Stereotype Formation as Trait Aggregation *

Burak Can and M. Remzi Sanver†
Istanbul Bilgi University

April, 2007

Abstract

We propose an aggregation model which explains stereotype formation under the attribution hypothesis. We show, under very mild axioms, that an observer can be thought of perceiving a group in terms her subjective opinion about the representativeness of subgroups as well as a possible prejudice she might have.

*This paper is part of a project entitled "Social Perception - A Social Choice Perspective", supported by Istanbul Bilgi University Research Fund. It has been completed while Remzi Sanver was visiting Ecole Polytechnique, Paris. We are grateful to both institutions. We owe our initiation to the related social psychology literature to Joachim I. Krueger and hope that the outcome is worthy of his valuable guidance. Michael Birnbaum, Zoltan Csigas, Todd Davies, Robin Dawes, Rosemary Gosling, Kutsi Nircan, Paul Price, Michel Regenwetter and Ramon Shaban made stimulating comments on relating our theoretical findings to the relevant literature. The paper benefited a lot from comments made by Goksel Asan, Murat Ali Cengelci, Bora Erdamar, Kevin Hasker, Jean-Jacques Herrings, Doruk Iris, Emin Karagozoglu, Murat Kayi, Bettina Klaus, Semih Koray, Ugur Ozdemir, Ipek Ozkal-Sanver, Selcuk Ozyurt, Andres Perea, Hans Peters, Ozer Selcuk, Taner Yigit. We thank all contributors and acknowledge the responsibility of all possible errors.

†Corresponding author: sanver@bilgi.edu.tr

1 Posing the Problem

A stereotype is an overall judgment brought over a given group of objects, such as “Princeton students are smart”, “French food is delicious” or “Muslim women wear scarf”. Understanding the formation of stereotypes is a central question of social psychology. As Krueger et al. (2003) eloquently discuss, a main strand of the literature rests on the *attribution hypothesis* which assumes a direct association between traits and groups. Under the attribution hypothesis, an observer judges a group according to the traits he observes in that group. For example, he looks at Muslim women; sees that some wear a scarf and some do not; his mental processing of that observation leads to some kind of a general judgment about Muslim women such as “Muslim women wear scarf” or as “Muslim women do not wear scarf”. Of course, bringing no judgment hence avoiding a stereotype is also possible. According to the attribution hypothesis, a trait which is “sufficiently prevalent” in a given group is associated with that group. To quote Zawadski (1948), “The popular conception of a group characteristic seems to be a characteristic which is present in the majority of the members of the group. According to this concept, it is a necessary and sufficient condition for a group characteristic to be represented in at least 51 per cent of the members of the group”.

We find this understandable, even when informally stated. However, the concepts which embrace the attribution hypothesis can be made further intelligible through a formal model. It goes without saying, for an economic theorist, that once a relationship between the subjective perception of a group and the individual traits prevailing among the members of that group is postulated, tools of social choice theory have something to contribute to our understanding of the problem. In fact, the main concepts of stereotype formation under the attribution hypothesis can be formally expressed through an aggregation model à la Arrow (1951).

We propose the following model: Take some group, e.g., Turkish citizens, and a certain trait, e.g., smoking. Some of the members of the group do and some do not possess this trait and a judgement such as “Turks do smoke” is an aggregation of individual traits into a social one. So we can speak of a *perception function* that maps individual traits into a subjective stereotype about the society. More formally, we have a finite set N of individuals with $\#N \geq 2$, to which we refer as a *group*. There is a trait which the members of the group may or may not possess. We write $t_i = 1$ when $i \in N$ possesses this trait and $t_i = -1$ otherwise. We let $T = \{-1, 1\}^N$ stand for the set of *trait profiles*. There is an observer[1] who looks at the group which exhibits

[1] To avoid confusion, we assume that the observer is not a member of the group. Al-

a trait profile $t = (t_1, ..., t_{\#N}) \in T$. Not necessarily all members of the group are visible to the observer. We write $V \subseteq N$ for the members of the group that are visible to the observer. An observer who sees $V \subset N$ is aware of the existence of the unobserved $N \backslash V$. On the other hand, we rule out the possibility of "wrong observation", i.e., the trait of every visible member of the group is observed as it truly is. We let $T_V = \{-1, 1\}^V$ stand for the set of trait profiles of the observable members. The observer has a subjective perception of the group as a function of the trait profile he is able to observe, which we express through a *(subjective) perception function* $\psi_V : T_V \rightarrow \{1, 0, -1\}$. So given any non-empty set $V \subseteq N$ of observed members and any prevailing trait profile $t \in T_V$ of these observed members, we write $\psi_V(t) = 1$ when the observer globally perceives the group N as possessing the trait in question. Similarly, we write $\psi_V(t) = -1$ when the observer globally perceives the group as not possessing the trait in question and $\psi_V(t) = 0$ refers to the observer's abstension of reaching a global perception of the group. We refer to the case $V = N$ as *perfect observation* and to $V \subset N$ as *imperfect observation.* Under perfect observation, we write ψ instead of ψ_N.

What kind of perception functions are used? We approach the problem axiomatically by considering the cases of perfect and imperfect observation separately.

2 Stereotype Formation Under Perfect Observation

2.1 Towards a Solution: Axioms for Perception

Being sensitive to individual traits is incorporated in the concept of a perception function. So, we wish to rule out imposed perceptions that are independent of individual traits such as "Muslims do not drink alchool because this is what the Quran says". Hence we posit that under perfect observation, the observer would say "Muslims do not drink alchool" if no Muslim drinks alchool and "Muslims do drink alchool" if every Muslim drinks alchool. We express these through the following axiom:

Non-Imposedness: A perception function $\psi : T \rightarrow \{1, 0, -1\}$ is *non-imposed* iff $\psi(1, 1, ..., 1) = 1$ and $\psi(-1, -1, ..., -1) = -1$.

Non-imposedness is a weak unanimity requirement which rules out imposed perceptions while it does not exclude biased ones such as saying "Mus-

though this has no effect to our model, belongingness of the observer to the observed group seems to actually matter, according to our intepretation of Krueger et al. (2003).

lims eat porc" if and only if every Muslim eats porc and saying "Muslims do not eat porc" even when there exists a single Muslim who does not eat porc. It is clear that such a perception is based on an unequal treatment of traits. Of course this may happen but when we wish to rule it out, we use the following axiom:

Impartiality: A perception function $\psi : T \to \{1, 0, -1\}$ satisfies *impartiality* iff $\psi(-t) = -\psi(t)\ \forall t \in T$.

Remark that given a trait profile t, the trait profile $-t$ stands for the reversal of every individual trait. So impartiality is an adaptation of the usual neutrality condition of social choice theory which ensures the equal treatment of alternatives. An observer with an impartial perception function is not prejudiced about the group's possessing or not possesing the trait: If the trait of every observed individual is reversed then so is the perception.

In contrast to what impartiality requires, one can perceive a society under an unequal treatment of the traits. For example, it is possible that the observer has a bias towards thinking that the society exhibits the trait in question. Such a bias is formally expressed through the following axiom:

Positive Prejudice: We say that a perception function $\psi : T \to \{1, 0, -1\}$ admits *positive prejudice* iff

(i) $\exists\ t \in T$ such that $\psi(t) = 1$ and $\psi(-t) \in \{0, 1\}$

and

(ii) $\psi(t) \in \{0, -1\} \implies \psi(-t) = 1\ \forall t \in T$.

So under a perception function admitting positive prejudice, there is a trait profile t such that the trait is rejected neither at t nor at $-t$. Moreover, there exists no trait profile t such that the observer rejects the trait or is undecisive both at t and $-t$.

Similarly, as expressed below, the observer can have a bias towards thinking that the society does exhibit the trait in question:

Negative Prejudice: We say that a perception function $\psi : T \to \{1, 0, -1\}$ admits *negative prejudice* iff

(i) $\exists\ t \in T$ such that $\psi(t) = -1$ and $\psi(-t) \in \{-1, 0\}$

and

(ii) $\psi(t) \in \{0, 1\} \implies \psi(-t) = -1\ \forall t \in T$

Another axiom we borrow from the social choice literature is a monotonicity condition: If a trait profile changes so that some individuals who did not possess the trait now possess it while this is the only change, then the perception should not affected in the opposite direction. We express this formally as follows:

Monotonicity: A perception function $\psi : T \to \{1, 0, -1\}$ is *monotonic* iff $\psi(t) \geq \psi(t')\ \forall t, t' \in T$ with $t_i \geq t'_i\ \forall i \in N$.

These axioms pave the way to the characterization of a class of perception functions which we call *subjective majority rules*. We have three main characterization results where we use the conjunction of non-imposedness and monotonicity with one of impartiality, positive prejudice and negative prejudice.[2] We close the section by establishing the logical independence of the axiom triples that we use.

Proposition 2.1 *Non-imposedness, monotonicity and impartiality are logically independent.*

Proof. To see that impartiality and non-imposedness do not imply monotonicity, let $\#N = 3$ and consider $\psi : T \to \{1, 0, -1\}$ which is defined for each $t \in T$ as $\psi(t) = 1$ when $\#\{i \in N : t_i = 1\} \in \{1, 3\}$ and $\psi(t) = -1$ otherwise. To see that impartiality and monotonicity do not imply non-imposedness, take $\psi(t) = 0$ for all $t \in T$. Finally, to see that non-imposedness and monotonicity do not imply impartiality, let $\psi(t) = -1$ if $t_i = -1$ $\forall i \in N$ and $\psi(t) = 1$ otherwise. ■

Proposition 2.2 *Non-imposedness, monotonicity and positive prejudice are logically independent.*

Proof. To see that positive prejudice and non-imposedness do not imply monotonicity, let $\#N = 3$ and consider $\psi : T \to \{1, 0, -1\}$ which is defined as $\psi(-1, -1, -1) = -1$, $\psi(1, 1, -1) = 0$ and $\psi(t) = 1$ otherwise. To see that positive prejudice and monotonicity do not imply non-imposedness, take $\psi(t) = 1$ for all $t \in T$. Finally, to see that non-imposedness and monotonicity do not imply positive prejudice, let $\psi(t) = 1$ if $t_i = 1$ $\forall i \in N$ and $\psi(t) = -1$ otherwise. ■

Proposition 2.3 *Non-imposedness, monotonicity and negative prejudice are logically independent.*

Proof. To see that negative prejudice and non-imposedness do not imply monotonicity, let $\#N = 3$ and consider $\psi : T \to \{1, 0, -1\}$ which is defined as $\psi(1, 1, 1) = 1$, $\psi(-1, -1, 1) = 0$ and $\psi(t) = -1$ otherwise. To see that negative prejudice and monotonicity do not imply non-imposedness, take $\psi(t) = -1$ for all $t \in T$. Finally, to see that non-imposedness and monotonicity do not imply negative prejudice, let $\psi(t) = -1$ if $t_i = -1$ $\forall i \in N$ and $\psi(t) = 1$ otherwise. ■

[2] Remark that impartiality, positive prejudice and negative prejudice are pairwise logically incompatible.

2.2 A Solution: Subjective Majority Rules

We first define a *(subjective) weight distribution* as a mapping $\omega : 2^N \to [0,1]$ such that $\omega(K)+\omega(N\backslash K) = 1$ for all $K \in 2^N$ while $\omega(N) = 1$. So ω expresses the subjective opinion of the observer about the representation weight of each subgroup of N. A weight distribution ω is *monotonic* iff $\omega(K) \leq \omega(L)$ for all $K, L \in 2^N$ with $K \subseteq L$. For the rest of the paper, we embed monotonicity into the definition of a weight distribution.

Given a weight distribution ω and any $q \in (0,1)$, *a subjective* $(\omega, q)-$*majority rule* is a perception function $\psi^{\omega,q} : T \to \{1, 0, -1\}$ defined for each $t \in T$ as follows:

$$\psi^{\omega,q}(t) = \left\{ \begin{array}{cc} 1 & \text{if } \omega(\{i \in N : t_i = 1\}) > q \\ -1 & \text{if } \omega(\{i \in N : t_i = -1\}) > 1 - q \\ 0 & \text{otherwise} \end{array} \right\}$$

So the observer looks at the group with some subjective opinion about how representative the subgroups are. If, according to this subjective opinion, the weight of those who possess the trait exceeds q, then the observer concludes that the group globally possesses that trait. Similarly, if the (subjective) weight of those who possess the trait is below q, then the observer concludes that the group globally does not possess that trait.[3] If neither of these two cases holds then no conclusion is derived.

Theorem 2.1 *A perception function $\psi : T \to \{1, 0, -1\}$ satisfies non-imposedness, monotonicity and impartiality iff ψ is a subjective $(\omega, \frac{1}{2})-$majority rule.*

Proof. We leave the "if" part to the reader. To see the "only if" part, take any perception function $\psi : T \to \{1, 0, -1\}$ that satisfies non-imposedness, monotonicity and impartiality. We define $W_\psi = \{K \in 2^N : \psi(t) = 1$ for $t \in T$ with $t_i = 1\ \forall i \in K$ and $t_i = -1\ \forall i \in N\backslash K\}$ and $L_\psi = \{K \in 2^N : \psi(t) = -1$ for $t \in T$ with $t_i = 1\ \forall i \in K$ and $t_i = -1\ \forall i \in N\backslash K\}$. As ψ is non-imposed, $N \in W_\psi$ and $\emptyset \in L_\psi$, hence W_ψ and L_ψ are each nonempty. Let $O_\psi = 2^N \backslash (W_\psi \cup L_\psi)$ be the (possibly empty) set of coalitions which are neither in W_ψ nor in L_ψ. Now consider a function $\omega : 2^N \to [0,1]$ defined for each $K \in 2^N$ as $\omega(K) = 1$ if $K \in W_\psi$, $\omega(K) = 0$ if $K \in L_\psi$ and $\omega(K) = \frac{1}{2}$ if $K \in O_\psi$. As ψ is impartial, for each $K \in 2^N$, we have $K \in L_\psi \iff N\backslash K \in W_\psi$ which implies $K \in O_\psi \iff N\backslash K \in O_\psi$. Thus $\omega(K)+\omega(N\backslash K) = 1\ \forall K \in 2^N$ while $\omega(N) = 1$. Moreover, the monotonicity

[3] Remark that $\omega(\{i \in N : t_i = -1\}) > 1 - q$ and $\omega(\{i \in N : t_i = 1\}) < q$ are equivalent requirements. However, we use the former statement to be coherent with our definition in Section 3, where we consider subjective majority rules under imperfect observation.

of ψ implies $\omega(K) \leq \omega(L)$ for all $K, L \in 2^N$ with $K \subseteq L$. So ω is a weight distribution. We complete the proof by showing that the subjective $(\omega, \frac{1}{2})-$majority rule $\psi^{\omega,\frac{1}{2}} : T \to \{1, 0, -1\}$ coincides with ψ. To see this, take any $t \in T$. If $\psi(t) = 1$, then $K = \{i \in N : t_i = 1\} \in W_\psi$, implying $\omega(K) = 1 > \frac{1}{2}$, which establishes $\psi^{\omega,\frac{1}{2}}(t) = 1$. If $\psi(t) = -1$, then $K = \{i \in N : t_i = 1\} \in L_\psi$, implying $\omega(K) = 0$, hence $\omega(N \backslash K) = 1 > \frac{1}{2}$, which establishes $\psi^{\omega,\frac{1}{2}}(t) = -1$. If $\psi(t) = 0$ then $K \in O_\psi$, implying $\omega(K) = \frac{1}{2}$, which establishes $\psi^{\omega,\frac{1}{2}}(t) = 0$. ■

Theorem 2.2 *A perception function $\psi : T \to \{1, 0, -1\}$ satisfies non-imposedness, monotonicity and positive prejudice iff ψ is a subjective $(\omega, q)-$majority rule with $q \in (0, \frac{1}{2})$ and $\omega(K) \in [q, 1-q]$ for some $K \in 2^N$.*

Proof. To see the "if" part, take any subjective $(\omega, q)-$majority rule $\psi^{\omega,q}$ with $q \in (0, \frac{1}{2})$ and $\omega(K) \in [q, 1-q]$ for some $K \in 2^N$. It is straightforward to check that $\psi^{\omega,q}$ satisfies non-imposedness and monotonicity. To show that $\psi^{\omega,q}$ satisfies positive prejudice, take some $K \in 2^N$ with $\omega(K) \in (q, 1-q]$. Remark that $\psi^{\omega,q}(t) = 1$ and $\psi^{\omega,q}(-t) \in \{0, 1\}$ for $t \in T$ with $t_i = 1\ \forall i \in K$ and $t_i = -1\ \forall i \in N \backslash K$. Now take any $t \in T$ with $\psi^{\omega,q}(t) \in \{0, -1\}$. Thus, letting $K = \{i \in N : t_i = 1\}$, we have $\omega(K) \leq q$, hence $\omega(N \backslash K) > q$, implying $\psi^{\omega,q}(-t) = 1$, showing that $\psi^{\omega,q}$ satisfies positive prejudice. To see the "only if" part, take any perception function $\psi : T \to \{1, 0, -1\}$ that satisfies non-imposedness, monotonicity and positive prejudice. Let W_ψ, L_ψ and O_ψ be defined as in the proof of Theorem 2.1. Note that $N \in W_\psi$ and $\emptyset \in L_\psi$ while O_ψ may be empty. Now pick some $q \in (0, \frac{1}{2})$ and consider a function $\omega : 2^N \to [0, 1]$ defined for each $K \in 2^N$ as $\omega(K) = 0$ if $K \in L_\psi$, $\omega(K) = q$ if $K \in O_\psi$. Moreover, if $K \in W_\psi$, then let $\omega(K) = 1$ when $N \backslash K \in L_\psi$; $\omega(K) = 1-q$ when $N \backslash K \in O_\psi$ and $\omega(K) = \frac{1}{2}$ when $N \backslash K \in W_\psi$. As ψ satisfies positive prejudice, for each $K \in L_\psi \cup O_\psi$ we have $N \backslash K \in W_\psi$. Thus $\omega(K) + \omega(N \backslash K) = 1\ \forall K \in 2^N$ while $\omega(N) = 1$. Moreover, the monotonicity of ψ implies $\omega(K) \leq \omega(L)$ for all $K, L \in 2^N$ with $K \subseteq L$. So ω is a weight distribution. Note also that by positive prejudice, $\exists K \in W_\psi$ such that $N \backslash K \in W_\psi \cup O_\psi$, implying $\omega(K) \in [q, 1-q]$. We complete the proof by showing that the subjective $(\omega, q)-$majority rule $\psi^{\omega,q} : T \to \{1, 0, -1\}$ coincides with ψ. To see this, take any $t \in T$. If $\psi(t) = 1$, then $K = \{i \in N : t_i = 1\} \in W_\psi$ and $\omega(K) \in \{\frac{1}{2}, 1-q, 1\}$ implying $\omega(K) > q$, which establishes $\psi^{\omega,q}(t) = 1$. If $\psi(t) = 0$, then $K = \{i \in N : t_i = 1\} \in O_\psi$ and $\omega(K) = q$, which establishes $\psi^{\omega,q}(t) = 0$. If $\psi(t) = -1$, then $K = \{i \in N : t_i = 1\} \in L_\psi$ and $\omega(K) = 0$, hence $\omega(N \backslash K) = 1 > 1-q$, which establishes $\psi^{\omega,q}(t) = -1$.
■

Theorem 2.3 *A perception function* $\psi : T \to \{1, 0, -1\}$ *satisfies non-imposedness, monotonicity and negative prejudice iff* ψ *is a subjective* $(\omega, q)-$*majority rule with* $q \in (\frac{1}{2}, 1)$ *and* $\omega(K) \in [1-q, q]$ *for some* $K \in 2^N$.

Proof. To see the "if" part, take any subjective $(\omega, q)-$majority rule $\psi^{\omega,q}$ with $q \in (\frac{1}{2}, 1)$ and $\omega(K) \in [1-q, q]$ for some $K \in 2^N$. It is straightforward to check that $\psi^{\omega,q}$ satisfies non-imposedness and monotonicity. To show that $\psi^{\omega,q}$ satisfies negative prejudice, take some $K \in 2^N$ with $\omega(K) \in (1-q, q]$. Remark that $\psi^{\omega,q}(t) = -1$ and $\psi^{\omega,q}(-t) \in \{-1, 0\}$ for $t \in T$ with $t_i = -1$ $\forall i \in K$ and $t_i = 1$ $\forall i \in N \backslash K$. Now take any $t \in T$ with $\psi^{\omega,q}(t) \in \{0, 1\}$. Thus, letting $K = \{i \in N : t_i = 1\}$, we have $\omega(K) \geq q > 1-q$, implying $\psi^{\omega,q}(-t) = -1$ showing that $\psi^{\omega,q}$ satisfies negative prejudice. To see the "only if" part, take any perception function $\psi : T \to \{1, 0, -1\}$ that satisfies non-imposedness, monotonicity and negative prejudice. Let W_ψ, L_ψ and O_ψ be defined as in the proof of Theorem 2.1. Note that $N \in W_\psi$ and $\emptyset \in L_\psi$ while O_ψ may be empty. Now pick some $q \in (\frac{1}{2}, 1)$ and consider a function $\omega : 2^N \to [0, 1]$ defined for each $K \in 2^N$ as $\omega(K) = 1$ if $K \in W_\psi$, $\omega(K) = q$ if $K \in O_\psi$. Moreover, if $K \in L_\psi$, then let $\omega(K) = 0$ when $N \backslash K \in W_\psi$; $\omega(K) = 1-q$ when $N \backslash K \in O_\psi$ and $\omega(K) = \frac{1}{2}$ when $N \backslash K \in L_\psi$. As ψ satisfies negative prejudice, for each $K \in W_\psi \cup O_\psi$ we have $N \backslash K \in L_\psi$. Thus $\omega(K) + \omega(N \backslash K) = 1$ $\forall K \in 2^N$ while $\omega(N) = 1$. Moreover, the monotonicity of ψ implies $\omega(K) \leq \omega(L)$ for all $K, L \in 2^N$ with $K \subseteq L$. So ω is a weight distribution. Note that by negative prejudice, $\exists K \in L_\psi$ such that $N \backslash K \in O_\psi \cup L_\psi$, implying $\omega(K) \in [1-q, q]$. We complete the proof by showing that the subjective $(\omega, q)-$majority rule $\psi^{\omega,q} : T \to \{1, 0, -1\}$ coincides with ψ. To see this, take any $t \in T$. If $\psi(t) = -1$, then $K = \{i \in N : t_i = 1\} \in L_\psi$ and $\omega(K) \in \{0, 1-q, \frac{1}{2}\}$ implying $\omega(K) < q$, hence $\omega(N \backslash K) > 1-q$, which establishes $\psi^{\omega,q}(t) = -1$. If $\psi(t) = 0$, then $K = \{i \in N : t_i = 1\} \in O_\psi$ and $\omega(K) = q$, which establishes $\psi^{\omega,q}(t) = 0$. If $\psi(t) = 1$, then $K = \{i \in N : t_i = 1\} \in W_\psi$ and $\omega(K) = 1 > q$, which establishes $\psi^{\omega,q}(t) = -1$. ■

Remark 2.1 *Monotonicity is a normatively appealing condition for perception functions and this is why we are keeping it throughout our analysis. However, it is clear from their proofs that Theorems 2.1, 2.2 and 2.3 can be stated by simultaneously dispensing with the monotonicity of the perception function and the monotonicity condition incorporated into the definition of a weight distribution.*

Until now, we did not bring any requirement for an equal treatment of individuals by the weight distribution ω. In fact, at one extreme, it is possible to have an observer who believes that a group is fully represented in the personality of one of its members $d \in N$ which would correspond to a weight distribution $\omega(K) = 1$ for all $K \in 2^N$ with $d \in K$. At the other extreme, we have $\omega^{=}(K) = \frac{\#K}{\#N}$ for all $K \in 2^N$ where all individuals are thought of having equal weight. Given the subjective nature of weight distributions (hence of stereotype formation), we do not think that an equal treatment of individuals should be required. However, we wish to explore the effects of imposing such a requirement. A perception function $\psi : T \rightarrow \{1, 0, -1\}$ is *anonymous* iff given any $t = (t_1, ..., t_{\#N}) \in T$ and any bijection $\pi : N \longleftrightarrow N$, we have $\psi(t_1, ..., t_{\#N}) = \psi(t_{\pi(1)}, ..., t_{\pi(\#N)})$. Given some $\alpha \in (0, 1)$, a weight distribution $\omega : 2^N \rightarrow [0, 1]$ is $\alpha-$*anonymous* iff given any $K, L \in 2^N$ with $\#K = \#L$ we have $\omega(K) > \alpha \iff \omega(L) > \alpha$ and $\omega(K) < \alpha \iff \omega(L) < \alpha$.[4]

Theorem 2.4 *A perception function $\psi : T \rightarrow \{1, 0, -1\}$ satisfies non-imposedness, monotonicity, impartiality and anonymity iff ψ is a subjective $(\omega, \frac{1}{2})-$majority rule for some $\frac{1}{2}-$anonymous ω.*

Proof. To show the "if" part, let $\psi : T \rightarrow \{1, 0, -1\}$ be a subjective $(\omega, \frac{1}{2})-$majority rule where ω is $\frac{1}{2}-$anonymous. We know by Theorem 2.1 that ψ satisfies non-imposedness, monotonicity and impartiality. To see the anonymity of ψ, take any $t = (t_1, ..., t_{\#N}) \in T$. Let $K = \{i \in N : t_i = 1\}$. Take any bijection $\pi : N \longleftrightarrow N$. Let $\pi(t) = (t_{\pi(1)}, ..., t_{\pi(\#N)})$ and $\pi(K) = \{\pi(i)\}_{i \in K}$. As π is a bijection, $\#K = \#\pi(K)$. Moreover, $\{i \in N : t_{\pi(i)} = 1\} = \pi(K)$. Thus $\psi(t_1, ..., t_{\#N}) = \psi(t_{\pi(1)}, ..., t_{\pi(\#N)})$ holds by the $\frac{1}{2}-$anonymity of ω.

To show the "only if" part, take any $\psi : T \rightarrow \{1, 0, -1\}$ satisfying non-imposedness, monotonicity, impartiality and anonymity. We know, by Theorem 2.1 that ψ is a subjective $(\omega, \frac{1}{2})-$majority rule. To see that ω is $\frac{1}{2}-$anonymous, take any $K, L \in 2^N \setminus \{\emptyset\}$ with $\#K = \#L$. Take $t = (t_1, ..., t_{\#N}) \in T$ with $\{i \in N : t_i = 1\} = K$. Take also some bijection $\pi : N \longleftrightarrow N$ such that $\{\pi(i)\}_{i \in K} = L$. Now let $\omega(K) > \frac{1}{2}$. So $\psi(t) = 1$. As ψ is anonymous, $\psi(t_{\pi(1)}, ..., t_{\pi(\#N)}) = 1$ as well, implying $\omega(\{i \in N : t_{\pi(i)} = 1\}) = \omega(L) > \frac{1}{2}$. One can similarly establish that letting $\omega(K) < \frac{1}{2}$ implies $\omega(L) < \frac{1}{2}$, thus showing the $\frac{1}{2}-$anonymity of ω. ■

Remark 2.2 *The mathematics of our model belongs to the literature on majority characterizations, which goes back to May (1952). This allows to make*

[4] Hence we also have $\omega(K) = \alpha \iff \omega(L) = \alpha$. Note that α-anonymity is weaker than a more standard anonymity condition which would require $\omega(K) = \omega(L)$ for all $K, L \in 2^N$ with $\#K = \#L$.

a remark about Theorem 2.4. Consider a set $A = \{x, y\}$ of alternatives and let each $i \in N$ have a preference $p_i \in \{{}^x_y, {}^y_x\}$ over A.[5] *Denoting xy for indifference between x and y, we conceive a social choice rule as a mapping $f : \{{}^x_y, {}^y_x\}^N \to \{{}^x_y, {}^y_x, xy\}$. Let n^* be the lowest integer exceeding $\frac{\#N}{2}$. Picking any $\mu \in \{n^*, ..., n\}$, we define a $\mu-$majority rule as a social choice rule $f_\mu : \{{}^x_y, {}^y_x\}^N \to \{{}^x_y, {}^y_x, xy\}$ where for any $p = (p_1, ..., p_{\#N}) \in \{{}^x_y, {}^y_x\}$ we have $f_\mu(p) = {}^x_y \iff \#\{i \in N : p_i = {}^x_y\} \geq \mu$ and $f_\mu(p) = {}^y_x \iff \#\{i \in N : p_i = {}^y_x\} \geq \mu$.*[6] *Theorem 3.2 of Asan and Sanver (2006) characterizes the set of Pareto optimal, anonymous, neutral and Maskin monotonic aggregation rules in terms of $\mu-$majority rules. In that abstract setting, Pareto optimality, anonymity and neutrality respectively coincide with our non-imposedness, anonymity and impartiality. On the other hand Maskin monotonicity is stronger than our monotonicity. So by Theorem 2.4, we can deduce that the class of $\mu-$majority rules is a subset of the class of subjective $(\omega, \frac{1}{2})-$majority rules with ω being $\frac{1}{2}-$anonymous. In other words, every aggregation rule that gives every coalition in the society its "objective" weight (i.e., letting the weight of $K \in 2^N$ be $\frac{\#K}{\#N}$) but possibly qualifies the required majority can alternatively be expressed by fixing majority as usual (i.e., as any coalition whose cardinality exceeds its complement) but assigning monotonic and $\frac{1}{2}-$anonymous (subjective) weights to coalitions.*

Remark 2.3 *As is for Theorem 2.1, Theorems 2.2 and 2.3 can be stated by simultaneously adding anonymity to the perception function and the corresponding $\alpha-$anonymity with $\alpha = q$ to the weight distribution.*

3 Stereotype Formation Under Imperfect Observation

3.1 Axioms Revisited

Throughout the section, we fix some non-empty set $V \subsetneq N$ of visible group members and consider the perception function $\psi_V : T_V \to \{1, 0, -1\}$. The existence of invisible group members entails a revision of the non-imposedness axiom. For, an observer who fails to observe some members of the group may be cautious to bring a global perception of the group, even when the prevailing trait profile is unanimous: Suppose a small group of Martians who

[5]where ${}^x_y$ is interpreted as x being preferred to y and ${}^y_x$ is interpreted as y being preferred to x. So individual preferences do not admit indifference between x and y.

[6]Thus $f_\mu(p) = xy \iff \#\{i \in N : p_i = {}^x_y\} < \mu$ and $\#\{i \in N : p_i = {}^y_x\} < \mu$.

are all aggressive attack the World. A human who knows the existence of other unobserved Martians may be prudent in qualifying Martians as aggressive, hence refusing to bring a global judgement about them under the hope that the unobservable crowd of Martians are peaceful. Thus, under imperfect observation, we propose the following version of the non-imposedness axiom:

Non-Imposedness: A perception function $\psi_V : T_V \to \{1, 0, -1\}$ is *non-imposed* iff $\psi_V(1, 1, ..., 1) \in \{0, 1\}$ and $\psi_V(-1, -1, ..., -1) \in \{-1, 0\}$.

Remark that the imperfect information version of non-imposedness is neither weaker nor stronger than its perfect information version. For, it is weakened by allowing the refusal of judgements but strenghtened by being imposed over the profiles where unanimity is reached among the members of V. Monotonicity, impartiality, positive prejudice and negative prejudice exhibit a strenghtening of the similar spirit, as we now impose them when the related changes in the trait profiles occur in the visible part of the group.

Monotonicity: A perception function $\psi_V : T_V \to \{1, 0, -1\}$ is *monotonic* iff $\psi_V(t) \geq \psi_V(t') \ \forall \ t, t' \in T_V$ with $t_i \geq t'_i \ \forall i \in V$.

Impartiality: A perception function $\psi_V : T_V \to \{1, 0, -1\}$ satisfies *impartiality* iff $\psi_V(t') = -\psi_V(t) \ \forall t, t' \in T_V$ such that $t'_i = -t_i \ \forall i \in V$.

Positive Prejudice: We say that a perception function $\psi_V : T_V \to \{1, 0, -1\}$ admits *positive prejudice* iff

(i) $\exists \ t \in T_V$ such that $\psi_V(t) = 1$ and $\psi_V(-t) \in \{0, 1\}$

and

(ii) $\psi_V(t) \in \{-1, 0\} \Longrightarrow \psi_V(-t) = 1 \ \forall \ t \in T_V$.

Negative Prejudice: We say that a perception function $\psi_V : T_V \to \{1, 0, -1\}$ admits *negative prejudice* iff

(i) $\exists \ t \in T_V$ such that $\psi_V(t) = -1$ and $\psi_V(-t) \in \{-1, 0\}$

and

(ii) $\psi_V(t) = \{0, 1\} \Longrightarrow \psi_V(-t) = -1 \ \forall \ t \in T_V$

To characterize perception under imperfect observation, we use the conjunction of non-imposedness and monotonicity with one of impartiality, positive prejudice and negative prejudice. The following proposition establishes the logical relationship between these axioms[7]:

Proposition 3.1 (i) *Monotonicity and impartiality imply non-imposedness.*

(ii)*Monotonicity and impartiality are logically independent.*

(iii)*Non-imposedness, monotonicity and positive prejudice are logically independent.*

(iv)*Non-imposedness, monotonicity and negative prejudice are logically independent.*

[7]As is in the perfect observation case (see Footnote 2), impartiality, positive prejudice and negative prejudice are pairwise logically incompatible.

Proof. *Proof of* (i)*: Let* $\psi_V : T_V \to \{-1, 0, 1\}$ *satisfy impartiality and fail non-imposedness. We have* $\psi_V (1, 1, ...1) = -1$ *or* $\psi_V (-1, -1, ..., -1) = 1$ *by the failure of non-imposedness which, by impartiality, implies* $\psi_V (1, 1, ...1) = -1$ *and* $\psi_V (-1, -1, ..., -1) = 1$*, contradicting monotonicity.*

Proof of (ii)*: Define* $\psi_V : T_V \to \{-1, 0, 1\}$ *as* $\psi_V (1, 1, ..., 1) = -1$, $\psi_V (-1, -1, ..., -1) = 1$ *and* $\psi_V (t) = 0$ $\forall t \in T_V$ *with* $t_i = 1$, $t_j = -1$ *for some* $i, j \in V$*. Check that* ψ_V *is impartial but not monotonic. Now let* $\psi_V (t) = 1$ $\forall t \in T_V$ *and check that* ψ_V *is monotonic but not impartial.*

Proof of (iii)*: To see that non-imposedness and monotonicity do not imply positive prejudice, let* $\psi_V (1, 1, ..., 1) = 1$ *and* $\psi_V (t) = -1$ *for any* $t \in T_V$ *with* $t_i \in \{-1, 0\}$ *for some* $i \in V$*. To see that non-imposedness and positive prejudice do not imply monotonicity, let* $\#V = 3$ *and let* $\psi_V (t) = 1$ *if* $\# \{i \in V : t_i = 1\} \in \{1, 3\}$*;* $\psi_V (t) = -1$ *if* $\# \{i \in V : t_i = 1\} = 2$ *and* $\psi_V (-1, -1, -1) = 0$*. To see that monotonicity and positive prejudice do not imply non-imposedness let* $\psi_V (t) = 1$ $\forall t \in T_V$*.*

Proof of (iv)*: To see that non-imposedness and monotonicity do not imply negative prejudice, let* $\psi_V (-1, -1, ..., -1) = -1$ *and* $\psi_V (t) = 1$ *for any* $t \in T_V$ *with* $t_i \in \{0, 1\}$ *for some* $i \in V$*. To see that non-imposedness and negative prejudice do not imply monotonicty let* $\#V = 3$ *and let* $\psi_V (t) = -1$ *if* $\# \{i \in V : t_i = 1\} \in \{0, 2\}$*;* $\psi_V (1, 1, 1) = 0$ *and* $\psi_V (t) = 1$ *if* $\# \{i \in V : t_i = 1\} = 1$*. To see that monotonicity and negative prejudice do not imply non-imposedness let* $\psi_V (t) = -1$ $\forall t \in T_V$*.* ■

3.2 Subjective Majority Rules Revisited

A *(subjective) weight distribution* as an ordered pair $\sigma = (\omega, p)$ where $\omega : 2^V \to [0, 1]$ is a mapping satisfying

(i) $\omega(K) + \omega(V \backslash K) = 1$ for all $K \in 2^V$

(ii) $\omega(V) = 1$

(iii) $\omega(K) \leq \omega(L)$ for all $K, L \in 2^V$ with $K \subseteq L$

and $p \in [0, 1]$ reflects the weight of V in N.[8]

Given a weight distribution $\sigma = (\omega, p)$ and any $q \in (0, 1)$, *a subjective* $(\sigma, q)-$*majority rule* is a perception function $\psi_V^{\sigma,q} : T_V \to \{1, 0, -1\}$ defined for each $t \in T_V$ as follows:

$$\psi_V^{\sigma,q}(t) = \left\{ \begin{array}{c} 1 \quad \text{if } p.\omega(\{i \in V : t_i = 1\}) > q \\ -1 \quad \text{if } p.\omega(\{i \in V : t_i = -1\}) > 1 - q \\ 0 \quad \text{otherwise} \end{array} \right\}$$

[8] Remark that under perfect observation, we used ω to express the weight distribution within N but now it expresses the (monotonic) weight distribution within V coupled with the parameter p which reflects the weight of V in N. Of course when $p = 1$, V can be conceived as the whole society, bringing us back to the case of perfect observation.

So the observer looks at V with some subjective opinion about how representative its subgroups are. Moreover, he has a subjective opinion about the representativenes of V within the whole society. If, according to these subjective opinions, the weight of those who possess the trait exceeds q, then the observer concludes that the group globally possesses that trait. Similarly, if the (subjective) weight of those who do not possess the trait exceeds $1-q$, then the observer concludes that the group globally does not possess that trait.[9] If neither of these two cases holds then no conclusion is derived.

Theorem 3.1 *A perception function $\psi_V : T_V \to \{1, 0, -1\}$ satisfies monotonicity and impartiality iff ψ_V is a subjective $(\sigma, \frac{1}{2})-$majority rule for some subjective weight distribution $\sigma = (\omega, p)$.*
Proof. *We leave the "if" part to the reader. To see the "only if" part, take any $\psi_V : T_V \to \{-1, 0, 1\}$ that satisfies monotonicity and impartiality. Let $W_\psi = \{K \in 2^V : \psi_V(t) = 1$ for $t \in T_V$ with $t_i = 1\ \forall i \in K$ and $t_i = -1\ \forall i \in V \backslash K\}$ and $L_\psi = \{K \in 2^V : \psi_V(t) = -1$ for $t \in T_V$ with $t_i = 1\ \forall i \in K$ and $t_i = -1\ \forall i \in V \backslash K\}$. We set $O_\psi = 2^V \backslash (W_\psi \cup L_\psi)$. Note that the perception function ψ_V defined as $\psi_V(t) = 0$ at each $t \in T_V$ is monotonic and impartial. So W_ψ and L_ψ can both be empty. However, by the impartiality of ψ_V, we have $W_\psi = \emptyset \iff L_\psi = \emptyset$. In fact, $W_\psi = \emptyset \iff L_\psi = \emptyset \iff \psi_V(t) = 0\ \forall t \in T_V$. First consider the case where $W_\psi = \emptyset$ and $L_\psi = \emptyset$. So $\psi_V(t) = 0\ \forall t \in T_V$. Take any subjective weight distribution $\sigma = (\omega, p)$ with $p \in [0, \frac{1}{2})$. It is straightforward to check that the subjective $(\sigma, \frac{1}{2})-$majority rule coincides with ψ_V. Now consider the case where neither W_ψ nor L_ψ is empty. Thus, $V \in W_\psi$ and $\emptyset \in L_\psi$. Consider the function $\omega : 2^V \to [0, 1]$ where $\omega(K) = 1\ \forall K \in W_\psi$, $\omega(K) = 0\ \forall K \in L_\psi$ and $\omega(K) = \frac{1}{2}\ \forall K \in O_\psi$. The impartiality of ψ_V ensures $K \in W_\psi \iff V \backslash K \in L_\psi\ \forall K \in 2^V$ and thus $K \in O_\psi \iff V \backslash K \in O_\psi\ \forall K \in 2^V$. Hence $\omega(K) + \omega(V \backslash K) = 1\ \forall K \in 2^V$ while $\omega(V) = 1$. Moreover, the monotonicity of ψ_V implies $\omega(K) \leq \omega(L)$ for all $K, L \in 2^V$ with $K \subseteq L$. Thus, any $p \in [0, 1]$ induces a subjective weight distribution (ω, p). Pick $p = 1$ and let $\sigma = (\omega, 1)$. We claim that the subjective $(\sigma, \frac{1}{2})-$majority rule $\psi_V^{\sigma, \frac{1}{2}} : T_V \to \{-1, 0, 1\}$ coincides with ψ_V. To see this, take any $t \in T_V$. If $\psi_V(t) = 1$, then $K = \{i \in V : t_i = 1\} \in W_\psi$ and $\omega(K) = 1$, implying $p.\omega(K) = 1 > \frac{1}{2}$, which establishes $\psi_V^{\sigma, \frac{1}{2}}(t) = 1$. If $\psi_V(t) = -1$, then $K = \{i \in V : t_i = 1\} \in L_\psi$ and $\omega(K) = 0$, hence $\omega(V \backslash K) = 1$, implying $p.\omega(V \backslash K) = 1 > \frac{1}{2}$, which establishes $\psi_V^{\sigma, \frac{1}{2}}(t) = -1$. If $\psi_V(t) = 0$, then $K = \{i \in V : t_i = 1\} \in O_\psi$ and $\omega(K) = \frac{1}{2}$, hence*

[9] Remark that $p.\omega(\{i \in V : t_i = -1\}) > 1 - q$ and $p.\omega(\{i \in V : t_i = 1\}) < q$ are equivalent requirements if and only if $p = 1$. See Footnote 3.

$\omega(V\backslash K) = \frac{1}{2}$. Thus neither $p.\omega(K) > \frac{1}{2}$, nor $p.\omega(V\backslash K) > \frac{1}{2}$ holds, which establishes $\psi_V^{\sigma,\frac{1}{2}}(t) = 0$. ■

Theorem 3.2 *Given any $V \subsetneq N$, a perception function $\psi_V : T_V \rightarrow \{1, 0, -1\}$ satisfies non-imposedness, monotonicity and positive prejudice iff ψ_V is a subjective $(\sigma, q)-$majority rule with $q \in (0, \frac{1}{2})$ while $\sigma = (\omega, p)$ is a weight distribution such that $p > \max\{2q, 1-q\}$ and $\omega(K) \in [\frac{q}{p}, \frac{1-q}{p}]$ for some $K \in 2^V$.*[10]

Proof. *To see the "if" part, let ψ_V be a subjective $(\sigma, q)-$majority rule as in the statement of the theorem. It is straightforward to check $\psi_V^{\sigma;q}$ satisfies non-imposedness and monotonicity. To show that $\psi_V^{\sigma,q}$ satisfies positive prejudice, take some $K \in 2^V$ with $\omega(K) \in (\frac{q}{p}, \frac{1-q}{p}]$. So $\psi_V^{\sigma,q}(t) = 1$ for $t \in T_V$ with $t_i = 1\ \forall i \in K$ and $t_i = -1\ \forall i \in V\backslash K$. Moreover, as $p.\omega(K) \leq 1-q$, we have $\psi_V^{\sigma,q}(-t) \in \{0, 1\}$. Now take any $t \in T_V$ with $\psi_V^{\sigma,q}(t) \in \{-1, 0\}$ and let $K = \{i \in V : t_i = 1\}$. If $\psi_V^{\sigma,q}(t) = -1$ then $p.\omega(V\backslash K) > 1-q > q$, implying $\psi_V(-t) = 1$. If $\psi_V^{\sigma,q}(t) = 0$ then $p.\omega(K) \leq q$. As $p > 2q$, we have $\omega(K) < \frac{1}{2}$, thus $\omega(V\backslash K) > \frac{1}{2}$ and $p.\omega(V\backslash K) > 2q$, implying $\psi_V(-t) = 1$ which shows that $\psi_V^{\sigma,q}$ satisfies positive prejudice. To see the "only if" part, take any $\psi_V : T_V \rightarrow \{1, 0, -1\}$ that satisfies non-imposedness, monotonicity and positive prejudice. We define W_ψ, O_ψ and L_ψ as in Theorem 3.1. Note that $V \in W_\psi$. Moreover, while one of O_ψ and L_ψ may be empty, $O_\psi \cup L_\psi$ is non-empty. Now pick some $q \in (0, \frac{1}{2})$ and consider the function $\omega : 2^V \rightarrow [0, 1]$ defined for each $K \in 2^V$ as $\omega(K) = \omega(V\backslash K) = \frac{1}{2}$ when K, $V\backslash K \in W_\psi$; $\omega(K) = 0$ and $\omega(V\backslash K) = 1$ when $K \in L_\psi$ and $V\backslash K \in W_\psi$; $\omega(K) = q$ and $\omega(V\backslash K) = 1-q$ when $K \in O_\psi$ and $V\backslash K \in W_\psi$. Note that positive prejudice ensures $K \in L_\psi \cup O_\psi \implies V\backslash K \in W_\psi$ for each $K \in 2^V$. Thus $\omega(K) + \omega(V\backslash K) = 1\ \forall K \in 2^V$ with $\omega(V) = 1$, while the monotonicity of ψ_V implies $\omega(K) \leq \omega(L)$ for all $K, L \in 2^V$ with $K \subseteq L$. Thus, any $p \in [0, 1]$ induces a subjective weight distribution (ω, p). Take any $p \in [0, 1]$ with $p >$ $\max\{2q, 1-q\}$. We will show that $\omega(K) \in [\frac{q}{p}, \frac{1-q}{p}]$ for some $K \in 2^V$. Recall that $O_\psi \cup L_\psi$ is non-empty. First let O_ψ be non-empty and take some $S \in O_\psi$. So $V\backslash S \in O_\psi$ and by construction of ω we have $\omega(V\backslash S) = 1-q$, thus $\omega(V\backslash S) \leq \frac{1-q}{p}$. Moreover, $1-q > q$ and $p > 2q$, thus $\omega(V\backslash S) = 1-q > \frac{q}{p}$, establishing $\omega(V\backslash S) \in (\frac{q}{p}, \frac{1-q}{p}]$. By definition of O_ψ, we have $p.\omega(S) \leq q$, thus $\omega(S) \leq \frac{q}{p} < \frac{1}{2}$ implying $\omega(V\backslash S) \geq 1 - \frac{q}{p} > \frac{1}{2} > \frac{q}{p}$. Again by definition of O_ψ, we have $p.\omega(V\backslash S) \leq 1-q$. Thus $\omega(V\backslash S) \in (\frac{q}{p}, \frac{1-q}{p}]$. Now let O_ψ be empty. By positive prejudice, $\exists K \in W_\psi$ such that $V\backslash K \in W_\psi$. Thus*

[10] Note that $p > \max\{2q, 1-q\}$ ensures $\frac{q}{p}$, $\frac{1-q}{p} \in (0, 1)$. Moreover, $q \in (0, \frac{1}{2})$ ensures $\frac{q}{p} < \frac{1-q}{p}$.

$\omega(K) = \frac{1}{2} \in [\frac{q}{p}, \frac{1-q}{p}]$, by the choice of p. Writing $\sigma = (\omega, p)$, we complete the proof by showing that the $(\sigma, q)-$majority rule $\psi_V^{\sigma;q}$ coincides with ψ_V. To see this, take any $t \in T_V$. If $\psi(t) = 1$, then $K = \{i \in N : t_i = 1\} \in W_\psi$ and $\omega(K) \in \{\frac{1}{2}, 1-q, 1\}$. Moreover, $p > 2q$. Thus, $p.\omega(K) > q$, establishing $\psi_V^{\sigma;q}(t) = 1$. If $\psi(t) = 0$, then $K = \{i \in N : t_i = 1\} \in O_\psi$ and $\omega(K) = q$, hence $\omega(V \backslash K) = 1-q$. Thus, neither $p.\omega(K) > q$, nor $p.\omega(V \backslash K) > 1-q$ holds, which establishes $\psi_V^{\sigma;q}(t) = 0$. If $\psi(t) = -1$, then $K = \{i \in N : t_i = 1\} \in L_\psi$ and $\omega(K) = 0$, hence $\omega(V \backslash K) = 1$ implying $p.\omega(V \backslash K) = p > 1-q$, which establishes $\psi_V^{\sigma;q}(t) = -1$. ■

Theorem 3.3 *Given any $V \subsetneq N$, a perception function $\psi_V : T_V \to \{1, 0, -1\}$ satisfies non-imposedness, monotonicity and negative prejudice iff ψ_V is a subjective $(\sigma, q)-$majority rule with $q \in (\frac{1}{2}, 1)$ while $\sigma = (\omega, p)$ is a weight distribution such that $p > \max\{2q, 1-q\}$ and $\omega(K) \in [\frac{1-q}{p}, \frac{q}{p}]$ for some $K \in 2^V$.*[11]

Proof. *To see the "if" part, let ψ_V be a subjective $(\sigma, q)-$majority rule as in the statement of the theorem. It is straightforward to check $\psi_V^{\sigma;q}$ satisfies non-imposedness and monotonicity. To show that $\psi_V^{\sigma;q}$ satisfies negative prejudice, take some $K \in 2^V$ with $\omega(K) \in (\frac{1-q}{p}, \frac{q}{p}]$. So $\psi_V^{\sigma;q}(t) = -1$ for $t \in T_V$ with $t_i = -1\ \forall i \in K$ and $t_i = 1\ \forall i \in V \backslash K$. Moreover, as $p.\omega(K) \leq q$, we have $\psi_V^{\sigma;q}(-t) \in \{-1, 0\}$. Now take any $t \in T_V$ with $\psi_V^{\sigma;q}(t) \in \{0, 1\}$ and let $K = \{i \in V : t_i = 1\}$. If $\psi_V^{\sigma;q}(t) = 1$ then $p.\omega(K) > q > 1-q$, implying $\psi_V(-t) = -1$. If $\psi_V^{\sigma;q}(t) = 0$ then $p.\omega(V \backslash K) \leq 1-q$. As $p > 2q$ and $q > 1-q$, we have $p > 2(1-q)$. So $\omega(V \backslash K) < \frac{1}{2}$, thus $\omega(K) > \frac{1}{2}$ and $p.\omega(K) > 1-q$, implying $\psi_V(-t) = -1$ which shows that $\psi_V^{\sigma;q}$ satisfies negative prejudice. To see the "only if" part, take any $\psi_V : T_V \to \{1, 0, -1\}$ that satisfies non-imposedness, monotonicity and negative prejudice. We define W_ψ, O_ψ and L_ψ as in Theorem 3.1. Note that $\emptyset \in L_\psi$. Moreover, while one of W_ψ and O_ψ may be empty, $W_\psi \cup O_\psi$ is non-empty. Now pick some $q \in (\frac{1}{2}, 1)$ and consider the function $\omega : 2^V \to [0, 1]$ defined for each $K \in 2^V$ as $\omega(K) = \omega(V \backslash K) = \frac{1}{2}$ when $K, V \backslash K \in L_\psi$; $\omega(K) = 0$ and $\omega(V \backslash K) = 1$ when $K \in L_\psi$ and $V \backslash K \in W_\psi$; $\omega(K) = q$ and $\omega(V \backslash K) = 1-q$ when $K \in O_\psi$ and $V \backslash K \in L_\psi$. Note that negative prejudice ensures $K \in W_\psi \cup O_\psi \implies V \backslash K \in L_\psi$ for each $K \in 2^V$. Thus $\omega(K) + \omega(V \backslash K) = 1\ \forall K \in 2^V$ with $\omega(V) = 1$, while the monotonicity of ψ_V implies $\omega(K) \leq \omega(L)$ for all $K, L \in 2^V$ with $K \subseteq L$. Thus, any $p \in [0, 1]$ induces a subjective weight distribution (ω, p). Take any $p \in [0, 1]$ with $p > \max\{2q,\ 1-q\}$. We will show that $\omega(K) \in [\frac{1-q}{p}, \frac{q}{p}]$ for some $K \in 2^V$. Recall that $W_\psi \cup O_\psi$ is non-empty. First let O_ψ be non-empty and*

[11]Note that $p > \max\{2q, 1-q\}$ ensures $\frac{q}{p}$, $\frac{1-q}{p} \in (0, 1)$. Moreover, $q \in (\frac{1}{2}, 1)$ ensures $\frac{1-q}{p} < \frac{q}{p}$.

take some $S \in O_\psi$. By construction of ω we have $\omega(S) = q$, thus $\omega(S) \leq \frac{q}{p}$. Moreover, $q > 1 - q$ and $p > 2(1-q)$, thus $\omega(S) = q > \frac{1-q}{p}$, establishing $\omega(S) \in (\frac{1-q}{p}, \frac{q}{p}]$. Now let O_ψ be empty. By negative prejudice, $\exists K \in L_\psi$ such that $V \backslash K \in L_\psi$. Thus $\omega(K) = \frac{1}{2} \in [\frac{1-q}{p}, \frac{q}{p}]$, by the construction of ω and the choice of p. Writing $\sigma = (\omega, p)$, we complete the proof by showing that the $(\sigma, q)-$majority rule $\psi_V^{\sigma,q}$ coincides with ψ_V. To see this, take any $t \in T_V$. If $\psi(t) = 1$, then $K = \{i \in N : t_i = 1\} \in W_\psi$ and $\omega(K) = 1$, implying $p.\omega(K) = p > q$, which establishes $\psi_V^{\sigma,q}(t) = 1$. If $\psi(t) = 0$, then $K = \{i \in N : t_i = 1\} \in O_\psi$ and $\omega(K) = q$, hence $\omega(V \backslash K) = 1 - q$. Thus, neither $p.\omega(K) > q$, nor $p.\omega(V \backslash K) > 1 - q$ holds, which establishes $\psi_V^{\sigma,q}(t) = 0$. If $\psi(t) = -1$, then $K = \{i \in N : t_i = 1\} \in L_\psi$ and $\omega(K) \in \{0, 1-q, \frac{1}{2}\}$, hence $\omega(V \backslash K) \in \{\frac{1}{2}, q, 1\}$. As $p > 2q$, $p.\omega(V \backslash K) > q > 1 - q$, establishing $\psi_V^{\sigma,q}(t) = -1$. ■

4 Concluding Remarks

We propose a formalization of the attribution hypothesis by modeling the formation of stereotypes as an aggregation of prevailing individual traits.[12] We show, under very mild axioms, that an observer can be thought of perceiving a group in terms of three parameters:

- In case the group is not perfectly visible, the subjective weight $p \in [0, 1]$ of the visible members within the whole group.
- A weight distribution ω over the visible part of the group (reflecting the subjective opinion of the observer about the representativeness of coalitions).
- A (dis)qualified majority $q \in (0, 1)$ (which is related to a possible prejudice that the observer may have).

The perception - compatible with the claim of Zawadski (1948) quoted in Section 1- proceeds as follows: An observer of the Turkish society perceives the group as smoker (resp., non-smoker) if and only if a "majority" of Turks

[12] Nevertheless, the tools we propose can be used to model the *categorization hypothesis* over which an alternative strand of the social psychology literature rises. The categorization hypothesis postulates a comparative reasoning in the formation of stereotypes: Whether a group is judged to possess a trait depends on the prevalance of that trait relative to some other comparison group. For gender stereotypes, Krueger et al. (2003) present empirical evidence in favor of the attribution hypothesis. For a more general and axiomatic analysis of categorization, see Azrieli and Lehrer (2007).

are smokers (resp. non-smokers). This approach differs from usual majoritarianism. For, whether the (non)-smoker Turks form a majority in the Turkish society depends on p and ω, which expresses the observer's subjective perception of how representative the observed (non)-smoker Turks are. Moreover, the observer may have a prejudice about the Turkish society being smoker or not, which is incorporated into the aggregation function through the choice of q.[13]

Our findings allow to keep track of three different possible generalizations (expressed through p, ω, q) underlying the formation of stereotypes. As a consequence, we are able to present a conceptual distinction between a stereotype and a prejudice. In fact, the existence of a stereotype does not necessarily imply the existence of a prejudice. As our theorems make it clear, a stereotype can be the outcome of an impartial aggregation of individual traits. In fact, the existence of a prejudice depends on one parameter (namely q) of the aggregation function and not on the outcome itself.

To be sure, our axiomatic characterizations do not allow to differentiate whether people form stereotypes by using subjective majority rules or as if they are using subjective majority rules.[14] However, the truth of the former raises interesting institutional design questions about affecting stereotype formation so as to ameliorate social outcomes. After all, the existence of stereotypes seems to be inevitable[15] and their desirability depends on the social outcomes they induce. For example, we may wish to overthrow stereotypes which foster racism but endorse those, such as the image of a brave fireman, which support the supply of public goods. Hence, establishing the relationship between existing or possible institutions and the parameters p, ω, q that individuals possess is certainly a matter of interest. So the simple but conceptually rich exercise we present, points to new possibilities in exploring questions of social psychology through the concepts of social choice theory - hence bringing a theoretical look at this area where an empirical literature abounds.[16]

---

[13] A dominant approach in social choice theory takes preferences as given and rules out qualifying social outcomes as "correct" or "wrong". This is reflected to our analysis and leaves no room to illusory correlation in stereotype formation (see Hamilton and Gifford (1976)).

[14] Olson and Fazio (2001) discuss the possibility of stereotype formation without awareness. This is reminiscent to our inability to discriminate, under the weak axiom of revealed preference, whether decision makers maximize a rational preference or act as if they are maximizing a rational preference.

[15] Lee et al. (1995) conceive stereotypes as tools used by the mind to navigate its complex environment.

[16] A final few words, just for a cheerful closing: This last section of the paper has been completed while Remzi Sanver was the guest of Ecole Polytechnique, Paris. As a result,

5 References

Arrow, K.J. (1951), Social Choice and Individual Values, John Wiley, New York.

Asan G. and M.R. Sanver (2006), Maskin monotonic aggregation rules. Economics Letters, 91(2), 179-183.

Azrieli, Y. and E. Lehrer (2007), Categorization generated by extended prototypes - An axiomatic approach. Journal of Mathematical Psychology, 51, 14-28.

Hamilton, D.L. and R.K. Gifford (1976), Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgements. Journal of Experimental Social Psychology, 12, 392-407.

Krueger, J.I., J.F. Hasman, M. Acevedo and P. Villano (2003), Perceptions of trait typicality in gender stereotypes: Examining the role of attribution and categorization processes. Personality and Social Psychology Bulletin, 29, 108-116.

Lee, Y., L.J. Jussim and C.R. McCauley (1995), Stereotype Accuracy: Toward Appreciating Group Differences. American Psychological Assocation, 1st Edition

May, K. (1952), A set of independent necessary and sufficient conditions for simple majority decision. Econometrica, 20, 680-684.

Olson, M.A and R.H. Fazio (2001), Implicit attitude formation through classical conditioning. Psychological Science, 12, 413-417.

Zawadski, B. (1948), Limitations of the scapegoat theory of prejudice. Journal of Abnormal and Social Psychology, 43, 127-141.

most of the work has been done in Parisian cafés. Interestingly, while thinking on how to conclude the paper, a waiter of les Deux Magots made the following remark: "You look much more intellectual than I thought Turks were."