A Simple Theory of the False Consensus Effect

Young Joon Park*

October 18, 2009

Abstract

People do overestimate the prevalence of own behavior or choices. I propose a simple model of information acquisition that describes this false consensus effect and several of its empirically identified qualitative properties. I show that, under mild assumptions, the tendency to believe one's own choice to be more informative about the behavior of a group than that of another's can be explained as a rational reaction. This 'generalized' false consensus effect has kept being considered irrational since the observation of Dawes (1989) that provides rational explanation for the false consensus effect. The paper shows that what is called 'truly false' by Dawes (1989) does not need to be false.

1 Introduction

In most strategic environments, it is important to be able to estimate the characteristics and behavior of other. For example, optimal bidding in auctions depends on not only on one's valuation of the good but also the valuation of other. In a bargaining situation, players' optimal strategies will depend on the valuations or time preferences of their opponents. Estimating the taste of potential customers is a key to estimate the demand for a new business. If the characteristics of other players are not known, estimating them is a critical task.

Psychological evidence suggests individuals possess systematic bias when they compare their attributes to others. In particular, the False Consensus Effect states that individuals overestimate the number of the people who possess the same attributes as they do. When there are two

*email: ypark@ucsd.edu

possible positions for people to choose, A and B, the false consensus effect exists when the average estimation of the fraction of people who chose position A made by those who hold position A is greater than the average estimation the fraction of people who chose position A made by those who hold position B. To the extent that the false consensus effect demonstrates an inability of individuals to process information rationally, it suggests that conventional economic analysis of strategic behavior, which assumes unbiased expectation formation, is deficient. This paper studies the extent to which the false consensus bias is consistent with a conventional model of information processing.

Ross et al. (1977) first reports the false consensus bias. The authors conducted a series of experiments that asked college students if they were willing to perform certain actions. Both the group who agreed to perform the action and the group who were unwilling to perform the action overestimated the size of the group who responded as they did. For example, in one of the experiments, college undergraduates were asked whether they would wear a sandwich sign saying "Eat at Joe's". Subjects' reaction were divided evenly[1], but their consensus estimation was biased toward overestimating their own cohort group. When asked to estimate the proportion of responders who agreed to wear the sign, the average subjects who agreed to wear the sign estimated that 64.6% of the subjects would wear the sign. The average estimate of the group by the subjects who would not wear the sign was 31.2%.

Dawes (1989) points out that the seemingly biased behavior can be described as a rational response given the fact that individuals are always able to observe their own attributes. This specific nature of own observation makes people project others' behavior from their own. He claimed that it is rational to use own endorsement as part of the information to estimate others' behavior. To illustrate, suppose individuals have a prior belief that everyone is equally likely to choose either alternative. If they only observe the choice they made and update the prior regarding their choice as a single sample, they will all consider themselves to belong to the majority group. Consequently, the population possesses the false consensus effect. This is called the induction hypothesis. With this observation, Dawes claims that the false consensus effect

[1] 53% answered that they would wear the sign and 47% answered they would't.

should not be considered irrational unless people consider own choice to be more informative than a choice made by another person in the process of estimating the overall composition of population choice. This 'generalized' false consensus effect has been considered as an example of irrational behavior even after Dawes's finding.

Although the induction hypothesis can explain the tendency for people to possess the false consensus effect, it overly simplifies agents' process of selecting their choices. Classical approaches by social psychologists focused on the estimation procedure once individuals' endorsements are already determined. Thus the analysis on how individuals reach the decision was overlooked. However, it is reasonable to believe that individuals have different preferences and beliefs even though they make the same choice. This observation motivates introducing another framework that describes individuals' procedures of making own decisions.

I propose a simple model of information acquisition in which individuals overestimate the size of the cohort group due to the imperfect information. In my model, an individual observes partial information about the true state of the world that determines the overall distribution of people's choices. This information is used not only to make her own decision but also to estimate the true state of the world and others' decisions. The estimate will be biased toward overestimating own cohort group because people will overestimate the probability of others receiving similar information. Consequently, the result suggests that the generalized false consensus effect can arise as the consequence of a rational individual's reaction to a limited information.

The rest of the paper is organized as follows. Section 2 reviews related literatures. I propose the model and its implications in Section 3. Section 4 summarizes several preliminary results that are used in following analysis. The stylized facts of the bias are introduced in Section 5. Section 6 relates the main results to the stylized facts. I attempt to illustrate the classical hypotheses of the false consensus effect from social psychology and compare their different implications on several situations in Section 7. Section 8 is the conclusion.

2 Related literature

Since the seminal finding of Ross et al. (1977), the false consensus effect was studied both experimentally and theoretically in social psychology. The experimental evidence indicates people possess this bias on various questions related to *personal traits and characteristics* (being optimistic, having brown or blue eyes (Ross et al, 1977), attitude on women's rights, (Judd and Johnson, 1981)), *personal preference* (eating white or brown bread (Gilovich et al, 1983)), *personal activities* (housekeeping activities (van der Pligt, 1984)), *political expectations* (preferred presidential candidate (Brown, 1982), performance of the political leaders (Manstead, 1982), prediction on the result of national vote (Baker et al, 1995)), *factual information* (knowledge of trivia questions (Mullen, 1983), the induction problem on the identification of suicide notes (Sherman et al, 1984)[2]), and so on.

Mullen et al. (1985) reviewed 115 experiment results and tested the robustness of the false consensus effect. By combining the results of multiple studies, they found the evidence of the false consensus effect in various attributes. Furthermore, their results indicate that the broadness of the reference group (eg. whether the estimation was made over the whole population or fellow students) and the difference between alternative choices in actual consensus (the difference between the agreed and the disagreed) do not have significant effect on the degree of the false consensus effect.[3]

Following Dawes (1989), who introduced a rational explanation of the false consensus effect, there were several studies conducted to confirm the existence of generalized false consensus effect. Alicke and Largo (1995, study 2) ran an experiment that lets subjects take a test consisted with a series of social sensitivity questions. After taking the test, subjects were informed the bogus result whether they passed or failed the test. Also, the subjects were informed about the result of 4 other subjects, which was manipulated to form a combination of either 4 pass and 1 fail result or 1 pass and 4 fail result. Then subjects were asked to provide the estimate of subjects who

[2]Subjects were provided two possible suicide notes, in which only one of them is an actual note. Subjects guessed the correct one and asked to estimate the proportion of people who made the same choice.

[3]On the other hand, they found the number of items in each experiment and the order of the report (whether the estimate is made first, or making of own choice is made first) has significant effect of the degree of the false consensus effect.

passed the test out of 1000 imaginary subjects. They measured the degree of the generalized false consensus effect by comparing the answers of subjects who share the same overall information sets, but distinguished by their own outcomes. The result indicates that subjects put higher weight on their own (fake) endorsement than those of others.

Engelmann and Strobel (2000) test the generalized false consensus effect by conducting an experiment where subjects estimate the consensus after observing randomly selected choices of others. They share the view of Dawes (1989) that the false consensus effect can arise as a result of rational reaction due to the nature of own observation and improve the experimental design of Alicke and Largo (1995) by providing monetary incentives and eliminating the manipulation. On various questions about own attributes, subjects in groups of 16 individuals answered their own endorsements. Then, they observed endorsements of 4 other randomly selected subjects. Finally, they estimated the consensus of the remaining 11 subjects. By comparing the estimate between the subjects who share the same overall observations yet opposite own choice, the authors test whether subjects consider their own choices more informative. The result shows the false consensus effect remains: after observing others' endorsements, the estimate of the prevalence of each alternative was still greater from the insider. However, the outcome does not indicate subjects possess the generalized false consensus effect.[4] In fact, out of 26 pairings, there were only 11 cases in which the results support the hypothesis of the generalized false consensus effect. Also, only 2 out of 11 supporting results were considered 'substantial' (significant with 5% level and at least 5 subjects from each group) while there were 10 substantial cases that indicates the opposite result.

Krueger and Clement (1994) introduced the egocentrism hypothesis to explain the phenomenon that individuals consider own decisions more informative than others'. They introduce the egocentric aspect of human behavior as a possible explanation of generalized false consensus effect. They emphasize that people tend to use different mental procedures when they process their own behavior rather than that of others. Thus people might consider own information more importantly. Consequently, the generalized false consensus effect can be explained under

[4]Engelmann and Strobel (2004) later found that the generalized false consensus effect reappears when individuals needs to put marginal effort to obtain information of others.

their assumptions.

Goeree and Großer (2007) and Vanberg (2008) introduce simple models that provide rationale of the false consensus effect. Their models characterize the observation of Dawes (1989) and use the false consensus effect to illustrate the behaviors of voters and experimental subjects respectively. Their simple characterization, however, does not fully describe the actual behavior as the simplified assumptions generate the extreme result such that everyone believes that she belongs to the majority. Agents receive a binary signal which is correlated with the binary state of the world. Given agents have uniform prior belief on the states, all agents put higher posterior probability on the state corresponding to their signal. Eventually, everyone believes there are more people who received the same signal they received than those who received the opposite signal.

3 Setup

The state of the world is $\theta \in [0,1]$. There is a continuum of agents. They are characterized by a preference parameter λ and a signal s. We assume that λ is distributed independently of θ and s. The density of λ on $[0,1]$ is $g(\cdot)$ with cdf G. We interpret $s \in [0,1]$ to be an idiosyncratic estimate of the true state θ. We assume that the signals received by individual agents are identically distributed and conditionally independent given θ. Let $f(s,\theta)$ be the density function of signals and states.; $f(\cdot)$ describes the information structure. Assume that s is an unbiased estimate of θ so that the conditional expectation of θ given s is equal to s. All distribution functions are common knowledge.

We assume that f satisfies the Monotone Likelihood Ration Property (MLRP). More precisely for $s < s'$ and $\theta < \theta'$,

$$\frac{f(s'|\theta')}{f(s|\theta')} \leq \frac{f(s'|\theta)}{f(s|\theta)}.$$

A well known consequence of this property (Krishna, 2009) is

1. $F(s|\theta)$ is decreasing in θ.

2. $F(\theta|s)$ is decreasing in s.

Given her characteristics, an agent makes a binary choice, picking either h or l. Assume that preferences are such that an (s, λ) type chooses h if and only if $s \geq \lambda$.[5] Let $\mathcal{H}$ be the set of agents in the economy who select h.

Let $H(\theta) \equiv \int_0^1 G(s)f(s|\theta)ds$. $H(\theta)$ is the fraction of the population that selects h when the true state of the world is θ. We are interested in how an agent's estimate of H depends on her membership in $\mathcal{H}$. Assume agents' prior belief of the true H is τ. i.e. $\int_0^1 H(\theta')f_\theta(\theta')d\theta' = \tau$ where f_θ is the marginal density of θ.

Agents can also observe other agents' endorsements. In this case, it is useful to define the conditional distribution of the state θ given all available information.

Definition 1 *Let $\Gamma(\theta|s, n, k)$ be the conditional distribution of θ when an agent receives the signal s, and observes n numbers of the population where k of them endorses h. The conditional distribution will be denoted as*

$$\Gamma(\theta|s, n, k) = \frac{\int_0^\theta f(s,t)H(t)^k(1 - H(t))^{n-k}dt}{\int_0^1 f(s,r)H(r)^k(1 - H(r))^{n-k}dr} \tag{1}$$

with the corresponding pdf, $\gamma(\theta|s, n, k)$.

Note $\Gamma(\theta|s, n, k)$ can be viewed as a generalization of the conditional distribution $F(\theta|s)$. Γ corresponds to the case in which individuals observe not only their own signal, but also the endorsements of the subset of the population.

Also notice that, regardless of agents' own signal, s, it is assumed that there is no bias observing other agents' endorsement. i.e. each agent observes the same number of endorsements and the probability of observing a certain number of h endorsers is the same for everyone. Also, assume agents do not change their own endorsement after they observe other agents' behavior. The observation of other agents' endorsement changes the belief of the true state. Therefore, agents might have incentive to change their own endorsement depending on the characteristics

[5]For example. an agent's preferences may be represented by $(\theta - \lambda)(x_h - x_l)$ where $x_i = 1$ if the agent takes action i and 0 otherwise.

of the problem.[6]

4 Preliminary Results

It is useful to state some preliminary results that are commonly used in the analysis. First of all, the higher the true state is, the more people choose h.

Lemma 1 *$H(\theta)$ is increasing in θ.*

proof) for $\theta < \theta'$,

$$\begin{aligned} H(\theta) - H(\theta') &= \int_0^1 G(s)[f(s|\theta) - f(s|\theta')]ds \\ &= -\int_0^1 g(s)[F(s|\theta) - F(s|\theta')]ds \end{aligned}$$

where the second equation is obtained by integration by parts. The expression is non positive due to the MLRP of F. . ■

In order to compare the differences of estimates between two endorsement groups, one can consider the probability distribution of signal s of agents in each endorsement group.

Definition 2 *Define the $\Phi_i(s)$ as the distribution of s given that the agent is in category i:*

$$\Phi_h(s) = \int_0^1 \frac{\int_0^s G(r)f(r,\theta)dr}{\int_0^1 G(t)f(t|\theta)dt} d\theta \tag{2}$$

and

$$\Phi_l(s) = \int_0^1 \frac{\int_0^s (1 - G(r))f(r,\theta)dr}{\int_0^1 (1 - G(t))f(t|\theta)dt} d\theta. \tag{3}$$

Also denote $\phi_i(s)$ for $i \in \{h, l\}$ as the corresponding density function. Φ_h first order stochastically dominates Φ_l.

Lemma 2 *$\Phi_h(s) \leq \Phi_l(s)$ for all s.*

[6] These two assumptions are revisited in Section 7 and the implications are analyzed when the assumptions are relaxed.

Proof)

$$\Phi_h(s) - \Phi_l(s) = \int_0^1 \left[\frac{\int_0^s G(t) f(t,\varphi) dt}{\int_0^1 G(t) f(t|\varphi) dt} - \frac{\int_0^s (1-G(t)) f(t,\varphi) dt}{\int_0^1 (1-G(t)) f(t|\varphi) dt} \right] d\varphi$$
$$= \int_0^1 \left[\frac{\int_0^s G(t) f(t|\varphi) dt - F(s|\varphi) \int_0^1 G(t) f(t|\varphi) dt}{\int_0^1 G(t) f(t|\varphi) dt \int_0^1 (1-G(t)) f(t|\varphi) dt} \right] f_\theta(\varphi) d\varphi$$

The numerator of the integrand can be rearranged to

$$\left[\int_0^s G(t) f(t|\varphi) dt - F(s|\varphi) \int_0^1 G(t) f(t|\varphi) dt \right] f_\theta(\varphi)$$
$$= \left[(1 - F(s|\varphi)) \int_0^s G(t) f(t|\varphi) dt - F(s|\varphi) \int_s^1 G(t) f(t|\varphi) dt \right] f_\theta(\varphi)$$
$$= G(s) f_\theta(\varphi) \left[(1 - F(s|\varphi)) \int_0^s \frac{G(t)}{G(s)} f(t|\varphi) dt - F(s|\varphi) \int_s^1 \frac{G(t)}{G(s)} f(t|\varphi) dt \right] \quad (4)$$

since

$$\int_0^s \frac{G(t)}{G(s)} f(t|\varphi) dt \leq F(s|\varphi), \text{ and } \int_s^1 \frac{G(t)}{G(s)} f(t|\varphi) dt \geq 1 - F(s|\varphi),$$

expression (4) is non-positive. ■

Lemma 2 states that, on average, agents in $\mathcal{H}$ group receive higher signals compared to agents in $\mathcal{L}$ group.

In order to compare the average estimate of H made by two endorsement groups, we further need to describe agents' estimation rule. There are two possible cases depending on the information set available to the agents. The first case is when individuals cannot observe any of the other agents' endorsements. The situation can be interpreted as the initial response to the task. In this situation, the estimate of the others' endorsements should be determined only by the signal of the true state, s.

Definition 3 *Define $A(s)$ as the expected estimation of H for an agent who received the signal*

s. More precisely,

$$A(s) = \int_0^1 H(\theta) f(\theta|s) d\theta \tag{5}$$

Another possible scenario is when agents are able to observe the endorsements of a subset of the population. After all agents determine their initial endorsements, they observe other agents' behaviors. This case is a generalization of the case where no observation is possible. The formal definition of agents' estimation rule follows.

Definition 4 *Define $\tilde{A}(s, n, k)$ as the expected estimation of H for an agent who received the signal s, and observed n number of other agents' endorsement where k of their endorsements are h. More precisely,*

$$\tilde{A}(s, n, k) = \int_0^1 H(\theta) \gamma(\theta|s, n, k) d\theta \tag{6}$$

If she observes n members of the population and k of them endorses h, her estimation of the fraction of the agents in $\mathcal{H}$ is denoted as $\tilde{A}(s, n, k)$. For simplicity, assume the number of the observed endorsement, n, is same for all agents.

5 Stylized Facts

This section describes several experimentally observed stylized facts about the false consensus effect in the language of the model.

5.1 Population possess the false consensus effect

In multiple experiments with various settings, subjects systematically show that they overestimate the size of their own cohort group. This is the false consensus effect.

Definition 5 (False Consensus Effect) *Given the true state and available information s, the* *population possesses the false consensus effect if*

$$\int_0^1 A(s)\phi_h(s)ds \geq \int_0^1 A(s)\phi_l(s)ds \tag{7}$$

with a strict inequality if the signal in non-trivial.

The average estimation of H made by agents in $\mathcal{H}$ is compared to the average estimation of H made by agents in $\mathcal{L}$. If the estimation made by agents in $\mathcal{H}$ is larger, then we call the population possess the false consensus effect.

5.2 Generalized false consensus effect

According to induction hypothesis, individuals might possess the false consensus because they consider their own endorsement decision as a single sample in the process of estimating the overall behavior of the population. The underlying logic is that there is no reason for individuals to undervalue the information drawn from their own decision. This observation further suggests that there should be no overvaluing of one's own endorsement either. Thus, once individuals are exposed to more observations of other individuals' choices, they should evaluate another individual's endorsement in the same manner as they evaluate their own information. In various experiments (Sherman et al. (1984), Krueger and Clement (1994), and Alicke and Largo (1995)), however, it has shown that subject put more weight on own choices than that of another when they estimate the behaviors of the population. Engelmann and Strobel (2000) report the result that this tendency disappears under the setting in which the subjects had proper incentives and was not deceived by the experimenter. However, they also found the result is fragile (Engelmann and Strobel, 2004) as the generalized false consensus effect reappears when subjects are required to put marginal effort to obtain the information of other subjects.

Definition 6 (Generalized False Consensus effect) *Given the true state and available information (s, n, k), the population possesses the generalized false consensus effect if*

$$\int_0^1 \tilde{A}(s,n,k)\phi_h(s)ds \geq \int_0^1 \tilde{A}(s,n,k)\phi_l(s)ds \tag{8}$$

with a strict inequality if the signal in non-trivial.

Similar to the case of the false consensus effect, the generalized false consensus effect is defined by comparing the average estimation of H made by agents in $\mathcal{H}$ is compared to the average estimation of H made by agents in $\mathcal{L}$. If the estimation made by agents in $\mathcal{H}$ is larger, then we say that the population possess the generalized false consensus effect.

5.3 Completeness

Another commonly observed phenomenon in the experimental results of the false consensus effect is that regardless of the endorsement group, there are agents who believe they are majorities as well as agents who believe they are minorities. Krueger (1998) reports a more detailed result from the experiment of Dawes and Mulford (1996) in terms of the subjects' accuracy of the estimation.[7] Among 68% of subjects who belonged to the majorities, about 23% of them estimated that they belong to the minorities. Also, 17% out of 32% who belong to the minorities believes they belong to the majorities.

Table 1: Composition of Estimation (Krueger, 1998)

		Estimated Consensus	
		Majority 0.62	Minority 0.38
Actual Consensus	Majority 0.68	**0.45**	**0.23**
	Minority 0.32	**0.17**	**0.15**

This evidence cannot be generated by the models of Goeree and Großer (2007) or Vanberg (2008) as they did not allow heterogeneity of agents in their models. Also, the signal agents receive is characterized in a simple binary way. Hence, according to their settings, every agent believes she belongs to the majority.

[7]Subjects were asked to estimate the consensus over 128 questions from the *California Psychological Inventory.* Since the authors had access to the data collected from a large population, the actual consensus of the population could be measured.

Definition 7 (Completeness) *The population is called complete if given* θ*, the lowest possible signal* $\underline{s}$ *and the highest possible signal* $\overline{s}$ *satisfy* $\tilde{A}(\underline{s}, n, k) < 1/2$ *and* $\tilde{A}(\overline{s}, n, k) > 1/2$*, and there exist* $\lambda^*, \lambda^{**} \in \Lambda$ *such that* $\lambda^* < \underline{s}$ *and* $\lambda^{**} \geq \overline{s}$*.*

The existence of λ^* implies that there is an agent who endorses h and believes herself to be a minority. The existence of λ^{**} implies that there is an agent who endorses l and believes herself to be a minority. Note that the existence of agents who believe they belong to the majority is automatically achieved.

6 Results

In this section, I introduce the implications of the model on the stylized facts described in the previous section.

Proposition 1 *The population possess the false consensus effect.*

The formal proof is omitted as it is a special case of Proposition (2). The agents in $\mathcal{H}$ group tend to receive higher signals and the higher signals lead them to make a higher estimate the size of $\mathcal{H}$. Thus, on average, the $\mathcal{H}$ group has a higher estimate of the size of the $\mathcal{H}$ group than does the $\mathcal{L}$ group.

6.1 Generalized false consensus effect

Now consider the case in which agents can observe some of the others' endorsements. Γ is the conditional distribution of θ given agents received their own signal, s and observed n other agents' choices where k of these choices are h. It can be shown that the MLRP of the conditional distribution F is preserved with extra information, (n, k).

Lemma 3 $\Gamma(\theta|s, n, k) \geq \Gamma(\theta|s', n, k)$ *for* $s < s'$*.*

Proof) For $\theta < \theta'$,

$$\frac{\gamma(\theta'|s, n, k)}{\gamma(\theta|s, n, k)} = \frac{f(\theta'|s)}{f(\theta|s)} \frac{H(\theta')^k (1 - H(\theta'))^{n-k}}{H(\theta)^k (1 - H(\theta))^{n-k}}$$

is decreasing in s due to the MLRP of F. Hence for $s < s'$,

$$\frac{\gamma(\theta'|s,n,k)}{\gamma(\theta|s,n,k)} \geq \frac{\gamma(\theta'|s',n,k)}{\gamma(\theta|s',n,k)}. \tag{9}$$

Thus Γ also satisfies the MLRP with respect to s. ■

Lemma (3) demonstrates that even with the observations of others' endorsements, the relationship between s and θ remains the same as the case without the observations. With the result that Γ behave in similar ways as F, I can show that population possess the generalized false consensus effect.

Proposition 2 *The population possess the Generalized False Consensus Effect.*

proof) $\tilde{A}(s,n,k)$ increases in s as for $s < s'$,

$$\begin{aligned}
\tilde{A}(s,n,k) - \tilde{A}(s'n,k) &= \int_0^1 H(\theta)\left[\gamma(\theta|s,n,k - \gamma(\theta|s'n,k))\right] d\theta \\
&= -\int_0^1 H'(\theta)\left[\Gamma(\theta|s,n,k - \Gamma(\theta|s'n,k))\right] d\theta \\
&\leq 0.
\end{aligned}$$

The result follows as

$$\begin{aligned}
\int_0^1 \tilde{A}(s,n,k)\phi_h(s)ds - \int_0^1 \tilde{A}(s,n,k)\phi_l(s)ds &= \int_0^1 \tilde{A}(s,n,k)[\phi_h(s) - \phi_l(s)]ds \\
&= -\int_0^1 \tilde{A}_1(s,n,k)[\Phi_h(s) - \Phi_l(s)]ds \\
&\geq 0
\end{aligned}$$

where $\tilde{A}_1(s,n,k)$ is the partial derivative of $\tilde{A}$ with respect to s. ■

The result is derived directly by the MLRP assumption. On average, the agents in $\mathcal{H}$ receives higher s than the ones in $\mathcal{L}$. Therefore, their average estimation of the fraction of the population who endorses h is larger among the members in $\mathcal{H}$.

6.2 Completeness

The population is complete if, in both endorsement groups, there are agents who believe they are majorities as well as agents who believe they are minorities. Intuitively, if agents' preferences are sufficient diverse, there will be diversity in agents' belief.

Proposition 3 *Define* $\underline{\eta} = max\{\underline{\lambda}, \underline{s}\}$ *and* $\overline{\eta} = min\{\overline{\lambda}, \overline{s}\}$ *where* $\underline{\lambda}$ *and* $\overline{\lambda}$ *are lowest and highest preference, and* $\underline{s}$ *and* $\overline{s}$ *are lowest and highest signal given* θ *respectively. The population is complete if for all* $s' \in \{s : A(s) < 1/2\}$ *and* $s'' \in \{s : A(s) > 1/2\}$*, the following two conditions hold.*

1. $\underline{\eta} < s'$
2. *There exists the preference value* λ^* *such that* $s'' < \lambda^* < \overline{\eta}$.

Proof) An agent whose preference is $\underline{\eta}$ and receives the same signal chooses h and believes herself to be a minority. Also, an agent whose preference is λ^* and receives the signal $\overline{\eta}$ chooses l and believes herself to be a minority. Also, note that the results with the observations, (n, k), is achieved automatically since any extreme observation is possible. ■

7 False Consensus Effect in Social Psychology

In this section, I compare some of the main explanations of the false consensus effect in social psychology with my model and describe their implications. I describe the extent the different predictions could be separated from each other. Krueger (1998) summarizes the hypotheses regarding the false consensus effect and related topics. There are five different categories of hypothesis: *sampling bias, motivational bias, induction, reversal of causality,* and *overconfidence in one's own belief.*

The hypothesis of *sampling bias* states that individuals receive or recognize biased samples about others' behavior. The reason for this may be 'selective exposure'- the tendency to associate with similar people. In this case, individuals could overestimate the prevalence of own choice as the available sample is drawn from a subset with more similar members. Also, psychological evidence shows that the similar behavior of others comes to mind more easily than dissimilar behavior. Individuals tend to view own attributes as more salient than the alternative. If individuals notice similar attributes more easily than dissimilar ones, then it is possible for them to systematically overestimate the number of agents who behave as they do.

The hypothesis of *motivational bias* assumes that people want to consider their own behaviors to be rational and appropriate responses to the situation. One mechanism by which the bias could arise is if agents believe that if others select an action, then there is a rational basis for selecting the action. Thus, individuals possess an intrinsic preference to overestimate the prevalence of own decisions in the population because behaving like others is a sign of rationality. The hypothesis is closely related to the implication of the self-serving bias - the tendency for people to attribute their successes to their own abilities but their failures to chance. People with a self-serving bias may overestimate the prevalence of their own bad behavior, but underestimate the prevalence when the behavior is desirable (Suls et al., 1988). They protect their own ego by perceiving that their own undesirable traits are more common and enhance their own ego by believing that their own desirable traits are rare. Also people tend to possess greater false consensus bias when the reference group is attractive than when the reference group is unattractive (Marks & Miller, 1982).

The *induction* hypothesis (Dawes, 1989) assumes that the information that forms the basis for an individual's choice will lead to systematic biases in the assessments of the behavior of others. If the information that induces one's own decisions is correlated with the information that motivates the decisions of others, some version of the false consensus effect will arise. When estimating an unknown attribute (others' endorsement), it is natural (and rational)to use the available information. This behavior is called projection, and the false consensus effect is well described with this behavior.

The *reversal of causality* hypothesis attributes the false consensus effect to people's tendency to choose their own endorsement following the majority or normative behavior. Most of the results of the false consensus effect measure the correlation between agents' endorsements and estimates. The causal direction is usually assumed but verified only with carefully conducted experiment. Hence it is possible that individuals think that the population is more like they are because the group's behavior determined their own behavior. Agents who sample different parts of the group may therefore end up behaving differently, but still maintain the belief that their behavior is representative.

The hypothesis of *overconfidence of own belief* assumes that people overestimate the precision of their own belief compared to that of others. This hypothesis shares the view of the induction hypothesis that people use their own endorsement as well as others' to estimate the overall behavior of the population. The hypothesis is introduced to provide rationale to the generalized false consensus effect. To illustrate, consider three cases that differ in the degree of how agents weigh their own endorsement compared to others'. In one extreme case, agents do not consider own endorsement as informative at all, and there will be no false consensus effect given agents observe an unbiased sample of the population. The intermediate situation is the case where agents consider their own endorsement as informative as any other endorsement of another agent. This is the case illustrated by induction hypothesis, and the false consensus effect arises in this case. In the other extreme case, agents do not consider others' endorsement as informative, and there is the highest possible false consensus effect. In each endorsement group, all agents use the information that led them to that specific endorsement to predict the overall behavior of the population. The hypothesis of *overconfidence of own belief* assumes that the agents' behavior is between the second and the third case illustrated above. Krueger and Clement (1994) suggest several assumptions that could give rise to this behavior. Firstly, people may not understand the nature of sampling procedure. If people understand how they make their own decision, but not how others make theirs, then individuals may view their own endorsement to be more informative than that of another since they understand the underlying logic of reaching their decision but not those of others. Secondly, the order of the observations can also explain the bias. The primacy effect indicates the tendency of people to put more weight on the information obtained earlier. Self-related information is an initial observation. Thus, it is possible for individuals to view other's information obtained later as less important. Finally, own information can be processed more automatically than that of others. The heuristic process takes less effort for own information. Thus in the process of estimating the behavior of the population, agents can lose some information obtained from other agents' endorsements due to the higher cost to recall them. This difference of mental cost in processing information might be able to explain the discrepancy between own-others' information.

The alternative hypotheses described above have their own merits. However, each of them predicts different outcomes in some situations depending on the characteristics of the attributes.

My model does not distinguish the bias between behaviors with different desirability. However, the hypothesis of motivational bias predicts that people might overestimate the consensus if the comparison is made on undesirable attributes and underestimate the consensus if the comparison is made on desirable ones. Also, my model does not assume any difference in reference groups. Thus, the result of heterogenous reference group is out of the paper's scope.

The hypothesis of reversed causal relation predicts that the false consensus effect arises not because people possess bias in assessing others' behavior, but because they choose their behavior as what the majority does. This result as well as the primacy hypothesis can possibly be tested by reversing the order of information acquisition. Suppose I run an experiment and tell subjects that they sampled a subset of a population and learned that $x\%$ chose option h. Then tell the subjects what option h is. The result supports the hypothesis of the reversed causal relation if subjects show behaviors that increase the false consensus bias when the treatment use the reversed order. The result supports the hypothesis of primacy if the subjects report different estimate in different order treatment.

8 Conclusion

In this paper, I introduce a simple model of information acquisition that describes the false consensus effect. The model extends the induction hypothesis (Dawes, 1989) that provides a rational explanation of the false consensus effect and describes the false consensus effect in its general form. I suggest that the generalized false consensus effect- people's tendency to overestimate the information of own choice- can arise if individuals can only observe partial information about the true state of the world.

Reference

Alicke, M. D. and E. Largo (1995). "The Role of the Self in the False Consensus Effect", *Journal of Experimental Social Psychology* vol. 31, 28-47.

Baker, L., R. Koestner, N. M. Worren, G. F. Losier, and R. J. Vallerand (1995). "False consensus effects for the 1992 Canadian referendum", *Canadian Journal of Behavioural Science* vol. 27, no. 2, 214-225.

Brown, C. E. (1982). "A false consensus bias in 1980 presidential preferences", *Journal of Social Psychology* vol. 118, 137-138.

Dawes, R. M. (1989). "Statistical Criteria for Establishing a Truly False Consensus Effect", *Journal of Experimental Social Psychology* vol. 25, 1-17.

Dawes, R. M. and M. Mulford (1996). "The False Consensus Effect and Overconfidence; Flaws in Judgement or Flaws in How We Study Judgement?", *Organizational Behavior and Human Decision Process* vol. 65, no. 3 March, 201-211

Engelmann, D. and M. Strobel (2000). "The False Consensus Effect Disappears if Representative Information and Monetary Incentives Are Given" *Experimental Economics* vol. 3, 241-60.

Engelmann, D. and M. Strobel (2004). "The False Consensus Effect: Deconstruction and Reconstruction of an Anomaly", *CERGE-EI Working Paper No. 233*.

Gilovich, T., D.L. Jennings, and S. Jennings (1983). "Causal Focus and Estimates Consensus: An Examination of the False Consensus Effect", *Journal of Personality and Social Psychology* vol. 45, 550-559.

Goeree J. K. and J. Großer (2007). "Welfare Reducing Polls", *Economic Theory* vol. 31, no. 1, 51-68.

Hoch, S. J. (1987). "Perceived Consensus and Predictive Accuracy: The Pros and Cons of Projection," *Journal of Personality and Social Psychology*, vol. 53, No. 2, 221-234.

Judd. C. M. and J. T. Johnson (1981). "Attitudes, Polarization and Diagnosticity: Exploring the Effect of Affect," *Journal of Personality and Social Psychology*, vol. 41, 26-36.

Krishna. V. (2009). “Auction Theory,” Academic Press.

Kreuger, J. (1998). “On the Perception of Social Consensus” *Advances in Experimental Social Psychology*, Edited by M. Zanna, vol 30, 163-240. Academic Press.

Krueger, J. and R. W. Clement (1994). “The Truly False Consensus: An Ineradicable and Egocentric Bias in Social Perception,” *Journal of Personality and Social Psychology*, vol. 67, No. 4, 596-610.

Manstead, A. S. R. (1982). “Perceived Support for Opinions: A Test of the Magnitude and Diversity Hypothesis,” *British Journal of Social Psychology*, vol. 21, 35-41.

Marks, G., and N. Miller (1982). “Target Attractiveness as a Mediator of Assumed Attitude Similarity,” *Personality and Social Psychology Bulletin*, vol. 8, 728-735

Mullen, B. (1983). “Egocentric bias in estimates of consensus,” *Journal of Social Psychology*, issue 121:1, 31-38

Mullen, B., J. L. Atkins, D. S. Champion, C. Edwards, D. Hardy, J. E. Story and M. Vanderklok (1985). “The False Consensus Effect: A Meta-analysis of 115 Hypothesis Tests,” *Journal of Experimental Social Psychology*, vol. 21, no 3, 261-283.

Ross, L., D. Greene, and P. House (1977). “The false consensus effect: An egocentric bias in social perception and attributional processes,” *Journal of Experimental Social Psychology*, vol. 13, 279-301.

Sherman, S. J., C. C. Presson, and L. Chassin (1984). “Mechanisms Underlying the False Consensus Effect: The Special Role of Threats to the Self,” *Personality and Social Psychology Bulletin*, vol. 10, 127-138.

Suls, J., C. K. Wan, and G. S. Sanders (1988). “False Consensus and False Uniqueness in Estimating the Prevalence of Health-Protective Behaviors,” *Journal of Applied Social Psychology*, vol. 18, 66-79.

Vanberg, C. (2008). “A Short Note on the Rationality of the False Consensus Effect” Mimeo.

van der Pligt, J. (1984). "Attributions, False Consensus, and Valence: Two Field Studies," *Journal of Personality and Social Psychology*, vol. 46, No. 1, 57-68.

Williams, J. (2009). "Do Financial Analysts Efficiently Aggregate their Information?," Mimeo.