Department of Economics

DISCUSSION PAPER SERIES IN ECONOMICS

DP No.06

Choice via Social Influence

December 2018

Abhinash Borah
Christopher Kops

https://www.ashoka.edu.in/ecodp

Plot No. 2, Rajiv Gandhi Education City,
National Capital Region
P.O.Rai,Sonepat Haryana-131029

Choice via Social Influence

Abhinash Borah, Christopher Kops*

December 2, 2018

Abstract

We introduce a theory of socially influenced individual choices. The source of social influence on an individual are his *reference groups* in society, formed of societal members he psychologically or contextually relates to. Choices made within an individual's reference groups have an influence on the choices he makes. Specifically, we propose a choice procedure under which, in any choice problem, he considers only those alternatives that he can identify with at least one of his reference groups. From this "consideration set," he chooses the best alternative according to his preferences. The procedure is an interactive one and captures the steady state of a process of mutual social influence. We behaviorally characterize this choice procedure. We also highlight the empirical content of the procedure by relating it to both experimental evidence and real world applications.

JEL codes: D01, D03

Keywords: Individual choice, social influence, reference groups, consideration sets, interactive behavioral choices

1 Introduction

Jamal is a tenth grader living in the Bronx. Jamal loves basketball and is a huge fan of the NBA. When Jamal goes to buy shoes, one of the brands that gets his consideration is Nike because it is popular with other NBA fans like him. Oliver is a recent college graduate living in the Upper East Side. He is hip and trendy and identifies with the young and happening crowd of his Upper East Side neighborhood. When Oliver goes shopping for clothes one of the outlets that he is naturally drawn

*Borah: Ashoka University, India (e-mail: abhinash.borah@ashoka.edu.in); Kops: University of Heidelberg, Germany (e-mail: christopher.kops@awi.uni-heidelberg.de). Correspondence address: Department of Economics, University of Heidelberg, Bergheimer Straße 58, 69115 Heidelberg, Germany. We are grateful to Yoram Halevy and Marco Mariotti for comments. We would also like to thank seminar participants at 12th BiGSEM Workshop on Economic Theory 2017 in Bielefeld, FUR Conference 2018 in York, Indian Statistical Institute Delhi, Indian Institute of Management Ahmedabad and Ashoka University.

to visit is Abercrombie and Fitch as he feels it represents the aspirations of members of this group. Both Jamal and Oliver are subject to social influence. Their *reference groups* in society that they relate to are strong sources of influence and serve as filters through which they organize their world and the choices they consider worth making. At the same time, they too influence others—when Jamal purchases the latest pair of Air Nikes, it may influence Elijah, his good friend at school, to consider doing likewise. This paper builds on this theme of *mutual social influence* and how this influence psychologically constrains choices of decision makers.

Our goal is to explicitly model this psychological constraint. Specifically, we consider decision makers who have tastes but are constrained in exercising them because of this constraint. The theoretical construct that we draw on to model this psychological constraint is that of *consideration sets* (Hauser and Wernerfelt 1990, Roberts and Lattin 1991, Masatlioglu, Nakajima, and Ozbay 2012). The concept of consideration sets is motivated by the observation that in any given choice problem, a decision maker (DM) may not end up considering all the possible alternatives that are available. Indeed, very often it may not even be clear to him what the set of all available alternatives is. At other times, even when this set is clear to him, he may want to consider only those alternatives that satisfy some normative criterion that he considers relevant. In other words, the alternatives that receive his attention and he considers in a choice problem may be a strict subset of the available alternatives. In the literature, this subset is referred to as the consideration set and consists of those alternatives that the DM finds psychologically salient in that choice problem.

In this paper, the psychological salience underlying a DM's consideration sets is determined by the working of social influence. Specifically, we imagine that our DM is socially influenced by certain reference groups that he relates to, e.g., NBA fans, people from his same neighborhood, people that he thinks of as similar to him in a given situation, etc. In keeping with the sociology and social psychology literature, we think of a reference group as a group that an individual uses as a standard or benchmark for evaluating himself and his behavior. It could be a group that he relates to psychologically, either by virtue of being a part of it or because he feels an emotional connect to it. At other times, no such psychological or emotional connect may be at work and this may simply be a group of individuals that a DM perceives to be situationally similar to him.[1] We hypothesize that, in any choice

[1] Think of a situation where our DM is walking along a street when he sees a couple of hooligans start to vandalize public property. Presumably, in this situation, amongst others, the following two alternatives are available to our DM. Either he confronts the hooligans or he turns a blind eye and walks away. Which of these alternatives he is willing to consider may depend on what the other individuals on the street choose to do, i.e., his consideration set is socially influenced. For instance, he may consider the first alternative, overcoming what is known as bystander apathy, if some of the other individuals choose to confront the hooligans. In this situation, the other individuals on the street form a reference group for the DM simply by virtue of the fact that he considers them

problem, an alternative enters a DM's consideration set only if he can identify it with one of his reference groups.[2]

When is it that a DM identifies an alternative with a reference group? Our answer to this question is a choice based one. To illustrate this, consider Oliver once again and the type of brands that he considers when he shops for clothes. As mentioned above, one of Oliver's reference groups in this context is trendy twenty year olds living in his Upper East Side neighborhood. We assume that Oliver can observe the brands that members of this group (as well as those outside it) choose. For Oliver to identify a brand, say, Abercrombie and Fitch, with this reference group, it has to signal a sense of homogeneity or similarity within the group. In the baseline model that we develop, we assume that this is the case if the alternative under consideration is a typical choice amongst members of this group in the sense of being chosen by "sufficiently many" within the group. We think of this benchmark or threshold for what constitutes "sufficiently many" to be a subjective element of his decision making process and this threshold may vary from one DM to another. In an extension of the baseline model, we consider the possibility that just this sense of ingroup similarity by itself may not be enough for a DM to identify an alternative with a reference group. What may also need to be true is that this alternative be an atypical choice among those not in this group so that observed patterns of choice with respect to this alternative *differentiate* this group and its members from those not in it.

Why is it that alternatives that a DM identifies with his reference groups are the ones likely to be part of his consideration set? We can think of several reasons for this. A DM may not be able to consider all available alternatives because of cognitive limitations. Very often it may not even be obvious to him what the set of available alternatives is. In situations like these, presumably, the DM engages in a search process to figure out what all are the alternatives from which he can make his choice. When doing so, those alternatives are likely to catch his attention that he identifies with his reference groups. In other situations, even when the set of available alternatives is transparent to him, he may still wish to present himself with a reason or rationale for whether an alternative is worth considering. When doing so, it is likely to be the case that alternatives that he identifies with his reference groups are the ones that would be easier for him to justify or reason as good alternatives to consider. Indeed, he may take such alternatives to be informative signals of quality.

In the model that we develop here, to theorize the above conceptualization of social

to be in a similar situation as him, even though they might be complete strangers otherwise.

[2]It is worth pointing out that the sociology and social psychology literature recognizes the fact that an individual may orient himself towards multiple reference groups at at time.

influence, we consider a society whose members may psychologically or contextually relate to groupings of other members of this society as their reference groups. Because of this, they may *mutually* influence each other's choices owing to the salience that choices made within their reference groups have in their decision making process. The mutuality or interactions arise because, for instance, just as Jamal may be in Elijah's reference group, Elijah may be in Jamal's. Hence, each of their choices may have an impact on the other's consideration set and, hence, choices. Specifically, we introduce a choice procedure that formally describes how such influence acts as a psychological constraint determining DMs' consideration sets and their choices from any such set. Under this procedure, which we call *choice via social influence (CSI)*, any DM has underlying tastes reflected by a strict preference ranking and, in any choice problem, he chooses that alternative from the consideration set that is best according to this preference ranking. As mentioned above, the key feature of the CSI procedure is that it is an interactive one with individuals' choices both being influenced by as well as influencing the choices of others. The procedure itself may be thought of as a fixed point or steady state of such a process of mutual social influence.

In this paper, our key theoretical goal is to provide a behavioral characterization of the CSI choice procedure. That is, we investigate the following question: can an outside observer look at the choices of a group of individuals and verify whether their choices are subject to the type of social influence that the CSI choice procedure lays out? We identify a single condition on DMs' choices that allows us to answer this question in the affirmative. We also investigate the question of the extent to which the key aspects of the CSI procedure—DMs' preferences and consideration sets—can be uniquely identified. We do this exercise under different assumptions on what is known to the outside observer about the DMs' reference groups.

We also show the empirical content of our theory by highlighting how it explains choices in well-known examples of social influence that have received attention in the literature. Our theory can account for the pattern of choices in the famous Asch conformity experiments (Asch 1955). These experimental results are a starting point for many theories of social influence in social psychology and, to the best of our knowledge, this is the first work that rationalizes this evidence within a choice-theoretic setting that economists adopt. Our theory can account for not just conformity in behavior, but also for nonconformity that can emerge when choices are socially influenced. For instance, a prominent experimental finding that has been reported in the literature is that we see greater diversity and variance in choices when these are made as part of a group as opposed to when these are made individually (e.g., Ariely and Levav (2000)). Our theory can accommodate such a finding. Moreover, we suggest how our theory can lend itself to analyzing interesting applications of social influence and interactions in choices.

We draw part of our motivation for working on this problem from a large body of work in economics that shows that social/peer influence and interactions have a discernible and quantitatively robust impact on outcomes in a wide range of economically relevant domains. For instance, such an impact has been reported, amongst others, in the areas of educational and academic outcomes (eg., Sacerdote (2001), Gaviria and Raphael (2001), Zimmerman (2003), Calvó-Armengol, Patacchini, and Zenou (2009)); crime and criminal behavior (eg., Damm and Dustmann (2014), Glaeser, Sacerdote, and Scheinkman (1996), Zimmerman and Messner (2010)); workers' productivity and labor market outcomes (eg., Clark (2003), Falk and Ichino (2006), Mas and Moretti (2009), Cornelissen, Dustmann, and Schönberg (2017)); social interactions in consumption patterns, including conspicuous consumption (eg. Charles, Hurst, and Roussanov (2009), Kaus (2013), Grinblatt, Keloharju, and Ikäheimo (2008), Kuhn, Kooreman, Soetevent, and Kapteyn (2011), Heffetz (2011)); adolescent and teenage choices (eg. Trogdon, Nonnemaker, and Pais (2008), Clark and Loheac (2007), Lundborg (2006)).

One observation from this literature worth noting is that when it comes to the underlying mechanism driving social influence, a direct impact through individual preferences and tastes need not be the only channel through which such influence operates. Take, for example, the fairly large literature documenting socially influenced consumption patterns—especially for goods whose consumption is visible—seen in phenomena like conspicuous consumption and positional concerns in consumption. Purchase decisions of cars is one leading example of such visible consumption that has received some attention. One message that emerges from this literature is that although there is considerable social influence in car purchase decisions, this influence need not necessarily work through the preference channel.[3]

[3] For example, Grinblatt, Keloharju, and Ikäheimo (2008) use a rich Finish panel data set from two major provinces and find that having neighbors who purchased a car, particularly those who purchased recently and are nearest in distance, increases the propensity of a consumer to purchase a car. They find that the influence is strongest in lower income groups, is more pronounced for used cars than new ones and is conformist in nature in the sense of the same makes and models being purchased. Such evidence rules out features of social preferences like envy as the mechanism driving the influence—if it were, then the influence should not be stronger for used cars than new ones and should not dampen with a rise in consumer income. Similarly, their results also makes it hard to reconcile the evidence with a signalling story driven by status concerns in preferences. Rather the influence seems to flow through other channels, with informational frictions being the leading candidate proposed by the authors. In a similar spirit, Kuhn, Kooreman, Soetevent, and Kapteyn (2011) provide evidence of social influence in the context of the Dutch Postcode lottery under which each week a randomly selected lottery participant in a randomly chosen postcode wins a BMW car. They report a statistically significant effect on the car consumption of neighbors of winners. This is true even though most BMW winners either choose to receive the cash prize in lieu of the car or sell their BMWs shortly after receiving them. Here too, the evidence makes it hard to maintain that envious or status-sensitive social preferences is the channel through which the influence operates. Indeed, the survey conducted by the authors confirms that having neighbors win the lottery does not reduce the happiness of non-winning households.

As such, given that the preference channel need not be the only one through which social influence and interactions impact economic outcomes, we need to conceptualize other possible mechanisms that may drive this influence. This is important both for understanding the theoretical differences that different mechanisms intermediating such influence produces as well as for providing better guidance to the "social econometrics" project devoted to quantitatively identifying the significance and magnitude of such influence. In this context, our mechanism, which highlights how the salience of reference groups in the decision making process constrains choice sets, may be an important channel through which such influence plays out. To further substantiate this claim and suggest possible applications, a few examples may be instructive at this point.

Example 1.1 (**Political Behavior**)**.** In politics, electoral or otherwise, while considering what positions to take or what alternatives or candidates to consider, individuals seem to put a great degree of importance on the views and behavior of those in their reference groups.[4] When such social influence is at play, it becomes vital to understand whether political choices can be attributed to deep preferences or are simply a reflection of such influence constraining the alternatives and positions that an individual would consider. Our model can potentially help to develop a better understanding of this.

Example 1.2 (**Groupthink**)**.** Groupthink is a psychological phenomenon whereby members of a group simply ignore their preferences and choose to act in accordance with some group norm. Groupthink often results in dysfunctional decision making, including making pathological moral judgments.[5] One of the reasons groupthink persists is because individuals within the group setting fail to consider alternatives that may be superior owing to the normative social influence cast by the group (Janis (1972), Janis (1982)). A phenomenon like this is very much consistent with the model we develop here.

Example 1.3 (**Minority Influence**)**.** The social psychology literature emphasizes that social influence is not simply about numbers and it need not emanate just from a majority view. At times, even a minority might be able to influence the views and behavior of a majority that disagrees with its position. For instance, for a long time in the history of western liberal democracies women were not given the right to vote because the popular view was that their domestic and reproductive roles

[4] For instance, Cohen (2003) presented conservative and liberal college students in the US with one of two welfare policies—a stringent one and a generous one. Along expected lines, conservatives preferred the former and liberals the latter. However, when the conservative (resp., liberal) subjects were made to believe that many Republican (resp., Democratic) politicians favored the latter (resp., former) policy—"95% of House Republicans (resp., Democrats) supported the policy"—many more favored this policy over the former (resp., latter).

[5] Prominent examples that have been cited in the literature of such decision-making include the United States' lack of precaution against the attack on Pearl Harbor, the faulty policy decision regarding the Vietnam war and the illegal practices related to the Watergate scandal.

made them incapable of participating in the public-political sphere. But, then, around the late nineteenth and early twentieth century a small group of suffragists were able to challenge and overturn this erroneous view. Our model can account for minority influence, precisely because it does not reduce the mechanics of influence to numbers, rather influence works through reference groups and their structure matters.

Example 1.4 (**Social Influence in Networks**)**.** The fact that people who we are connected to in social networks influence our behavior is quite well understood. What is also well recognized is that influence in networks critically depends on one's position in the network. Sometimes a small minority of nodes in the network may end up having a disproportionate level of influence in behavior seen within the overall network. Our language of reference groups provides a general framework to study social influence in networks, including understanding minority influence, as mentioned above.

Example 1.5 (**Positional Concerns**)**.** It is well documented that individuals care not just about their absolute position (w.r.t., wealth, income, consumption etc.) but also their relative position. This, of course, raises the question–relative to whom? Here research points to the fact that reference groups matter. For instance, in regards to consumption, people may want to consider the patterns of consumption of those considered at par or above them in a social hiearchy, where such hierarchies could be based on economic dimensions like income and wealth or social ones like class, caste and race (Frank, Levine, and Dijk (2014), Bertrand and Morse (2016)). Here too, our model can help understand how reference groups constrain, for instance, the consumption alternatives that an individual might consider. In so doing, it may provide an alternative approach to modeling social status concerns—one that maintains a parsimonious structure on deep preferences as well as does not rely on features like asymmetric information for the explanation.

From a methodological point of view, our paper relates to the recent literature on theories of behavioral choice. Like many of the papers in this area, ours too explicitly spells out a psychologically motivated choice procedure by which an individual makes choices and provides a behavioral characterization of this procedure. Our procedure has a sequential structure under which a cognitive phenomenon—that of social influence—constrains the available set of alternatives and it is from this constrained set that the decision maker chooses. In terms of this structure, our model bears resemblance to Cherepanov, Feddersen, and Sandroni (2013), Manzini and Mariotti (2012), Masatlioglu, Nakajima, and Ozbay (2012), and Lleras, Masatlioglu, Nakajima, and Ozbay (2017), amongst others. Of particular interest to us in this context is the paper by Cuhadaroglu (2017), who too models social influence within the behavioral choice paradigm. We provide a comprehensive comparison of our work to this body of research in the concluding section of this paper.

The rest of the paper is organized as follows. Section 2 lays out the primitives. Section 3 formally defines the CSI choice procedure. Section 4 provides a behavioral characterization of the procedure. Section 5 shows the extent to which parameters of the model can be identified based on observable choices. Sections 6 and 7 provide extensions of the baseline model. Section 8 provides an application of the model that shows how it can rationalize the famous experimental findings of Asch (1955). Finally, in Section 9 we relate our work to the existing behavioral choice theory literature.

2 Primitives

Let X be a finite set of alternatives with typical elements denoted by x, y, z etc. $\mathcal{P}(X)$ denotes the set of non-empty subsets of X with typical elements R, S, T etc. A choice function, $c : \mathcal{P}(X) \to X$, is a mapping that for any choice problem $S \in \mathcal{P}(X)$ picks an element $c(S) \in S$. Let $\mathcal{I} = \{1, \dots, n\}$ be a set of individuals in society, with typical individuals denoted by i, j etc. Each individual $i \in \mathcal{I}$ has a choice function c_i on X. This profile of choice functions $(c_i)_{i \in \mathcal{I}}$ is the primitive of our model. In the way of notation, for any $i \in \mathcal{I}$, $\mathcal{I}_{-i}$ denotes the set $\mathcal{I} \setminus \{i\}$. Further, c_{-i} denotes the vector of choice functions $(c_j)_{j \in \mathcal{I}_{-i}}$ and, for any $S \in \mathcal{P}(X)$, $c_{-i}(S)$ denotes the vector $(c_j(S))_{j \in \mathcal{I}_{-i}}$.

3 Choice via Social Influence

Our starting point in analyzing how any particular DM's choices (say, i's) are socially influenced by those of the others is the observation that i has certain reference groups in society that he either psychologically or contextually relates to. For the purpose of this paper, a reference group is any group of societal members that an individual uses as a standard for evaluating himself and his behavior.[6] Formally, any such group G is a non-empty subset of $\mathcal{I}_{-i}$. Let $\mathcal{G}_i$ be the collection of such reference groups that i relates to.[7] For each i, $\mathcal{G}_i \neq \emptyset$. We assume for now that, for any such $i \in \mathcal{I}$, the collection $\mathcal{G}_i$ is commonly known and a researcher can directly observe it. In Section 7, we discuss how our framework can be extended to

[6] As suggested in the Introduction, this interpretation closely follows that used in the sociology and social psychology literature.

[7] It is worth pointing out here that we do not impose any a priori structure on the set $\mathcal{G}_i$. This, we think, adds to the model's flexibility as it allows the description of the reference groups to be guided by the particular application we are looking at, thus, increasing the scope of its applicability.

accommodate the case when reference groups are not directly observable.

The process of social influence that we model involves any DM, $i \in \mathcal{I}$, using his reference groups as a standard or benchmark and, in any choice problem, an alternative receives his consideration if he *identifies* it with one of these groups. Of course, we need to formalize this notion of what it means for a DM to identify an alternative with a reference group. We do this below, but before we get to that, it is worth providing an overview first of the channel through which social influence operates in the model. For any choice problem $S \in \mathcal{P}(X)$, the process of social influence involves i considering only a subset $\Gamma_i(S) \subseteq S$ of alternatives, which are precisely the ones he identifies with at least one of his reference groups. It is from this subset that he makes his choice. In the literature, such a correspondence, $\Gamma_i : \mathcal{P}(X) \to \mathcal{P}(X)$, is referred to as a consideration set mapping. In our model, it is the consideration set mapping which captures the psychological constraint that social influence imposes on what a DM chooses.

Having laid out the broad structure, we now outline the exact way in which social influence determines the consideration set mapping. We start with a formalization of what it means to say a DM identifies some alternative in a choice problem with one of his reference groups. In our model, this process of identification is a choice based one. Individual i identifies an alternative with a group if it is chosen by sufficiently many individuals within this group. As a way of formalizing this idea, take c_{-i} as given,[8] and define, for any choice problem $S \in \mathcal{P}(X)$, a mapping $\gamma_{i,S} : S \times \mathcal{G}_i \to [0, 1]$ given by:[9]

$$\gamma_{i,S}(x, G) = \frac{\#\{j \in G : c_j(S) = x\}}{\#G}$$

That is, $\gamma_{i,S}(x, G)$ is the proportion of individuals in G who choose x in choice problem S. We may interpret $\gamma_{i,S}(x, G)$ as a measure of the degree to which there is agreement among members of G about the choice of x. The greater the degree of this agreement, the closer would be the association of x with group G in choice problem S. This motivates our definition of what it means to say a DM identifies some alternative in a choice problem with one of his reference groups.

Definition 3.1. *Given $c_{-i}(S)$, individual i identifies alternative $x \in S$ with group $G \in \mathcal{G}_i$ in choice problem $S \in \mathcal{P}(X)$ if $\gamma_{i,S}(x, G) \geq \tau_i \in [0, 1]$. An alternative $x \in S$ receives i's consideration via social influence in choice problem $S \in \mathcal{P}(X)$ if he identifies x with some group $G \in \mathcal{G}_i$.*

That is, i identifies alternative x with group G in choice problem S if x is chosen

[8]Note that c_{-i} is a given only from the perspective of i. As we will see, from the perspective of the choice procedure that we propose below, c_{-i} is an endogenous object.

[9]For any finite set H, $\#H$ denotes the number of elements in H.

by sufficiently many individuals in G in this choice problem. The (social influence) threshold for what constitutes sufficiently many is determined by the parameter τ_i. This parameter is a subjective element of the DM's decision making process and may vary from one DM to another. Note that, in our model, DMs who are not socially influenced are characterized by a threshold value of 0. We denote the set of all alternatives that receive i's consideration via social influence in choice problem S by $\Gamma_i(S; c_{-i}(S), \tau_i)$.

We can now introduce the choice procedure that we are proposing in this paper. The key feature of this procedure is that it is an interactive one with individuals' choices both being influenced by as well as influencing the choices of others. Our procedure may be thought of as a fixed point or steady state of such a process of mutual social influence.

Definition 3.2. *Given $(\mathcal{G}_i)_{i\in\mathcal{I}}$, the profile of choice functions $(c_i)_{i\in\mathcal{I}}$ is a choice via social influence (CSI) if for each i, there exists a linear order $\succ_i$ on X and a social influence threshold $\tau_i \in [0,1]$ such that for any $i \in \mathcal{I}$ and $S \in \mathcal{P}(X)$,*

1. *$c_i(S) \in \Gamma_i(S; c_{-i}(S), \tau_i)$, and*
2. *If $y \in \Gamma_i(S; c_{-i}(S), \tau_i)$, $y \neq c_i(S)$, then $c_i(S) \succ_i y$*

The CSI choice procedure is characterized by two parameters for each individual i: a strict preference ranking $\succ_i$ over the alternatives in X and a threshold $\tau_i \in [0,1]$. The profile of choice functions are in accordance with this choice procedure if for each choice problem S, the profile of choices $(c_1(S), \ldots, c_n(S))$ satisfies the following two conditions for each i. First, taking $c_{-i}(S)$ as given, individual i identifies $c_i(S)$ with some reference group $G \in \mathcal{G}_i$ in choice problem S (based, of course, on what his τ_i is). Hence, this alternative is in his consideration set for this choice problem. Second, if some other alternative y also receives consideration via social influence in this choice problem given what $c_{-i}(S)$ is, then by his preferences $\succ_i$, $c_i(S)$ should be strictly preferred to y. We had mentioned above that any i with $\tau_i = 0$ is not socially influenced. Such an individual in our setting is, therefore, a standard rational agent and what his reference groups are have no bearing on choices.

Next, we provide a few examples to illustrate the concept of CSI choice profiles better.

Example 3.1 (Menu Dependence in Choices)**.** Menu dependence in choices is a prominent behavioral "anomaly" that involves choices of the following type: $c(\{x,y,z\}) = y$, $c(\{x,y\}) = x$, $c(\{y,z\}) = y$, $c(\{x,z\}) = x$. Here z serves as a "decoy" for y. That is, when only x and y are in the menu, the DM chooses x, but when z is added to the menu, the choice shifts to y. Menu dependent choices have received interest in the literature because they are often empirically observed, but they violate the weak axiom of revealed preferences (WARP)–the benchmark for rational choice behavior. We now show that a CSI choice profile may exhibit such menu dependence. Consider a family, where Mary (M) and John (J) are Liz's (L) parents. Over her summer vacation, Liz has the option of either taking extra math classes (x), taking ballet lessons (y) or going for soccer practice (z). Suppose preferences of the three are $z \succ_J x \succ_J y$; $y \succ_M x \succ_M z$; and $x \succ_L y \succ_L z$.[10] Further, thresholds are $\tau_J = \tau_M = 0$, $\tau_L \in (0, \frac{1}{2}]$ and $\mathcal{G}_L = \{\{1, 2\}\}$.[11] Under these assumptions, Mary and John's choices over these alternatives would be according to their preference orderings and, hence, satisfy WARP. However, as Table 1 shows, Liz's choices under a CSI profile are menu dependent. The menu dependence arises here because in any choice problem, Liz considers only those alternatives that she feels have the "approval" of at least one of her parents.

Table 1

	$\{x, y\}$	$\{y, z\}$	$\{x, z\}$	$\{x, y, z\}$
$c_J(.)$	x	z	z	z
$c_M(.)$	y	y	x	y
$\Gamma_L(.)$	$\{x, y\}$	$\{y, z\}$	$\{x, z\}$	$\{y, z\}$
$c_L(.)$	x	y	x	y

Example 3.2 (Downsian Politics and Platform Divergence)**.** We next consider the Downsian model of electoral competition with two political parties and a continuum of voters uniformly distributed on the $[-1, 1]$ ideological spectrum. The two parties have to decide where to position themselves on this spectrum with the goal of maximizing their vote share. All voters' preferences are "Euclidean" and, in the absence of any other consideration, they would vote for the party that's closest to their ideological position. Because of this, in this standard set-up, the unique Nash equilibrium involves both parties positioning themselves at 0, which is the preferred ideological position of the median voter. We will now illustrate how this conclusion may change with CSI-type voters. Refer to the set of voters in the set $L = [-1, 0]$ as left-wingers and those in the set $G_L = [-1, -0.75]$ as left wing elites. Similarly, we refer to the set of voters in $R = (0, 1]$ as right-wingers, and those in the set $G_R = [0.75, 1]$ as the right wing elites. Further, suppose, $\tau_i = 0$, $\forall i \in G_L \cup G_R$; $\tau_i > 0$, $\forall i \in (G_L \cup G_R)^c$; $\mathcal{G}_i = \{G_L\}$, $\forall i \in L$; and $\mathcal{G}_i = \{G_R\}$, $\forall i \in R$. That is, the elites on both sides of the ideological spectrum are not socially influenced. However, the non-elites on either side are socially influenced by the elites on their side of the spectrum. Under these assumptions, we can verify that unlike in the standard Downsian set-up, what we get with CSI-type decision makers is a situation where

[10] John and Mary's preferences have a similar flavor to those in a battle of sexes game.

[11] As noted above, since $\tau_J = \tau_M = 0$, what Mary's and John's reference groups are has no bearing on choices.

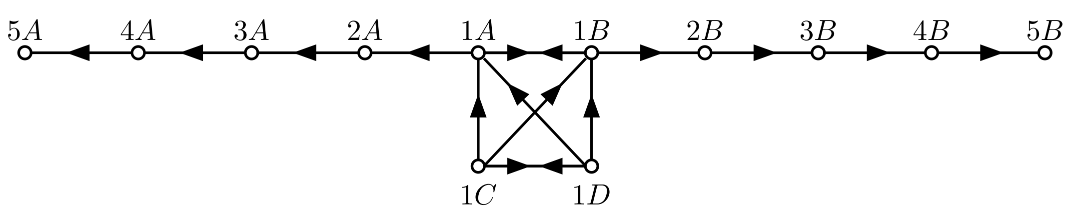
there may be policy divergence. Observe that the two political parties positioning themselves at -0.75 and 0.75 is a Nash equilibrium.

Example 3.3 (Societal Networks and Minority Influence)**.** Consider the societal network structure represented by the directed line graph in Figure 1.

Figure 1

Initially, there are 10 individuals–the 5 A-types and the 5 B-types–represented as nodes on the graph. Each individual has a unique reference group, specifically, $\mathcal{G}_i = \{\{j : (j, i) \text{ is an edge of the graph}\}\}$. E.g., $\mathcal{G}_{1A} = \{\{1B\}\}$, $\mathcal{G}_{3B} = \{\{2B\}\}$, etc. There are two available alternatives: x, which provides universal suffrage and y, which gives voting rights only to men. All of these individuals strictly prefer y to x and each individual's threshold is $\tau_i = \frac{1}{2}$. In this scenario, everyone choosing y is a CSI choice profile and one that, presumably, is quite stable in this society, given that every individual strictly prefers y to x. Now assume that two suffragists, $1C$ and $1D$, come along, who both strictly prefer x to y. Both of them are convinced of their position on universal suffrage and cannot be socially influenced on the matter, i.e., $\tau_{1C} = \tau_{1D} = 0$. Assume that the new network structure formed in society with these two individuals joining is given by the directed graph in Figure 2. Given this network structure, only $1A$ and $1B$'s reference groups change—$\mathcal{G}_{1A} = \{\{1B, 1C, 1D\}\}$ and $\mathcal{G}_{1B} = \{\{1A, 1C, 1D\}\}$—that of the other 8 individuals remains the same. In this case, $1C$ and $1D$ choosing x is sufficient to ensure that everyone choosing x is the unique CSI choice profile in this society. This alludes to the process of minority influence whereby a dedicated minority, by effectively positioning themselves in a social network, might be able to influence the views and attitudes of a majority that disagrees with its position.

Figure 2



Example 3.4 (Relative Concerns in Consumption)**.** There are 3 households, $i = 1, 2, 3$, in a relatively poor neighborhood with income levels $m_1 = 30, m_2 = 25$ and $m_3 = 20$. Households 2 and 3 each have a unique reference group which consists of household/s with a higher income level than it, i.e., $\mathcal{G}_3 = \{\{1, 2\}\}$, $\mathcal{G}_2 = \{\{1\}\}$

and let $\tau_3 > 0$, $\tau_2 > 0$. For household 1, we simply set $\tau_1 = 0$ to reflect the fact that, in this setting, it is not socially influenced. Households spend their income on necessities (x) and (money left over for) luxuries (y). Social influence operates through the luxury good. Households' preferences are identical and represented by the quasi-linear utility function, $u(x, y) = 20 \ln x + y$. We take y to be the numeraire commodity and normalize its price to be equal to 1 and denote the price of x as p. Further, we suppose that 1 unit of the luxury good can buy 20 units of necessities, i.e., $p = \frac{1}{20}$. Under these assumptions, it is straightforward to see that if households were to solve their consumption problem without any social influence all households would consume 400 units of the necessity; households 1 would consume 10 units, household 2, 5 units and household 3, 0 units of the luxury good. However, in the unique CSI choice profile all households will consume 10 units of the luxury good. Now, whereas household 1 continues to consume 400 units of the necessity, households 2 and 3 consume only 300 and 200 units of the necessity, respectively.

Example 3.5 (CSI choice profile does not exist)**.** Let $X = \{x, y, z\}$ be the set of alternatives and $\mathcal{I} = \{1, 2, 3\}$ be the set of individuals in society. Consider them as being positioned in the circle $1 - 2 - 3 - 1$ and let each individual have a single reference group consisting of her two immediate neighbors on her left and right, i.e., $\mathcal{G}_1 = \{\{2, 3\}\}$, $\mathcal{G}_2 = \{\{1, 3\}\}$, and $\mathcal{G}_3 = \{\{1, 2\}\}$. Let their preferences be $x \succ_1 y \succ_1 z$, $y \succ_2 z \succ_2 x$, and $z \succ_3 x \succ_3 y$, and let the threshold values be $\tau_1 = \tau_2 = 0$ and $\tau_3 > \frac{1}{2}$. For these primitives, there exists no choice profile over $\mathcal{P}(X)$ that is a CSI. To see this, consider the set $\{x, y\}$ and verify that $c_1(\{x, y\}) = x$ and $c_2(\{x, y\}) = y$. Clearly, $\gamma_{3,\{x,y\}}(x, \{1, 2\}) = \gamma_{3,\{x,y\}}(y, \{1, 2\}) = \frac{1}{2} < \tau_3$. So, by Definition 3.1 no alternative in $\{x, y\}$ receives individual 3's consideration.

Remark 3.1. Although we might have cases like in the example above where a CSI choice profile does not exist, it is not hard to provide sufficient conditions for the existence of such profiles. For instance, $\#\{i \in \mathcal{I} | \tau_i = 0\} \leq 1$ is one such sufficient condition.[12]

4 Behavioral Foundation

We now provide a behavioral (i.e., choice-based) characterization of the CSI procedure. Such a characterization enables any outside observer to verify whether choice data coming out of social settings that resembles our's is consistent with the CSI procedure or not. As it turns out, this procedure can be characterized by just one

[12] If $\#\{i \in \mathcal{I} | \tau_i = 0\} = 0$, then, for any $S \in \mathcal{P}(X)$, pick any $x \in S$ and let $c_i(S) = x$ for all $i \in \mathcal{I}$. This defines a CSI choice profile. On the other hand, if $\{i \in \mathcal{I} | \tau_i = 0\} = \{j\}$, then, for any $S \in \mathcal{P}(X)$, let $x \in S$ be the $\succ_j$-best element in S and let $c_i(S) = x$ for all $i \in \mathcal{I}$. This again defines a CSI choice profile.

axiom, which adapts the well known WARP condition to our current setting. To understand our axiom, recall that the standard WARP condition can be written in the following way.

Axiom 4.1 (WARP)**.** *The choice function c_i satisfies WARP if for any $S \in \mathcal{P}(X)$, there exists $x^* \in S$ such that, for any T including x^*, if $c_i(T) \in S$, then $c_i(T) = x^*$.*

That is, the standard WARP condition says that in any set $S \in \mathcal{P}(X)$, there exists a "most preferred element" x^* such that if this element is available in another set $T \in \mathcal{P}(X)$ and the chosen element from T is in S, then this chosen element has to be x^*. The reason this condition might fail in our set-up is because there is no guarantee that x^* receives consideration in T. For instance, consider Example 3.1. x is Liz's most preferred alternative in the set $\{x, y, z\}$ but it does not receive her consideration in it.

However, if we can identify those situations in which x^* does receive consideration in T, then the content of the axiom should still apply to these situations. When are we guaranteed that x^* receives consideration in T? Suppose there exists some choice problem $R \in \mathcal{P}(X)$ and reference group $\widetilde{G} \in \mathcal{G}_i$ such that

$$\gamma_{i,T}(x^*, \widetilde{G}) \geq \gamma_{i,R}(c_i(R), G), \text{ for all } G \in \mathcal{G}_i.$$

Since i chooses $c_i(R)$ in R, it must receive his consideration in this set. That is, there exists some group $G' \in \mathcal{G}_i$, such that i identifies the alternative $c_i(R)$ with this group in choice problem R. The inequality above, then, implies that $\gamma_{i,T}(x^*, \widetilde{G}) \geq \gamma_{i,R}(c_i(R), G')$. That is, the proportion of individuals who choose x^* in T in reference group $\widetilde{G}$ is at least as large as the proportion choosing $c_i(R)$ in R in reference group G'. Therefore, given how social influence operates in our setting, it stands to reason that if i identifies $c_i(R)$ with group G' in choice problem R, then he must identify x^* with group $\widetilde{G}$ in choice problem T. Accordingly, x^* must receive his consideration in T when the above inequality holds. The appropriate adaptation of WARP to the current context, therefore, is the following condition, which characterizes the CSI choice procedure.[13] Note that this axiom applies collectively to a choice profile $(c_i)_{i \in \mathcal{I}}$ and not piecewise to individual choice functions.

Axiom 4.2 (WARP-SI)**.** *For any $i \in \mathcal{I}$ and $S \in \mathcal{P}(X)$ there exists $x^* \in S$ such that, for any $T \in \mathcal{P}(X)$, if*

1. *$x^* \in T$, $c_i(T) \in S$, and*
2. *there exists $R \in \mathcal{P}(X)$ and $\tilde{G} \in \mathcal{G}_i$ satisfying $\gamma_{i,T}(x^*, \tilde{G}) \geq \gamma_{i,R}(c_i(R), G)$, for all $G \in \mathcal{G}_i$*

[13] The condition and the reasoning leading to it has a similar flavor as the WARP-LA condition of Masatlioglu, Nakajima, and Ozbay (2012) that characterizes their CLA model.

then $c_i(T) = x^*$.

Theorem 4.1. *The choice profile* $(c_i)_{i \in \mathcal{I}}$ *is a CSI if and only if it satisfies WARP-SI.*

Proof: Please refer to Section A.2.

We now present a couple of examples which show how, for a given choice data set, we can verify whether WARP-SI is satisfied.

Example 4.1 (Not a CSI)**.** Let $X = \{x, y, z\}$ be the set of alternatives and $\mathcal{I} = \{1, 2, 3, 4\}$ be the set of individuals in society. Consider them as being positioned in the circle $1 - 2 - 3 - 4 - 1$ and let each individual have a single reference group consisting of her two immediate neighbors on her left and right, i.e., $\mathcal{G}_1 = \{\{2, 4\}\}$, $\mathcal{G}_2 = \{\{1, 3\}\}$, $\mathcal{G}_3 = \{\{2, 4\}\}$, and $\mathcal{G}_4 = \{\{1, 3\}\}$. For this structure of reference groups we can verify that the choice profile specified in Table 2 does not satisfy WARP-SI.

Table 2

	$\{x, y\}$	$\{y, z\}$	$\{x, z\}$	$\{x, y, z\}$
$c_1(.)$	y	y	z	y
$c_2(.)$	y	z	x	y
$c_3(.)$	x	z	x	x
$c_4(.)$	x	y	z	x

Consider individual 4 and the set $\{x, y, z\}$. We establish that there does not exist an $x^* \in \{x, y, z\}$ of the type WARP-SI requires. To see this, first suppose that $x^* = x$. Consider $T := \{x, z\}$; clearly $c_4(T) \in \{x, y, z\}$. Further, $\mathcal{G}_4$ is a singleton consisting of just the group $\{1, 3\}$ and, for it, $\frac{1}{2} = \gamma_{4,T}(x, \{1, 3\}) \geq \gamma_{4,\{x,y\}}(c_4(\{x, y\}), \{1, 3\})$, but $c_4(T) = z \neq x$. Next, suppose that $x^* = y$. For $T := \{x, y\}$, $c_4(T) \in \{x, y, z\}$ and $\frac{1}{2} = \gamma_{4,T}(y, \{1, 3\}) \geq \gamma_{4,\{y,z\}}(c_4(\{y, z\}), \{1, 3\})$, but $c_4(T) = x \neq y$. Finally, suppose that $x^* = z$. For $T := \{y, z\}$, $c_4(T) \in \{x, y, z\}$ and $\frac{1}{2} = \gamma_{4,T}(z, \{1, 3\}) \geq \gamma_{4,\{x,z\}}(c_4(\{x, z\}), \{1, 3\})$, but $c_4(T) = y \neq z$. Hence, there exists no $x^* \in \{x, y, z\}$ such that WARP-SI is satisfied.

Example 4.2 (A CSI)**.** Let $X = \{x, y, z\}$ be the set of alternatives and $\mathcal{I} = \{1, 2, 3, 4, 5\}$ be the set of individuals in society. Consider them as being positioned in the circle $1 - 2 - 3 - 4 - 5 - 1$ and let each individual have a single reference group consisting of all her neighbors, i.e., $\mathcal{G}_i = \mathcal{I}_{-i}$, for all $i \in \mathcal{I}$. For this structure of reference groups, the choice profile specified in Table 3 satisfies WARP-SI.

To establish that WARP-SI holds, we need to identify, for each $i \in \mathcal{I}$ and $S \in \mathcal{P}(X)$, the alternative x^* that "does the job." For non-singleton $S \in \mathcal{P}(X)$,

Table 3

	$\{x,y\}$	$\{y,z\}$	$\{x,z\}$	$\{x,y,z\}$
$c_1(.)$	y	y	z	y
$c_2(.)$	y	z	z	y
$c_3(.)$	x	z	z	x
$c_4(.)$	x	y	z	x
$c_5(.)$	x	y	z	x

– for individual 1, let

$$x^* = \begin{cases} y, & \forall S \in \mathcal{P}(X) \text{ with } y \in S \\ z, & \text{otherwise} \end{cases}$$

– for individual 2, let

$$x^* = \begin{cases} z, & \forall S \in \mathcal{P}(X) \text{ with } z \in S \\ y, & \text{otherwise} \end{cases}$$

– for individual 3, let

$$x^* = \begin{cases} x, & \forall S \in \mathcal{P}(X) \text{ with } x \in S \\ z, & \text{otherwise} \end{cases}$$

– for individuals 4 and 5, let

$$x^* = \begin{cases} x, & \forall S \in \mathcal{P}(X) \text{ with } x \in S \\ y, & \text{otherwise} \end{cases}$$

With these specifications for x^*, it is straightforward to verify that WARP-SI holds.

We now want to propose a procedure for checking WARP-SI which can be conveniently used even for large data sets. To that end, we first define, for each i, a binary relation P_i on X as follows: For any distinct $x, y \in X$, xP_iy if there exists $S, R \in \mathcal{P}(X)$, with $x, y \in S$ and $c_i(S) = x$, and $\tilde{G} \in \mathcal{G}_i$ satisfying

$$\gamma_{i,S}(y, \tilde{G}) \geq \gamma_{i,R}(c_i(R), G), \forall G \in \mathcal{G}_i, \tag{1}$$

P_i can be interpreted as a revealed preference relation that can be *directly* elicited from choice data. This is because x is chosen in set S and, hence, considered. Further, we know from the discussion preceding the statement of WARP-SI that the inequality above implies that y too must have received consideration in set S.

Next, we propose an equivalent way of defining the binary relation P_i. Let

$$\bar{\tau}_i := \min_{S\in\mathcal{P}(X)} \max_{G\in\mathcal{G}_i} \gamma_{i,S}(c_i(S), G)$$

$\bar{\tau}_i$ can be interpreted as a revealed threshold specifying a conservative estimate from the choice data as to what proportion of individuals in a reference group has to choose an alternative for it to be identified with the group. This is because in any choice problem S, the chosen alternative, $c_i(S)$ must receive i's consideration. Hence, the threshold can be no higher than the maximal proportion of individuals in a reference group in $\mathcal{G}_i$ choosing this alternative. Since this has to be true in all choice problems, the elicited thresholds must be the minimum of all these maxima. Finally, based on these elicited thresholds, we can define revealed consideration sets for each $i \in \mathcal{I}$, given by

$$\underline{\Gamma}_i(S; c_{-i}, \bar{\tau}_i) := \{z \in S | \gamma_{i,S}(z, G) \geq \bar{\tau}_i, \text{ for some } G \in \mathcal{G}_i\}$$

As such the binary relation P_i can equivalently be defined by: xP_iy if there exists $S \in \mathcal{P}(X)$ such that $c_i(S) = x$ and $y \in \underline{\Gamma}_i(S; c_{-i}, \bar{\tau}_i)$. The following result establishes the relationship of P_i to WARP-SI.

Lemma 4.1. *P_i is acyclic, for all $i \in \mathcal{I}$, if and only if the choice profile $(c_i)_{i\in\mathcal{I}}$ satisfies WARP-SI.*

Proof: Please refer to Section A.1.

As such, to verify whether c_i satisfies WARP-SI, we can define $\bar{\tau}_i$ as above, derive the sets $\underline{\Gamma}_i(S, c_{-i}, \bar{\tau}_i)$ for each set S and elicit P_i. Then, all we need to do, is to check whether P_i is acyclic or not. For instance, in Example 4.1, $\bar{\tau}_4 = \min_{S\in\mathcal{P}(X)} \gamma_{4,S}(c_4(S), \{1, 3\}) = \frac{1}{2}$ and accordingly $\underline{\Gamma}_4(\{x, y\}) = \{x, y\}$, $\underline{\Gamma}_4(\{y, z\}) = \{y, z\}$, $\underline{\Gamma}_4(\{x, z\}) = \{x, z\}$, and $\underline{\Gamma}_4(\{x, y, z\}) = \{x, y\}$. Therefore, this implies that xP_4y, yP_4z, and zP_4x which is a cycle.

On the other hand, in Example 4.2, $\bar{\tau}_1 = \bar{\tau}_2 = \bar{\tau}_3 = \frac{1}{4}$ and $\bar{\tau}_4 = \bar{\tau}_5 = \frac{1}{2}$. Therefore, $\underline{\Gamma}_i$ is as specified in Table 4, for $i = 1, 2, 3, 4, 5$. Accordingly, $P_1 = \{(y, x), (y, z)\}$, $P_2 = \{(z, y), (y, x)\}$, $P_3 = \{(x, y), (z, y)\}$, $P_4 = P_5 = \{(x, y), (y, z)\}$. Observe that P_i is acyclic for all $i = 1, 2, 3, 4, 5$ and, thus, WARP-SI is satisfied.

5 Identification

The CSI choice procedure is based on two important aspects of DMs' decision-making processes. The first aspect concerns their preferences and the second pertains to their consideration sets determined by their social influence thresholds. In

Table 4

	$\{x,y\}$	$\{y,z\}$	$\{x,z\}$	$\{x,y,z\}$
$\underline{\Gamma}_1(.)$	$\{x,y\}$	$\{y,z\}$	$\{z\}$	$\{x,y\}$
$\underline{\Gamma}_2(.)$	$\{x,y\}$	$\{y,z\}$	$\{z\}$	$\{x,y\}$
$\underline{\Gamma}_3(.)$	$\{x,y\}$	$\{y,z\}$	$\{z\}$	$\{x,y\}$
$\underline{\Gamma}_4(.)$	$\{x,y\}$	$\{y,z\}$	$\{z\}$	$\{x,y\}$
$\underline{\Gamma}_5(.)$	$\{x,y\}$	$\{y,z\}$	$\{z\}$	$\{x,y\}$

the last section, we identified a testable condition (WARP-SI) that can be applied to any given choice profile to determine whether the profile can be thought of as resulting from a CSI-like cognitive process of mutual social influence. Suppose we have a choice profile that does satisfy WARP-SI and, therefore, is consistent with the CSI logic. The question that we now address is about the extent to which the two key aspects of the CSI procedure—DMs' preferences and consideration sets—can be uniquely identified from such a choice profile. We first consider the question of identification of preferences.

Definition 5.1. *Let $(c_i)_{i\in\mathcal{I}}$ be a CSI choice profile. We say that x is revealed to be preferred to y by individual i if for any preference ranking $\succ_i$ that is part of a CSI representation of these choices, we have $x \succ_i y$.*

In other words, the notion of revealed to be preferred captures the extent to which preferences can be uniquely identified from a CSI choice profile, since the preference information captured by it is invariant under (possibly) alternative CSI representations of this choice data. Now, consider once again the binary relation P_i defined above. It is straightforward to verify that if xP_iy, then x is revealed to be preferred to y. To see why, suppose xP_iy with $S, R \in \mathcal{P}(X)$ and (c_i, c_{-i}) as in Statement (1) above. Further, let $(\succ_i, \tau_i)$ be part of a CSI representation. Then, $c_i(R) \in \Gamma_i(R; c_{-i}(R), \tau_i)$. That is, there exists $G' \in \mathcal{G}_i$ such that $\gamma_{i,R}(c_i(R), G') \geq \tau_i$. But, this then implies that $\gamma_{i,S}(y, \tilde{G}) \geq \gamma_{i,R}(c_i(R), G') \geq \tau_i$. Hence, $y \in \Gamma_i(S; c_{-i}(S), \tau_i)$ and, by the definition of CSI, it follows that $c_i(S) = x \succ_i y$. Next, define P_i^* to be the transitive closure of P_i. It also follows that if xP_i^*y, then x is revealed to be preferred to y. Loosely speaking, this is true because if xP_iz and zP_iy (and hence, xP_i^*y) for some z then, since the underlying preference relations defining a CSI representation are transitive, it follows that x is revealed to be preferred to y by i even when xP_iy is not directly revealed from choices. A natural question that arises is whether all revealed preferences are contained in P_i^*. That is, could there exist x and y such that x is revealed to be preferred to y by i, but $\neg[xP_i^*y]$. The following result establishes that this cannot be the case. Therefore, the binary relation P_i^* characterizes the extent to which i's preference rankings can be uniquely identified.

Proposition 5.1. *Let $(c_i)_{i\in\mathcal{I}}$ be a CSI choice profile. Then x is revealed to be* *preferred to y by individual i if and only if xP_i^*y.*

Proof: Please refer to Section A.3.

Example 4.2 (cont'd) We derived above that $P_1 = \{(y,x),(y,z)\}$, $P_2 = \{(z,y),(y,x)\}$, $P_3 = \{(x,y),(z,y)\}$, and $P_4 = P_5 = \{(x,y),(y,z)\}$. As such $P_1^* = P_1$, $P_2^* = \{(z,y),(y,x),(z,x)\}$, $P_3^* = P_3$, and $P_4^* = P_5^* = \{(x,y),(y,z),(x,z)\}$. Hence, in any CSI representation $(\succ_i, \tau_i)_{i=1}^5$, $\succ_2$, $\succ_4$, and $\succ_5$ are uniquely determined with $\succ_2 = P_2^*$, $\succ_4 = P_4^*$, and $\succ_5 = P_5^*$. However, $\succ_1$ and $\succ_3$ are not and their identification is restricted to $P_1 \subseteq \succ_1$ and $P_3 \subseteq \succ_3$. In particular, both the linear orders $y \succ_1 x \succ_1 z$ and $y \succ_1' z \succ_1' x$ for 1, and $x \succ_3 z \succ_3 y$ and $z \succ_3' x \succ_3' y$ for 3 are consistent with a CSI-representation.

We now address the question of the extent to which the DM's consideration sets can be identified.

Definition 5.2. *Let $(c_i)_{i\in\mathcal{I}}$ be a CSI choice profile. We say that x is revealed to receive i's consideration in choice problem S if for every threshold τ_i that is part of a CSI representation of these choices, we have $x \in \Gamma_i(S; c_{-i}(S), \tau_i)$*

That is, the notion of revealed consideration identifies those alternatives in a choice problem which are guaranteed to receive consideration irrespective of what the social influence thresholds under the CSI-representation of these choices are.

Proposition 5.2. *Let $(c_i)_{i\in\mathcal{I}}$ be a CSI choice profile. Then, for any $i \in \mathcal{I}$, x is revealed to receive consideration in choice problem S if and only if there exists $T \in \mathcal{P}(X)$ (possibly equal to S) and a $\tilde{G} \in \mathcal{G}_i$ such that $\gamma_{i,S}(x,\tilde{G}) \geq \gamma_{i,T}(c_i(T), G)$ for all $G \in \mathcal{G}_i$.*

Proof: Please refer to Section A.4.

Going back to Example 4.2, for any $S \in \mathcal{P}(X)$, all $x \in \underline{\Gamma}_i(S, c_{-i}, \bar{\tau}_i)$ shown in Table 4 are revealed to receive i's consideration.

6 Difference and (not just) Similarity

In the analysis thus far, we have maintained that for a DM to identify an alternative with one of his reference groups, the alternative has to be a typical choice amongst members of this group in the sense of being chosen by sufficiently many in the group. That is, the alternative has to signal a sense of similarity within the group. However, in some situations, just this sense of ingroup similarity by itself may not be enough for a DM to identify an alternative with a reference group. What may also need to be true is that this alternative should be an atypical choice among those not in this group so that observed patterns of choice with respect to this alternative differentiate this group and its members from those not in it. This is likely to be the case especially when the alternative under consideration has a snob or signalling value, or has partisan appeal when it comes to expressing the identity of certain groups or individuals. Examples of such effects can be found with respect to alternatives in the realm of partisan politics, high-end fashion, luxury goods, expressing distinct or nouveau tastes in things like music etc. For instance, think of Oliver once again. It may just so be that if a significant number of people outside Oliver's aforementioned reference group were to start choosing Abercrombie and Fitch, he may no longer identify this brand with this reference group, even if significant numbers within this group were to continue choosing it. This may be so because others outside this group using this brand may dilute the signalling value that this brand has with respect to this group. The following incident narrated by Berger (2016) illustrates this point.

Between 2009 and 2012, MTV had run a reality show by the name of Jersey Shore which followed the lives of eight housemates. The show itself was quite offensive as it played up the worst kind of Italian-American stereotypes. One of the housemates was an individual named Micheal "The Situation" Sorrentino, who became quite well known from the show but who also elicited very strong reactions based on the way he conducted himself. Certain sections found his mannerisms and personality quite repulsive. As it turns out, this strong reaction that he evoked was a concern for Abercrombie and Fitch when it found Sorrentino wearing their label in public. Sorrentino, in terms of his sense of fashion and personality, could not have been more different than the traditional patrons of this brand. This left the brand in a fix as it did not want its carefully crafted brand image to be identified with Sorrentino. However, it did end up finding a rather ingenious way of resolving the situation. In 2010, the brand signed a contract with Sorrentino under which it ended up paying him a significant sum of money in return for the guarantee that he will *not* be seen wearing their clothes in the future! It is instructive to quote from the press release that Abercrombie and Fitch released on the matter: "We are deeply concerned that Mr. Sorrentino's association with our brand could cause significant damage to our image . . . this association is contrary to the aspirational nature of our brand, and may be distressing to many of our fans." In other words, the very fact that Sorrentino is seen wearing this brand may be enough to dilute its image as representing the identity and attitude of hip and trendy youngsters who are its traditional patrons.

We want to take this kind of reasoning about group identification seriously. Therefore, in the extension of our baseline model that we are proposing here, we entertain the possibility that along with ingroup similarity, the process of group identification may also be about outgroup difference. In terms of its empirical content, this approach allows us to consider the case of non-conformity, which is another way in which social influence may impact behavior.

We can now extend our earlier definitions to capture the situation where along with ingroup similarity, outgroup difference plays a role in determining the alternatives that a DM considers in any choice problem. We continue to work here with the assumption that DMs' reference groups are commonly known and directly observable. This analysis can be extended to the case where reference groups are not directly observed along the lines suggested in the next section. Now, for any $i \in \mathcal{I}$ and $G \in \mathcal{G}_i$, let G^c denote the complement of G in $\mathcal{I}_{-i}$. We may think of G^c as the outgroup of the reference group G. Further, let $\mathcal{G}_i^c = \{G^c \subseteq \mathcal{I}_{-i} : G \in \mathcal{G}_i\}$ denote the set of all such outgroups of i's reference groups. Finally, extend the domain of the γ-functions defined earlier to include $\mathcal{G}_i^c$: define, for any choice problem $S \in \mathcal{P}(X)$, a mapping $\hat{\gamma}_{i,S} : S \times (\mathcal{G}_i \cup \mathcal{G}_i^c) \to [0,1]$ given by:

$$\hat{\gamma}_{i,S}(x, G) = \begin{cases} \frac{\#\{j \in G : c_j(S) = x\}}{\#G} & \text{if } G \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

That is $\hat{\gamma}_{i,S}(x, .)$ measures the proportion of individuals choosing alternative x in S, not just in each of i's reference groups but also in the outgroups of these reference groups. If $\mathcal{I}_{-i} \in \mathcal{G}_i$ and, therefore, $\emptyset \in \mathcal{G}_i^c$, we adopt the convention that $\hat{\gamma}_{i,S}(x, \emptyset) = 0$.

We can now define what it means to say a DM identifies some alternative in a choice problem with one of his reference groups in this extended sense.

Definition 6.1. *Given $c_{-i}(S)$, individual i strongly identifies alternative x with group $G \in \mathcal{G}_i$ in choice problem $S \in \mathcal{P}(X)$ if $\hat{\gamma}_{i,S}(x, G) \geq \tau_i \geq \tau_i' \geq \hat{\gamma}_{i,S}(x, G^c)$, $\tau_i, \tau_i' \in [0,1]$. An alternative $x \in S$ receives i's consideration via strong social influence in choice problem $S \in \mathcal{P}(X)$ if he strongly identifies x with some group $G \in \mathcal{G}_i$.*

Strong identification is captured by two social influence threshold parameters for each i, $\tau_i, \tau_i' \in [0,1]$, with $\tau_i \geq \tau_i'$. For any $i \in \mathcal{I}$ to strongly identify some alternative x in a choice problem S with a reference group $G \in \mathcal{G}_i$, not only must the proportion of individuals in G who choose x be at least as great as the threshold τ_i, but the proportion choosing x in the corresponding outgroup G^c must be no greater than τ_i'. We denote the set of all alternatives that receive i's consideration via strong social influence in a choice problem S by $\hat{\Gamma}_i(S; c_{-i}(S), \tau_i, \tau_i')$. This naturally leads us to the following definition of the CSI procedure in this context, which we refer to as choice via strong social influence.

Definition 6.2. *The profile of choice functions* $(c_i)_{i \in \mathcal{I}}$ *is a choice via strong social influence (CSI-S) if for each* i*, there exists a linear order* $\succ_i$ *on* X *and a pair of social influence thresholds* $\tau_i, \tau_i' \in [0,1]$*,* $\tau_i \geq \tau_i'$*, such that for any* $i \in \mathcal{I}$ *and* $S \in \mathcal{P}(X)$*,*

1. $c_i(S) \in \hat{\Gamma}_i(S; c_{-i}(S), \tau_i, \tau_i')$
2. *If* $y \in \hat{\Gamma}_i(S; c_{-i}(S), \tau_i, \tau_i')$*,* $y \neq c_i(S)$*, then* $c_i(S) \succ_i y$

CSI-S works just like CSI except for the way the DMs' consideration sets are constituted.

To understand better the empirical content of CSI-S, consider the following example which highlights an experimental finding that has received some attention in the social influence literature. The finding underscores the point that, in many contexts, we see greater diversity and variance in choices when these are made as part of a group as opposed to when they are made individually. For instance, Ariely and Levav (2000) report results from several experiments establishing this observation. In one of their experiments, for instance, patrons at a local microbrewery were given the option of getting a free beer from a menu of four beers. Two different experimental groups were created. The first comprised of patrons who were asked to report their choice over the four beers in an individual setting with no other patron around them when making the choice. The second comprised of patrons who were randomly assigned to tables with two or more other patrons and their orders were taken in this group setting. The authors found that there was a greater variance in the orders in the group setting as compared to the individual setting. A post experiment survey found that, on average, patrons in the individual setting were much happier with their orders than patrons who made these orders in the group setting. Presumably, this is because, owing to social influence—specifically the desire to be different from the others—the latter group of patrons ended up deviating from their preferred choice that their tastes would have dictated. The example below replicates the spirit of these experimental findings in the context of the CSI-S model.

Example 6.1 (Greater Diversity in Choices)**.** Let $X = \{v, w, x, y, z\}$ be the set of alternatives and $\mathcal{I} = \{1, 2, 3, 4, 5, 6\}$, be the set of individuals in society. Let $x \succ_1 y \succ_1 z \succ_1 w \succ_1 v$, $x \succ_2 z \succ_2 y \succ_2 v \succ_2 w$, $v \succ_3 y \succ_3 w \succ_3 z \succ_3 x$, $v \succ_4 w \succ_4 z \succ_4 y \succ_4 x$, $v \succ_5 w \succ_5 y \succ_5 z \succ_5 x$, and $v \succ_6 z \succ_6 w \succ_6 y \succ_6 x$ denote these individuals' preferences. Observe that with these preferences, if choices were to be made individually by people, then we will see just two choices being made from

X: x, which would be made by individuals 1 and 2, and v which would be made by individuals 3 to 6. However, suppose that instead of these choices being made individually, they were made in a group setting. Further, imagine that we are in a setting like that of the beer tasting experiment in Ariely and Levav (2000) where the experimental subjects have no reason to differentiate between individuals. As such, it is reasonable to assume that their reference groups are the set of all other individuals viewed as a homogenous unit and each individual viewed individually. That is $\mathcal{G}_i = \mathcal{I}_{-i} \cup \{\{j\} : j \neq i\}$. Further, to keep matters simple, let $\tau_i = \tau_i' = 1/2$ for all $i \in \mathcal{I}$. It is then straightforward to verify that Table 5 illustrates a choice profile in X that can be supported as a CSI-S choice profile. To see this, observe that $\hat{\Gamma}_i(X; c_{-i}(X), \tau_i, \tau_i') = \{x, y, z\}$ for all i. Here, individuals 1 and 2 are choosing their most preferred alternative, but no other individual does so, i.e., 3, 4, 5, 6 are not choosing v.

	$\{v, w, x, y, z\}$
$c_1(.)$	x
$c_2(.)$	x
$c_3(.)$	y
$c_4(.)$	z
$c_5(.)$	y
$c_6(.)$	z

Table 5

Next, we come to the question of behavioral characterization of the CSI-S model. As it turns out a single axiom characterizes this model. To present this axiom, first, a definition needs to be stated. We refer to a collection $(G_S)_{S \in \mathcal{P}(X)}$, $G_S \in \mathcal{G}_i$, as a *salient collection of reference groups* for $i \in \mathcal{I}$ if

1. $\hat{\gamma}_{i,S}(c_i(S), G_S) \geq \hat{\gamma}_{i,S}(c_i(S), G_S^c)$, for all $S \in \mathcal{P}(X)$; and
2. $\bigcap_{S \in \mathcal{P}(X)} [\hat{\gamma}_{i,S}(c_i(S), G_S^c), \hat{\gamma}_{i,S}(c_i(S), G_S)] \neq \emptyset$

The import of a salient collection of reference groups, $(G_S)_{S \in \mathcal{P}(X)}$, $G_S \in \mathcal{G}_i$, in the context of the CSI-S model is the following. Note that since identification here is in the strong sense, a *necessary* condition for an alternative x to receive i's consideration via strong social influence in a choice problem S is that there exists some reference group $G \in \mathcal{G}_i$ such that $\hat{\gamma}_{i,S}(x, G) \geq \hat{\gamma}_{i,S}(x, G^c)$. We know that if the profile $(c_i)_{i \in \mathcal{I}}$ is a CSI-S, then for each i and each choice problem S, $c_i(S)$ must receive i's consideration via strong social influence in that choice problem. That is, there must exist $G_S \in \mathcal{G}_i$ such that $\hat{\gamma}_{i,S}(c_i(S), G_S) \geq \hat{\gamma}_{i,S}(c_i(S), G_S^c)$.

Moreover, for each such S, this set G_S must be such that $\hat{\gamma}_{i,S}(c_i(S), G_S) \geq \tau_i \geq \tau'_i \geq \hat{\gamma}_{i,S}(c_i(S), G^c_S)$. In other words, $\bigcap_{S \in \mathcal{P}(X)} [\hat{\gamma}_{i,S}(c_i(S), G^c_S), \hat{\gamma}_{i,S}(c_i(S), G_S)] \neq \emptyset$. That is, if $(c_i)_{i\in\mathcal{I}}$ is a CSI-S, then such a salient collection of reference groups must exist. The axiom that we introduce posits the existence of such a salient collection of reference groups and uses it to identify when an alternative receives consideration. In so doing, it makes it possible to retrieve a WARP-like condition, similar in spirit to WARP-SI.[14]

Axiom 6.1 (WARP-SSI)**.** *For any $i \in \mathcal{I}$, there exists (i) a salient collection of reference groups $(G_S)_{S\in\mathcal{P}(X)}$, $G_S \in \mathcal{G}_i$, and (ii) for any $S \in \mathcal{P}(X)$, an alternative $x^* \in S$ such that, for any $T \in \mathcal{P}(X)$ if*

1. *$x^* \in T$, $c_i(T) \in S$, and*
2. *there exists $R, R' \in \mathcal{P}(X)$ and $\tilde{G} \in \mathcal{G}_i$ satisfying $\hat{\gamma}_{i,T}(x^*, \tilde{G}) \geq \hat{\gamma}_{i,R}(c_i(R), G_R)$ and $\hat{\gamma}_{i,T}(x^*, \tilde{G}^c) \leq \hat{\gamma}_{i,R'}(c_i(R'), G^c_{R'})$*

then $c_i(T) = x^*$.

WARP-SSI inherits the logic of WARP-SI with the appropriate modification made for the stronger notion of what it means for a DM to identify an alternative with one of his reference groups in a choice problem. Condition 2 in the statement of WARP-SSI indeed implies that x^* must receive i's consideration via strong social influence in choice problem T, provided the salient collection of reference groups $(G_S)_{S\in\mathcal{P}(X)}$ has been consistently identified w.r.t. the two social influence threshold parameters, τ_i and τ'_i. If this is the case, then, as noted above, $\tau_i \leq \hat{\gamma}_{i,R}(c_i(R), G_R) \leq \hat{\gamma}_{i,T}(x^*, \tilde{G})$ and $\tau'_i \geq \hat{\gamma}_{i,R}(c_i(R), G^c_R) \geq \hat{\gamma}_{i,T}(x^*, \tilde{G}^c)$. Hence, x^* receives consideration in T.

The following result establishes that WARP-SSI forms the behavioral foundation of the CSI-S model.

Theorem 6.1. *The choice profile $(c_i)_{i\in\mathcal{I}}$ is a CSI-S if and only if it satisfies WARP-SSI.*

Proof: Please refer to Section A.5.

[14]Once again, this axiom applies to a choice profile $(c_i)_{i\in\mathcal{I}}$ and not piecewise to individual choice functions.

7 Unobserved Reference Groups

So far in the analysis, we have assumed that the DMs' reference groups are commonly known and the researcher can directly observe them. With the benefit of such observability, we identified a condition on choices that allowed us to determine when such choices are consistent with the CSI (resp., CSI-S) procedure. In the process, we deduced the extent to which the DMs' preferences and social influence thresholds determining consideration sets can be backed out from the choice data. We can extend this analysis to the case where these groups are unobserved. That is, we can attempt to identify the reference groups from the choice data along with the other parameters. However, given that we use the same data to identify more parameters, the extent to which we can uniquely identify them reduces. We substantiate these observations in this section.

We do the analysis here in the context of the baseline CSI model. A similar set of observations can be made for the CSI-S model. To develop this analysis, we can start by considering, for each i, a non-empty collection of non-empty subsets of $\mathcal{I}_{-i}$, denoted by $\bar{\mathcal{F}}_i$. Think of $\bar{\mathcal{F}}_i$ as a collection of potential reference groups from which it is known that i's actual set of reference groups, $\mathcal{F}_i^*$, is drawn, i.e., it is known that $\mathcal{F}_i^* \subseteq \bar{\mathcal{F}}_i$. The set $\bar{\mathcal{F}}_i$ represents the extent of knowledge that the outside observer has about i's reference groups. It is conceivable that this set is equal to the set of all non-empty subsets of $\mathcal{I}_{-i}$ in which case the outside observer, a priori, knows nothing about the structure of i's reference groups. However, it is also possible that the outside observer has some a priori knowledge about these groups ($\bar{\mathcal{F}}_i$ is a *strict* subset of the set of all non-empty subsets of $\mathcal{I}_{-i}$) including knowing them completely, which was the case considered above. Observe that, in this setting, a CSI representation specifies for each i, a triple $(\succ_i, \tau_i, \mathcal{F}_i^*)$, i.e., along with any individual's preference ranking and social influence threshold, the representation should also specify her collection of reference groups. We call such a representation a CSI* representation. Formally,

Definition 7.1. *Given $(\bar{\mathcal{F}}_i)_{i\in\mathcal{I}}$, the profile of choice functions $(c_i)_{i\in\mathcal{I}}$ is a choice via (unobserved) social influence (CSI*) if for each i, there exists a collection $\mathcal{F}_i^* \subseteq \bar{\mathcal{F}}_i$, a linear order $\succ_i$ on X and a social influence threshold $\tau_i \in [0,1]$ such that for any $i \in \mathcal{I}$ and $S \in \mathcal{P}(X)$,*

1. *$c_i(S) \in \Gamma_i(S; c_{-i}(S), \tau_i, \mathcal{F}_i^*)$,[15] and*
2. *If $y \in \Gamma_i(S; c_{-i}(S), \tau_i, \mathcal{F}_i^*)$, $y \neq c_i(S)$, then $c_i(S) \succ_i y$*

[15] The consideration set $\Gamma_i(.)$ depends on what the collection of reference groups $\mathcal{F}_i^*$ is. That is, $x \in \Gamma_i(S; c_{-i}(S), \tau_i, \mathcal{F}_i^*)$ if there exists $G \in \mathcal{F}_i^*$ such that $\gamma_{i,S}(x, G) \geq \tau_i$.

Based on our earlier work, it is straightforward to establish that the following axiom characterizes a CSI* representation in this setting.

Axiom 7.1 (WARP-SI*). *For any $i \in \mathcal{I}$, there exists (i) a collection $\mathcal{F}_i \subseteq \bar{\mathcal{F}}_i$ and (ii) for any choice problem $S \in \mathcal{P}(X)$, an alternative $x^* \in S$ such that, for any $T \in \mathcal{P}(X)$, if*

1. *$x^* \in T$, $c_i(T) \in S$, and*
2. *there exists $R \in \mathcal{P}(X)$ and $\tilde{G} \in \mathcal{F}_i$ satisfying $\gamma_{i,T}(x^*, \tilde{G}) \geq \gamma_{i,R}(c_i(R), G)$, for all $G \in \mathcal{F}_i$*

then $c_i(T) = x^*$.

We now want to highlight the extent to which the model parameters, say, preferences, can be uniquely identified under a CSI* representation. To that end, consider a choice profile $(c_i)_{i\in\mathcal{I}}$ for which a CSI* representation $(\succ_i, \tau_i, \mathcal{F}_i^*)_{i\in\mathcal{I}}$ exists. For such a choice profile and for any i, let $\mathcal{F}_i^m \subseteq \bar{\mathcal{F}}_i$, $m = 1, \ldots, M_i$, denote the exhaustive list of reference group structures for i which satisfies WARP-SI*. Clearly, if $(\succ_i, \tau_i, \mathcal{F}_i^*)_{i\in\mathcal{I}}$ is a CSI representation in this setting, then $\mathcal{F}_i^* \in \{\mathcal{F}_i^1, \ldots, \mathcal{F}_i^{M_i}\}$. Now, for each $m = 1, \ldots, M_i$, define the binary relation P_i^m on X as follows: For any distinct $x, y \in X$, $xP_i^m y$ if there exists $S, R \in \mathcal{P}(X)$, with $x, y \in S$ and $c_i(S) = x$, and $\tilde{G} \in \mathcal{F}_i^m$ satisfying

$$\gamma_{i,S}(y, \tilde{G}) \geq \gamma_{i,R}(c_i(R), G), \forall G \in \mathcal{F}_i^m$$

Finally, define the binary relation $\bar{P}_i$ on X by $\bar{P}_i = \bigcap_{m=1}^{M_i} P_i^m$. What can be shown along similar lines as the proof of Proposition 5.1 is that, under a CSI* representation in this setting, i's preferences can be uniquely identified up to the transitive closure of $\bar{P}_i$, denoted by $\bar{P}_i^*$. Clearly, for the same choice profile $(c_i)_{i\in\mathcal{I}}$ that satisfies WARP-SI in the baseline setting with known reference groups $\mathcal{G}_i$ and WARP-SI* in the current setting, if $\mathcal{G}_i \subseteq \bar{\mathcal{F}}_i$, then $\bar{P}_i^* \subseteq P_i^*$ (where P_i^* refers to the revealed preferences of the baseline model). We illustrate this below in the context of Example 4.2.

Example 4.2 (cont'd) Recall that, in this example, the five individuals in society are positioned in a circle: $1-2-3-4-5-1$. Suppose that the only knowledge that the outside observer has about each individual i's reference groups is that they comprise i's neighbors. However, it is not clear to this observer, whether this includes the two immediate neighbors alone, or, whether it also includes the second immediate neighbors, or, both. That is, in this case the outside observer's knowledge about reference groups is captured by $\bar{\mathcal{F}}_1 = \{\{2,5\}, \mathcal{I}_{-1}\}$,

$\bar{\mathcal{F}}_2 = \{\{1,3\},\mathcal{I}_{-2}\}$, $\bar{\mathcal{F}}_3 = \{\{2,4\},\mathcal{I}_{-3}\}$, $\bar{\mathcal{F}}_4 = \{\{3,5\},\mathcal{I}_{-4}\}$, and $\bar{\mathcal{F}}_5 = \{\{1,4\},\mathcal{I}_{-5}\}$. For each i, $\#\{\mathcal{F}_i \neq \emptyset : \mathcal{F}_i \subseteq \bar{\mathcal{F}}_i\} = 3$ and it can be shown that each of these three sets satisfies WARP-SI*. For $i = 1,2,3$ and any $\mathcal{F}_i^m \subseteq \bar{\mathcal{F}}_i$, $m = 1,2,3$, we can show that $P_1^m = \{(y,x),(y,z)\}$, $P_2^m = \{(z,y),(y,x)\}$, and $P_3^m = \{(x,y),(z,y)\}$. Hence, $\bar{P}_1 = \{(y,x),(y,z)\}$, $\bar{P}_2 = \{(z,y),(y,x)\}$, and $\bar{P}_3 = \{(x,y),(z,y)\}$. On the other hand, for $i = 4$, if $\mathcal{F}_4^1 := \{\{3,5\}\}$, then $P_4^1 = \{(y,z)\}$; and, if $\mathcal{F}_4^2 := \{\mathcal{I}_{-4}\}$, or, $\mathcal{F}_4^3 := \bar{\mathcal{F}}_4$, then $P_4^2 = P_4^3 = \{(x,y),(y,z)\}$. Hence, $\bar{P}_4 = \{(y,z)\}$. Finally, for $i = 5$, if $\mathcal{F}_5^1 := \{\{1,4\}\}$, then $P_5^1 = \{(x,y)\}$; and, if $\mathcal{F}_5^2 := \{\mathcal{I}_{-5}\}$, or, $\mathcal{F}_5^3 := \bar{\mathcal{F}}_5$, then $P_5^2 = P_5^3 = \{(x,y),(y,z)\}$. Hence, $\bar{P}_5 = \{(x,y)\}$. Accordingly, the transitive closures of these binary relations are given by $\bar{P}_1^* = \bar{P}_1 = \{(y,x),(y,z)\}$, $\bar{P}_2^* = \{(z,y),(y,x),(z,x)\}$, $\bar{P}_3^* = \bar{P}_3 = \{(x,y),(z,y)\}$, $\bar{P}_4^* = \bar{P}_4 = \{(y,z)\}$, and $\bar{P}_5^* = \bar{P}_5 = \{(x,y)\}$. Recall that $P_1^* = \{(y,x),(y,z)\}$, $P_2^* = \{(z,y),(y,x),(z,x)\}$, $P_3^* = \{(x,y),(z,y)\}$, and $P_4^* = P_5^* = \{(x,y),(y,z),(x,z)\}$. Clearly, $\bar{P}_i^* \subseteq P_i^*$, for all i, and this inclusion is strict for $i = 4,5$.

8 Uni-directional Social Influence: A Famous Example

A special case of the model that we have developed is where the social influence is uni-directional. That is, choices of others influence a given decision maker by impacting his consideration sets, but not vice versa. As such, the interactive component in decision making is not there and it becomes a "pure" choice problem in the classical sense. To illustrate this case and to further highlight the empirical content of our theory, we relate it to a famous experiment by Asch (1955) on conformity and group influence. We focus on this experiment here because it is often presented as the leading example in the social psychology literature on social influence. To the best of our knowledge, our's is the first work that rationalizes this evidence within a choice-theoretic setting that economists adopt. Using it, we also highlight, once again, how our model can account for minority influence.

In these experiments, groups of 7 to 9 individual participants were faced with tasks that, on its face, appeared to assess their visual judgment skills. Participants were required to match a single line with its identical twin out of a set of three lines of different lengths. The actual intention of Asch's study, however, was to analyze whether and when individuals succumb to the view of a majority group. To that end, each experimental group consisted of just one real subject, while the rest were "confederates" who were instructed to give pre-specified answers which happened to be correct on some trials and false on others. Participants were asked to announce their answers sequentially in public with the real subject typically being the last in the sequence. That is, he heard the responses of all the other participants before providing his own answer.

In the baseline treatment, the confederates were instructed to all agree on an answer in each of the 18 trials and to agree on one of the wrong answers in 12 of these trials. The unanimous view in this baseline condition induced the real subjects to agree with a wrong judgment in 36.8% of all trials. In the control condition where subjects had to match the lines individually and in isolation, the frequency of wrong answers was less than 1%. This suggests that on average about 1/3 of subjects succumb or conform to the wrong view provided that it is unanimous. In a variation of the baseline treatment, one of the confederates was instructed to always state the correct answer while the rest of the confederates all agreed on an answer. When the set of trials in the treatment in which all but one of the confederates gave the wrong answer was compared to the set of trials in the baseline treatment in which all the confederates gave the wrong answer, the experimental subject's propensity to give the wrong answer reduced by 75%. This illustrates that sometimes the critical influence in behavior or cognition can come from a small minority and influence need not be simply majoritarian in nature. That is why Asch's experiments are often a reference point for theories of minority influence in social psychology (e.g., see Moscovici (1976)).

We now analyze the observed conformity behavior in Asch's experiments against the predictions of our model. We will conduct this exercise using the CSI model. Note that both the real subject in the experiment and the active DM in the CSI model, say $i = n$, can take the choices of the remaining individuals as given. Further, suppose that $\mathcal{G}_n = \{\{1\}, \ldots, \{n-1\}, \{1, \ldots, n-1\}\}$. Such a specification of reference groups makes sense for, given the setup of the experiment, n has no basis to take a strict non-singleton subset of $\mathcal{I}_{-n}$ as a reference group as all the other $n-1$ individuals are ex ante identical from his perspective. Therefore, it seems plausible that his reference groups are either any individual viewed in isolation or the set of all individuals viewed as a homogenous unit. To make social influence salient, let $\tau_n > 0$. In the way of notation, denote the two wrong answers (choices) by w and w' and the correct answer by c, such that $X = \{w, w', c\}$ denotes the set of alternatives.

First, consider the baseline treatment in which all the confederates unanimously picked one of the wrong answers, say, w. Observe that, in this choice problem with the given choices, c_{-n}, individual n identifies w with all his reference groups whereas neither of the other two alternatives receive consideration. Hence, his consideration set in this choice problem is $\Gamma_n(X; .) = \{w\}$. Accordingly, our theory implies that irrespective of his preference ranking he must conform and chooses w. This is in line with 1/3 of the subjects in the experimental treatment who succumb to the unanimous majority view.

Next, turn to Asch's variation of the baseline treatment in which there is one truthful confederate who deviates from the choices of the other confederates and chooses the correct answer c. Suppose this confederate is $m \in \{1, \ldots, n-1\}$. Given these choices, observe that, now, n identifies c with the singleton group $\{m\}$. He still identifies w with all the other singleton reference groups and does likewise with the group $\{1, \ldots, n-1\}$ provided $\tau_n \leq \frac{n-2}{n-1}$. Hence, his consideration set in this choice problem given these choices is given by $\Gamma_n(X; .) = \{w, c\}$. Given the evidence from the control condition that essentially rules out any wrong answers, it is reasonable to assume that the correct answer is the top element of n's preference ranking. As such, our theory predicts that in this situation the experimental subject chooses the correct alternative, thus supporting the evidence of a significant reduction in conformity in the alternative treatment. Our theory, therefore, provides an illustration of the fact that group influence does not necessarily require a majority group as the exerting force. A minority's choices may also sway the DM to choose differently. The fact that our theory can account for minority influence allows us to distinguish it from other theories of social influence where influence is solely of the majoritarian type.

9 Related Literature

Our model in this paper relates to the recent literature on theories of behavioral choice that have sought to explicitly model the psychology of decision making procedures and departures from the standard rational choice paradigm. Our attempt at modeling decision makers who are socially influenced is one such departure from the rational choice paradigm.

A major distinguishing feature of the CSI model in relation to this literature is that it is not a single person choice model but captures an interactive choice procedure under which individuals mutually socially influence each other. The only other paper in this literature that we know of that incorporates others' behavioral details in a DM's choice procedure is the choice on mutual influence (CMI) model of Cuhadaroglu (2017), in which two DMs, 1 and 2, influence each other's choices, specified by choice correspondence C_1 and C_2, in the following way.

Definition 9.1. *A pair of choice correspondences (C_1, C_2) is a CMI if there exists a pair of asymmetric and transitive binary relations $(\succ_1, \succ_2)$ such that for all $S \in \mathcal{P}(X)$*

$$C_1(S) = \max(\max(S, \succ_1), \succ_2)^{16}$$
$$C_2(S) = \max(\max(S, \succ_2), \succ_1)$$

In other words, like in the CSI model, in the CMI model too the social context matters in the moment of decision-making. This relates these two models to other ones in the behavioral choice theory literature which emphasize that the context in which choices are made matters, e.g., Rubinstein and Salant (2006) and Salant and Rubinstein (2008)

Observe that the primitives of the CMI model are choice correspondences, whereas that of the CSI model are choice functions. Besides that another significant way in which the CSI procedure differs from the CMI procedure is in that, under CMI, the DM first applies her preferences to the set of available alternatives and only consults other individuals' preferences if her preferences are not decisive. That is, the individual element of the procedure comes in the first stage and social influence in the second. In the CSI procedure on the other hand, social influence shows up in the first stage in terms of shortlisting the set of available alternatives and the DM's preferences enters the picture in the second stage. The following short examples illustrate that CMI and CSI are distinct models and neither is a special case of the other.

Example 9.1 (a CMI but not a CSI)**.** Let $X = \{x, y, z\}$ be the set of alternatives and $\mathcal{I} = \{1, 2\}$ be the set of individuals in society. Further, let individual choices be such that $c_1(x, y) = c_2(x, y) = x$ and $c_1(X) = x$, $c_2(X) = y$. These choices are a CMI for $\succ_1 = \{(x, y), (x, z), (y, z)\}$ and $\succ_2 = \{(z, x)\}$. To see this, note that $\max(\max(X, \succ_2), \succ_1) = \{y\}$, $\max(\max(\{x, y\}, \succ_2), \succ_1) = \{x\}$, and $\max(X; \succ_1) = \max(\{x, y\}; \succ_1) = \{x\}$. On the other hand, the primitives of our model demand that $\mathcal{G}_i \neq \emptyset$, for all $i = 1, 2$. Hence, $\mathcal{G}_1 = \{2\}$ and $\mathcal{G}_2 = \{1\}$. Further, since c_2 violates WARP, τ_2 has to be greater than zero. Hence, if the choice profile (c_1, c_2) is a CSI, then $c_2(X) = c_1(X)$. But, under the specified choices, $c_2(X) \neq c_1(X)$.

Example 9.2 (a CSI but not a CMI)**.** Let $X = \{x, y, z\}$ be the set of alternatives and $\mathcal{I} = \{1, 2\}$ be the set of individuals in society. Further, let individual choices be such that $c_i(x, y) = c_i(x, z) = x$ and $c_i(X) = z$, for $i = 1, 2$. Observe that these choices are a CSI for $\mathcal{G}_1 = \{2\}$, $\mathcal{G}_2 = \{1\}$, $\tau_i > 0$, for $i = 1, 2$, and $x \succ_1 z \succ_1 y$, $z \succ_2 x \succ_2 y$. On the other hand, no matter what individual preferences entail these choices cannot be a CMI. To see this, observe that $c_1(x, z) = x$ implies that $x \succ_i z$, for some $i \in \{1, 2\}$, such that $z \notin \max(X, \succ_i)$ and, thus, $c_i(X) \neq z$, for that i.

There are other papers in this literature that have productively employed the modeling construct of consideration sets. For instance, the choice with limited attention (CLA) model of Masatlioglu, Nakajima, and Ozbay (2012) uses consideration sets

[16]For any $T \in \mathcal{P}(X)$ and binary relation $\succ$ on X, $\max(T, \succ) := \{x \in T | \nexists y \in T \text{ such that } y \succ x\}$.

to model decision makers who have limited attention and therefore consider only a subset of available alternatives. The overwhelming choice model (OC) of Lleras, Masatlioglu, Nakajima, and Ozbay (2017) use consideration sets to model decision makers who may be overwhelmed by too many choices—a phenomenon referred to as choice overload—and, therefore, may not be able to consider all alternatives on offer. In the CLA model, the consideration set mapping satisfies a property referred to as attention filter. A consideration set mapping $\Gamma : \mathcal{P}(X) \rightarrow \mathcal{P}(X)$ is an attention filter if $\Gamma(S) = \Gamma(S \backslash \{x\})$ whenever $x \notin \Gamma(S)$. That is, a consideration set mapping is an attention filter if whenever an alternative that does not receive consideration in a set is removed from it, the consideration set does not change. The following example shows that the consideration set mapping in the CSI model need not be an attention filter.

Example 9.3. Let $S = \{x, y, z\}$ be a choice problem and $\mathcal{I} = \{1, 2\}$ be the set of individuals in society. Let $\mathcal{G}_i = \{\mathcal{I}_{-i}\}$ be each i's set of reference groups and let $\tau_i = 1/2$ be their thresholds. Consider the choice problems $\{x, y\}$ and $\{x, y, z\}$ with choices of individuals 1 and 2 over these sets as specified in Table 6. Then, verify that for each i, $z \notin \Gamma_i(\{x, y, z\}; .)$, but $\Gamma_i(\{x, y, z\}; .) = \{x\} \neq \{y\} = \Gamma_i(\{x, y\}; .)$. Hence, neither Γ_1, nor Γ_2 is an attention filter. On the other hand, the choice profile in Table 6 satisfies WARP-SI.

	$\{x, y\}$	$\{x, y, z\}$
$c_1(.)$	y	x
$c_2(.)$	y	x

Table 6

In that sense, the theoretical structure of the consideration set mapping in our model is different from that in the CLA model. On the other hand, in the OC model, the consideration set mapping satisfies a property referred to as competition filter. A consideration set mapping $\Gamma : \mathcal{P}(X) \rightarrow \mathcal{P}(X)$ is a competition filter if $x \in \Gamma(T)$ whenever $x \in \Gamma(S)$ and $T \subseteq S$. In other words, for a DM who is subject to choice overload, if an alternative receives consideration in a larger set, then it must receive consideration in a subset of it as well. It is straightforward to show that the consideration set mapping in the CSI model need not be a competition filter either.

In an insightful comparison, Lleras, Masatlioglu, Nakajima, and Ozbay (2017) also show that their model is closely related to the categorize then choose model of Manzini and Mariotti (2012a) and the model of rationalization of Cherepanov, Feddersen, and Sandroni (2013). To this end, they define for these two-stage choice procedures which also involve the idea of shortlisting at the first stage the appropriate consideration set mapping and compare these consideration set mappings to that in the OC model. In so doing, it becomes apparent that the three models are indistinguishable from each other on the basis of choice data alone even though they capture very different positive models of behavior. The axiom that characterizes these models is that of W(eak)WARP, i.e., if $c(\{x, y\}) = c(T) = x$ for some $T \in \mathcal{P}(X)$, then $c(S) \neq y$ for all $S \in \mathcal{P}(X)$ such that $\{x, y\} \subseteq S \subseteq T$. The following example illustrates a choice situation where WWARP is violated and shows that such violations can be accommodated by the CSI model.

Example 9.4 (Violation of WWARP)**.** Let $X = \{x, y, d_x, d_y\}$ be the set of alternatives and $\mathcal{I} = \{1, 2\}$ be the set of individuals in society. In the spirit of the attraction effect (Huber, Payne, and Puto 1982), let d_x (respectively, d_y) be clearly inferior to x (respectively, y). Suppose preferences of the two individuals are $x \succ_1 y \succ_1 d_x \succ_1 d_y$ and $y \succ_2 x \succ_2 d_y \succ_2 d_x$. Further, for $i = 1, 2$, thresholds are $0 < \tau_i \leq 1$ and $\mathcal{G}_i = \mathcal{I}_{-i}$. Then, the choices specified in Table 7 reveal the extended attraction effect and show that it can be accommodated by the CSI model. These choices clearly violate WWARP.

Table 7: Extended Attraction Effect

	$\{x, y\}$	$\{x, y, d_y\}$	$\{x, y, d_x, d_y\}$
$c_1(.)$	x	y	x
$c_2(.)$	x	y	x

A Proofs

A.1 Proof of Lemma 4.1

Necessity: Suppose P_i is acyclic. To establish that WARP-SI holds, consider any $S \in \mathcal{P}(X)$. Since this set is finite, there exists $x^* \in S$ such that there is no $y \in S$ with yP_ix^* or, else, P_i has a cycle. This means that for any $y \in S$, $y \neq x^*$, there does not exist $T \in \mathcal{P}(X)$ with $x^*, y \in T$ and $c_i(T) = y$, $R \in \mathcal{P}(X)$ and $\tilde{G} \in \mathcal{G}_i$ satisfying

$$\gamma_{i,T}(x^*, \tilde{G}) \geq \gamma_{i,R}(c_i(R), G), \forall G \in \mathcal{G}_i.$$

Accordingly, whenever $c_i(T) \in S$, it must be that $c_i(T) = x^*$, which establishes WARP-SI.

Sufficiency: Suppose P_i is not acyclic, i.e., there exists a sequence $(x_m)_{m=1}^M$ in X, $M \geq 3$, with $x_1 = x_M$ such that for each $m \in \{1, \ldots, M-1\}$, $x_mP_ix_{m+1}$. Of course, if a cycle exists, we can pick the sequence in a way such that $x_m \neq x_{m'}$, for all $m, m' \in \{1, \ldots, M-1\}$, $m \neq m'$. By definition of P_i in Section 4, this means that for each $m \in \{1, \ldots, M-1\}$, there exists $T_m \in \mathcal{P}(X)$, with $x_m, x_{m+1} \in T_m$ and $c_i(T_m) = x_m$, $S_m \in \mathcal{P}(X)$ and $\tilde{G} \in \mathcal{G}_i$ such that $\gamma_{i,T_m}(x_{m+1}, \tilde{G}) \geq \gamma_{i,S_m}(c_i(S_m), G)$ for all $G \in \mathcal{G}_i$. Now, consider the set $S := \{x_1, \ldots, x_{M-1}\}$. Then, for any $x \in S$, there exists $T \in \mathcal{P}(X)$, with $x \in T$ and $c_i(T) \in S$, $R \in \mathcal{P}(X)$ and $\tilde{G} \in \mathcal{G}_i$ such that $\gamma_{i,T}(x, \tilde{G}) \geq \gamma_{i,R}(c_i(R), G)$ for all $G \in \mathcal{G}_i$, but $c_i(T) \neq x$. That is, WARP-SI is violated.

A.2 Proof of Theorem 4.1

Proof. Necessity: Let $(c_i)_{i \in \mathcal{I}}$ be a CSI. That is, for each i there exist the objects $(\succ_i, \tau_i)$ that define a CSI representation. To show that any such c_i satisfies WARP-SI, in any set S, let the desired x^* be such that $x^* \succ_i y$ for all $y \in S \setminus \{x^*\}$. Now, consider any $T \in \mathcal{P}(X)$ such that $x^* \in T$, $c_i(T) \in S$ and x^* such that there exists $R \in \mathcal{P}(X)$ and $\tilde{G} \in \mathcal{G}_i$ satisfying

$$\gamma_{i,T}(x^*, \tilde{G}) \geq \gamma_{i,R}(c_i(R), G), \quad \forall G \in \mathcal{G}_i$$

Since $c_i(R) \in \Gamma_i(R; c_{-i}(R), \tau_i)$, it implies that $\gamma_{i,R}(c_i(R), G') \geq \tau_i$, for some $G' \in \mathcal{G}_i$. Accordingly, $\gamma_{i,T}(x^*, \tilde{G}) \geq \gamma_{i,R}(c_i(R), G') \geq \tau_i$. Hence, $x^* \in \Gamma_i(T; c_{-i}(T), \tau_i)$ and $c_i(T) = x^*$.

Sufficiency: Suppose that c_i satisfies WARP-SI for all $i \in \mathcal{I}$. To show that the choice profile $(c_i)_{i \in \mathcal{I}}$ is a CSI, we need to identify, for each i, the objects $(\succ_i, \tau_i)$ that define a CSI representation. We first define

$$\tau_i := \min_{S \in \mathcal{P}(X)} \max_{G \in \mathcal{G}_i} \gamma_{i,S}(c_i(S), G)$$

Next, consider the binary relation P_i over X which we defined as follows: For any distinct $x, y \in X$, $x P_i y$ if there exists $S \in \mathcal{P}(X)$, with $x, y \in S$, such that $c_i(S) = x$ and $\gamma_{i,S}(y, G) \geq \tau_i$, for some $G \in \mathcal{G}_i$. The proof of Lemma 4.1 establishes that P_i is asymmetric and acyclic.

Let P_i^* be the transitive closure of P_i. P_i^* is transitive and asymmetric, i.e., a partial order. By Szpilrajn's Theorem, we know that it can be extended to a linear order $\succ_i$ on X. We now verify that the objects $((\succ_i, \tau_i))_{i \in \mathcal{I}}$ represent the choice profile $(c_i)_{i \in \mathcal{I}}$ as a CSI. To that end, pick any $i \in \mathcal{I}$, $S \in \mathcal{P}(X)$ and let $c_i(S) = x$. By our definition of τ_i above, it follows that there exists $G \in \mathcal{G}_i$ such that $\gamma_{i,S}(x, G) \geq \tau_i$. So, $x \in \Gamma_i(S; c_{-i}(S), \tau_i)$. It remains to show that there is no alternative $y \in \Gamma_i(S; c_{-i}(S), \tau_i)$ with $y \succ_i x$. Suppose, to the contrary, that there is such a $y \in S$. Then, there exists $F \in \mathcal{G}_i$ such that $\gamma_{i,S}(y, F) \geq \tau_i$. By our definition of τ_i above, there exists $T \in \mathcal{P}(X)$ such that $\gamma_{i,S}(y, F) \geq \gamma_{i,T}(c_i(T), G)$, for all $G \in \mathcal{G}_i$. By our definition of P_i above, this implies that xP_iy. But, $P_i \subseteq \succ_i$ which implies that $x \succ_i y$, thus contradicting $y \succ_i x$, since $\succ_i$ is asymmetric. This establishes our desired conclusion. □

A.3 Proof of Proposition 5.1

Proof. Necessity: Suppose xP_i^*y does not hold. Then, the following two cases remain possible: Either yP_i^*x or $\neg[yP_i^*x]$. Consider the first case and let $\succ_i$ be a preference ranking (linear order) that is part of a CSI representation. Since P_i^* is defined as the transitive closure of P_i, yP_i^*x implies that there exists a sequence $(z_m)_{m=1}^M$ in X such that yP_iz_1, $z_1P_iz_2$, $\ldots$, z_MP_ix. Further, for any such $\succ_i$, since $P_i \subseteq \succ_i$ and $\succ_i$ is transitive, it follows that $y \succ_i x$. In the second case, where $\neg[yP_i^*x]$, there exists no sequence $(z_m)_{m=1}^M$ in X such that yP_iz_1, $z_1P_iz_2$, $\ldots$, z_MP_ix. Hence, it is as well possible to extend P_i^* to a linear order $\succ_i$ with $y \succ_i x$. The proof of Theorem 4.1 establishes that any such linear order $\succ_i$ can be part of a CSI representation. Therefore, in either case, x is not revealed to be preferred to y by i.

Sufficiency: We have already shown in Section 5 that if xP_iy, then x is revealed to be preferred to y by i. Now, consider the case when xP_i^*y. Since P_i^* is defined as the transitive closure of P_i, this implies that there exists a sequence $(z_m)_{m=1}^M$ in X such that xP_iz_1, $z_1P_iz_2$, $\ldots$, z_MP_iy. In this case, we know that for any $\succ_i$ that is part of a CSI representation, $P_i \subseteq \succ_i$ and, hence, $x \succ_i z_1$, $z_1 \succ_i z_2$, $\ldots$, $z_M \succ_i y$. Further, since $\succ_i$ is transitive it follows that $x \succ_i y$ and, hence, x is revealed to be preferred to y by i. □

A.4 Proof of Proposition 5.2

Proof. Necessity: First, note that the proof of Theorem 4.1 shows that there exists a CSI representation with

$$\bar{\tau}_i := \min_{S \in \mathcal{P}(X)} \max_{G \in \mathcal{G}_i} \gamma_{i,S}(c_i(S), G)$$

Now, consider some $S \in \mathcal{P}(X)$ and suppose that there exists no $T \in \mathcal{P}(X)$ and no $G \in \mathcal{G}_i$ such that $\gamma_{i,S}(x, G) \geq \gamma_{i,T}(c_i(T), G')$, for all $G' \in \mathcal{G}_i$. Then, $\gamma_{i,S}(x, G) < \max_{G' \in \mathcal{G}_i} \gamma_{i,T}(c_i(T), G')$, for all $T \in \mathcal{P}(X)$ and for all $G \in \mathcal{G}_i$, and it follows that $\gamma_{i,S}(x, G) < \bar{\tau}_i$, for all $G \in \mathcal{G}_i$. Hence, $x \notin \Gamma(S; c_{-i}(S), \tau_i)$. That is, it is not the case that x is revealed to receive i's consideration at S.

Sufficiency: If there exists $T \in \mathcal{P}(X)$ (possibly equal to S) and a $G \in \mathcal{G}_i$ such that $\gamma_{i,S}(x, G) \geq \gamma_{i,T}(c_i(T), G')$ for all $G' \in \mathcal{G}_i$, then $x \in \Gamma(S; c_{-i}(S), \bar{\tau}_i)$. But, by the definition of $\bar{\tau}_i$ above, it clearly holds that $\bar{\tau}_i \geq \tau_i$ for all τ_i that are part of a CSI representation. Hence, $x \in \Gamma_i(S; c_{-i}(S), \tau_i)$ for any such τ_i. That is, x is revealed to receive consideration at S. □

A.5 Proof of Theorem 6.1

Necessity of WARP-SSI for the representation is straightforward to establish and we omit the details here.

To establish the sufficiency of the axioms for the representation, first, note that according to WARP-SSI, for any $i \in \mathcal{I}$, we have a collection of salient reference groups $(G_S)_{S \in \mathcal{P}(X)}$ w.r.t. which the conclusion of the axiom follows. Now, for each such $i \in \mathcal{I}$, define :

- $T_i = \bigcap_{S \in \mathcal{P}(X)} [\hat{\gamma}_{i,S}(c_i(S), G_S^c), \hat{\gamma}_{i,S}(c_i(S), G_S)]$
- $\tau_i = \max T_i$ and $\tau_i' = \min T_i$
- a binary relation Q_i on X by xQ_iy if there exists $S \in \mathcal{P}(X)$ with $x, y \in S$, $c_i(S) = x$, and $\tilde{G} \in \mathcal{G}_i$, $R, R' \in \mathcal{P}(X)$ s.t. $\hat{\gamma}_{i,S}(y, \tilde{G}) \geq \hat{\gamma}_{i,R}(c_i(R), G_R)$ and $\hat{\gamma}_{i,S}(y, \tilde{G}^c) \leq \hat{\gamma}_{i,R'}(c_i(R'), G_{R'})$; equivalently, xQ_iy if there exists $S \in \mathcal{P}(X)$ with $x, y \in S$, $c_i(S) = x$, and $\tilde{G} \in \mathcal{G}_i$ s.t. $\hat{\gamma}_{i,S}(y, \tilde{G}) \geq \tau_i$ and $\hat{\gamma}_{i,S}(y, \tilde{G}^c) \leq \tau_i'$.

We can show along similar lines as in the proof of Lemma 4.1 that WARP-SSI implies that Q_i is acyclic. Further, let Q_i^* be the transitive closure of Q_i. Since Q_i^* is a partial order, by Szpilrajn's Theorem, it can be extended to a linear order $\succ_i$ on X as well.

Now consider any $S \in \mathcal{P}(X)$. Clearly, $\hat{\gamma}_{i,S}(c_i(S), G_S) \geq \tau_i$ and $\hat{\gamma}_{i,S}(c_i(S), G_S^c) \leq \tau_i'$. Hence, $c_i(S) \in \hat{\Gamma}_i(S; c_{-i}, \tau_i, \tau_i')$. Further, consider any other $y \in \hat{\Gamma}_i(S; c_{-i}, \tau_i, \tau_i')$, $y \neq c_i(S)$. That is, there exists $\tilde{G} \in \mathcal{G}_i$ s.t. $\hat{\gamma}_{i,S}(y, \tilde{G}) \geq \tau_i$ and $\hat{\gamma}_{i,S}(y, \tilde{G}^c) \leq \tau_i'$. In other words, there exists $R, R' \in \mathcal{P}(X)$ s.t. $\hat{\gamma}_{i,S}(y, \tilde{G}) \geq \hat{\gamma}_{i,R}(c_i(R), G_R)$ and $\hat{\gamma}_{i,S}(y, \tilde{G}^c) \leq \hat{\gamma}_{i,R'}(c_i(R'), G_{R'})$. Hence, it follows that $c_i(S)Q_iy$ and, accordingly, $c_i(S) \succ_i y$.

References

Ariely, D., and J. Levav (2000): "Sequential Choice in Group Settings: Taking The Road Less Traveled and Less Enjoyed," *Journal of Consumer Research*, 27(3), 279–290.

Asch, S. E. (1955): "Opinions and Social Pressure," *Scientific American*, 193, 31–35.

Berger, J. (2016): *Invisible Influence: The Hidden Forces that Shape Behavior*. Simon and Schuster.

Bertrand, M., and A. Morse (2016): "Trickle-down consumption," *Review of Economics and Statistics*, 98(5), 863–879.

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009): "Peer Effects and Social Networks in Education," *The Review of Economic Studies*, 76(4), 1239–1267.

Charles, K. K., E. Hurst, and N. Roussanov (2009): "Conspicuous Consumption and Race," *The Quarterly Journal of Economics*, 124(2), 425–467.

Cherepanov, V., T. Feddersen, and A. Sandroni (2013): "Rationalization," *Theoretical Economics*, 8(3), 775–800.

Clark, A. E. (2003): "Unemployment as a Social Norm: Psychological Evidence from Panel Data," *Journal of Labor Economics*, 21(2), 323–351.

Clark, A. E., and Y. Loheac (2007): "It wasnt me, it was them! Social Influence in Risky Behavior by Adolescents," *Journal of Health Economics*, 26(4), 763–784.

Cohen, G. L. (2003): "Party Over Policy: The Dominating Impact of Group Influence on Political Beliefs," *Journal of Personality and Social Psychology*, 85(5), 808–822.

Cornelissen, T., C. Dustmann, and U. Schönberg (2017): "Peer Effects in the Workplace," *The American Economic Review*, 107(2), 425–56.

Cuhadaroglu, T. (2017): "Choosing on Influence," *Theoretical Economics*, 12(2), 477–492.

Damm, A. P., and C. Dustmann (2014): "Does Growing up in a High Crime Neighborhood Affect Youth Criminal Behavior?," *The American Economic Review*, 104(6), 1806–1832.

Falk, A., and A. Ichino (2006): "Clean Evidence on Peer Effects," *Journal of Labor Economics*, 24(1), 39–57.

Frank, R. H., A. S. Levine, and O. Dijk (2014): "Expenditure Cascades," *Review of Behavioral Economics*, 1(1–2), 55–73.

Gaviria, A., and S. Raphael (2001): "School-Based Peer Effects and Juvenile Behavior," *The Review of Economics and Statistics*, 83(2), 257–268.

Glaeser, E. L., B. Sacerdote, and J. A. Scheinkman (1996): "Crime and Social Interactions," *The Quarterly Journal of Economics*, 111(2), 507–548.

Grinblatt, M., M. Keloharju, and S. Ikäheimo (2008): "Social Influence and Consumption: Evidence from the Automobile Purchases of Neighbors," *The review of Economics and Statistics*, 90(4), 735–753.

Hauser, J. R., and B. Wernerfelt (1990): "An Evaluation Cost Model of Consideration Sets," *Journal of Consumer Research*, 16(4), 393–408.

Heffetz, O. (2011): "A Test of Conspicuous Consumption: Visibility and Income Elasticities," *Review of Economics and Statistics*, 93(4), 1101–1117.

Huber, J., J. W. Payne, and C. Puto (1982): "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis," *Journal of Consumer Research*, 9(1), 90–98.

Janis, I. L. (1972): *Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascoes.* Houghton Mifflin.

Janis, I. L. (1982): *Groupthink: Psychological Studies of Policy Decisions and Fiascoes*, vol. 349. Houghton Mifflin Boston.

Kaus, W. (2013): "Conspicuous Consumption and Race: Evidence from South Africa," *Journal of Development Economics*, 100(1), 63–73.

Kuhn, P., P. Kooreman, A. Soetevent, and A. Kapteyn (2011): "The Effects of Lottery Prizes on Winners and their Neighbors: Evidence from the Dutch Postcode Lottery," *American Economic Review*, 101(5), 2226–47.

Lleras, J. S., Y. Masatlioglu, D. Nakajima, and E. Y. Ozbay (2017): "When More is Less: Limited Consideration," *Journal of Economic Theory*, 170, 70–85.

Lundborg, P. (2006): "Having the Wrong Friends? Peer Effects in Adolescent Substance Use," *Journal of Health Economics*, 25(2), 214–233.

Manzini, P., and M. Mariotti (2012): "Categorize Then Choose: Boundedly Rational Choice and Welfare," *Journal of the European Economic Association*, 10(5), 1141–1165.

——— (2012a): "Categorize Then Choose: Boundedly Rational Choice and Welfare," *Journal of the European Economic Association*, 10(5), 1141–1165.

Mas, A., and E. Moretti (2009): "Peers at Work," *The American Economic Review*, 99(1), 112–145.

Masatlioglu, Y., D. Nakajima, and E. Y. Ozbay (2012): "Revealed Attention," *American Economic Review*, 102(5), 2183–2205.

Moscovici, S. (1976): *Social Influence and Social Change*. Academic Press Inc.

Roberts, J. H., and J. M. Lattin (1991): "Development and Testing of a Model of Consideration Set Composition," *Journal of Marketing Research*, 28(4), 429–440.

Rubinstein, A., and Y. Salant (2006): "A Model of Choice from Lists," *Theoretical Economics*, 1(1), 3–17.

Sacerdote, B. (2001): "Peer Effects with Random Assignment: Results for Dartmouth Roommates," *The Quarterly Journal of Economics*, 116(2), 681–704.

Salant, Y., and A. Rubinstein (2008): "(A, f): Choice with Frames," *The Review of Economic Studies*, 75(4), 1287–1296.

Trogdon, J. G., J. Nonnemaker, and J. Pais (2008): "Peer Effects in Adolescent Overweight," *Journal of Health Economics*, 27(5), 1388–1399.

Zimmerman, D. J. (2003): "Peer Effects in Academic Outcomes: Evidence From a Natural Experiment," *The Review of Economics and Statistics*, 85(1), 9–23.

Zimmerman, G. M., and S. F. Messner (2010): "Neighborhood Context and the Gender Gap in Adolescent Violent Crime," *American Sociological Review*, 75(6), 958–980.