Good Information Cascades

Elyashiv Wiedman – HUJI

Schedule: December 18-2014, 11:15-12:30

Place: Bldg. 72, room 465

Good Information Cascades

Elyashiv Wiedman

Abstract

In 2008 the FDA changed the guidelines for advisory committees, replacing sequential voting procedures by simultaneous voting procedures. As clarified by the FDA, this was in response to the extensive literature on information cascades arising from sequential procedures. In a simple model I show the advantages of a sequential procedure arising from an information cascade. Contrary to the above-mentioned literature, I show that a sequential procedure may yield more information than a simultaneous procedure. In particular, I compare the behavior of heterogeneous experts in committees using simultaneous and sequential voting procedures. In addition I analyze the aggregate information a decision maker can accumulate under each procedure. I show that under a sequential voting procedure only incompetent experts cascade while competent experts follow their private information. Cascading by incompetent experts makes it possible to identify competent experts, which in turn results in better aggregate information.

Keywords: Information cascade; Information Aggregation; Committees; Experts.

JEL Classification: D71; D82; D83.

1 Introduction

Using a model of information acquisition from a committee comprised of heterogeneous experts, I study whether more information is expected to be achieved under a sequential voting procedure or under a simultaneous voting procedure. The literature on social learning (e.g., Banerjee 1992, Bikhchandani, Hirshleifer and Welch 1992, Smith and Sorensen 2000) demonstrates the pathological outcomes of an information cascade under a sequential procedure, and in particular the ensuing loss of information. An immediate conclusion is that a decision maker who wishes to elicit information from a committee should avoid the use of sequential procedures. Contrary to this understanding, I show that a sequential procedure resulting in an information cascade may yield more information than a simultaneous procedure. Moreover, it is the cascade itself that yields the additional information. To characterize the preferred procedure, I show that the prevalence of competent experts is a decisive factor in the choice of a voting procedure. In particular, a sequential voting procedure yields more information than a simultaneous voting procedure when competent experts are neither too rare nor too prevalent.

Whether a committee of heterogeneous experts should cast their votes sequentially or simultaneously in order to maximize the aggregate information that arise from a committee, is not only theoretically interesting but also has practical significance, as illustrated in the following example. The U.S. Food and Drug Administration (FDA) uses a large number of advisory committees (comprised of experts) in order to get better information on a variety of issues, such as benefit-risk assessments of new drugs and medical devices, before making a regulatory decision. In 2008 the FDA published new guidelines for advisory committees (Voting Procedure at Advisory Committee Meetings, 2008) regarding the voting procedure used by these committees. A significant policy change was the shift from a sequential voting procedure to a simultaneous voting procedure. According to the FDA, this change was in response to the economics literature and was meant to prevent "voting momentum" in sequential procedures, where experts ignore their private information and vote according to the voting of the experts who precede them in the sequence. In this paper I show that the FDA's amendment is not necessarily correct: cascading behavior of incompetent experts may be beneficial when aggregating information from a committee.

In this paper I analyze a model of information acquisition from a committee comprised of heterogeneous experts and explore their incentives to provide an honest vote or to cascade. In particular, I assume that each expert privately observes a signal over an uncertain state of the world that depends on his type: a competent expert observes a signal that is more accurate than that observed by an incompetent expert. Experts, who seek to provide a vote that coincides with the true state, are asked to vote simultaneously or sequentially. Clearly, the experts' behavior in providing votes depends on the implemented voting procedure. Before taking an action a bayesian decision maker aggregates experts' votes. I assume that the final decision is not delegated to the committee and the decision maker is not obligated to any voting rule. Therefore, a vote may be seen as an individual recommendation.[1]

I begin the analysis in Section II, where experts vote simultaneously. As one may expect, under a simultaneous voting procedure all the experts vote honestly, i.e., according to their signals. Nevertheless, a simultaneous voting procedure may lead to a wrong decision: any two votes that contradict each other do not contribute any information, and so the decision may ultimately be based on a small number of experts. Although the literature on social learning, starting with Condorcet (1785), implies that adopting the recommendation arising from the votes of experts will lead to the correct decision being taken almost surely, this is not necessarily the case for committees of experts. Whereas results in the past were based on the votes of an infinite number of experts, committees are usually comprised of only a few experts.

[1]This is exactly the case for the FDA advisory committees. Although the FDA uses the term voting procedure, the FDA is not obligated to follow the committees recommendation and the voting rule is not specified, see Urfalino and Costa (2010).

In Section III I consider the case of a sequential voting procedure. Under such a procedure, in addition to his private signal, each expert observes the votes of preceding experts. I show that the effect of this additional information differs for competent experts and incompetent experts. While an incompetent expert is heavily influenced by the expert preceding him, a competent expert may well be uninfluenced even by multiple preceding experts who vote at odds with his own private information. The effect of preceding votes on an incompetent expert translates into a cascading behavior of incompetent experts, while competent experts stick to their private information and vote honestly. The different behavior of the two types of experts is a key to the advantage of the sequential voting procedure: it allows to identify the competent ones. Since incompetent experts always cascade, only a competent expert may break a sequence of identical votes. Whenever an expert vote in contradiction to the expert preceding him, he is revealed to be a competent expert for sure. It is the cascading behavior of incompetent experts that allows to identify competent experts.

The possibility of identifying competent experts may be valuable in terms of information acquisition, especially when the number of votes is small. Nevertheless, under a sequential voting procedure the information possessed by incompetent experts is lost by their cascading behavior. Hence, a decision maker should evaluate the gain of information arising from the identification of competent experts, against the loss of information caused by the cascading behavior of incompetent experts. When competent experts are very rare, the gain of information does not outweigh the loss of information.

Even though a competent expert is not as prone to influence by preceding votes as is an incompetent expert, eventually a competent expert will cascade. A competent expert will cascade after observing enough votes that contradict his own information and when his information is not completely accurate. The number of contradicting votes that induce a competent expert to ignore his own information depends on either the accuracy of his information or, equivalently, the prevalence of competent experts. When all types of experts cascade, no information can be inferred from their votes. Thus, when competent experts are very prevalent and their information is not completely accurate, a cascade of both types of experts is possible, and the gain of information from identifying the competent experts who refrain from cascading does not outweigh the loss of information from cascading of other experts.

1.1 Literature

The phenomenon of information cascades has been well studied, starting with Scharfstein and Stein (1990), Banerjee (1992), Bikhchandani, Hirshleifer, and Welch (1992), and Smith and Sörensen (2000). Information cascades arise when agents take actions in sequential order, and subsequent agents in the sequence are conformationally influenced by preceding experts in the sequence. This idea lies at the heart of a vast literature on social learning. These papers analyze the informational effect on subsequent agents in a sequence caused by the actions of preceding agents: when agents act sequentially, at some point all agents abandon their private information and cascade after preceding agents. They demonstrate the negative effect of an information cascade on efficiency: with positive probability all agents take the wrong action. In contrast, in this paper I show the positive effect of an information cascade: the cascading of incompetent experts serves to identify competent experts to the benefit of the aggregate information.

It is not surprising that decisions are taken after consultation with many experts. Condorcet (1785) in his seminal contribution showed that the aggregate information of many individuals is better than the information provided by a single individual. Condorcet's jury theorem in its simplest formulation features a committee of size n that is faced with a binary policy, which in turn depends on an unknown state of the world. Each committee member receives some private information about the state and policy is determined by a voting rule (such as majority). The main result of this model is that when the number of voters, n, grows, the correct decision will be made almost surely.[2] A crucial assumption of the present paper has to do with the number of members of the committee: we assume, realistically, committees are comprised of a few experts; therefore, there is a positive probability of taking the incorrect decision. Thus, the question which procedure yields more information is important. Although the superiority of a sequential voting procedure may still hold for large committees, as I show in Section 6, the important contribution of the paper concerns small committees.

Sequential and simultaneous procedures are common instruments in practice, and therefore constitute a fertile field for economic research, particularly for the literature on strategic voting. Dekel and Piccione (2000) found the sequential procedure to be as good as the simultaneous procedure. Gershkov and Szentes (2009) found the sequential procedure to be optimal. As pointed out by Dekel and Piccione, in their model (as well as in Gershkov and Szentes') an information cascade does not arise. Dekel and Piccione assumed that agents care about the final decision but not about their own vote, while Gershkov and Szentes assumed that each agent does not observe the votes of preceding agents. In this paper I assume that experts care about their own vote coinciding with the eventually revealed state of the world, which gives rise to an information cascade.

Since many economic and political decisions are taken by committees, the literature on decision making by committees is well developed. Consider an action that must be taken under uncertainty and a committee comprised of experts who each possess some information about the unknown state. In some cases the decision on what action to execute is delegated to the committee, while in others the committee only advises a decision maker. In the latter case, an important question is how to aggregate information emerging from the committee. Glazer and Rubinstein (1998) analyzed different mechanisms to accumulate information when the experts have the same preferences as the social planner or when the experts' desire is to have their advice imple-

[2]The social learning literature can be thought of as claiming that this result is not true for sequential procedures.

mented. They show that sometimes selfish experts lead to a better decision. An additional point to consider is the order of speaking, since the order of speaking in a committee also affects the aggregate information. Ottaviani and Sörensen (2001) analyzed the optimal order of speaking for a committee of experts. The importance of different orders of speaking arises when experts are known to be of varying expertise. Unlike Ottaviani and Sörensen, I assume throughout this paper that since the decision maker is not aware of each expert's competence, he refers to the experts as being (ex-ante) of the same competence. The present paper may be seen as complementary to their paper. Another point to consider is whether to impose a transparent procedure or to impose a procedure that ensures the secrecy of the experts' votes. Levy (2007) analyzed the effect of votes' transparency on experts' incentives to vote honestly, i.e., according to their private information. She found that a transparent procedure may induce experts to ignore their own information. In contrast to the present paper, Levy considered only the case of a simultaneous procedure.

2 Committees of experts

A decision maker must take an action, $a \in \{a_h, a_l\}$, under equiprobable states of the world, $\omega \in \{\omega_h, \omega_h\}$. The decision maker receives a positive utility if he takes an action corresponding to the true state of the world, i.e., $u_{DM}(a_h, \omega_h) = u_{DM}(a_l, \omega_l) = 1$, and a negative utility otherwise, i.e., $u_{DM}(a_l, \omega_h) = u_{DM}(a_h, \omega_l) = -1$. The focus of this paper is the information that can be accumulated from a committee of experts, and so the symmetry in the decision maker's utility does not play a role. Since the true state of the world is unknown to the decision maker, he may ask for advice from a committee comprised of N experts. I analyze the smallest committee that yields interesting results, which is a three-member committee. The intuition of the main results can be extended to larger committees, as I show in Section 6.

Experts do not observe the true state of the world but privately observe a noise signal, $s_i \in \{h, l\}$ (where the experts are denoted by $i \in \{1, 2, 3\}$), with the symmetric prior probability $Pr(h|\omega_h) = Pr(l|\omega_l) = p \geq \frac{1}{2}$. Experts' signals are assumed to be conditionally independent drawn among experts. The parameter p reflects the ability of an expert to observe the true state of the world. For a competent expert, denoted by $\mathcal{C}$, p equals g, and for an incompetent expert, denoted by $\mathcal{I}$, p equals b, where $\frac{1}{2} \leq b < g \leq 1$.

Each expert is aware of his own competence, $t_i \in \{\mathcal{C}, \mathcal{I}\}$, while all the others share the same prior beliefs that expert i is competent with probability θ, which is identical ex ante for all experts.[3] The parameter θ plays a key role in the analysis and can be considered as the share of competent experts in the population of experts. Although experts' abilities are only privately known, signals' qualities are common knowledge.

As a member of a committee an expert is asked to give his recommendation, i.e., to vote, $m_i \in \{\overline{h}, \overline{l}\}$, simultaneously or sequentially; i.e., an expert may cast his vote privately or in front of other experts. Experts' votes have a dual purpose in our setting: first, the information that is accumulated by the decision maker is based on the votes, and second, the utility of the experts is assumed to be derived from the votes they provide and the (eventually) revealed state of the world. I assume that each expert has a utility of $u_i(\overline{h}, \omega_h) = u_i(\overline{l}, \omega_l) = 1$, i.e., his utility is positive if his vote coincides with the revealed state of the world, and a utility of $u_i(\overline{h}, \omega_l) = u_i(\overline{l}, \omega_h) = -1$ otherwise.[4] Thus, experts are interested in voting in such a way that their vote will coincide with the true (unknown) state of the world.

I refer to expert i's voting decision as his *strategy*, $\sigma_i : \{h, l\} \rightarrow \Delta\{\overline{h}, \overline{l}\}$. The analysis will focus on two important strategies: when experts vote ac-

[3] If the beliefs over the different experts are not identical, under a sequential procedure the question of optimal order arises, as in Ottaviani and Sorensen (2001).

[4] In a previous version of the paper I included a section in which experts had career concerns: future employers who evaluate the experts' competency based on the votes of the experts and the revealed state of the world, as in Scharfstein and Stein (1990). In Section 6, i discuss such a possibility.

cording to their signals and when experts disregard their signals and follow the votes of preceding experts.[5]

DEFINITION **1.** *An expert votes* ***honestly*** *if:*

$$\sigma_i = \begin{cases} \overline{h} & if \quad s_i = h \\ \bar{l} & if \quad s_i = l. \end{cases}$$

DEFINITION **2.** *An expert* ***cascades*** *if:*

$$\sigma_i = \begin{cases} \overline{h} & if \quad m_{i-1} = h \\ \bar{l} & if \quad m_{i-1} = l. \end{cases}$$

where $i-1$ is the expert who precedes expert i.[6]

Before taking an action, the decision maker aggregates experts' votes and updates his posterior probability over the state of the world. In doing so, he takes into account the behavior of experts and his beliefs over the experts' competency. Since experts' behavior is different under simultaneous and sequential procedures, the decision maker is faced with the problem of which procedure is expected to yield more accurate information about the state of the world. As the next proposition suggests, this depends non-monotonically on the prevalence of competent experts, θ.

Proposition 1. *There exist $\underline{\theta}(g,b)$ and $\overline{\theta}(g,b)$ such that:*

- *if $\theta \in [\underline{\theta}, \overline{\theta}]$ then $E[u_{DM}(simultaneous)] < E[u_{DM}(sequential)]$.*
- *if $\theta \notin [\underline{\theta}, \overline{\theta}]$ then $E[u_{DM}(sequential)] < E[u_{DM}(simultaneous)]$.*

When competent experts are very rare or very common, a simultaneous procedure should lead to more correct decisions than a sequential procedure. As will be clarified below, under a sequential procedure incompetent experts always cascade while competent experts cascade only under certain circumstances. The intuition of the proposition is that under a sequential procedure it may be possible to identify competent experts, which results in more accurate information. When competent experts are rare the probability of identifying the competent experts is low and does not offset the loss

[5] I do not restrict experts' strategies to only these two possibilities. However, as will be seen below, only these two strategies take place in equilibrium.

[6] As one can easily see, a cascade may arise only under a sequential procedure.

of information from incompetent experts' cascading behavior. On the other hand, when competent experts are common then they too cascade, which results in loss of information.

To prove the proposition, one should first analyze the behavior of experts under simultaneous and sequential procedures and the information that the decision maker can aggregate from the experts' votes.

3 Simultaneous Procedure

When experts vote simultaneously, their private signal is the only information they possess. Therefore, a vote by an expert is based solely on his signal. Since considerations of strategy are identical for all experts $\{1, ..., n\}$, I will dispense with the name index i, and analyze the strategy for each possible type and signal.

After observing his signal each expert updates his posterior over the state of the world using Bayes' rule, and so the posterior probabilities are given by:

- $Pr(\omega_h|h) = Pr(\omega_l|l) = p$
- $Pr(\omega_h|l) = Pr(\omega_l|h) = 1 - p$.

Because the signals of both types are (weakly) informative, i.e., $\frac{1}{2} \leq b < g$, the more likely state is the one in accordance with the expert's signal, regardless of his type. Thus, an expert who is interested in having his vote coincide with the unknown state of the world should follow his signal. This result is summarized in the following lemma.

Lemma 1. *Under a simultaneous procedure, in a unique equilibrium all experts vote honestly regardless of their type.*[7]

[7]The equilibrium is unique only when $\frac{1}{2} < b$; when $\frac{1}{2} = b$ any voting strategy of incompetent experts can be supported in equilibrium, and so the above equilibrium is not unique.

Proof. See Appendix □

The intuition of the lemma is simple: since the only information an expert possesses is his own signal, voting according to the signal maximizes the probability of being correct about the state of the world. Hence, all the experts follow their signals and vote *honestly*.

3.1 Simultaneous Procedure: Information Aggregation

When aggregating information the decision maker uses the insight of Lemma 1. Since all experts vote *honestly* i.e., according to their signal, the decision maker can refer to each vote as the signal itself. However, the decision maker is not aware of the competency of the experts. To elicit information from an individual vote, the decision maker uses the probability that it was given by a competent expert or by an incompetent expert. Let $z = \theta g + (1-\theta)b$ denote the informativeness of an individual vote, $Pr(\omega_j|m_i = \overline{j})$, that was given by an expert of an unknown type. One can refer to z as the expected correctness of a vote made by an expert of an unknown type. It is straightforward to see that the aggregate information is a Bayesian updating using the votes and z. Let $m = (m_1, ..., m_n)$ denote a voting profile, $\#_h(m)$ the number of $\overline{h}$ in m and $\#_l(m)$ the number of $\overline{l}$ in m. Assuming $\#_h(m) = k$ and $\#_l(m) = N-k$, using Bayes' rule we get:

$$Pr(\omega_h|m) = \frac{z^k(1-z)^{N-k}}{z^k(1-z)^{N-k} + (1-z)^k z^{N-k}}.$$

Since $\frac{1}{2} < z$ we get $\frac{1}{2} < Pr(\omega_h|m)$ if and only if $\frac{N}{2} < k$; in other words, for the decision maker to maximize his utility he must follow the majority.

The drawback of a simultaneous procedure is the disability to discern between the different types of experts. Thus, the information arise from the committee, as well as the corresponding decision, depends only on the difference between the number of experts who voted $\overline{h}$ and the number of experts who voted $\overline{l}$ as expressed in the following lemma.

Lemma 2. *Let m be a voting profile such that* $\#_h(m) - \#_l(m) = d > 0$, *then*

$$Pr(\omega_h|m) = \frac{z^d}{z^d + (1-z)^d}$$

Proof. See Appendix □

The lemma suggests that if there are d more experts that voted $\overline{h}$ than experts that voted $\overline{l}$, the aggregate information is equivalent to the information arising from a committee of d experts, with a unanimous vote of $\overline{h}$, and vice versa if there are more $\overline{l}$ votes than $\overline{h}$ votes. The important message is that under simultaneous procedures the decision may be based on a small number of experts, smaller than the committee's size.

4 Sequential Procedure

Under sequential procedures each expert is called upon to vote in the presence of all other experts. For convenience I will assume that the voting order is done according to the experts' names, i.e., expert 1 vote first, expert 2, second and so forth. When experts vote sequentially, subsequent experts in the sequence may acquire some additional information by observing the votes of preceding experts. Therefore, when a given expert considers his voting, in addition to his private signal he may use this additional information. Denote by ϕ_i the additional information for expert i, excluding his private signal. This information contains the votes of preceding experts, i.e., $\phi_1 = \phi$ (empty set), $\phi_2 = (m_1)$,..., $\phi_n = (m_1, ..., m_{n-1})$.

After observing his signal, s_i, and the additional information ϕ_i, an expert updates his posterior probability over the states of the world in order to determine his voting. Denote by $\alpha_i = Pr(\omega_h|s_i, \phi_i)$ the probability that an expert with signal s_i and information ϕ_i attributes to the state ω_h. The additional information may affect the expert's perception significantly: his posterior beliefs over the state of the world may not be the same if he observes only his own signal. This effect is not the same for the different types of experts because of the differences in signal quality.

To introduce the equilibrium under a sequential procedure, we define the following voting strategies:

$$m_i^{\mathcal{C}}(s_i, \phi_i) = \begin{cases} \overline{h} & if \;\; s_i = h \\ \overline{l} & if \; s_i = l \end{cases}$$

$$m_i^{\mathcal{I}}(s_i, \phi_i) = \begin{cases} \overline{h} & if \;\; \frac{1}{2} < Pr(\omega_h|\phi_i) \; or \; \phi_i = \phi \; \& \; s_i = h \\ \overline{l} & if \; Pr(\omega_h|\phi_i) < \frac{1}{2} \; or \; \phi_i = \phi \; \& \; s_i = l \end{cases}$$

While in the first voting strategy, $m_i^{\mathcal{C}}$, the vote relies solely on the signal, in the second voting strategy, $m_i^{\mathcal{I}}$, the vote relies solely on the signal only if no other information exists. Whenever additional information exists it determines the vote $m_i^{\mathcal{I}}(s_i, \phi_i)$.

In equilibrium each type of expert uses one of those two strategies, as expressed in the following proposition.

Proposition 2. *If* $\frac{z}{1-z} \cdot \frac{\widetilde{z}}{1-\widetilde{z}} < \frac{g}{1-g}$, *then in a unique equilibrium competent experts vote according to* $m_i^{\mathcal{C}}(s_i, \phi_i)$ *and incompetent experts vote according to* $m_i^{\mathcal{I}}(s^i, \phi_i)$, *where* $z = b + \theta(g - b)$ *and* $\widetilde{z} = \frac{1+\theta g - \theta}{2-\theta}$.

Proof. See Appendix □

Under a sequential procedure incompetent experts *cascade* while competent experts vote *honestly*, i.e., follow their signals, if the condition in Proposition 2 holds. The condition requires the signal of each competent expert to be sufficiently accurate. The intuition for the different behavior of the different types is the different effect of the additional information, ϕ_i, on an expert's perception of the more likely state.

Incompetent experts rely on any additional information given by a more informed expert (on average). Even one vote that contradicts an incompetent expert's signal is enough to change his perception of the more likely state.

Lemma 3. *Let $\phi_i = (m_1)$ be the additional information observed by an incompetent expert, then for any signal s_i:* $\begin{cases} \frac{1}{2} < \alpha_i & if\ \phi_i = \overline{h} \\ \alpha_i < \frac{1}{2} & if\ \phi_i = \overline{l} \end{cases}$

Proof. See Appendix □

As noted above, an incompetent expert perceives himself as being less informed than any of the other experts. Thus, an incompetent expert will cascade whenever he observes another expert's vote. On the other hand, the perception of a competent expert of the more likely state is according to his own signal, even when it is at odds with all the other votes.[8]

Lemma 4. *If $\frac{z}{1-z} \cdot \frac{\widetilde{z}}{1-\widetilde{z}} < \frac{g}{1-g}$ then for any competent expert, i, with signal s_i and information ϕ_i, then:* $\begin{cases} \frac{1}{2} < \alpha_i & if\ s_i = h \\ \alpha_i < \frac{1}{2} & if\ s_i = l \end{cases}$

Proof. See Appendix □

If a competent expert's signal is accurate enough, then not only does he perceive himself as possessing better information than each of the preceding experts, but he also perceives his information as more accurate than the aggregate information arising from the votes of the preceding experts.

Alternatively, the condition may be considered in terms of the prevalence of competent experts, θ, as expressed in the following lemma.

Lemma 5. *There exists $\overline{\theta}$ such that $\theta < \overline{\theta} \Leftrightarrow \frac{z}{1-z} \cdot \frac{\widetilde{z}}{1-\widetilde{z}} < \frac{g}{1-g}$.*

Proof. See Appendix □

Since experts are conscious of their ability, when θ is low a competent expert evaluates himself as possessing better information on average than the other two experts who are less informed.[9] Thus, in this case a competent

[8] This holds even for a committee of more than 3 experts. Note that the condition needs to be change to $\frac{z}{1-z}(\frac{\widetilde{z}}{1-\widetilde{z}})^{n-2} < \frac{g}{1-g}$, which is always true for g close to 1.

[9] This can be easily extended to a committee of more than 3 experts.

expert will rely on his own signal. On the other hand, when $\overline{\theta} < \theta$ the other two experts are likely to be competent experts, and so the third expert will evaluate the aggregate information arising from the votes of the first two experts as being better than his own signal, even though each of them is on average less informed.[10] In other words, when a competent expert observes two votes that contradict his signal, the more likely state is according to these votes and not according to his signal. Hence, the equilibrium above no longer holds. In equilibrium a competent expert cascades if there are enough votes that contradict his signal.

Proposition 3. *If $\overline{\theta} < \theta$, then $m_3 = m_2$ whenever $\phi_3 = (m_1, m_2)$ and $m_1 = m_2$.*
Otherwise, $m_i = \begin{cases} m_i^{\mathcal{C}}(s_i, \phi_i) & if \ t_i = \mathcal{C} \\ m_i^{\mathcal{I}}(s^i, \phi_i) & if \ t_i = \mathcal{I} \end{cases}$

Proof. See Appendix □

When an expert updates his posterior probability α_i using preceding votes ϕ_i, he evaluates each vote and the probability that it was given by each type. Since in equilibrium incompetent experts cascade, a vote that agrees with the vote that precedes it is likely to be given by an incompetent expert. On the other hand, a vote that contradicts the vote that precedes it is given by a competent expert for sure since incompetent experts always cascade in equilibrium. Therefore, the probability that a vote is made by a competent expert is:

$$Pr(t_i = \mathcal{C}) = \begin{cases} 1 & if \quad m_i \neq m_{i-1} \\ \frac{\theta}{2-\theta} & otherwise \end{cases}$$

The key point here is that the cascading behavior of incompetent experts serves to identify a competent expert as anyone who contradicts the vote that precedes his.

Corollary 1. (Identifying competent experts) *Under a sequential procedure it may be possible to identify competent experts, but it is not possible to*

[10]This holds only when $g < 1$. For $g = 1$, a competent expert should always follow his signal.

identify incompetent experts.

Although it is possible to identify competent experts since in equilibrium they are the only ones who contradict preceding votes, it is impossible to identify incompetent experts. When a vote agrees with the vote that precedes it, it is still possible that it was given by competent expert who observed a corresponding signal.

4.1 Sequential Procedure: Information Aggregation

When aggregating information from a committee, the decision maker uses the different behavior of the possible types to evaluate the information arising from each vote. As was noted above, when a vote contradicts the vote that precedes it, the decision maker can deduce that it was given by a competent expert. On the other hand, when a vote agrees with the vote that precedes it, the probability that it was given by a competent expert is not θ any more but $\frac{\theta}{2-\theta}$.

When competent experts vote *honestly* while incompetent experts *cascade*, the information arising from a vote is:

$$Pr(\omega_h|m_i=\overline{h}) = \begin{cases} z=\theta g+(1-\theta)b & if\ i=1 \\ \widetilde{z}=\frac{1-\theta(1-g)}{2-\theta} & if\ m_i=m_{i-1}\ and\ i\neq 1 \\ 1 & if\ m_i\neq m_{i-1}\ and\ i\neq 1. \end{cases}$$

When all experts *cascade*, regardless of their type, no information can be inferred from the third vote.[11]

The aggregate information for the decision maker is a bayesian updating over the states, given the votes and experts' equilibrium behavior. That means that the information of a vote is also determined by the vote that precedes it. A voting profile that contains a vote that contradicts a preceding vote reveals a competent expert and hence this vote is assigned greater weight in the aggregate information.

[11] For a committee of N experts there exist $3\leq k\leq N$ such that only after k identical votes will s competent expert cascade.

Corollary 2. (Wisdom of the minority) *Under a sequential procedure, it may be better for the decision maker to follow the minority.*

Proof. See Appendix □

To see this, consider the voting profile $m = (m_1 = \bar{l}, m_2 = \bar{l}, m_3 = \overline{h})$. As noted above, the decision maker can deduce that the third expert is competent since he did not cascade. Since a signal of a competent expert is quite accurate, it overweighs the other two votes that were given by experts of an unknown type.

5 Simultaneous vs. Sequential: An Ex Ante Decision

Up to now I analyzed experts' behavior under simultaneous and sequential voting procedures and the posterior updating over the state of the world by a decision maker after observing a voting profile. To choose the preferred procedure, $x \in \{Sim, \ Seq\}$, it must be asked which procedure is *expected* to yield more information for the decision maker ex ante. Clearly, the choice of procedure is made before experts vote. Hence, in order to determine the expected information from each procedure, the decision maker must evaluate all possible voting profiles, $\mathcal{M} = \{m | m = (m_1, m_2, m_3)\}$, the information he can infer from each profile, and the probability of each profile appearing.

Denote by $V(m, x) = Max_\omega Pr(\omega | m, x)$ the informativeness of a voting profile $m \in \mathcal{M}$ under the procedure x and the equilibrium behavior of the experts under this procedure. $V(m, x)$ is the posterior probability of the more likely state according to the Bayesian updating after observing the voting profile m under the procedure x. Alternatively, $V(m, x)$ can be considered as the probability of being correct about the state when following the aggregate recommendation arising from the profile m. Note that the aggregate recommendation may not be that of the majority, as pointed out above.

Denote by $Q(m, x) = Pr(m \text{ to be realized}|x)$ the probability of a voting profile m appearing when procedure x is implemented. Since the decision maker can anticipate the experts' behavior under each profile, he can assign a probability to each possible vote to be observed, and hence the probability for each voting profile to be realized. For example, under a sequential procedure, if the first vote is $\overline{h}$, then the second vote is very likely to be $\overline{h}$ rather than $\overline{l}$ since the latter can be given only by competent experts.

Knowing the information that can be inferred from each profile and the probability of this profile allows us to calculate the expected information under each procedure, denoted by $\Phi(x) = \sum_m V(m, x)Q(m, x)$. The decision maker will choose the procedure that maximizes $\Phi(x)$. As clarified above, under a sequential procedure incompetent experts cascade, and so when competent experts vote *honestly* it may be possible to identify competent experts. Although the information of the incompetent experts is lost, there is a gain from identifying the competent experts. Thus, a sequential procedure will be preferred by the decision maker only when there is such a gain of information. Nevertheless, when competent experts also cascade, no information can be obtained from a vote. The key parameter for this decision is θ as will be shown below.

I start with the case where competent experts may cascade. In our setting the third expert will cascade regardless of his type if $\overline{\theta} < \theta$. The next lemma suggests that in this case a simultaneous procedure yields more information than a sequential procedure.

Lemma 6. *If $\overline{\theta} < \theta$ then $\Phi(Seq) < \Phi(Sim)$.*

Proof. See Appendix □

Although under a sequential procedure it is still possible to identify competent experts that do not cascade, there is a big loss of information because of the cascade of the third expert. Moreover, under a simultaneous procedure each vote is quite likely to be given by a competent expert since θ is high.

The more interesting case is $\theta < \overline{\theta}$ since under a sequential procedure competent experts always vote *honestly.* On the other hand, incompetent experts always *cascade* and so the information they possess is lost. A sequential procedure yields more information only when the information arising from the identification of the competent experts outweighs the loss of information from the cascade of the incompetent experts. When competent experts are rare, i.e., θ is low, the chance of identifying a competent experts is low; that is, the chance of gaining additional information is low and hence would not outweigh the loss of information caused by the cascading of the incompetent experts. The next lemma clarifies when a sequential procedure is better than a simultaneous procedure, and vice versa.

Lemma 7. *There exist* $[\underline{\theta}, \overline{\theta}]$ *such that:*

- *if* $\theta \in [\underline{\theta}, \overline{\theta}] \;\Rightarrow\; \Phi(Sim) < \Phi(Seq)$
- *if* $\theta \notin [\underline{\theta}, \overline{\theta}] \;\Rightarrow\; \Phi(Seq) < \Phi(Sim)$

Moreover, $[\underline{\theta}, \overline{\theta}]$ *is not empty if* $b \leq \frac{3}{4}g$*.*

Proof. See Appendix □

When θ is relatively small, i.e., competent experts are rare, the chance of identifying competent experts is low while the information of incompetent experts, who are common, is lost because of their cascade.[12] A simultaneous procedure yields more information since the common incompetent experts vote *honestly.* On the other hand, when competent experts are very common, i.e., θ is relatively high, there is not much gain from identifying competent experts since a random expert (as in a simultaneous procedure) is very likely to be competent, and there is still a loss of information because of the cascade of incompetent experts.

For a sequential procedure to yield more information than a simultaneous procedure, there are to requests: first, competent experts must be moderately

[12] If $b = \frac{1}{2}$, i.e., incompetent experts are not informed at all, there is no loss of information from a cascade of incompetent experts.

prevalent, and second, competent experts must possess significantly better information than incompetent experts. The latter requirement assures that when identifying a competent expert the decision maker will gain valuable information. If competent experts do not significantly differ from incompetent experts, identifying them is worthless.

6 Discussion

I conclude by discussing some assumptions and extensions of the model, which illustrate the robustness of the results, and introduce some ideas for future work.

6.1 Committees of many experts

The main results of this paper can be extended to a committee of any size; that is, a sequential procedure may yield more information than a simultaneous procedure in a committee of any size. As one can observe, the cascade of an incompetent expert does not depend on the size of the committee. The condition under which a competent expert votes honestly must be modified as a function of the size of the committee. Nevertheless, it is easy to see that for $g = 1$ a competent expert always votes honestly. Thus, if the signal of a competent expert differ significantly from the signal of an incompetent expert, a sequential procedure yields more information than a simultaneous procedure, as expressed in the following proposition.

Proposition 4. *If* $g = 1,\ b = \frac{1}{2}$ *then for any size of committee* N*,* $\Phi(Sim) \leq \Phi(Seq)$*.*

Proof. See Appendix □

6.2 Career concerns

I assumed that experts care only about providing a vote that will coincide with the revealed state of the world. A more general case is experts who care for their reputation; i.e., an expert wishes to appear competent to others. In such a case the strategies by which an incompetent expert always cascades while a competent expert votes honestly do not constitute an equilibrium, the reason being that an incompetent expert will deviate and contradict the vote of the expert preceding him in order to appear competent, since contradicting the preceding vote implies that the expert is competent.

Indeed, if experts care about appearing well informed, in equilibrium some anti-herding behavior should arise among both types; cf. Levy (2004). In equilibrium an incompetent expert will balance between being correct about the true state, which is more likely when he follows preceding votes, and appearing as well informed by contradicting the preceding vote, which is more likely to be done by a competent expert. A competent expert on the other hand will contradict a preceding vote for sure if he observed the corresponding signal. Although a clear identification of competent experts is not possible in this case, the experts' behavior implies for their competency. Thus, under a sequential procedure different votes are assigned different weights in the aggregate information and a sequential procedure may result in better aggregate information.

6.3 Discussion of other assumptions

In the model I made several simplifying assumptions, most of which are not important. The assumptions of binary states and binary signals do not significantly matter for the main results; using alternative assumptions would only slightly modify the results. I assumed that the probability of the states is a priori the same. The results continue to hold for small perturbations, i.e., when the states are close to being balanced. If one state is way more likely, e.g., $Pr(\omega_h) = \alpha >> \frac{1}{2}$, and the signal of an incompetent expert is hardly informative, then an incompetent expert will vote according to the more likely state even under a simultaneous procedure. Therefore, identifying competent experts is also possible under a simultaneous procedure.

I have considered the case with only two possible types. The case with a multiple type space should lead to similar results. Equilibrium behavior will change to a threshold strategy whereby an expert cascades or votes honestly if his signal accuracy is below/above some threshold. The threshold differs for the different types and allows us to identify some of them.

I assumed that experts care only about giving the “correct” vote, but do not care about the final decision, nor whether their own vote is implemented. This assumption gives rise to the phenomenon of an information cascade. If experts also care about the final decision to some extent, the results hold as long as an information cascade exists. On the other hand, assuming that experts care only about the final decision should be an interesting question for future work.

7 Concluding Remarks

Many decisions are taken after consultation in committees. In this paper I analyzed consultation with a committee of experts, the incentives of heterogeneous experts to advise according to their private information, and the information that can be accumulated in committees with sequential and simultaneous orders of speaking.

Social learning literature has demonstrated the disadvantage of an information cascade arising from a sequential procedure. As a result, organizations such as the FDA changed their guidelines and moved from sequential procedures to simultaneous ones. I demonstrate how an asymmetric information cascade may be beneficial when the intention is to increase the aggregation of information, and hence the advantage of sequential over simultaneous procedures when competent experts are neither too rare nor too common.

Under a sequential order of speaking experts are exposed to additional information other than their private signal. I show that the effect of this additional information is not symmetric for competent and incompetent experts. While for incompetent experts this additional information results in cascading from the very beginning, competent experts may cascade if their information is not sufficiently accurate. In this sense, incompetent experts are more prone to cascade than competent experts. As a result, when aggregating the information from a committee of experts, a decision maker can use the different behavior of the experts to elicit their competency, and hence assign better weights to their votes in the aggregate information. It is the cascading of incompetent experts that allows for the identification of competent experts. To my knowledge, this is the first paper to shed light on the good side of an information cascade.

References

[1] Ali, S. N., Goeree, J. K., Kartik, N., Palfrey, T. R. (2008). Information aggregation in standing and ad hoc committees, *American Economic Review* 98(2), 181-186.

[2] Austen-Smith, D., Banks, J. S. (1996). Information aggregation, rationality, and the Condorcet jury theorem. *American Political Science Review*, 34-45.

[3] Banerjee, A.V., 1992. A simple model of herd behavior. *Quarterly Journal of Economics* 107 (3), 797817.

[4] Bikhchandani, S., Hirshleifer, D., Welch, I., 1992. A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy*, 100 (5), 992-1026.

[5] Callander, S., 2007. Bandwagons and momentum in sequential voting. *Review of Economic Studies* 74, 653-684.

[6] Callander, S., and Hrner, J. "The wisdom of the minority." Journal of Economic theory 144.4 (2009): 1421-1439.

[7] Condorcet, M., 1785. Essai sur lápplication de lánalyse à la probabilité des décisions rendues à la pluralité des voix. In: Baker, K.M. (Ed.), *Condorcet: Selected Writings*. Bobbs-Marrill, Indianapolis, pp. 3370, Translation 1976.

[8] Dekel, E., Piccione, M., 2000. Sequential voting procedures in symmetric binary elections. *Journal of Political Economy* 108 (1), 3455.

[9] FDA Guidance for FDA Advisory Committee Members and FDA Staff: Voting Procedure for Advisory Committee Meetings. U.S. Food and Drug Administration, 2008. Availible at: *http://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM125641.pdf*

[10] Gerling, Kerstin, Hans Peter Grner, Alexandra Kiel and Elisabeth Schulte, 2004, Information Acquisition and Decision Making in Committees: A Survey, European *Journal of Political Economics*, 21 (3), 563-597.

[11] Gerardi, D., Yariv, L. (2007). Deliberative voting. *Journal of Economic Theory*, 134(1), 317-338.

[12] Gerardi, D., Yariv, L. (2008). Information acquisition in committees. *Games and Economic Behavior*, 62(2), 436-459.

[13] Gershkov, A., Szentes, B. (2009) "Optimal voting schemes with costly information acquisition." *Journal of Economic Theory* 144.1, 36-68.

[14] Glazer, J., Rubinstein, A., 1998. Motives and implementation: On the design of mechanisms to elicit opinions. *Journal of Economic Theory* 79 (2), 157173.

1999. *Review of Economic Studies* 66 (1), 169182.

[15] Levy, G. (2004). Anti-herding and strategic consultation. *European Economic Review*, 48(3), 503-525.

[16] Levy, G. (2007) Decision making in committees: transparency, reputation and voting rules. *American Economic Review* 97:150168

[17] Ottaviani, M. and P. Srensen (2001), Information Aggregation in Debate: Who Should Speak First?, *Journal of Public Economics*, pp. 393-421.

[18] Scharfstein, D., Stein, J., 1990. Herd behavior and investment. *American Economic Review* 80 (3), 465479.

[19] Smith, L., Srensen, P., 2000. "Pathological outcomes of observational learning." *Econometrica* 68.2 ,371-398.

[20] Urfalino, P., Costa, P. (2010). Public and Oral Voting in FDA's Advisory Committees, Working paper.

[21] Visser, B., Swank, O. H., 2007. On committees of experts. *Quarterly Journal of Economics* 122, 337-372.

8 Appendix A

Proof of Lemma 1. As noted, the only information an expert possesses is his private signal. Thus, the posterior probability of an expert is given by:

$$\begin{cases} Pr(\omega_h|h) = Pr(\omega_l|l) = p \\ Pr(\omega_h|l) = Pr(\omega_l|h) = 1-p. \end{cases}$$

Therefore, when voting according to his signal, an expert's expected utility is:

$$E[u_i(m_i = \overline{s_i})|s_i] = 1 \cdot p + (-1)(1-p) = 2p - 1. \tag{1}$$

On the other hand, when voting in opposition to his signal, his expected utility is:

$$E[u_i(m_i \neq \overline{s_i})|s_i] = 1 \cdot (1-p) + (-1)p = 1 - 2p. \tag{2}$$

Since $\frac{1}{2} \leq p$, for any p the expected utility in (1) is greater than the expected utility in (1).
Note that when $b = \frac{1}{2}$ (1) is weakly greater than (2), and so the equilibrium is not unique. ■

Proof of Lemma 2. Let d denote the difference between the number of each possible vote. Let m be a voting profile and assume $\#_h(m) = k$ and $\#_l(m) = N - k$ such that $\frac{N}{2} < k$; i.e., there are more votes by $\overline{h}$ than $\overline{l}$. Let $\#_h(m) - \#_l(m) = d$, and note that $d = N - 2k$.
Given m the decision maker updates his posterior probability over the state:

$$Pr(\omega_h|\underbrace{\overline{l}, ..., \overline{l}}_{k\,times}, \underbrace{\overline{h}, ..., \overline{h}}_{N-k\,times}) \tag{3}$$

Since all experts vote according to their signal, a vote is a one-to-one mapping from votes to signals. Hence we get:

$$Pr(\omega_h|\underbrace{\overline{l}, ..., \overline{l}}_{k\,times}, \underbrace{\overline{h}, ..., \overline{h}}_{N-k\,times}) = Pr(\omega_h|\underbrace{l, ..., l}_{k\,times}, \underbrace{h, ..., h}_{N-k\,times}).$$

Since each signal is of average informativeness, by Bayes' rule we get:

$$
\begin{aligned}
Pr(\omega_h | \underbrace{l, ..., l}_{k\,times}, \underbrace{h, ..., h}_{N-k\,times}) &= \frac{z^{N-k}(1-z)^k}{z^{N-k}(1-z)^k + (1-z)^{N-k}z^k} \\
&= \frac{z^{N-2k}}{z^{N-2k} + (1-z)^{N-2k}} \\
&= \frac{z^d}{z^d + (1-z)^d} = Pr(\underbrace{h, ..., h}_{d\,times}).
\end{aligned}
\tag{4}
$$

The case where $\#_h(m) < \#_l(m)$ is similar. ■

Proof of Proposition 2. In order to prove that $m_i^{\mathcal{C}}$ and $m_i^{\mathcal{I}}$ constitute an equilibrium we first should prove Lemmas 3 and 4.

Proof of Lemma 3. First, note that the first expert, 1, always votes according to his signal, as in a simultaneous procedure, since the only information he possesses is his signal.
Assume that $\phi_2 = (\overline{h})$; since 1 always votes according to his signal, when updating α_2 we can replace $\overline{h}$ by h.
If in addition $s_2 = h$, using Bayes' rule we get:

$$\alpha_2 = Pr(\omega_h | s_2 = h, m_1 = \overline{h}) = \frac{zb}{zb + (1-z)(1-b)}. \tag{5}$$

Since $\frac{1}{2} < b$ and $\frac{1}{2} < z$ we get $\frac{1}{2} < \alpha_2$.
On the other hand, if $s_2 = l$ we get:

$$\alpha_2 = Pr(\omega_h | s_2 = l, m_1 = \overline{h}) = \frac{z(1-b)}{z(1-b) + (1-z)b}. \tag{6}$$

From (6) we get:

$$\alpha_2 < \frac{1}{2} \;\Leftrightarrow\; z > b. \tag{7}$$

Since $0 \leq \theta \leq 1$ we get $z > b$ for all g and b.
The case when $\phi_2 = \overline{l}$ is similar. ■

Proof of Lemma 4. Clearly, if ϕ_i contain only votes that point to the same state as s_i, we get $\begin{cases} \frac{1}{2} < \alpha_i & if s_i = h \\ \alpha_i < \frac{1}{2} & if s_i = l \end{cases}$ since $\frac{1}{2} \leq b < g \leq 1$.
Therefore, ϕ_i should affect α_i only if it contains some contradiction to s_i.
First, note that if the lemma holds for $i = 3$ it also holds for $i < 3$. The reason is that the third expert observes more information than the others, and so may observe more votes that contradict his own signal. Formally,

$$\begin{aligned} Pr(\omega_h|s_3 = h, \phi_3 = (\bar{l}, \bar{l}), t_3 = \mathcal{C}) &< Pr(\omega_h|s_2 = h, \phi_2 = (\bar{l}), t_2 = \mathcal{C}) \\ &< Pr(\omega_h|s_1 = h, t_1 = \mathcal{C}) \end{aligned}$$

and similarly for ω_l, $s_i = l$, etc.
In addition, it is easy to see that $Pr(\omega_h|s_3 = h, \phi_3 = (\bar{l}, \bar{l}) < Pr(\omega_h|s_3 = h, \phi_3 = (\bar{l}, \bar{h})$. ; i.e., the more votes there are that contradict an expert's signal, the more the expert will be affected.
Hence, it is sufficient to prove the lemma for $i = 3$ while assuming it holds for $i = 2$.
Assume w.l.o.g. that $s_3 = h$ and $\phi_3 = (\bar{l}, \bar{l})$. Since, as noted above, expert 1 always votes according to his signal, we get:

$$Pr(\omega_h|m_1 = \bar{l}) = 1 - \theta g - (1 - \theta)b. \tag{8}$$

Since expert 2 cascades if incompetent (according to Lemma 3) but votes honestly if competent (according to the assumption of the lemma), we get:

$$Pr(\omega_h|m_2 = \bar{l}, m_1 = m_2) = 1 - \frac{1 + g\theta - \theta}{2 - \theta}. \tag{9}$$

Updating over the state of the world we get:

$$\begin{aligned} \alpha_3 &= Pr(\omega_h|s_3 = h, m_1 = \bar{l}, m_2 = \bar{l}) \\ &= \frac{(1 - \theta g - (1 - \theta)b)(1 - \frac{1+g\theta-\theta}{2-\theta})g}{(1 - \theta g - (1 - \theta)b)(1 - \frac{1+g\theta-\theta}{2-\theta})g + (\theta g + (1 - \theta)b)\frac{1+g\theta-\theta}{2-\theta}(1 - g)}. \end{aligned} \tag{10}$$

Applying $z = \theta g - (1-\theta)b$ and $\widetilde{z} = \frac{1+g\theta-\theta}{2-\theta}$ to (10). we get:

$$\alpha_3 = \frac{(1-z)(1-\widetilde{z})g}{(1-z)(1-\widetilde{z})g + z\widetilde{z}(1-g)}. \tag{11}$$

From (11) we can get:

$$\frac{1}{2} < \alpha_3 \iff \frac{z\widetilde{z}}{(1-z)(1-\widetilde{z})} < \frac{g}{1-g}. \tag{12}$$

The case when $\alpha_3 < \frac{1}{2}$, $s_3 = l$ is similar. ■

Proof of Lemma 5.

$$\begin{aligned} &\tfrac{z\widetilde{z}}{(1-z)(1-\widetilde{z})} < \tfrac{g}{1-g} \\ \Leftrightarrow\ &\tfrac{(\theta g-(1-\theta)b)(\frac{1+g\theta-\theta}{2-\theta})}{(\theta g-(1-\theta)b)(1-\frac{1+g\theta-\theta}{2-\theta})} < \tfrac{g}{1-g} \\ \Leftrightarrow\ &\tfrac{g^2+2gb-2g^2b-b}{(g-b)(2g-1)^2} - \sqrt{\tfrac{2gb-g^2-10g^2b+4g^2b^2+4g^3+12g^3b-8g^3b^2-4g^4b+4g^4b^2}{(g-b)^2(2g-1)^2}} < \theta. \end{aligned} \tag{13}$$

Define the left hand side of (13) as $\overline{\theta}$ so that we get $\overline{\theta} < \theta$.
One can verify that $0 < \overline{\theta} < 1$ since $\frac{1}{2} \leq b < g \leq 1$. ■

Proof of Proposition 2. In order to maximize his utility an expert should vote according to the more likely state since such a vote has a greater probability of coinciding with the revealed state of the world. Formally, assume that $\frac{1}{2} < \alpha_i$; then $E[u_i(m_i = \overline{l})] < E[u_i(m_i = \overline{h})]$, and vice versa for $\alpha_i < \frac{1}{2}$.
For expert 1 of any type, the more likely state is according to his signal, since $\frac{1}{2} \leq b < g$ and $\phi_1 = \phi$.
For expert $1 < i$, if he is incompetent, according to Lemma 3 the more likely state is according to ϕ_i but not necessarily to s_i. Hence, in order to maximize his utility an incompetent expert should cascade, as in $m_i^{\mathcal{I}}$.
For expert $1 < i$, if he is competent and the condition of Proposition 2 holds, according to Lemma 4 the more likely state is according to s_i. Hence, in order to maximize his utility a competent expert should vote honestly as in $m_i^{\mathcal{C}}$. ■

Proof of Proposition 3. As shown above, an expert will vote $m_i = \overline{h}$ when $\frac{1}{2} < \alpha_i$ and vice versa when $\alpha_i < \frac{1}{2}$.

I first show that $m_3 = m_2$ if $\phi_3 = (m_1, m_2)$ and $m_1 = m_2$.

Assume that $s_3 = h$ and $\phi_3 = (\bar{l}, \bar{l})$. Since $\theta < \overline{\theta}$, from Lemmas 4 and 5 we get $\alpha_3 < \frac{1}{2}$, and so $m_3 = \bar{l}$ will maximize the third expert's utility. The case when $s_3 = l$ and $\phi_3 = (\overline{h}, \overline{h})$ is identical.

Since Lemma 3 is still valid, for an incompetent expert $m_i^{\mathcal{I}}$ is optimal.

Although the condition of Lemma 4 does not hold (by assumption), a competent expert will vote according to his signal unless there are two votes that contradict his signal.

For expert 1, voting according to his signal is straightforward as been showed above.

For expert 2 with $s_2 = h$ and $\phi_2 = \bar{l}$, voting $m_2 = \overline{h}$ is optimal since using Bayes' rule we get:

$$\alpha_2 = \frac{(1-\theta g-(1-\theta)b)g}{(1-\theta g-(1-\theta)b)g+(\theta g+(1-\theta)b)(1-g)} > \frac{1}{2} \Leftrightarrow g > b. \tag{14}$$

which always holds. The case when $s_2 = l$ and $\phi_2 = \overline{h}$ is similar.

In a similar way for expert 3 with $s_3 = h$ and $\phi_3 = (\overline{h}, \bar{l})$, voting $m_3 = \overline{h}$ is optimal since:

$$\alpha_3 = \frac{(\theta g+(1-\theta)b)(1-g)g}{(\theta g+(1-\theta)b)(1-g)g+(1-\theta g+(1-\theta)b)g(1-g)} = \theta g+(1-\theta)b > \frac{1}{2}. \tag{15}$$

and similarly for when changing h and l. ■

Proof of Corollary 2. Consider the following voting profile, $m = (\bar{l}, \bar{l}, \overline{h})$. If the condition of Proposition 2 holds, then:

$$\begin{aligned} Pr(\omega_h|m) = & \frac{(1-\theta g-(1-\theta)b)(1-\frac{1+g\theta-\theta}{2-\theta})g}{(1-\theta g-(1-\theta)b)(1-\frac{1+g\theta-\theta}{2-\theta})g+\frac{1+g\theta-\theta}{2-\theta}(\theta g-(1-\theta)b)(1-g)} \\ & = \frac{(1-z)(1-\widetilde{z})g}{(1-z)(1-\widetilde{z})g+z\widetilde{z}(1-g)} \end{aligned} \tag{16}$$

Applying lemma 4 to (16) we get:

$$\frac{1}{2} < Pr(\omega_h|m). \tag{17}$$

Hence, following the minority is expected to yield higher utility for the decision maker. ■

Proof of Lemma 6. Since the two states are symmetric it is enough to analyze only voting profiles that point toward state ω_h. For a simultaneous voting procedure these voting profiles are $(\overline{h}, \overline{h}, \overline{h})$, $(\overline{l}, \overline{h}, \overline{h})$, $(\overline{h}, \overline{l}, \overline{h})$, $(\overline{h}, \overline{h}, \overline{l})$. The informativeness of each of these profile is:

$$V(\overline{h}, \overline{h}, \overline{h}) = \frac{z^3}{z^3 + (1-z)^3}. \tag{18}$$

$$V(\overline{l}, \overline{h}, \overline{h}) = V(\overline{h}, \overline{l}, \overline{h}) = V(\overline{h}, \overline{h}, \overline{l}) = \frac{z^2(1-z)}{z^2(1-z) + (1-z)^2 z} = z. \tag{19}$$

The probabilities of these profiles to be realized are

$$\begin{aligned} &P(\overline{h}, \overline{h}, \overline{h}) = Pr(\overline{h}, \overline{h}, \overline{h}|\omega_h)Pr(\omega_h) + Pr(\overline{h}, \overline{h}, \overline{h}|\omega_l)Pr(\omega_l) \\ &= \tfrac{1}{2}(z^3 + (1-z)^3). \end{aligned} \tag{20}$$

$$\begin{aligned} &P(\overline{l}, \overline{h}, \overline{h}) = P(\overline{h}, \overline{l}, \overline{h}) = P(\overline{h}, \overline{h}, \overline{h}) \\ &= Pr(\overline{l}, \overline{h}, \overline{h}|\omega_h)Pr(\omega_h) + Pr(\overline{l}, \overline{h}, \overline{h}|\omega_l)Pr(\omega_l) \\ &= \tfrac{1}{2}(z^2(1-z) + z(1-z)^2). \end{aligned} \tag{21}$$

Combining (18), (19), (20), and (21), we get:

$$\begin{aligned} &\Phi(Sim) = \textstyle\sum_m P(m)V(m) = \\ &= 2(\tfrac{z^3}{z^3+(1-z)^3}\tfrac{1}{2}(z^3 + (1-z)^3) + 3z\tfrac{1}{2}(z^2(1-z) + z(1-z)^2)) \\ &= 3z^2 - 2z^3. \end{aligned} \tag{22}$$

For a sequential voting procedure, since the third expert cascades, the possible voting profiles are $(\overline{h}, \overline{h}, \overline{h})$, $(\overline{l}, \overline{h}, \overline{h})$, $(\overline{h}, \overline{l}, \overline{h})$. Note that the profile $(\overline{l}, \overline{l}, \overline{h})$ has zero probability of emerging.

The informativeness of each of these profile is:

$$V(\overline{h},\overline{h},\overline{h}) = \frac{z\widetilde{z}^2}{z\widetilde{z}^2+(1-z)(1-\widetilde{z})^2}. \tag{23}$$

$$V(\overline{l},\overline{h},\overline{h}) = \frac{(1-z)g\widetilde{z}}{(1-z)g\widetilde{z}+z(1-g)(1-\widetilde{z})}. \tag{24}$$

$$V(\overline{h},\overline{l},\overline{h}) = \frac{z(1-g)g}{z(1-g)g+(1-z)g(1-g)} = z. \tag{25}$$

The probabilities of these profiles being realized are:

$$\begin{aligned} &P(\overline{h},\overline{h},\overline{h}) = \tfrac{1}{2}(z(\theta g+1-\theta)+(1-z)(\theta(1-g)+1-\theta)) \\ &= \tfrac{1}{2}(1-\theta z-\theta g(1-2z)). \end{aligned} \tag{26}$$

$$\begin{aligned} &P(\overline{l},\overline{h},\overline{h}) = \tfrac{1}{2}(1-z)\theta g(\theta g+1-\theta)+\tfrac{1}{2}z\theta(1-g)(\theta g+1-\theta) \\ &= \tfrac{1}{2}\theta(z+g-2zg-(1-g)\theta g). \end{aligned} \tag{27}$$

$$P(\overline{h},\overline{l},\overline{h}) = \tfrac{1}{2}z\theta(1-g)\theta g+\tfrac{1}{2}(1-z)\theta g\theta(1-g) = \tfrac{1}{2}(1-g)g\theta^2. \tag{28}$$

Combining (23)-(28) we get:

$$\begin{aligned} &\Phi(Seq) = \textstyle\sum_m P(m)V(m) = \\ &= 2(\tfrac{1}{2}((1-z)(1-\theta g)+z(1+\theta g-\theta))\tfrac{z\widetilde{z}^2}{z\widetilde{z}^2+(1-z)(1-\widetilde{z})^2} \\ &+\tfrac{1}{2}\theta(z+g-2zg-(1-g)\theta g)\tfrac{(1-z)g\widetilde{z}}{(1-z)g\widetilde{z}+z(1-g)(1-\widetilde{z})}+\tfrac{1}{2}(1-g)g\theta^2 z) \\ &= \tfrac{g\widetilde{z}(1-z)\theta(g+z-2gz-\theta g(1-g))}{z(1-g)+\widetilde{z}(g-z)}+\tfrac{z\widetilde{z}^2(1-\theta z+\theta g(2z-1))}{(1-\widetilde{z})^2+z(2\widetilde{z}-1)}+\theta^2 zg(1-g). \end{aligned} \tag{29}$$

Since $\frac{1}{2} < \widetilde{z} < z < g \leq 1$, one can verify that $\Phi(Seq) < \Phi(Sim)$. ■

Proof of Lemma 7. First note that $\Phi(Sim) = 3z^2 - 2z^3$, as in Lemma 6. On the other hand, $\Phi(Seq)$ is different from the one given in Lemma 6 since a competent expert always votes according to his signal.
Under a sequential voting procedure the voting profiles that point toward ω_h are $(\overline{h},\overline{h},\overline{h})$, $(\overline{l},\overline{h},\overline{h})$, $(\overline{h},\overline{l},\overline{h})$, $(\overline{l},\overline{l},\overline{h})$.

The informativeness of each of these profiles is:

$$V(\overline{h},\overline{h},\overline{h}) = \frac{z\widetilde{z}^2}{z\widetilde{z}^2 + (1-z)(1-\widetilde{z})^2}. \tag{30}$$

$$V(\bar{l},\overline{h},\overline{h}) = \frac{(1-z)g\widetilde{z}}{(1-z)g\widetilde{z} + z(1-g)(1-\widetilde{z})}. \tag{31}$$

$$V(\overline{h},\bar{l},\overline{h}) = \frac{z(1-g)g}{z(1-g)g + (1-z)g(1-g)} = z. \tag{32}$$

$$V(\bar{l},\bar{l},\overline{h}) = \frac{(1-z)(1-\widetilde{z})g}{(1-z)(1-\widetilde{z})g + z\widetilde{z}(1-g)}. \tag{33}$$

The probabilities of these profiles being realized are:

$$\begin{aligned} &P(\overline{h},\overline{h},\overline{h}) = \tfrac{1}{2}(z(\theta g + 1 - \theta)^2 + (1-z)(\theta(1-g) + 1 - \theta)^2) \\ &= \tfrac{1}{2}(1 - \theta(z(2-\theta) + 2g(1 - z(2-\theta)) - \theta g^2)). \end{aligned} \tag{34}$$

$$\begin{aligned} &P(\bar{l},\overline{h},\overline{h}) = \tfrac{1}{2}((1-z)\theta g(\theta g + 1 - \theta) + z\theta(1-g)(\theta(1-g) + 1 - \theta)) \\ &= \tfrac{1}{2}(\theta(g + z - 2gz - \theta g(1-g))). \end{aligned} \tag{35}$$

$$\begin{aligned} &P(\overline{h},\bar{l},\overline{h}) = \tfrac{1}{2}(z\theta(1-g)\theta g + (1-z)\theta g\theta(1-g)) \\ &= \tfrac{1}{2}(\theta^2 g(1-g)). \end{aligned} \tag{36}$$

$$\begin{aligned} &P(\bar{l},\bar{l},\overline{h}) = \tfrac{1}{2}((1-z)(\theta(1-g) + 1 - \theta)\theta g + z(\theta g + 1 - \theta)\theta(1-g)) \\ &= \tfrac{1}{2}(\theta(g(1-\theta g) + z(1-\theta) - 2gz(1-\theta))). \end{aligned} \tag{37}$$

In order to prove the lemma we will use the following lemmas:

Lemma 8. *$\Phi(Sem)$ and $\Phi(Seq)$ are both concave functions.*

Proof of Lemma 8. As shown above, $\Phi(Sem) = 3z^2 - 2z^3$; hence:

$$\frac{d\Phi(Sem)^2}{d^2\theta} = 6z''_\theta - 12z \cdot (z'_\theta)^2 - 6z^2 z''_\theta. \tag{38}$$

Since $z'_\theta = g - b$ and $z''_\theta = 0$, we get:

$$\frac{d\Phi(Sem)^2}{d^2\theta} = -12(\theta g + (1-\theta)b)(g-b)^2 < 0 \;\Leftrightarrow\; b < g. \tag{39}$$

Hence $\Phi(Sem)$ is a concave function.
In order to show that $\Phi(Seq)$ is concave I show that for each of the above voting profiles $P(m)V(m)$ is concave, and since the sum of concave functions is also concave, $\Phi(Seq)$ is concave.
First, for any voting profile:

$$(P(m)V(m))''_\theta = (P(m))''_\theta V(m) + 2(P(m))'_\theta (V(m))'_\theta + P(m)(V(m))''_\theta \tag{40}$$

and clearly, $0 < P(m)$ and $0 < V(m)$.
For $m = (\overline{h}, \overline{h}, \overline{h})$:

$$\begin{aligned}
V(\overline{h},\overline{h},\overline{h})'_\theta &= \frac{\left((1-2g)^2\theta^2-1\right)\left((2g-1)\theta^2(b-g)(4b-6g+1)+4(2b-1)(2g-1)\theta(g-b)+b(8(b-1)g-4b+5)-g\right)}{((2g-1)\theta(4b(\theta-1)-6g\theta+\theta+2)-1)^2} \\
V(\overline{h},\overline{h},\overline{h})''_\theta &= \frac{3(2b-1)(2g-1)\theta^2(2b-4g+1)(b-g)+b(4b(-2bg+b+3g-2)-2g+3)-g}{(1-(2g-1)\theta(4b(\theta-1)-6g\theta+\theta+2))^3} \\
&\quad - \frac{8(2g-1)\left(3\theta\left(b^3(8g-4)+b^2\left(4-12g^2\right)+b(2g(6g-5)+1)+g^2\right)\right)}{(1-(2g-1)\theta(4b(\theta-1)-6g\theta+\theta+2))^3} \\
&\quad - \frac{\left((2g-1)\theta^3\left(4b^3+4b^2(g+1)(g(2g-5)+1)+b(3-4(g-6)g(2g-1))+g((13-22g)g-2)\right)\right)}{(1-(2g-1)\theta(4b(\theta-1)-6g\theta+\theta+2))^3} \\
P(\overline{h},\overline{h},\overline{h})'_\theta &= \tfrac{1}{2}(b(2g-1)(3(\theta-2)\theta+2)+g(\theta(g(10-6\theta)+3\theta-4)-2)) \\
P(\overline{h},\overline{h},\overline{h})''_\theta &= 3b(2g-1)(\theta-1)+g(g(5-6\theta)+3\theta-2)
\end{aligned} \tag{41}$$

Since $\frac{1}{2} \le b < g \le 1$ and $0 \le \theta \le 1$ we get that $P(\overline{h},\overline{h},\overline{h})V(\overline{h},\overline{h},\overline{h})$ is concave.
For $m = (\bar{l}, \overline{h}, \overline{h})$:

$$\begin{aligned}
V(\bar{l},\overline{h},\overline{h})'_\theta &= \frac{(2b-1)(g-1)g\left((2g-1)\left(b(\theta-1)^2-g(\theta-2)\theta\right)-g\right)}{\left((2g-1)\theta^2(b-g)-2bg+b+g\right)^2} \\
V(\bar{l},\overline{h},\overline{h})''_\theta &= \frac{2(2b-1)(g-1)g(2g-1)(b-g)\left(b(2g-1)(\theta-1)^3+g\theta(\theta(-2g(\theta-3)+\theta-3)-3)+g\right)}{\left((2g-1)\theta^2(g-b)+2bg-b-g\right)^3} \\
P(\bar{l},\overline{h},\overline{h})'_\theta &= \tfrac{1}{2}\left(-2\theta\left(-2bg+b+g^2\right)-2bg+b+g\right) \\
P(\bar{l},\overline{h},\overline{h})''_\theta &= b(2g-1)-g^2
\end{aligned} \tag{42}$$

Since $\frac{1}{2} \le b < g \le 1$ and $0 \le \theta \le 1$ we get that $P(\bar{l},\overline{h},\overline{h})V(\bar{l},\overline{h},\overline{h})$ is concave.
For $m = (\overline{h}, \bar{l}, \overline{h})$:

$$\begin{aligned}
V(\overline{h},\bar{l},\overline{h})'_\theta &= g-b \\
V(\overline{h},\bar{l},\overline{h})''_\theta &= 0 \\
P(\overline{h},\bar{l},\overline{h})'_\theta &= (1-g)g\theta \\
P(\overline{h},\bar{l},\overline{h})''_\theta &= (1-g)g
\end{aligned} \tag{43}$$

Since $\frac{1}{2} \leq b < g \leq 1$ and $0 \leq \theta \leq 1$ we get that $P(\overline{h}, \overline{l}, \overline{h})V(\overline{h}, \overline{l}, \overline{h})$ is concave.
For $m = (\overline{l}, \overline{l}, \overline{h})$:

$$\begin{aligned}
V(\overline{l}, \overline{l}, \overline{h})'_\theta &= \frac{(g-1)g\big((2g-1)\theta^2(-(2b-4g+1))(b-g)+2(2b-1)(2g-1)\theta(b-g)+b(b(2-4g)+4g-3)+g\big)}{(g-(2g-1)(b(\theta-1)^2-g(\theta-2)\theta))^2} \\
V(\overline{l}, \overline{l}, \overline{h})''_\theta &= \frac{2(1-g)g(2g-1)(g-b)\big(2b^2(2g-1)(\theta-1)^3-b(\theta-1)\big(12g^2\theta^2+g(4-8\theta(\theta+1))+(\theta-1)(\theta+5)\big)\big)}{((2g-1)(b(\theta-1)^2-g(\theta-2)\theta)-g)^3} \\
&\quad + \frac{2(1-g)g(2g-1)(g-b)(g(\theta((2g-1)\theta((4g-1)\theta-3)-3)+3))}{((2g-1)(b(\theta-1)^2-g(\theta-2)\theta)-g)^3} \\
P(\overline{l}, \overline{l}, \overline{h})'_\theta &= \tfrac{1}{2}(1-\theta)(g+(2g-1)(3b\theta-b-3g\theta)) \\
P(\overline{l}, \overline{l}, \overline{h})''_\theta &= -3(2g-1)\theta(b-g)+4bg-2b-3g^2+g
\end{aligned} \tag{44}$$

Since $\frac{1}{2} \leq b < g \leq 1$ and $0 \leq \theta \leq 1$ we get that $P(\overline{l}, \overline{l}, \overline{h})V(\overline{l}, \overline{l}, \overline{h})$ is concave.
■

Lemma 9. $\lim_{\theta \to 0} \Phi(Sim) - \Phi(Seq) \geq 0$, *with strict inequality if* $\frac{1}{2} < b$.

Proof of Lemma 9. Note that for $\theta \to 0$ the only voting profile that will emerge under a sequential voting procedure is $(\overline{l}, \overline{h}, \overline{h})$, since for any other voting profile $P(m) \to 0$. The informativeness of this profile is:

$$V(\overline{l}, \overline{h}, \overline{h}|Seq) = \frac{z\widetilde{z}^2}{z\widetilde{z}^2 + (1-z)(1-\widetilde{z})^2} < \frac{z^3}{z^3 + (1-z)^3} = V(\overline{l}, \overline{h}, \overline{h}|Sim). \tag{45}$$

where the inequality holds since $\widetilde{z} < z$.
Hence, for small θ we get $\Phi(Seq) < \Phi(Sim)$. ■

Since two concave functions can intersect each other at most twice, the only possible cases are as follows.
<u>Case 1</u> $\Phi(Sem)$ and $\Phi(Seq)$ do not intersect. Hence, from Lemma 9 $\Phi(Seq) < \Phi(Sim)$ and so $[\underline{\theta}, \overline{\theta}]$ is empty.
<u>Case 2</u> $\Phi(Sem)$ and $\Phi(Seq)$ intersect only once; denote this intersection by $\underline{\theta}$.
If for $\underline{\theta} < \theta$ $\Phi(Sim) < \Phi(Seq)$ then $\overline{\theta} = 1$.
If for $\underline{\theta} < \theta$ $\Phi(Seq) < \Phi(Sim)$ then $\overline{\theta} = \underline{\theta}$.
<u>Case 3</u> $\Phi(Sem)$ and $\Phi(Seq)$ intersect twice; denote these intersections by

$\underline{\theta}, \overline{\theta}$. Obviously, because of Lemma 9, we get that if $\theta \in [\underline{\theta}, \overline{\theta}]$ then $\Phi(Sim) < \Phi(Seq)$. ■

Proof of Proposition 1. The proof of Proposition 1 will follows the proofs of Lemmas 6 and 7.

Proof of Proposition 4. As noted $\Phi(m)$ is the probability of taking the correct action when following the recommendation of m.
To prove Proposition 4, we calculate the probability of being wrong when following the recommendation of m under simultaneous and sequential voting procedures.
Note that under a sequential voting procedure a competent expert never cascade since $g = 1$. For the same reason, whenever a competent expert is identified $V(m, Seq) = 1$.
The only case of being wrong under a sequential procedure is when all the experts are incompetent and the first vote is wrong, hence we get:

$$\Phi(seq) = 1 - (\frac{(1-\theta)^N}{2}) \tag{46}$$

Under a simultaneous procedure the decision maker will take the wrong action whenever the majority point toward the wrong state, which is:

$$\Phi(Sim) = 1 - \begin{pmatrix} (1-\theta)^N \left[\left(\frac{1}{2}\right)^N \binom{N}{N} + \left(\frac{1}{2}\right)^N \binom{N}{N-1} + ... + \left(\frac{1}{2}\right)^N \binom{N}{\frac{N+1}{2}} \right] \\ +(1-\theta)^{N-1}\theta[...] \\ +(1-\theta)^{N-2}\theta^2[...] \end{pmatrix} \tag{47}$$

Assume N is odd, then from the binomial theorem we get:

$$2^N = \sum_{k=0}^{N} \binom{N}{k} = \binom{N}{N} + \binom{N}{N-1} + ... + \binom{N}{0} \tag{48}$$

Since $\binom{N}{k} = \binom{N}{N-k}$ we get:

$$2^{N-1} = \binom{N}{N} + \binom{N}{N-1} + \binom{N}{N-1} ... + \binom{N}{\frac{N+1}{2}} \tag{49}$$

Using (49), the first part of the summation in (47) is equal to $(\frac{(1-\theta)^N}{2})$, and all the other parts are positive therefore, $\Phi(Sim) < \Phi(Seq)$.
The case when N is even is similar. ■