{a_n} = \underline{u}) < \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I})$. Therefore, $F(V_{\mathrm{I}}(\underline{u}))$ is not identified by equation (C.7). To see that $\widehat{F(V_{\mathrm{I}}(\underline{u}))}_N$ is inconsistent, and search costs are underestimated, note that

$$\widehat{F(V_{\mathrm{I}}(\underline{u}))}_N \xrightarrow{a.s.} \mathbb{E}\Big[\widehat{F(V_{\mathrm{I}}(\underline{u}))}_N\Big]$$

$$= F(V_{\mathrm{I}}(\underline{u})) + (1-\gamma)(1-\alpha F(V_{\mathrm{I}}(\underline{u})))F(V_{\mathrm{S}}(\underline{u})) > F(V_{\mathrm{I}}(\underline{u})),$$

where the equality holds by equation (C.6). ■

A.3 Potential Remedies

A.3.1 Agent-Level Data on Social Information

Suppose the researcher observes

$$\underline{u}_N^a(\mathrm{I}) := \frac{\sum_{n=1}^N \mathbb{1}_{\{u_n^{a_n}=\underline{u}\}}\mathbb{1}_{\{\theta_n=\mathrm{I}\}}}{\sum_{n=1}^N \mathbb{1}_{\{\theta_n=\mathrm{I}\}}} \quad \text{and} \quad \underline{u}_N^a(\mathrm{S}) := \frac{\sum_{n=1}^N \mathbb{1}_{\{u_n^{a_n}=\underline{u}\}}\mathbb{1}_{\{\theta_n=\mathrm{S}\}}}{\sum_{n=1}^N \mathbb{1}_{\{\theta_n=\mathrm{S}\}}}.$$

In this case, the researcher can implement a two-step procedure to identify and consistently estimate both $F(V_{\mathrm{I}}(\underline{u}))$ and $F(V_{\mathrm{S}}(\underline{u}))$.

Proposition C.2. *Suppose the researcher observes data on choice and agent-level data on social information. Then, $F(V_I(\underline{u}))$ and $F(V_S(\underline{u}))$ can be identified and consistently estimated.*

Proof. By the strong law of large numbers,

$$\underline{u}_N^a(\mathrm{I}) \xrightarrow{a.s.} \mathbb{E}[\underline{u}_N^a(\mathrm{I})] = \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I}). \tag{C.9}$$

By equation (C.4), $F(V_{\mathrm{I}}(\underline{u}))$ is identified by

$$\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I}) = 1 - \alpha - \alpha(1-\alpha)F(V_{\mathrm{I}}(\underline{u})). \tag{C.10}$$

Replacing $\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I})$ with its sample analog $\underline{u}_N^a(\mathrm{I})$ in equation (C.10), we obtain

$$\widehat{F(V_{\mathrm{I}}(\underline{u}))}_N = \frac{1}{\alpha} - \frac{\underline{u}_N^a(\mathrm{I})}{\alpha(1-\alpha)},$$

which, by the convergence in (C.9), is a consistent estimator of $F(V_{\mathrm{I}}(\underline{u}))$.

Once $F(V_{\mathrm{I}}(\underline{u}))$ is identified and consistently estimated, it is possible to identify and consistently estimate $F(V_{\mathrm{S}}(\underline{u}))$. By the strong law of large numbers,

$$\underline{u}_N^a(\mathrm{S}) \xrightarrow{a.s.} \mathbb{E}[\underline{u}_N^a(\mathrm{S})] = \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S}). \tag{C.11}$$

By equation (C.5), $F(V_{\mathrm{S}}(\underline{u}))$ is identified by

$$\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S}) = [1 - \alpha - \alpha(1-\alpha)F(V_{\mathrm{I}}(\underline{u}))][1 - \alpha F(V_{\mathrm{S}}(\underline{u}))]. \tag{C.12}$$

Replacing $\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S})$ with its sample analog $\underline{u}_N^a(\mathrm{S})$ and $F(V_{\mathrm{I}}(\underline{u}))$ with its estimator $\widehat{F(V_{\text{I}}(\underline{u}))}_N$ in equation (C.12), we obtain

$$\widehat{F(V_{\text{S}}(\underline{u}))}_N = \frac{1}{\alpha} - \frac{\underline{u}_N^a(\text{S})}{\alpha\Big[1-\alpha-\alpha(1-\alpha)\widehat{F(V_{\text{I}}(\underline{u}))}_N\Big]},$$

which, by the convergences in (C.9) and (C.11), is a consistent estimator of $F(V_{\text{S}}(\underline{u}))$. ■

A.3.2 Partial Identification Estimating γ Offline

Suppose the researcher can identify and consistently estimate γ offline. By equation (C.6), it is clear that neither $F(V_{\text{I}}(\underline{u}))$ nor $F(V_{\text{S}}(\underline{u}))$ can be point-identified. However, the researcher can rely on a partial identification approach. Let $p\colon (0,1)\times(0,1)\times(0,1)\times(0,1]\to[0,1]$ be defined pointwise as

$$p(\alpha, F(V_{\text{I}}(\underline{u})), F(V_{\text{S}}(\underline{u})), \gamma) \coloneqq [1-\alpha-\alpha(1-\alpha)F(V_{\text{I}}(\underline{u}))][1-(1-\gamma)\alpha F(V_{\text{S}}(\underline{u}))].$$

The joint identified set for $(F(V_{\text{I}}(\underline{u})), F(V_{\text{S}}(\underline{u})))$ given γ, denoted by $B^{F^{\text{I}},F^{\text{S}}}(\gamma)$, consists of all $\Big(\widetilde{F(V_{\text{I}}(\underline{u}))}, \widetilde{F(V_{\text{S}}(\underline{u}))}\Big) \in (0,1)\times(0,1)$ compatible with $\mathbb{P}(u_n^{a_n}=\underline{u})$ as a model prediction for the given γ.

Definition C.1. *Suppose the researcher observes data on choice and estimates γ offline. The joint identified set for $(F(V_{\text{I}}(\underline{u})), F(V_{\text{S}}(\underline{u})))$ given γ is*

$$\begin{aligned} B^{F^{\text{I}},F^{\text{S}}}(\gamma) \coloneqq \Big\{\Big(\widetilde{F(V_{\text{I}}(\underline{u}))}, \widetilde{F(V_{\text{S}}(\underline{u}))}\Big) \in (0,1)\times(0,1) : \\ p\Big(\alpha, \widetilde{F(V_{\text{I}}(\underline{u}))}, \widetilde{F(V_{\text{S}}(\underline{u}))}, \gamma\Big) = \mathbb{P}(u_n^{a_n}=\underline{u}) \text{ and } F(V_{\text{S}}(\underline{u})) \le F(V_{\text{I}}(\underline{u}))\Big\}. \end{aligned} \tag{C.13}$$

The restriction $F(V_{\text{S}}(\underline{u})) \le F(V_{\text{I}}(\underline{u}))$ in definition (C.13) captures that, conditional on the first searched alternative having utility $\underline{u}$, the expected gain from the second search for an agent with social information is lower than that for an isolated agent.

The identified set for $F(V_{\text{I}}(\underline{u}))$ (resp., $F(V_{\text{S}}(\underline{u}))$) given γ, denoted by $B^{F^{\text{I}}}(\gamma)$ (resp., $B^{F^{\text{S}}}(\gamma)$), consists of all $\widetilde{F(V_{\text{I}}(\underline{u}))} \in (0,1)$ (resp., $\widetilde{F(V_{\text{S}}(\underline{u}))} \in (0,1)$) compatible with $\mathbb{P}(u_n^{a_n}=\underline{u})$ as a model prediction for the given γ and some $\widetilde{F(V_{\text{S}}(\underline{u}))} \le \widetilde{F(V_{\text{I}}(\underline{u}))}$ (resp., $\widetilde{F(V_{\text{I}}(\underline{u}))} \ge \widetilde{F(V_{\text{S}}(\underline{u}))}$).

Definition C.2. *Suppose the researcher observes data on choice and estimates γ offline. The identified set for $F(V_{\text{I}}(\underline{u}))$ (resp., $F(V_{\text{S}}(\underline{u}))$) given γ, denoted by $B^{F^{\text{I}}}(\gamma)$ (resp., $B^{F^{\text{S}}}(\gamma)$), is the projection of $B^{F^{\text{I}},F^{\text{S}}}(\gamma)$ along its first (resp., second) dimension.*

Replacing $\mathbb{P}(u_n^{a_n} = \underline{u})$ with its sample analog $\underline{u}_N^a$ in the definitions of $B^{F^{\mathrm{I}},F^{\mathrm{S}}}(\gamma)$, $B^{F^{\mathrm{I}}}(\gamma)$, and $B^{F^{\mathrm{S}}}(\gamma)$, we obtain the set estimators $\widehat{B}_N^{F^{\mathrm{I}},F^{\mathrm{S}}}(\gamma)$, $\widehat{B}_N^{F^{\mathrm{I}}}(\gamma)$, and $\widehat{B}_N^{F^{\mathrm{S}}}(\gamma)$. The next proposition follows from the convergence in (15).

Proposition C.3. *As $N \to \infty$, $(F(V_I(\underline{u})), F(V_S(\underline{u}))) \in B^{F^I,F^S}(\gamma)$ a.s., $F(V_I(\underline{u})) \in B^{F^I}(\gamma)$ a.s., and $F(V_S(\underline{u})) \in B^{F^S}(\gamma)$ a.s.*

A.3.3 Partial Identification without Knowing γ

Suppose the researcher only knows that $\gamma \in (0, 1]$. The joint identified set for $(F(V_{\mathrm{I}}(\underline{u})), F(V_{\mathrm{S}}(\underline{u})))$, denoted by$B^{F^{\mathrm{I}},F^{\mathrm{S}}}$, consists of all $\left(F(\widetilde{V_{\mathrm{I}}(\underline{u})}), F(\widetilde{V_{\mathrm{S}}(\underline{u})})\right) \in (0, 1) \times (0, 1)$ compatible with $\mathbb{P}(u_n^{a_n} = \underline{u})$ as a model prediction for some $\gamma \in (0, 1]$.

Definition C.3. *With data on choice and no assumptions on γ, the joint identified set for $(F(V_{\mathrm{I}}(\underline{u})), F(V_{\mathrm{S}}(\underline{u})))$ is*

$$B^{F^{\mathrm{I}},F^{\mathrm{S}}} := \bigcup_{\gamma \in (0,1]} B^{F^{\mathrm{I}},F^{\mathrm{S}}}(\gamma). \tag{C.14}$$

The identified set for $F(V_{\mathrm{I}}(\underline{u}))$ (resp., $F(V_{\mathrm{S}}(\underline{u})))$ given γ, denoted by $B^{F^{\mathrm{I}}}$ (resp., $B^{F^{\mathrm{S}}}$), consists of all $F(\widetilde{V_{\mathrm{I}}(\underline{u})}) \in (0, 1)$ (resp., $F(\widetilde{V_{\mathrm{S}}(\underline{u})}) \in (0, 1)$) compatible with $\mathbb{P}(u_n^{a_n} = \underline{u})$ as a model prediction for $\gamma \in (0, 1]$ and some $F(\widetilde{V_{\mathrm{S}}(\underline{u})}) \leq F(\widetilde{V_{\mathrm{I}}(\underline{u})})$ (resp., $F(\widetilde{V_{\mathrm{I}}(\underline{u})}) \geq F(\widetilde{V_{\mathrm{S}}(\underline{u})})$).

Definition C.4. *With data on choice and no assumptions on γ, the identified set for $F(V_{\mathrm{I}}(\underline{u}))$ (resp., $F(V_{\mathrm{S}}(\underline{u}))$), denoted by $B^{F^{\mathrm{I}}}$ (resp., $B^{F^{\mathrm{S}}}$), is the projection of $B^{F^{\mathrm{I}},F^{\mathrm{S}}}$ along its first (resp., second) dimension.*

Replacing $\mathbb{P}(u_n^{a_n} = \underline{u})$ with its sample analog $\underline{u}_N^a$ in the definitions of $B^{F^{\mathrm{I}},F^{\mathrm{S}}}$, $B^{F^{\mathrm{I}}}$, and $B^{F^{\mathrm{S}}}$, we obtain the set estimators $\widehat{B}_N^{F^{\mathrm{I}},F^{\mathrm{S}}}$, $\widehat{B}_N^{F^{\mathrm{I}}}$, and $\widehat{B}_N^{F^{\mathrm{S}}}$. The next proposition follows from the convergence in (15).

Proposition C.4. *As $N \to \infty$, $(F(V_{\mathrm{I}}(\underline{u})), F(V_{\mathrm{S}}(\underline{u}))) \in \widehat{B}_N^{F^{\mathrm{I}},F^{\mathrm{S}}}$ a.s., $F(V_{\mathrm{I}}(\underline{u})) \in \widehat{B}_N^{F^{\mathrm{I}}}$ a.s., and $F(V_{\mathrm{S}}(\underline{u})) \in \widehat{B}_N^{F^{\mathrm{S}}}$ a.s.*

The analysis for the joint partial identification of $(F(V_{\mathrm{I}}(\underline{u})), F(V_{\mathrm{S}}(\underline{u}), \gamma)$ mimics that above and in Section 5.4.

B Online Appendix: Simultaneous Search

In this section, we replace sequential search with simultaneous search. Otherwise, the model is as in Section 2. We summarize our findings below. We refer to the rest of Appendix B for the formal analysis.

Summary of Results. Table D.1 illustrates when neglecting social information leads to non-identification of search cost distributions, whether search costs are under or overestimated, and how this depends on the dataset.[5]

Table D.1: Identification and Estimation in a Simultaneous Search Model.

	Under- (<0) or Over-Estimation (>0)	
	Choice	Number of Searches
$\underline{c} < \Delta(V_{\mathrm{S}})$	< 0	NO
$\underline{c} > \Delta(V_{\mathrm{S}})$	NO	> 0

The main insights are consistent with our previous analysis, with only one difference worth noting. In a simultaneous search model, if $\underline{c} < \Delta(V_{\mathrm{S}})$, β is identified with data on the number of searches, even neglecting social information. In contrast, in a sequential search model, with data on the number of searches, β is never identified with social information.

The remedies that we discuss in Section 5 for a sequential search model also apply to a simultaneous search model, with the obvious changes.

B.1 Model

Basic Setting. Consider (a stylized version of) the canonical simultaneous search model by Stigler (1961) as developed by the empirical search literature (see Honka et al., 2019). The basic setting is as in Section 2, but now agent n collects information about the realized utilities via costly simultaneous search:

1. Agent n commits to searching a fixed set of alternatives $S_n \in 2^X \setminus \{\emptyset\}$, where 2^X denotes the power set of the set X. By searching alternative x, agent n perfectly learns its realized utility u_n^x.

2. After searching the alternatives in S_n, agent n chooses an alternative $a_n \in S_n$.

Searching one alternative is free. Searching both alternatives costs $c_n \in \{\underline{c}, \overline{c}\}$, where $0 < \underline{c} < \alpha(1-\alpha)\Delta u < \overline{c}$. Absent this assumption, a search problem of type I is trivial: an agent would always commit to searching either both or only one alternative, irrespective of her search cost.

Social Information. Social information is as in Section 2, with one exception: all search problems, including that of agent n_0, are simultaneous search problems.

[5] As optimal stopping is not defined for simultaneous search, we do not consider such data.

If $\theta_n = \mathrm{S}$, agent n observes neither agent n_0's search cost nor agent n_0's decisions of how many and which alternatives to search.

B.2 Optimal Decisions

B.2.1 Simultaneous Search Problem of Type I

Search. Agent n commits to searching a fixed set of alternatives $S_n \in 2^X \setminus \{\emptyset\}$ that maximizes her expected utility from searching that set net of search costs. That is,

$$S_n \in \arg\max_{S \in 2^X \setminus \{\emptyset\}} [V_{\mathrm{I}}(S) - c_n(|S| - 1)],$$

where, for all $S \in 2^X \setminus \{\emptyset\}$,

$$V_{\mathrm{I}}(S) := \mathbb{E}\left[\max_{x \in S}\{u_n^x\}\right].$$

Clearly,

$$V_{\mathrm{I}}(S) = \begin{cases} \underline{u} + \alpha \Delta u & \text{if } S = \{0\} \text{ or } S = \{1\} \\ \overline{u} - (1-\alpha)^2 \Delta u & \text{if } S = X \end{cases}. \tag{D.1}$$

For all $x \in X$, $V_{\mathrm{I}}(X) - c_n > V_{\mathrm{I}}(\{x\}) \iff c_n < \alpha(1-\alpha)\Delta u$. Thus, since $0 < \underline{c} < \alpha(1-\alpha)\Delta u < \overline{c}$,

$$S_n = \begin{cases} \frac{1}{2} \circ \{0\} + \frac{1}{2} \circ \{1\} & \text{if } c_n = \overline{c} \\ X & \text{if } c_n = \underline{c} \end{cases},$$

where $\sum_x \xi(x) \circ \{x\}$ denotes the mixture that assigns probability $\xi(x)$ to set $\{x\}$. Hereafter, we denote by s_n the alternative in S_n whenever $|S_n| = 1$.

Since utilities are i.i.d.:

- If $S_n = \frac{1}{2} \circ \{0\} + \frac{1}{2} \circ \{1\}$,

$$u_n^{s_n} = \begin{cases} \overline{u} & \text{with probability } \alpha \\ \underline{u} & \text{with probability } 1-\alpha \end{cases}.$$

- If $S_n = X$,

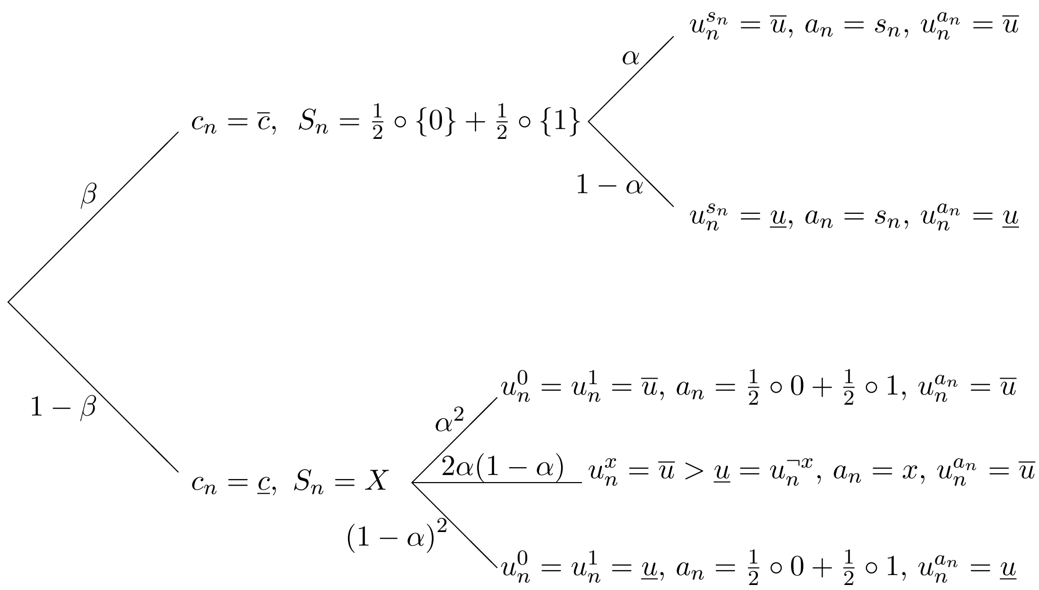
$$\begin{cases} u_n^0 = u_n^1 = \overline{u} & \text{with probability } \alpha^2 \\ u_n^x = \overline{u} > \underline{u} = u_n^{\neg x} & \text{with probability } 2\alpha(1-\alpha) \\ u_n^0 = u_n^1 = \underline{u} & \text{with probability } (1-\alpha)^2 \end{cases}.$$

Choice. Agent n chooses the best alternative among those she sampled, randomizing uniformly if indifferent:

$$a_n = \begin{cases} s_n & \text{if } S_n = \frac{1}{2} \circ \{0\} + \frac{1}{2} \circ \{1\} \\ x & \text{if } S_n = X \text{ and } u_n^x > u_n^{\neg x} \\ \frac{1}{2} \circ 0 + \frac{1}{2} \circ 1 & \text{if } S_n = X \text{ and } u_n^x = u_n^{\neg x} \end{cases}.$$

Decision Tree. Agent n's decisions in a simultaneous search problem of type I are in Figure D.1.

Figure D.1: Decision Tree for a Simultaneous Search Problem of Type I.



B.2.2 Simultaneous Search Problem of Type S

Search. Agent n commits to searching a fixed set of alternatives $S_n \in 2^X \setminus \{\emptyset\}$ that maximizes her expected utility from searching that set net of search costs. That is,

$$S_n \in \arg\max_{S \in 2^X \setminus \{\emptyset\}} [V_{\mathrm{S}}(S) - c_n(|S| - 1)],$$

where, for all $S \in 2^X \setminus \{\emptyset\}$,

$$V_{\mathrm{S}}(S) \coloneqq \mathbb{E}\left[\max_{x\in S}\{u_n^x\} \mid a_{n_0}\right]. \tag{D.2}$$

In a simultaneous search problem of type S, there are two possibilities, each having positive probability:

1. Agent n_0 searched only one alternative. If so, agent n_0's choice is uninformative about the utility of alternative $\neg a_{n_0}$.
2. Agent n_0 searched both alternative. If so, since agent n_0 chose alternative a_{n_0}, it must be that $u_n^{a_{n_0}} \geq u_n^{\neg a_{n_0}}$ and, with positive probability $u_n^{a_{n_0}} > u_n^{\neg a_{n_0}}$.

Agent n's belief about the utility of alternative a_{n_0} strictly first-order stochastically dominates her belief about the utility of alternative $\neg a_{n_0}$, and so, $\mathbb{E}\left[u_n^{a_{n_0}} \mid a_{n_0}\right] > \mathbb{E}\left[u_n^{\neg a_{n_0}} \mid a_{n_0}\right]$. Thus, if agent n commits to searching only one alternative, she will search alternative a_{n_0}: $S_n = \{a_{n_0}\}$. This is the first difference between types of simultaneous search problems.

Since

$$u_{n_0}^{a_{n_0}} = \begin{cases} \overline{u} & \text{with probability } \alpha + \alpha(1-\alpha)(1-\beta) \\ \underline{u} & \text{with probability } \alpha\beta(1-\alpha) + (1-\alpha)^2 \end{cases},$$

where the probabilities are calculated from Figure D.1, in a simultaneous search problem of type S, we have

$$V_{\mathrm{S}}(S) = \begin{cases} \beta[\underline{u} + \alpha\Delta u] + (1-\beta)\left[\overline{u} - (1-\alpha)^2\Delta u\right] & \text{if } S = \{a_{n_0}\} \\ \overline{u} - (1-\alpha)^2\Delta u & \text{if } S = X \end{cases}. \tag{D.3}$$

The value of committing to search only one alternative (namely, alternative a_{n_0}) for an agent with social information is larger than that for an isolated agent: $V_{\mathrm{S}}(\{a_{n_0}\}) > V_{\mathrm{I}}(\{x\})$ for all $x \in X$ (compare equations (D.1) and (D.3) for $|S| = 1$). This is the second difference between types of simultaneous search problems.

Since $V_{\mathrm{S}}(X) - c_n > V_{\mathrm{S}}(\{a_{n_0}\}) \iff c_n < V_{\mathrm{S}}(X) - V_{\mathrm{S}}(\{a_{n_0}\}) = \alpha(1-\alpha)\beta\Delta u$, by defining $\Delta(V_{\mathrm{S}}) \coloneqq V_{\mathrm{S}}(X) - V_{\mathrm{S}}(\{a_{n_0}\})$:[6]

- If $\underline{c} < \Delta(V_{\mathrm{S}})$,

$$S_n = \begin{cases} \{a_{n_0}\} & \text{if } c_n = \overline{c} \\ X & \text{if } c_n = \underline{c} \end{cases},$$

- If $\underline{c} > \Delta(V_{\mathrm{S}})$,

$$S_n = \{a_{n_0}\}.$$

[6]Since non-generic in the parameter space, we ignore $\underline{c} = \Delta(V_{\mathrm{S}})$.

Now we have:

- If $S_n = \{a_{n_0}\}$,

$$u_n^{s_n} = u_{n_0}^{a_{n_0}}.$$

- If $S_n = X$,

$$\begin{cases} u_n^0 = u_n^1 = \overline{u} & \text{with probability } \alpha^2 \\ u_n^x = \overline{u} > \underline{u} = u_n^{\neg x} & \text{with probability } 2\alpha(1-\alpha) \\ u_n^0 = u_n^1 = \underline{u} & \text{with probability } (1-\alpha)^2 \end{cases}.$$

Choice. Optimal choice is as in a simultaneous search problem of type I.

Decision Tree. Agent n's decisions in a simultaneous search problem of type S are in Figure D.2: Panel A if $\underline{c} < \Delta(V_{\mathrm{S}})$ and Panel B if $\underline{c} > \Delta(V_{\mathrm{S}})$.

B.2.3 Comparison of Optimal Decisions

The differences between types of search problems have the following implications.

Choice. The probability with which an agent with social information chooses an alternative with utility $\underline{u}$ is smaller than that with which an isolated agent does so if $\underline{c} < \Delta(V_{\mathrm{S}})$,

$$\underline{c} < \Delta(V_{\mathrm{S}}) \Longrightarrow \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S}) < \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I}), \tag{D.4}$$

and the same as that one if $\underline{c} > \Delta(V_{\mathrm{S}})$,

$$\underline{c} > \Delta(V_{\mathrm{S}}) \Longrightarrow \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S}) = \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I}). \tag{D.5}$$

Suppose $\theta_n = \mathrm{S}$.

- If $\underline{c} < \Delta(V_{\mathrm{S}})$, there are two cases:

  (i) If $c_n = \overline{c}$, agent n commits to searching only alternative a_{n_0}, and so $\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S}) = \mathbb{P}(u_{n_0}^{a_{n_0}} = \underline{u}) = \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I})$.

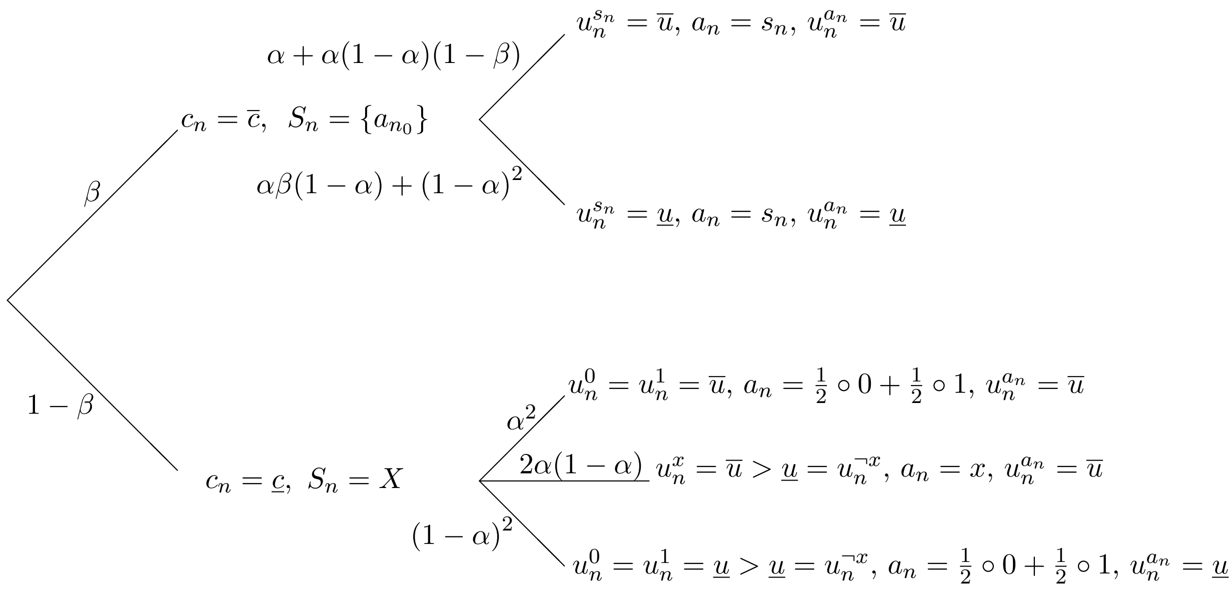
  (ii) If $c_n = \underline{c}$, agent n commits to searching both alternatives; with positive probability, $u_n^{\neg a_{n_0}} = \overline{u} > \underline{u} = u_n^{a_{n_0}}$, in which case agent n chooses alternative $a_n = \neg a_{n_0}$, so that $u_n^{a_n} = \overline{u}$. Thus, $\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S}) < \mathbb{P}(u_n^{a_{n_0}} = \underline{u}) = \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I})$.

  Since (i) and (ii) occur with positive probability, implication (D.4) follows.

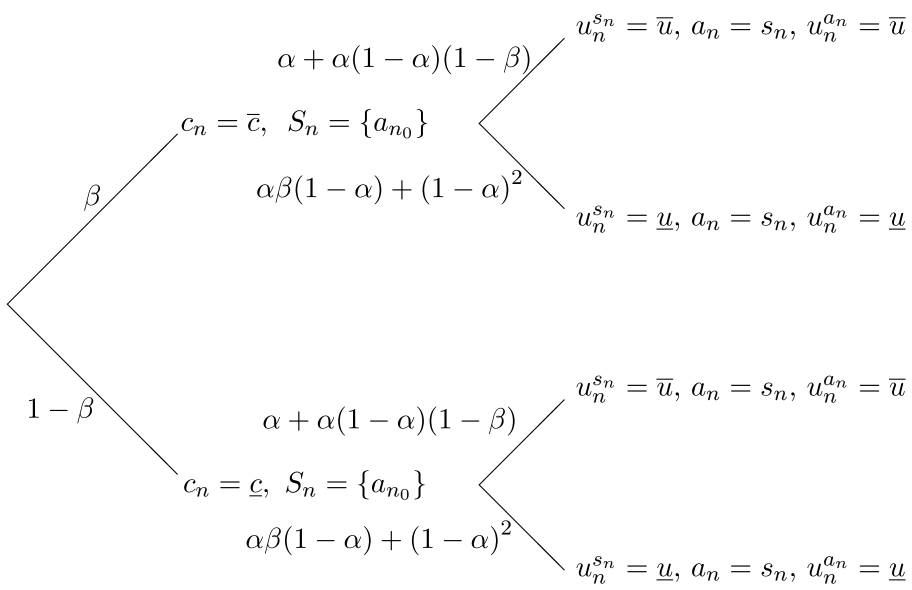
- If $\underline{c} > \Delta(V_{\mathrm{S}})$, agent n always commits to searching only alternative a_{n_0}, and so $\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S}) = \mathbb{P}(u_{n_0}^{a_{n_0}} = \underline{u}) = \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I})$. Implication

Figure D.2: Decision Trees for a Simultaneous Search Problem of Type S.

Panel A: $\underline{c} < \Delta(V_{\text{S}})$.



Panel B: $\underline{c} > \Delta(V_{\text{S}})$.



(D.5) follows.

Number of Searches. The probability with which an agent with social information commits to searching only one alternative is the same as that with which an isolated agent does so if $\underline{c} < \Delta(V_{\mathrm{S}})$,

$$\underline{c} < \Delta(V_{\mathrm{S}}) \Longrightarrow \mathbb{P}(|S_n| = 1 \mid \theta_n = \mathrm{S}) = \mathbb{P}(|S_n| = 1 \mid \theta_n = \mathrm{I}), \tag{D.6}$$

and larger than that one if $\underline{c} > \Delta(V_{\mathrm{S}})$,

$$\underline{c} > \Delta(V_{\mathrm{S}}) \Longrightarrow \mathbb{P}(|S_n| = 1 \mid \theta_n = \mathrm{S}) > \mathbb{P}(|S_n| = 1 \mid \theta_n = \mathrm{I}). \tag{D.7}$$

Note the following.

- If $\underline{c} < \Delta(V_{\mathrm{S}})$, agent n commits to searching only one alternative if and only if $c_n = \bar{c}$ independently of whether $\theta_n = \mathrm{S}$ or $\theta_n = \mathrm{I}$. Implication (D.6) follows.
- If $\underline{c} > \Delta(V_{\mathrm{S}})$, agent n always commits to searching only one alternative when $\theta_n = \mathrm{S}$, but commits to searching only one alternative if and only if $c_n = \bar{c}$ when $\theta_n = \mathrm{I}$. Implication (D.7) follows.

B.3 Identification and Social Information

B.3.1 Data on Choice

Data on choice are the same as in Section 4.

Preliminary Observations for Identification. We characterize the probability with which an agent chooses an alternative with utility $\underline{u}$ in the data generating process. By the law of total probability,

$$\mathbb{P}(u_n^{a_n} = \underline{u}) = \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I})\gamma + \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S})(1-\gamma). \tag{D.8}$$

By Figures D.1 and D.2,

$$\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I}) = \alpha\beta(1-\alpha) + (1-\alpha)^2 \tag{D.9}$$

and

$$\mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{S}) = \begin{cases} \alpha\beta^2(1-\alpha) + (1-\alpha)^2 & \text{if } \underline{c} < \Delta(V_{\mathrm{S}}) \\ \alpha\beta(1-\alpha) + (1-\alpha)^2 & \text{if } \underline{c} > \Delta(V_{\mathrm{S}}) \end{cases}. \tag{D.10}$$

By equations (D.8)–(D.10),

$$\mathbb{P}(u_n^{a_n} = \underline{u}) = \begin{cases} \alpha\beta(1-\alpha)[\gamma + (1-\gamma)\beta] + (1-\alpha)^2 & \text{if } \underline{c} < \Delta(V_{\mathrm{S}}) \\ \alpha\beta(1-\alpha) + (1-\alpha)^2 & \text{if } \underline{c} > \Delta(V_{\mathrm{S}}) \end{cases}. \tag{D.11}$$

By the strong law of large numbers,

$$\underline{u}_N^a \xrightarrow{a.s.} \mathbb{E}[\underline{u}_N^a] = \mathbb{P}(u_n^{a_n} = \underline{u}). \tag{D.12}$$

All Agents Are Isolated. Suppose $\gamma = 1$. By equations (D.8) and (D.9), β is identified by

$$\mathbb{P}(u_n^{a_n} = \underline{u}) = \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I}) = \alpha\beta(1-\alpha) + (1-\alpha)^2. \tag{D.13}$$

Replacing $\mathbb{P}(u_n^{a_n} = \underline{u})$ with its sample analog $\underline{u}_N^a$ in equation (D.13), we obtain

$$\hat{\beta}_N^1 = \frac{\underline{u}_N^a}{\alpha(1-\alpha)} - \frac{1-\alpha}{\alpha}, \tag{D.14}$$

which, by the convergence in (D.12), is a consistent estimator of β.

Social Information. The next proposition summarizes the identification of β with data on choice when $\gamma < 1$, but the researcher assumes $\gamma = 1$.

Proposition D.1. *Let $\gamma < 1$. Suppose the researcher observes data on choice and assumes $\gamma = 1$. Then:*

(i) If $\underline{c} < V_S(\underline{u})$, the parameter β is not identified by equation (D.13), the estimator $\hat{\beta}_N^1$ in equation (D.14) is inconsistent, and search costs are underestimated.

(ii) If $\underline{c} > V_S(\underline{u})$, the parameter β is identified by equation (D.13), and the estimator $\hat{\beta}_N^1$ in equation (D.14) is consistent.

Proof. [Part(i)] If $\underline{c} < \Delta(V_{\mathrm{S}})$, by equations (D.4) and (D.8), $\mathbb{P}(u_n^{a_n} = \underline{u}) < \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I})$. Therefore, β is not identified by equation (D.13). To see that $\hat{\beta}_N^1$ is inconsistent, and search costs are underestimated, note that

$$\hat{\beta}_N^1 \xrightarrow{a.s.} \mathbb{E}\big[\hat{\beta}_N^1\big] = \gamma\beta + (1-\gamma)\beta^2 < \beta,$$

where the equality holds by equation (D.11) for $\underline{c} < \Delta(V_{\mathrm{S}})$.

[Part(ii)] If $\underline{c} > \Delta(V_{\mathrm{S}})$, by equations (D.5) and (D.8), $\mathbb{P}(u_n^{a_n} = \underline{u}) = \mathbb{P}(u_n^{a_n} = \underline{u} \mid \theta_n = \mathrm{I})$. ■

B.3.2 Data on the Number of Searches

Data on the number of searches are the same as in Section 4. Formally, the researcher now observes

$$d_N := \frac{\sum_{n=1}^N \mathbb{1}_{\{|S_n|=1\}}}{N}.$$

Preliminary Observations for Identification. We characterize the probability with which an agent commits to search only one alternative in the data generating process. By the law of total probability,

$$\mathbb{P}(|S_n|=1)=\mathbb{P}(|S_n|=1\mid\theta_n=\mathrm{I})\gamma+\mathbb{P}(|S_n|=1\mid\theta_n=\mathrm{S})(1-\gamma). \tag{D.15}$$

By Figures D.1 and D.2,

$$\mathbb{P}(|S_n| = 1 \mid \theta_n = \mathrm{I}) = \beta \tag{D.16}$$

and

$$\mathbb{P}(|S_n| = 1 \mid \theta_n = \mathrm{S}) = \begin{cases} \beta & \text{if } \underline{c} < \Delta(V_\mathrm{S}) \\ 1 & \text{if } \underline{c} > \Delta(V_\mathrm{S}) \end{cases}. \tag{D.17}$$

By equations (D.15)–(D.17),

$$\mathbb{P}(|S_n| = 1) = \begin{cases} \beta & \text{if } \underline{c} < \Delta(V_\mathrm{S}) \\ \gamma\beta + (1-\gamma) & \text{if } \underline{c} > \Delta(V_\mathrm{S}) \end{cases}. \tag{D.18}$$

By the strong law of large numbers,

$$d_N \xrightarrow{a.s.} \mathbb{E}[d_N] = \mathbb{P}(|S_n| = 1). \tag{D.19}$$

All Agents Are Isolated. Suppose $\gamma = 1$. By equations (D.15) and (D.16), β is identified by

$$\mathbb{P}(|S_n| = 1) = \mathbb{P}(|S_n| = 1 \mid \theta_n = \mathrm{I}) = \beta. \tag{D.20}$$

Replacing $\mathbb{P}\left(s_n^2 = d \mid u_n^{s_n^1} = \underline{u}\right)$ with its sample analog $d_N^{\underline{u}}$ in equation (D.20), we obtain

$$\hat{\beta}_N^2 := d_N \tag{D.21}$$

which, by the convergence in (D.19), is a consistent estimator of β.

Social Information. The next proposition summarizes the identification of β with data on optimal stopping when $\gamma < 1$ but the researcher assumes $\gamma = 1$.

Proposition D.2. *Let $\gamma < 1$. Suppose the researcher observes data on the number of searches and assumes $\gamma = 1$. Then:*

(i) If $\underline{c} < V_S(\underline{u})$, the parameter β is identified by equation (D.20), and the estimator $\hat{\beta}_N^2$ in equation (D.21) is consistent.

(ii) If $\underline{c} > V_S(\underline{u})$, the parameter β is not identified by equation (D.20), the estimator $\hat{\beta}_N^2$ in equation (D.21) is inconsistent, and search costs are overestimated.

Proof. [Part(i)] If $\underline{c} < \Delta(V_\mathrm{S})$, by equations (D.6) and (D.15), $\mathbb{P}(|S_n| = 1) = \mathbb{P}(|S_n| = 1 \mid \theta_n = \mathrm{I})$.

[Part(ii)] If $\underline{c} > \Delta(V_{\mathrm{S}})$, by equations (D.7) and (D.15), $\mathbb{P}(|S_n| = 1) > \mathbb{P}(|S_n| = 1 \mid \theta_n = \mathrm{I})$. Therefore, the parameter β is not identified by equation (D.20). To see that $\hat{\beta}_N^2$ is inconsistent, and search costs are overestimated, note that

$$\hat{\beta}_N^2 \xrightarrow{a.s.} \mathbb{E}\left[\hat{\beta}_N^2\right] = \gamma\beta + (1-\gamma) > \beta,$$

where the equality holds by equation (D.18) for $\underline{c} > \Delta(V_{\mathrm{S}})$. ■