# Minority Opinion and Herd Behaviour

Daniel Sgroi

April 2004

CWPE 0421

**Not to be quoted without permission**

# Abstract

Is majority opinion a better guide to action than a minority view? This paper demonstrates that a direct application of rational herding theory to this novel area can produce a surprisingly counter-intuitive result: given (i) the minority has a clear conformist view and (ii) decision-makers learn through observation as in a herding model, then size does not matter when evaluating whether some groups make better decisions than others. Extending this further we argue that it may be advantageous for risk averse agents to support a form of positive discrimination, that new generations have a largely ambiguous impact, and that the use of electoral colleges can be supported on informational grounds.



*Keywords*: minorities, majorities, conformity, observational learning, herding.

*JEL Classification*: D83, D79, D71, D69

# Minority Opinion and Herd Behavior

Daniel Sgroi *
Churchill College & Department of Applied Economics
University of Cambridge

April 2004

## I. Introduction

More people use PCs than Apple Macs, though Macs seem very prominent in certain sectors, so what should I choose? Different political views seem especially prominent in some geographical regions, is this any help to my decision about how to vote? In the country where I live most people cook with electricity, but in this town most go for gas, what should I use? Most people eat meat, but in certain communities the notion is viewed with abhorrence. Some societies seem based around the notion that God exists, though there are invariably small subgroups who argue otherwise. If the majority in a society seem to believe that God does not exist, then equally there seem to be small subgroups convinced of the opposite. In many of life's difficult choices, whether trivial or important, there may be a conformist view among the majority, and quite possibly other views prevalent within minority groups. For an individual decision-maker with little presumption in either direction, it can be difficult to make a decision on the basis of a small amount of partial information, and the tendency is to look to others for affirmation. The majority conformist line may hold an attractive magnetic pull, at least when forming an initial leaning. So belief in God may go hand in hand with what others in your community believe, as might the choice of whether to cook with gas. In each case we might believe that there is a clear correct choice, but we may never be certain whether we have made that choice. In that sense some of these problems have a *credence good* element to them that makes it hard to know, even after the event, whether a good choice has been made, as originally described by Darby and Karni (1973). Experimentation may be difficult to

* The author is grateful to Rupert Gatti, David Gill and Andrew Temple for useful comments.

justify in this case. Alternatively, the correct decision may become apparent in time, but only after an irreversible choice has been made.

The herding literature pioneered by Banerjee (1992) and Bikchandani, Hirshleifer and Welch (1992) made headway into how those with some private information, and a desire to get such decisions right, might look to others to provide inspiration. This literature indicates that a sensible economic agent will often disregard his own private information when public information seems overwhelming. Perhaps the simplest description of the findings in this paper result from the addition of a meta-question to the standard herding model, asking what can be inferred about the true state of the world by observing the equilibrium in a herding model. In one community PCs may have been used by enough people that the vast majority simply opt for PCs without giving Macs much thought. We ask what can be inferred from this choice.

Added to this we note that many agents in society are disproportionately concerned with the views of only a subset of the total population. Be they neighbors, work colleagues, those of a similar age, family or friends, these are the people that may be observed for examples of what to do when difficult choices have to be made. Under such circumstances we might imagine selective herding behaviour in which those within a community mimic the actions of those also within their community, while in other communities very different actions might dominate. For someone outside either community it might be difficult to see why one state, county or constituency has a history of always voting in a particular way, while a neighboring state is so different. Or why most industry sectors rely on PCs, but some believe Apple Macs to be the superior machine, and so on.

Returning to the individual who stands outside any particular community, we ask the question: should he follow the conformist line in the majority group because the weight of opinion is behind it, or does a minority view have anything to offer? This leading question is based on a simple and even intuitive idea: if 95% of the population believes $X$ and only 5% believes $Y$ (and $X$ and $Y$ are mutually exclusive) then surely $X$ is much more likely to be true? Possibly even 19 times more likely? The answer which this paper gives to this question is a qualified "no". In fact, *if the minority and majority are conforming to different choices, the minority is as likely to give the right result as the majority.* We can also show that there exist circumstances when a social planner might wish to strongly support minority views over the majority, particularly when utility is transferable and agents risk averse. We find counter-intuitive patterns emerging when we consider the actions of new generations born within a population characterized by many disjoint subgroups. Finally, we can support existing institutions such as the electoral college system, the U.S. Senate, or the unilateral nation-state veto in the E.U., or suggest a way to go forward with the reform of the House of Lords in the

U.K., on the basis of the same underlying model. In each case a remarkably simple argument produces a surprisingly counter-intuitive result.

The next section describes a simple example of a herd and the simplest possible model which can encompass the notion of rational herding within disjoint subgroups in society. Sections 3 through 6 go through results and applications: Section 3 demonstrates the equivalence of minority and majority opinion; Section 4 examines the implications of transferable utility and risk averse agents; Section 5 looks at the behaviour of future generations; and Section 6 finds support for the use of electoral colleges. The final section offers some conclusions.

# II. Defining A Herd

Here we will first motivate the paper with an example, then go on to provide a simple model capable of generalizing our example.

## A. An Example

Consider a simple example. A sequence of agents must decide whether to opt for choice $A$ or choice $B$. Agents all have a private signal. They consider all private information to be of the same quality so their private signal is not *a priori* any better than anyone else's. They can add to the information contained within their signal by observing the action (not the signal) of their predecessors. The first agent will follow the choice dictated solely by his private signal, say into $A$. The next agent has both his private signal and the public information relating to the action of the first agent, which in this case perfectly reveals his signal. Let us assume that agent 2 also has a signal favoring $A$. He will therefore also opt for $A$. The third agent observes the decisions of the first two agents. What if his signal suggests that the second choice is superior? He can infer that the first agent had a signal favoring $A$. The second may have had a signal favoring $A$, which would account for his actions or he may have had a signal favoring $B$, but he might have gone for $A$ in indifference. We assume that an agent randomizes if indifferent, though this is not important. The third player knows that the net information revealed by agent 2's decision is in favor of $A$ which swamps his own signal and a *herd* on $A$ begins. With no new information, agent 4 will also go for $A$, as will agent 5, and so on. It is of course possible that $B$ is in some sense superior. For example, $B$ may literally be the objectively correct decision, or $B$ may be the decision suggested by the majority of private signals, both of which are perfectly consistent with the incorrect herd on $A$ in the example.

Expanding the example consider the initial herd having taken place in a large community, resulting in choice $A$ becoming the conformist action. Consider an alternative much smaller community with the same distribution of signals considering the same two choices. Now it might easily emerge that the second smaller community herds on choice $B$, and that becomes the norm or conformist action in that community. The key point is that each community is not aware of (or does not care about) the other community's sequence of choices. From the perspective of an impartial observer glancing at conformity across both communities, it might appear that the first choice has the weight of considerably more private signals behind it (virtually all of the entire larger community have adopted that choice, whereas fashion $B$ is popular only in the much smaller community). However, the dominant position of choice $A$ need not in fact be based on any greater quantity of signals than choice $B$'s more modest following.

## B. A Simple Framework

Smith and Sorensen (2000) provide probably the most general framework in which to study rational herding, by adding noise and heterogeneous preferences to the seminal work by Banerjee (1992) and Bikchandani, Hirshleifer and Welch (1992). However, our concern is less with the likelihood of a herd under different assumptions, and more with normative policy suggestions given that a herd is likely. Therefore, we consider the simplest possible model which allows herding behaviour, and choose to follow Bikchandani, Hirshleifer and Welch (1992), but incorporating a simple notion of heterogeneity through the use of disjoint sets in the population. More on the assumptions required to allow herding given homogeneous agents can be found in Lee (1993), and Kalai and Jackson (1998) which provides a good discussion on the relationship between herding and repeated games.

A finite sequence of individuals $1, 2, ..., N$ act in an exogenously given order. There are two possible states of the world, $\theta \in \Theta = \{A, B\}$. The stream of actions can be defined as $\{a_i\}_{i=1}^{N}$, where $a_i \in \mathbb{A}_i = \{A, B\}$. Utilities are denoted by $\{u_i\}_{i=1}^{N}$, the stream of utilities gained by the population of size $N$, where we simply allocate a utility of 1 if $i$ correctly matches $a_i$ with $\theta$, otherwise 0. All individuals hold a prior of 0.5 : 0.5 over these states but receive private conditionally *i.i.d.* signals $\sigma_i$ with values $\sigma_A$ or $\sigma_B$ distributed:

$$\Pr(\sigma = \sigma_A \mid \theta = A) = \Pr(\sigma = \sigma_B \mid \theta = B) = p$$

We restrict $p \in (0.5, 1)$ so there are no perfectly revealing signals. Any agent $x$ can only observe action sequence $\{a_i\}_{i=1}^{x-1}$ and will Bayesian update his own private signal with $\{a_i\}_{i=1}^{x-1}$ to produce a decision. We adopt the convention of assuming a fair coin flip on indifference, so $\Pr(A) = \Pr(B) = 0.5$. Note that the qualitative results throughout are not dependent upon this assumption, and for example assuming that agents follow their private information on indifference makes no substantive difference.

**Definition 1** *An agent who follows the behaviour of the preceding individual, without regard to his own information, is said to be in a herd.*

Return to the scenarios in the introduction and consider spliting the population into mutually disjoint sets of individuals. Initially, as in the examples of PC vs. Macintosh, Republican vs. Democrat, gas vs. electricity, etc. we will confine ourselves to two groups. Define the size of the majority $\alpha$-group as $N_\alpha$ and the minority $\beta$-group as $N_\beta < N_\alpha$. Each group is disjoint in the sense that members care only about the actions of other members of their group, so any member $x$, of group $\tau$, denoted $x^\tau$ is modeled as observing only the actions of agents $\{a_i^\tau\}_{i=1}^{x-1}$ where $\tau \in \{\alpha, \beta\}$. Potentially we can add further groups, $\delta, \gamma$, etc., and do so below. Some definitions follow.

**Definition 2** *A sequence of agents $\{x_1^\tau, x_2^\tau, ..., x_{N_\tau}^\tau\}$ indexed by $\tau \in \{\alpha, \beta, \delta, ...\}$ is denoted the "$\tau$-group" and is of measure $N_\tau$.*

**Definition 3** *The $\tau$-group is called observationally disjoint if no element of the sequence $\{a_i^\tau\}_{i=1}^{N_\tau}$ can be observed by agent $x^{-\tau}$, where $-\tau \neq \tau$.*

**Definition 4** *The $\tau$-group is called self-observing if typical group member $x^\tau$ can observe the sequence $\{a_i^\tau\}_{i=1}^{x-1}$ as well as his own private signal $\sigma_x$.*

# III. The Equivalence of Majority and Minority Opinion

Consider a typical observationally disjoint but self-observing group. Denote the probability of a herd starting by agent $N$, and thus an end to any new public information in such a group, as $Z(p, N)$. The probability of no herd by agent $N$ is therefore $1 - Z(p, N)$.

Summarizing the calculations in the Appendix, the probability of no herd by agent $N$ can be simply given as:

$$1 - Z(p, N) = [p(1-p)]^{\frac{N}{2}}$$

So the probability of a herd is:

$$Z(p, N) = 1 - [p(1-p)]^{\frac{N}{2}}$$

More on the calculation and properties of these probabilities can be found in the Appendix and in Sgroi (2002). Now let us define, the conformist action (or view), as follows:

**Definition 5** *The "conformist action (or view) for the $\tau$-group" or simply "$\tau$ conformist action (or view)" is the action followed (or view taken) after the start of a herd in the $\tau$ group.*

Next follow some results, including the equivalence of the conformist action to the action followed by the majority in the group.

**Lemma 1** *The conformist action in the $\tau$-group will be followed by the majority in the group.*

**Proof.** Up to the start of a herd, an identical number must be following each action, and then from the herd all must be following the conformist action. ■

**Remark 1** *Even if the minority group is small, the probability of there being a conformist action is very high.*

Here we need that the probability of a herd starting is high even for a small $N_\beta$. We have $Z(p, N_\beta) = 1 - [p(1-p)]^{N_\beta/2}$, and $p \in (0.5, 1)$. Furthermore since $dZ(p, N_\beta)/dp > 0$, we know that:

$$\arg\sup_p \{1 - Z(p, N_\beta) \mid p \in (0.5, 1)\} = 0.5$$

We can therefore minimize the chance of a herd by taking the limit of $Z(p, N_\beta)$ as $p \to 0.5$. Even using this limit we still have the probability of a herd, and thus the probability of a conformist action arising as close to 1 for very small populations, for example:

$$\begin{aligned}
\lim_{p\to 0.5} Z(p, 10) &\approx 0.999 \text{ (to 3 decimal places)} \\
\lim_{p\to 0.5} Z(p, 20) &\approx 0.999999 \text{ (to 6 decimal places)} \\
\lim_{p\to 0.5} Z(p, 52) &\approx 1 \text{ (to 30 decimal places)}
\end{aligned}$$

**Remark 2** *We can tell whether the group has a conformist action just by looking at the total split in the actions of that group (which is a sufficient statistic for the information revealed).*

We know that the chance of a conformist action being prevalent is very high for all but very small groups, and should we wish to verify that there is indeed a conformist action we can do so easily by simply looking at the total split of actions in the group. Anything significantly different from a $50:50$ split indicates a conformist action has arisen, and as $N$ rises this action will dominate an increasingly high proportion of the total population's actions.

Finally, we can summarize the main result in this section as follows:

**Proposition 1** *Consider two groups of measure $N_\alpha$ and $N_\beta$. Given that each group has a conformist action, the majority view in the $\beta$-group is as likely to be correct, as the majority view in the $\alpha$-group.*

**Proof.** Immediate from Definition 5 and Lemma 1. ■

Recall from Remark 1 that the probability of obtaining a conformist action is extremely high even for a small population. Despite this, if the conformist actions in the two groups are identical then we have a trivial result. If however they differ, so we have a conformist view in the smaller minority group and a *different* conformist view in the larger majority group, then this provides the counter-intuitive result that *a minority view even if different from that of the far greater majority view, is as likely to be correct.*

# IV. Transferable Utility and Risk Aversion

In this section and those that follow we build on the results in the last section with a sequence of extensions. Here we consider a social planner who wishes to maximize overall utility, and is able to do two things:

(a) Transfer utility between agents costlessly via a lump sum tax;
(b) Force adherence by some people to the conformist action in the other $\tau$-group.

We might think of (b) as "negative discrimination" if the planner forces a member of the minority to follow the conformist line in the $\alpha$-group (majority) or "positive discrimination" if the planner forces some in the majority to opt for the conformist action in the $\beta$-group (minority).

Start with a thought experiment. Consider a society with two groups ($\alpha$ and $\beta$) opting for different alternatives ($\alpha$-group's conformist action is $A$ and $\beta$-group's conformist action is $B$) with each equally likely to be correct by Proposition 1. By making the $\alpha$-group expand to the full size of the population we are left with a single option which might return 1 to all agents, or 0, with equal probability, so 0.5 *in expectation.* By making each group exactly 50% of the population (setting $N_\alpha = N_\beta$) we ensure that 50% of the population gets 1 and 50% gets 0, allowing through transfers a *certain* 0.5 to be obtained.

Now return to the social planner. Let us assume that the social planner wishes to maximize total utility. The result in this section is not specifically dependent upon a form of utility, so consider a very general welfare function $W\left(\phi(\{u_i\}_{i=1}^N)\right)$, though crucially since all agents are risk averse, the social planner faces an increasing and concave function, so $W'(\phi) > 0$ and $W''(\phi) < 0$. The social planner has no bias towards membership of either the $\alpha$-group or $\beta$-group, and the social planner observes the conformist choice in both $\tau$-groups. We carry through all other assumptions from the last section, so for example, both groups are observationally disjoint and self-observing.

Like any risk averse economic agent the social planner prefers the insurance of diversification. Allow the social planner to be able to force adherence to a particular action, so we can think of the planner as observing the result of the herds over groups $\alpha$ and $\beta$ and then being able to force a subset of one group to adhere to the conformist view of the other. In particular, we then have the following proposition:

**Proposition 2** *Consider two groups of measure $N_\alpha$ and $N_\beta$. So long as the conformist actions in both groups differ, the risk averse social planner will force a subset of the majority to adhere to the minority conformist action.*

**Proof.** With prior $50:50$ and given two different conformist actions in each group, applying Proposition 1, the social planner has posterior neutral beliefs over the superiority of any action in $\mathbb{A}_i$. Since $W''(x) < 0$ the risk averse social planner prefers a $50:50$ split of the population over the actions $\mathbb{A}_i$ which is the only division of society which ensures a certain return to society of $0.5N$ as well as an expected return of $0.5N$. Since $N_\alpha > N_\beta$ and through a direct application of Proposition 1 and Lemma 1 we know that the $\alpha$-group action is undertaken by in excess of $0.5N$, so the social planner will wish to force a subset of the $\alpha$-group to switch to the conformist action in the $\beta$-group. ■

This is again a remarkably counter-intuitive result, though the proof is an extremely straightforward application of the results in the previous section. Since both conformist actions are as likely to be correct, and the social planner would rather have a $50:50$ split across these two actions, the planner will force movement to the minority view, effectively ignoring the observational learning within groups. Returning to our initial description, the social planner prefers the society in which $N_\alpha = N_\beta$ which ensures that 50% of the population gets 1 and 50% gets 0, allowing a certain 0.5 to be obtained after transfers.

By making agents risk-loving we would have the reverse result. Alternatively, with risk neutrality we would have indifference. We focus on the risk averse outcome since risk aversion on the part of economic agents seems most likely, while the suggested positive discrimination seems most counter-intuitive.

The irony is of course that a stylized view of history might suggest that minority views have at best been tolerated, and rarely aggressively supported to the point of working to suggest that some in the majority should actively be forced to switch allegiance. Historical exceptions to this have usually come where the social planner believes that he or she knows better than the majority and wishes to implement total compliance to his or her preferred action, as in the enforcement of religious conformity in the era surrounding the Reformation in Europe. However, this paper suggests than a social planner representing risk averse agents, concerned only about maximizing some social welfare function, might aim to bolster the minority view on information-based insurance grounds. Of course a social planner, president or prime minister has many more factors to incorporate than those based on observation learning alone.

# V. New Generations

Imagine the two groups, $\alpha$ and $\beta$ from Section 2, have settled into their respective conformist actions. The two actions differ, but there is no intervention from any social planner. So here we have a dominant group in society following one "social norm", and a minority group, following another. This persists since members of each group are concerned only about the actions of others in their own group, i.e. they are observationally disjoint. We also do not allow the agents to judge whether their own social norm is superior to that in the other group, imagining that social norms are akin to *credence goods* in which there is no way of deriving quality after the event. Historically, we might imagine that the two groups have settled down, and but for occasional frictions, this is how it has been for some time.

Now consider the arrival of a new generation. To motivate this, consider the initial population to be from two sources. The first is a larger group and quickly conforms to a particular set of social norms, the second is a smaller "ethnic" group which conforms to a different norm (possibly in both cases the norms were determined long before migration to the country in question). Now a generation later there is more interaction between groups, and the new generation of agents who appear, the $\delta$-group of measure $N_\delta$, have the option of following the $\alpha$-group's conformist action, or the $\beta$-group's conformist action. From Proposition 1 they know that despite the fact that the $\beta$-group's conformist action is in the minority in the population as a whole, it is as likely to be optimal. How should this new generation react?

## A. With Observation

We can model the new generation as reacting just as the older generations did. Once again each agent $x$ receives private information $\sigma_x$, correct with probability $p$, and once again acts in a strict exogenous sequence. The new agents begin with the same prior as the older generations, and once again the herding model is applied. Now since each agent in the stream of new ones, which we will call the $\delta$-group, can observe the actions of previous members of the $\delta$-group, and since the public information contained in the actions of the $\alpha$-group and the $\beta$-group cancel out (each conformist action is equally likely by Proposition 1), the stream of $\delta$-agents acts just like in the standard herding setup of Section B.

So, depending on which way the herd goes, either the minority action will catch up with the majority (and then overtake it if the stream of new agents is big enough), or the majority view will take an even bigger lead. Because neither the majority nor minority are more likely to be correct, and there is a simple *ex ante* 50 : 50 on which of these two scenarios occurs. This is summarized in the following proposition:

**Proposition 3** *Consider the conformist action in both $\alpha$ and $\beta$ groups to be different. Introducing a new, self-observing $\delta$-group, which can observe the conformist action in both $\alpha$ and $\beta$ groups, results in an ex ante equal probability of the conformist action in the $\delta$-group corresponding to the conformist action in the larger $\alpha$-group or smaller $\beta$-group.*

Yet, the action followed by the new agents is more likely to be correct (as a herd is more likely to start on the correct action). If both $\alpha$ and $\beta$ groups follow identical actions then this will instantly swamp the private information available to each member of the $\delta$-group, and so they can be expected to follow the conformist action supported by both old generations. We therefore find a remarkable knife-edge corollary to Proposition 3:

**Corollary 1** *If both $\alpha$ and $\beta$ groups share the same conformist action then the $\delta$-group will also follow this action. If the smaller $\beta$-group conforms to an action, no matter how small the group, the $\delta$-group is ex ante equally likely to follow either action.*

**Proof.** If both $\alpha$ and $\beta$ groups share the same conformist action, then this reveals that the net amount of private signals from groups $\alpha$ and $\beta$ support this action, which is enough to swamp the private signal of any $\delta$-group individual, which proves the first part. The second part is immediate from Proposition 3. ■

So a minority group with no conformist action will be ignored by the new generation, but one with a conformist action will change the overall social outcome. Given that the likelihood of a conformist action existing is increasing in $N_\beta$, there is likely to be a discontinuity at some population size.

## B. Without Observation

This section looks at the case when the new generation or $\delta$-group is not self-observing. We need to justify this since we have assumed that both the $\alpha$-group and $\beta$-group are self-observing. We might argue that we are examining this case simply for completeness, but there might also be a case in which is makes sense for the later generation to not concern itself with observing the actions of other members, while earlier generations did. The first two groups faced a 50 : 50 prior, and so had a natural need to obtain more information, especially via costless observation. Possibly it is now more difficult to costlessly observe since the new generation has to disentangle the conformist actions of the majority and minority from the newly decided actions of another member of the $\delta$-group. If I observe someone purchasing product $A$ is it because they come from the $\alpha$-group whose conformist action is to purchase $A$, or are they a member of the $\delta$-group who has a signal indicating that $A$ is the best choice? The very different results which follow the assumption of no observation give a final justification.

Begin by assuming that the conformist actions in both the $\alpha$-group and $\beta$-group differ. If the $\delta$-group is not self-observing then since the information from the previous groups cancels out, they all act on their own signal with no potential for herding. Therefore, a proportion $p$ of the new $\delta$-group agents will choose the correct action, with $1 - p$ choosing the incorrect action. Overall, the proportion of people choosing the right action will tend to $p$ as $N_\delta/N \to 1$.

**Proposition 4** *If the $\delta$-group is not self-observing then (i) a proportion $p$ of the $\delta$-group will choose the correct action, and (ii) the overall proportion of the total population choosing the right action tends to $p$, as $N_\delta/N \to 1$.*

**Proof.** Immediate from Definition 4. ■

We finish with some remarks.

**Remark 3** *Assume that the proportion of the population that hold the $\alpha$ conformist view exceeds $p$, then over time, if the $\alpha$ conformist view was wrong, the $\beta$ conformist view will slowly increase over time (towards $p$ as $N_\delta/N \to 1$), while the $\alpha$ view will slowly decrease over time (towards $1-p$ as $N_\delta/N \to 1$). If the $\alpha$ conformist view was correct, it will still decrease over time (towards $p$), while the $\beta$ view will increase over time towards $(1-p)$.*

So observing a reduced proportion adhering to the majority social norm does not indicate that it must be incorrect. In fact, if the $\alpha$ conformist view is followed by a proportion above $p$ before the new generation appears, the result of the new generation's actions must result in the proportion following the $\alpha$-view falling in the long run. If the initial proportion of the population following the conformist action endorsed by the $\alpha$-group did not exceed $p$ before the arrival of the new generation, then Remark 3 needs to be altered.

**Remark 4** *Assume that the proportion of the population that follow the $\alpha$ conformist view does not exceed $p$, but does exceed $1-p$, then over time, if the $\alpha$ conformist view was wrong, the $\beta$ conformist view will slowly increase over time (towards $p$ as $N_\delta/N \to 1$), while the $\alpha$ view will slowly decrease over time (towards $1-p$ as $N_\delta/N \to 1$). If the $\alpha$ conformist view was correct, it will increase over time (towards $p$), while the $\beta$ view will decrease over time towards $(1-p)$.*

The key feature in this section is that should the new generation act without observation, since it will effectively ignore the older groups, we can expect the population to move towards a proportion $p$ undertaking the correct action and $(1-p)$ undertaking the incorrect action. It is likely that the $\alpha$-group conformist action exceeded proportion $p$ so most likely the result will be a negative impact on the proportion undertaking the majority conformist action regardless of whether it is correct or not.

# VI. Electoral Colleges

A final application concerns the notion of *electoral colleges*: the system in which several groups, often very different in size, receive an equal say in political decision-making. The idea is relatively old, dating back at least to the 1793 French revolutionary constitution which specifically discusses the role of electoral colleges, and the 1787 Federal Constitution of the United States which enshrines the policy of electing two senators for each state regardless of the level of population in each state. A plausible critique of the notion of an electoral college is that it makes more sense in a democracy to award representation in a political process on the basis of raw numbers of voters. The electoral college system instead allocates a representative on the basis of a geographical region, some historical division, or many other (sometimes seemingly arbitrary) criteria. In the U.S.A. this is most famously exhibited through the Senate. The U.K. had parliamentary representation for the two oldest universities (Oxford and Cambridge) until relatively recently, indeed J. M. Keynes was reportedly asked to become the "Member of Parliament for the University of Cambridge", but refused, as detailed in Skidelsky (2000). In many cases, the E.U. allows any one nation to veto major decisions, regardless of how small the nation. With a small modification the underlying model in this paper can support this form of decision-making on informational grounds alone.

Consider a number of $\tau$-groups each of which has a conformist action (or is of sufficient size to make a conformist action or herd almost certain), is observationally disjoint and self-observing. Once again define actions and utilities as in Section B. Now, following Proposition 1 any two such groups regardless of relative size, have much the same probability of correctly selecting an action. Now consider two voting cases:

**Case 1** *("Senate") Each $\tau$-group selects a representative, who then votes on a senate to select which action to be made law, based on the conformist action in his constituent $\tau$-group.*

**Case 2** *("House of Representatives") The entire population of size $N$ simply votes, or equivalently, each $\tau$-group wields a block vote proportionate to its size.*

Case 2 will favor larger $\tau$-groups, while Case 1 accords equal voting power to each $\tau$-group, though appears less democratic in the sense that a much smaller $\tau$-group is weighted equal to a larger group. Despite our democratic concerns about Case 1, from Proposition 1 it is clear that the natural unit of representation is in fact the $\tau$-group, not the individual voter. Note that the conformist action in each $\tau$-group is identical to the action that would be selected through a sequential and public majority vote in that $\tau$-group, following Lemma 1. The following remark makes clear the argument in favor of Case 1.

**Remark 5.** *If we observe $\eta$ $\tau$-groups with a conformist action $a_\eta$, and $\lambda < \eta$ $\tau$-groups with a different conformist action $a_\lambda \neq a_\eta$ then the presumption should be in favor action $a_\eta$.*

More concretely, we might imagine an elected delegate from each group simply supporting the conformist action from his constituency, and all such elected delegates selecting the national choice through a majority vote. This would select option $a_\eta$, which we know to be more likely to be correct via a direct application of Proposition 1, which tells us that we can cancel out opposing conformist views for different groups, which would leave one view predominant.

To motivate this, simply rename action $a_\eta$ as "Democrat", and $a_\lambda$ as "Republican" (or vice versa). Now call the $\tau$-groups "states", arguing that historically state divisions are based on a degree of observational disjointedness, and we have an argument for the existence of a Senate. This suggests at least one rationale for the importance of the Senate in the face of the more obvious democratic arguments for the superiority of the House of Representatives. We might also take this as an argument in favor of a similar institution in the U.K. in the light of the planned reforms of the U.K.'s second chamber, the House of Lords.

Rephrased, this argument makes sense as an information-based theory of democracy: when bestowing electoral power it justifies giving equal weighting to each unit which embodies an identical amount of information concerning the decision to be made, rather than just groups of identical size. Of course, there is a good deal more to think about than the quantity of embodied information when designing an electoral system, but we might wish to add this to our list of worthwhile criteria.

# VII. Conclusions

While the preference for one viewpoint over another may have many and varied elements, observational learning allows an additional consideration which individual decision-makers and policy-makers alike should bear in mind. The findings in this paper indicate, that if the minority achieves a conformist action (or equivalently is of a sufficient size) then it is as likely to be correct as the majority conformist view, as both are based on the same amount of revealed public information. This result is strongly counter-intuitive, but has considerable implications and applications.

While earlier pioneering work by Banerjee (1992) and Bikchandani, Hirshleifer and Welch (1992) suggested that herding on the wrong action, through the crowding out of private by public information, is a possibility, it is still of course still sensible to follow the actions of the majority since that embodies more information than your own private signal. This paper takes this further by showing that where two separate herds develop the natural inclination to consider the view of the larger, majority herd, to be more reasonable does not withstand closer scrutiny. The huge loss of information that occurs after any herd takes place eliminates the informational advantage, which is potentially held by the larger majority herd. So, having observed two herds develop, it might be better to let your own private information guide you rather than simply conform to the outcome prevalent in the larger group. We might even go so far to suggest that policy makers should place no greater weight on the views of large majorities over minorities of a reasonable size when determining which course of action to take where there is an objective correct action, since either group might have found the best action to take with equal probability. Perhaps the most forceful way to stress the counter-intuitive nature of the result in this paper is through one of the simple numerical examples in Section III: if we observe one hermetically sealed community of only 20 herding in unison on a given action or choice, and a neighboring community of 180 opting for another action, given the assumptions of the model, we cannot argue on information grounds alone that the larger group (a majority of 90%) is more likely to be correct.

Furthermore, transferable utility plus risk aversion might suggest a strong degree of positive discrimination towards minority opinion. A social planner wishing to insure society, might wish to induce some who follow the majority line to switch to the minority viewpoint. The rationale for this is once again entirely informational, and need not require any biased preference on the part of the social planner.

The introduction of new generations of agents who act after the initial groups have made their opinions known provides a way to model the behaviour of new generations in society, who face different background influences, but are open to alternative modes of behaviour. The paper finds that if the new generation received private information in much the same way as the older groups then they will decide independent of the older groups and may move overall opinion in either direction. If the new generation does not observe each other's decisions then in the long run the correct decision will rise to encompass a percentage $p$ if the population which may well result in an decrease in the proportion following the majority line even if that happens to be the correct viewpoint.

Finally, the equivalence of minority and majority opinion also provides a strong informational rationale for the use of U.S. Senate style political institutions in a democracy. Possibly, the U.K. might consider a similar institution in the current debate on the reform of the House of Lords.

It should be noted that it is straightforward to generalize to multiple initial groups, as in the application to electoral colleges, though we would need to be careful about using notions such as majority and minority, since even the largest group may not exceed 50% of the total population in size. However, following similar logic to that expressed in this paper, various immediate results obtain. For example, with three groups and two actions, the action endorsed by two groups will most likely be the correct one, a social planner might consider trying to induce the other group to follow the action of the two in agreement, and new generations will most likely all follow the line taken by the two groups. To say something more concrete would require a re-evaluation of the assumptions, and given the wide variety of combinations of number of groups and actions, that is left for the reader to consider on a case-by-case basis.

## References


Banerjee, A. V. (1992). 'A simple model of herd behaviour', *Quarterly Journal of Economics* 107, pp 797-817.

Bikchandani, S., Hirshleifer, D. and Welch, I. (1992). 'A theory of fads, fashion, custom and cultural change as informational cascades', *Journal of Political Economy* 100, pp 992-1026.

Darby, M., and Karni, E. (1973). 'Free competition and the optimal amount of fraud', *Journal of Law and Economics 16,* pp *67-88.*

Lee, I. H. (1993). 'On the convergence of information cascades', *Journal of Economic Theory* 61, pp 295-411.

Jackson, M. O., and Kalai, E. (1997). 'Social learning in recurring games', *Games and Economic Behavior* 21, pp 102-134.

Skidelsky, R. (2000). *John Maynard Keynes, Volume 3.* New York: Penguin Putnam, Viking.

Sgroi, D. (2002). 'Optimizing information in the herd: guinea pigs, profits and welfare', *Games and Economic Behavior 39*, pp 137-166.

Smith, L. and Sorensen P. (2000). 'Pathological outcomes of observational learning', *Econometrica* 68, pp 371-398.

## Appendix

In this Appendix the probability of no herd having started, after an even number of $N$ agents, is calculated.

From the model specifications we can derive the unconditional *ex ante* probabilities of a herd on action $A$, a herd on action $B$, or no herd after $N$ agents in a straightforward iterative manner. Recall from the main text that we denote the probability of a herd starting by agent $N$ as $Z(p,N)$, and the probability of no herd having been initiated after $N$ agents is therefore $1-Z(p,N)$. Define $A(N)$ to be a herd on action $A$ after $N$ agents and define $B(N)$ to be a herd on action $B$ after $N$ agents. So $1-\Pr[A(N)]-\Pr[B(N)] = 1-Z(p,N)$ for all $N$.

Starting with 2 agents we have $\Pr[A(2) \mid \theta = A] = p^2 + 0.5(1-p)p$ and $\Pr[A(2) \mid \theta = B] = (1-p)^2 + 0.5p(1-p)$. Therefore $\Pr[A(2)] = 0.5(1-p+p^2)$. Similarly, we have $\Pr[B(2)] = 0.5(1-p+p^2)$. No herd by agent 2 will occur with probability $1-\Pr[A(2)]-\Pr[B(2)]$, therefore $1-Z(p,2) = p-p^2$. Note of course that this can be alternatively calculated as the occurrence of $\{\sigma_A, \sigma_B\}$ or $\{\sigma_B, \sigma_A\}$ and a coin flip by agent 2, so $1-Z(p,2) = 0.5p(1-p) + 0.5(1-p)p$. Further note that $\Pr[A(2)]$ and $\Pr[B(2)]$ are not conditional on $\theta$ since they are fully symmetric so $\Pr[A(N)] = 0.5Z(p,N)$. Now note that $\Pr[A(4)] = \Pr[A(2)] + [1-Z(p,2)]\Pr[A(2)]$ and similarly for $\Pr[B(4)]$. Further $1-Z(p,4) = [1-Z(p,2)]^2$. Using this we can easily deduce the general probabilities after an even number of $N$ agents to be $1-Z(p,2) = [1-Z(p,2)]^N = (p-p^2)^{N/2}$ for no herd after $N$ agents, and $\Pr[A(N)] = \Pr[B(N)] = 0.5Z(p,N) = 0.5[1-(p-p^2)]^{N/2}$ for a herd on action $A$ or action $B$ after $N$ agents. Now note that as $p \to 1$ herds tend to start sooner, so more precise signals raise the probability of histories that lead to the correct herd where we define correct herd as a herd on $A$ if $\theta = A$ or a herd on $B$ if $\theta = B$.