Over-Caution of Large Committees of Experts*

Rune Midjord, Tomás Rodríguez Barraquer, and Justin Valasek†

November 12, 2013

Abstract

In this paper, we demonstrate that payoffs linked to a committee member's individual vote may explain over-cautious behavior in committees. A committee of experts must decide whether to approve or reject a proposed innovation on behalf of society. In addition to a payoff linked to the adequateness of the committee's decision, each expert receives a disesteem payoff if he/she voted in favor of an ill-fated innovation. An example is FDA committees, where committee members can be exposed to a disesteem (negative) payoff if they vote to pass a drug that proves to be fatal for some users. We show that no matter how small the disesteem payoffs are, information aggregation fails in large committees: under any majority rule, the committee rejects the innovation almost surely. We then show that this inefficiency can be mitigated by pre-vote information pooling, but only if the decision is take under unanimity: in the presence of disesteem payoffs, committee members will only vote efficiently if they are *all* responsible for the final decision.

Keywords: Committees, Information aggregation, Disesteem payoffs.
JEL Classification Codes: D71, D72

*We thank Sivan Frenkel, Andrea Galeotti, Volker Hahn, Sergiu Hart, Steffen Huck, Macartan Humphreys, Nagore Iriberri, Annick Laruelle, Marcos Nakaguma, Henry Mak, Andrea Matozzi, Rebecca Morton, Peter Norman Sorensen, Jesús Vázquez, Fernando Vega-Redondo, Bauke Visser, Elyashiv Wiedman, Huseyin Yildirim and seminar participants at EUI (Villa San Paolo and Villa La Fonte), Public University of Navarra, the University of the Basque Country, and the Center for Rationality at Hebrew University, University of Bern, and IGIER-Bocconi. Tomás Rodríguez Barraquer acknowledges funding for the research leading to these results from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n. [249159]. Furthermore, he would specially like to thank the hospitality and support of the Center for the Study of Rationality at Hebrew University during the development of this project. All errors are our own.

†University of Copenhagen, Hebrew University of Jerusalem, and WZB Berlin respectively. Contact e-mail: rune.midjord@econ.ku.dk.

1 Introduction

The logic for allocating a social decision to a group of experts rather than an individual is clear: committees aggregate multiple sources of information and expertise, and therefore allow for more informed decisions. The question we ask here is, does a committee effectively utilize the information held by its members when, in addition to caring about making the right decision, committee members also care about individually voting for the right decision? Surprisingly, even though committee members still have an incentive to choose the right option, this kind of idiosyncratic payoff can significantly bias the committee's decision. Key to this insight is the fact that when decisions are made by groups, each individual's ability to impact the final decision is diluted, and this dilution leads to biased decisions in large committees.

A particularly salient example is the FDA advisory committees: the FDA currently has 50 distinct standing boards, or committees of medical experts, that are called upon to decide whether or not to approve a new medical product, such as a pharmaceutical drug, for general use. Presumably, each committee member, just like each individual in society, prefers to accept safe drugs and reject bad drugs. However, if the committee passes a drug that proves to have unexpected severe side-effects, committee members will receive a negative (disesteem) payoff if they personally voted to approve the drug.[1] For example, when Posicor, a drug to relieve high blood pressure, resulted in the death of over 140 people, numerous newspaper articles (including an article that received the prestigious Pulitzer Prize) singled out individual committee members based on their vote: while the committee as a whole made the wrong decision, only committee members who personally voted for the drug were scrutinized.[2]

On one hand, it is unsurprising that FDA committees are cautious, since committee members who face negative disesteem payoffs are more cautious than society at large. The main

[1]This payoff can be purely intrinsic (self-esteem), or as in Brennan and Pettit (2004) and Ellingsen and Johannesson (2008), esteem payoffs can reflect an agent's payoff from their general regard by other members of society (also see the discussion of the relevant psychological and classical literature in Brennan and Pettit). We argue that committee members are exposed to esteem payoffs to the extent that their voting decision is made salient ex-post: If the committee votes to approve the innovation there is some probability of a 'disastrous' event, such as the death of patients, which causes ex-post scrutiny of the committee's decision. This scrutiny makes the committee members' individual votes salient, either through media attention, social and professional networks, or by causing internal deliberation. Moreover, since a committee's decision to correctly approve an innovation is unlikely to become salient, we consider a negative disesteem payoff to be the relevant payoff in these applications.

[2]Other examples include hiring committees and juries. Hiring committee members might be held responsible for a bad hire only if they voted for the candidate. Jury members might receive a negative intrinsic payoff if they vote to free a suspect who then goes on to commit another crime.

contribution of our work, however, is to show that the collective nature of the decision-making process can magnify the caution of individual committee members. This mechanism could explain claims that committees tend to be overly cautious when considering a change from the status quo (see Li (2001)), and our empirical finding that larger FDA committees are more likely to reject new drug applications: a simple OLS regression suggests that an additional committee member decreases the likelihood that any member votes for approval by 1.3 percent, a decrease of 30 percent from the smallest to largest committee in our sample.[3] Theoretically, our main result shows that this over-caution is particularly stark in the limit, and no matter how small the disesteem payoffs are, a very large committee will always reject the drug regardless of the information held by its members. As the size of the committee increases, the chances that a single vote is pivotal diminishes and the incentive to avoid potential disesteem magnifies.

Our analysis highlights that committees of experts are subject to a variant of a familiar problem: decisions by groups require an aggregate decision-making approach and, as is often the case when collective action is required to achieve a socially desirable result, the process is susceptible to collective-action problems (as discussed in Olson (1965) and the subsequent literature on collective action). Idiosyncratic payoffs in committees, such as disesteem payoffs, can create a situation in which each member prefers a certain collective action be taken (pass the innovation given a minimum number of signals to accept), but lacks an individual motivation to contribute to the preferred result. Therefore, for large committees, voting to accept given a signal of accept is a public good: all benefit from the increased probability that good drugs are passed, but only the individual is subject to the risk of disesteem payoffs. In committees, just as in society at large, public goods are generally under-provided (as in the seminal contributions of Samuelson (1954) and Bergstrom et al. (1986)), leading to over-caution of large committees of experts.

In our framework, a committee is composed of n experts who must vote simultaneously to approve or reject an innovation using a q-rule, which specifies that the innovation is approved only if more than a fraction q of the committee members vote for approval. Whether the innovation is beneficial to society or not depends on an unobservable binary state of the

[3] We present this empirical finding in detail, complete with a discussion of alternative explanations, in Appendix B. We have voting data on approve/disapprove decisions from 174 meetings spread over twenty-one topical FDA committees. Each of the FDA panels in our sample consists of 11-15 regular members, but for any particular decision, the size of the committee varies (in the range 3-26) due to two main factors. (1) Absenteeism: permanent members frequently cancel on the meetings (members serve on a voluntary basis and most of them are physicians and professors of medicine). (2) Invited members: often, individuals who are not regular committee members, but who have expertise particularly relevant to the drug in question, are invited to participate.

world, which is revealed only if the innovation is approved.[4] If the innovation is rejected (status quo) committee members get a payoff of zero. If the innovation is rightfully approved, each expert gets a positive payoff of W. However, if the innovation is wrongfully approved then all committee members receive a negative payoff of C, and the committee members that supported the approval receive an additional penalty of K.

We model the opinion (signal) of each expert as a noisy version of society's *state of the art* with respect to his field of expertise. Each expert's opinion of whether a drug is safe or not is the result of applying a small measure of white noise to a hypothetical ideal dictamen by the *state of the art*, which in itself is a noisy reflection of the true state of the world, with exogenous accuracy.[5] This approach differs from the standard model of committee behavior, where signals are generated by the state of the world, and hence a large committee that aggregates the signals efficiently will never make an error.

We show that for each set of values of the exogenous parameters there is essentially a unique equilibrium. We characterize this unique equilibrium and study the comparative statics. Of most interest, by increasing K (the disesteem penalty) the committee acceptance rate decreases. The relation between the acceptance rate and committee size, however, is non-monotonic. As more experts join the committee there is potential for more information aggregation, which may make the experts more confident about accepting the innovation. On the other hand, the probability of being pivotal decreases, which exacerbates the free riding problem. Eventually, this latter effect dominates. Similarly, we find that a decrease in the noise of the experts' signals generated by the state of the art may not necessarily increase the committee's acceptance rate, since less noise implies that agents can better predict the actions of their peers, which can decrease their ex ante probability of being pivotal.

In the extensions of the model, we show that our main result persists even when committee members pool their information before voting, unless the decision is made by unanimity. The contrast between unanimity and other decision rules is particularly stark with an initial round of cheap talk: with decision rules other than unanimity, committee members may have an incentive to lie about their signal to induce other committee members to vote for the innovation, resulting in non-truthful communication. Under a unanimity rule, however, the innovation is only approved when all agents are exposed to the disesteem payoffs, which implies that payoffs to all committee members are homogenous, resulting in truthful communication and efficiency. Finally, we study a variation of the model in which the disesteem payoffs get diluted as the committee's size increases. We provide sufficient conditions on the

[4] In our FDA example, information on harmful side-effects of a drug is only generated if it is made generally available.

[5] The state of the art can be thought of as the decision that an ideal computer, programmed with the best available decision procedures and criteria for classifying all the evidence, would arrive at.

speed of dilution of the disesteem payoffs for the main result to hold.

The paper is organized as follows. Section 2 introduces the payoff structure and the process that generates each expert's opinion (signal). Section 3 characterizes the symmetric equilibria of the game, establishes the main result of the paper, and provides comparative statics. Finally, Section 4 presents the extensions of the model. All proofs are relegated to Appendix A. Appendix B contains a detailed discussion of the empirical analysis noted in footnote 3. Lastly, for completeness, a supplementary Appendix[6] provides a general characterization of information aggregation under the state of the art view of expertise, without disesteem payoffs.[7]

Related Literature

This paper contributes to the game theoretic literature on information aggregation in committees (see Austen-Smith and Banks (1996) for an early reference and recent surveys by Gerling et al. (2005) and Li and Suen (2009)). Our paper is most closely related to a subset of the committee literature that considers information aggregation when voters have a common interest in making the right decision and additional "idiosyncratic" payoffs that condition on the individuals' votes.

One branch of this literature studies information aggregation when committee members receive an idiosyncratic (reputation) payoff whenever their vote matches the state, which is revealed ex post. In Ottaviani and Sorensen (2001), the committee members only receive reputation payoffs, which are independent of the decision adopted by the committee. Each expert is of unknown ability type, where a "smart" expert receives a more precise signal than a "dumb" expert. When the experts speak sequentially, the reputation payoffs give rise to informational herding. When voting is simultaneous, the first best can be achieved if the probability distribution over the binary state variable is not too skewed. Relatedly, Levy (2007) considers the issue of transparency when committee members care about their reputation for expertise.

In Visser and Swank (2007), committee members deliberate, prior to voting, on whether to accept a project. The members are concerned about the value of the project and their reputation for being well informed. The market, whose judgement the experts care about, does not observe the value of the project, only the decision taken. Visser and Swank show that reputation concerns make the a priori unconventional decision more attractive and lead

[6]Available online at http://mwpweb.eu/JustinValasek/.

[7]All the results in the absence of disesteem payoff are analogous to those of the literature on the Condorcet jury theorem with strategic voters (see Austen-Smith and Banks (1996)), McLennan (1998) and Feddersen and Pesendorfer (1998)). For the most general version of the Condorcet jury theorem, see Peleg and Zamir (2012).

committees to show a united front. As the number of committee members grows, however, converging on the unconventional decision becomes a weaker indicator of signal concurrence, which in turn lowers the reputation concerns and leads to overall better decisions.

While disesteem payoffs also realize according to the correctness of the individual's vote, our analysis differs from the existing literature on reputation concerns in that the idiosyncratic payoffs condition on both state of the world *and* the committee's decision. This allows us to analyze behavior in situations where the information on (or the salience of) the accuracy of the individual's vote depends on the committee's decision.

Another set of recent papers considers idiosyncratic payoffs in elections. Callander (2008) analyzes voting under simple majority rule when voters wish for the the better candidate to be elected, but want to vote for the winner. The idiosyncratic payoff for voting for the winner (independently of the winning candidate's quality) creates multiple symmetric equilibria, some with unusual properties. When considering optimal equilibria as the population becomes large, Callander (2008) shows that in elections without a dominant front-running candidate the better candidate is almost surely elected, whereas information cannot be fully aggregated in races with a clear front-runner.

Morgan and Várdy (2012) study a model in which voters are driven by both instrumental and purely expressive idiosyncratic payoffs. That is, a voter receives some consumption utility if he/she votes in a pre-defined way (e.g. in accordance with one's norms) that is irrespective of the correct outcome and the implemented decision. Some voters will receive a signal that is in conflict with their expressive motive. If the degree of conflict is low and thus the expressive preferences are mostly shaped by facts (the signals) then Condorcet's (1785) jury theorem holds and large voting bodies make correct decisions. However, when expressive preferences are relatively impervious to facts, then large voting bodies do no better than a coin flip.

While Callander (2008) and Morgan and Várdy (2012) both demonstrate that idiosyncratic payoffs can lead to a failure of information aggregation in large committees, the mechanism we present here is quite different. In both of the above papers, idiosyncratic payoffs give agents a direct incentive to vote for, say, candidate A regardless of the state of the world; that is, information aggregation fails because the idiosyncratic payoffs run counter to the common value payoff of electing the better candidate. In our analysis, however, information aggregation fails despite idiosyncratic payoffs that reinforce common value payoffs: disesteem payoffs realize only when the committee approves a bad drug.

Lastly, Li (2001) shows that committees might have an incentive to adopt a more conservative decision rule, in the sense of requiring a higher information threshold, to induce members to individually invest more in information gathering. Our results give a complementary explanation for why, even in situations where the committee decision rule is based on votes rather than quantifiable evidence, committee members have an incentive to *vote* conservatively. Interestingly, although we consider a different setting, in the comparative statics section we detail a result that is related to Li (2001) in spirit: in some cases, increasing the number of votes required for approval may, in equilibrium, *increase* the probability that the committee passes good innovations.

2 The Model

An innovation is submitted for approval by a committee of n experts that operates according to a q-rule: If strictly more than a fraction q of the committee members $i \in \{1, 2, ..., n\}$ vote in favor of approval then the innovation is approved, and otherwise it is rejected. We denote the votes of each committee members $i \in \{1, 2, ..., n\}$ by $v_i \in \{a, r\}$ and the decision of the committee by $X \in \{a, r\}$, where a indicates *accept* and r indicates *reject*. The payoff to each expert i depends on the decision of the committee, an underlying state of the world $\omega \in \{A, R\}$, and the expert's vote v_i:

$$U(v_i, X, \omega) = \begin{cases} 0 & \text{if } X = r \\ W & \text{if } X = a,\ \omega = A \\ -C & \text{if } X = a,\ \omega = R,\ v_i = r \\ -(C+K) & \text{if } X = a,\ \omega = R,\ v_i = a \end{cases}$$

where $W, C, K > 0$.

One interpretation of the structure of the payoffs is as follows: if the innovation is rejected, then payoffs to all agents in the committee are zero, since the status quo is preserved and no further information about the innovation's quality is generated. If the innovation is approved, then the quality of the innovation is revealed and the committee members receive a common payoff and, depending on the state of the world and their vote, an individual disesteem payoff. The common payoff is W or $-C$ depending on whether the committee has made the right decision with respect to the state of the world. The individual disesteem payoff is only awarded in the case that the committee has made the wrong decision, and is non-zero ($-K$) only for the agents that supported that wrong decision.[8] If K is small these payoffs represent a seemingly small departure from a pure common values situation,

[8] K can be thought of as the probability that the decision is disastrously wrong, e.g. side effects exist and are fatal, multiplied by the negative payoff that accrues to committee members who supported the decision to approve the drug.

in which the payoffs to all committee members are identical in all possible events. However, as our main result shows, for a sufficiently large committee this small departure implies a large difference in equilibrium behavior.

2.1 The state of the art and expert's opinions (signals)

We denote by $p_A \equiv p(\omega = A)$ society's prior belief on the state of the world. We think of the committee members as experts in a relevant discipline for the decision at hand. We model the knowledge of each member of the committee as an idiosyncratic departure from the *state of the art* of that discipline. We denote the *state of the art* by $t \in \{a, r\}$ and let α denote the probability that the state of the art is wrong when it indicates that the innovation should be rejected ($\alpha = p(\omega = A|t = r), 0 < \alpha < \frac{1}{2}$), and let β denote the probability that the state of the art is wrong when it indicates that the innovation should be accepted ($\beta = p(\omega = R|t = a), 0 < \beta < \frac{1}{2}$). Put in terms of our example of the FDA advisory committees, there is a commonly available collection of evidence on the efficacy and safety of the drug–a whole battery of data from clinical trials. The state of the art, t, can be thought of as the decision which an *ideal computer*, programmed with the ideal decision procedures of medical science and state of the art criteria for evaluating all data, would arrive at.

The state of the art is not directly observable to the experts. Instead, we think of an expert as a coarse embodiment of the state of the art. The coarseness reflects idiosyncrasies at the individual decision making level, such as possible errors of interpretation, conceptual misunderstandings, lapses of attention (all these often classified as "human error"), but also inspired hunches and extraordinary insights. We further assume that these individual differences with respect to the state of the art are purely idiosyncratic, in the sense that conditioning on t, the sincere opinions of different experts (which we henceforth refer to as signals) are independent. Concretely, with probability $1 - \varepsilon$ the signal of expert i, s_i, coincides with the state of the art ($p(s_i = t|t) = 1 - \varepsilon$, $\varepsilon < \frac{1}{2}$), and with probability ε it differs with respect to the state of the art ($p(s_i \neq t|t) = \varepsilon$).[9]

2.1.1 Equilibrium Concept

In what follows we will use $\sigma_i : \{a, r\} \to [0, 1]$ to denote the possibly-mixed strategy according to which member i sets $v_i = a$ with probability $\sigma_i(a)$ after receiving signal $s_i = a$, and sets $v_i = a$ with probability $\sigma_i(r)$ after receiving signal $s_i = a$. Throughout the analysis we

[9]The state of the art can be thought of in an alternative, more constructive way. Rather than thinking of the opinions of the experts as idiosyncratic distortions of a pre-existing state of the art, we can think of the state of the art as the probability limit of the average of the signals $\frac{1}{n} \lim_{n\to\infty} \sum_{=1}^{n} s_i$ and explicitly set forth conditions which would deem the signals conditionally independent given this limit.

rely on the concept of Bayesian Nash equilibrium and focus on symmetric strategies only; that is, conditioning on signals, all members use the same decision rule.[10] Assuming that all members other than i play according to strategy $\sigma = (\sigma(a), \sigma(r))$ we denote i's expected payoff from using strategy σ_i by: $E_\sigma[U(\sigma_i, X, \omega)|s_i]$.[11]

DEFINITION 1 **(Symmetric Equilibrium)** *A strategy, $\sigma = (\sigma(a), \sigma(r))$, is a symmetric equilibrium if and only if for all $i \in \{1, 2, ..., n\}$, $s_i \in \{r, a\}$ and, strategy of expert i, σ_i:*

$$E[U_\sigma(\sigma, X, \omega)|s_i] \geq E[U_\sigma(\sigma_i, X, \omega)|s_i]$$

3 Analysis

We first characterize the equilibria of the model and then present the main result and comparative statics. Denote by $G^K_{n,q}$ the game with disesteem payoffs K, decision rule q, and n players. We show that, other than the babbling equilibrium in which all agents vote to reject, each game $G^K_{n,q}$ has at most one equilibrium. We let piv_i denote the event that among all experts other than i, there are exactly $\lfloor nq \rfloor$ votes for approval. Assuming that all other members are using strategy σ, expert i finds it optimal to set $v_i = a$ upon observing signal s_i if, and only if, his *willingness to vote to reject* the innovation $R_{s_i}(p_A, \alpha, \beta, \varepsilon, q, n, K, W, C, \sigma)$ is nonpositive:[12]

$$R_{s_i}(n, \sigma) = Kp_\sigma(X = a, \omega = R|s_i) - Wp_\sigma(piv_i, \omega = A|s_i) + Cp_\sigma(piv_i, \omega = R|s_i) \leq 0 \qquad (1)$$

Note that if $\sigma(a) = \sigma(r) = 0$, then all probabilities in this inequality vanish. It follows that it is always an equilibrium for the members to reject the innovation regardless of their signal (referred to as the babbling equilibrium). However, in contrast to $K = 0$, when $q < \frac{n-1}{n}$, then $\sigma(a) = \sigma(r) = 1$ is not an equilibrium.[13] As is also the case with $K = 0$, non-babbling equilibria often involve mixed strategies.

[10] Restricting attention to symmetric strategies is common in the voting literature when voting is simultaneous; see for example Palfrey and Rosenthal (1985) and Feddersen and Pesendorfer (1997).

[11] Specifically

$$\begin{aligned} E_\sigma[U(\sigma_i, X, \omega)|s_i] &= \sigma_i(s_i) \sum_{X \in \{a,r\}} \sum_{\omega \in \{A,R\}} p_\sigma(X, \omega | v_i = a, s_i) U(a, X, \omega) \\ &+ (1 - \sigma_i(s_i)) \sum_{X \in \{a,r\}} \sum_{\omega \in \{A,R\}} p_\sigma(X, \omega | v_i = r, s_i) U(r, X, \omega) \end{aligned}$$

[12] Henceforth, we use the abbreviated notation $R_{s_i}(n, \sigma) \equiv R_{s_i}(p_A, \alpha, \beta, \varepsilon, q, n, K, W, C, \sigma)$ unless we need to stress the dependence of R on the other parameters.

[13] When $K = 0$, $\sigma(a) = \sigma(r) = 1$ is an equilibrium as long as $q \geq \frac{1}{n}$. With every member voting to accept, the innovation is accepted by the committee for sure and expert i's action has no impact on his payoff.

Relying on the following Lemma and Corollary, we are able to fully characterize the members' *willingness to reject* functions (1), prove uniqueness of non-babbling equilibria, and demonstrate how equilibria respond to changes in the exogenous parameters of the model.

LEMMA **1**
Suppose that at least one of $\sigma(r)$ or $\sigma(a)$ is strictly positive. If $R_r(n,\sigma) \leq 0$, then $R_a(n,\sigma) < R_r(n,\sigma)$.

Lemma 1 implies that if an expert weakly prefers to set $v_i = a$ upon receiving signal r (i.e. (1) holds when $s_i = r$) he will strictly prefer to set $v_i = a$ upon receiving signal a (i.e. (1) holds strictly when $s_i = a$).[14]

Corollary 1 follows immediately from Lemma 1 and shows that in any other equilibrium of $G^K_{n,q}$, behavior is ordered in the sense that $\sigma(a) > \sigma(r)$, and that a properly mixed action is used after receiving at most one of the signals.

COROLLARY **1**
Any equilibrium of any game $G^K_{n,q}$ has the following form: $\sigma(r) = 0, \sigma(a) \geq 0$, or $0 < \sigma(r)$, $\sigma(a) = 1$.

By virtue of Lemma 1, equilibria of the form $\sigma(r) = 0, 0 < \sigma(a) < 1$ are fully characterized by solutions to the equation $R_a(n,(\sigma(a),0)) = 0$ and equilibria of the form $0 < \sigma(r) < 1, \sigma(a) = 1$ are fully characterized by solutions to the equation $R_r(n,(1,\sigma(r))) = 0$.[15]

This allows us to characterize the equilibria of the model using the following function:

$$R(n,z) = \begin{cases} R_a(n,(z,0)) & \text{if } z \leq 1 \\ R_r(n,(1,z-1)) & \text{if } z > 1 \end{cases}$$

Where $z = \sigma_a + \sigma_r$. Importantly, in contrast to R_a and R_r, the last argument of R is one-dimensional. Therefore, with all parameters other than z being held constant, the equilibria of $G^K_{n,q}$ correspond to the values of z that are roots of R when $z \neq 1$, as the function is continuous for all $z \neq 1$, or to a discontinuous crossing in case $z = 1$, which corresponds to the equilibrium $\sigma = (1,0)$. We can now present the proposition characterizing the non-babbling equilibrium:

[14] When $K = 0$ a stronger relation holds: Specifically, with the exception of the case in which everyone votes to reject, agents always have a strictly smaller willingness to reject after observing $s_i = a$ than after observing $s_i = r$.

[15] The reason is that by Lemma 1, if $R_a(n,(\sigma(a),0)) = 0$, it must be the case that $R_r(n,(\sigma(a),0)) > 0$ so $(\sigma(a),0)$ is an equilibrium. Similarly if $R_r(n,(1,\sigma(r))) = 0$ then $R_a(n,(1,\sigma(r))) < 0$ so $(1,\sigma(r))$ is an equilibrium.

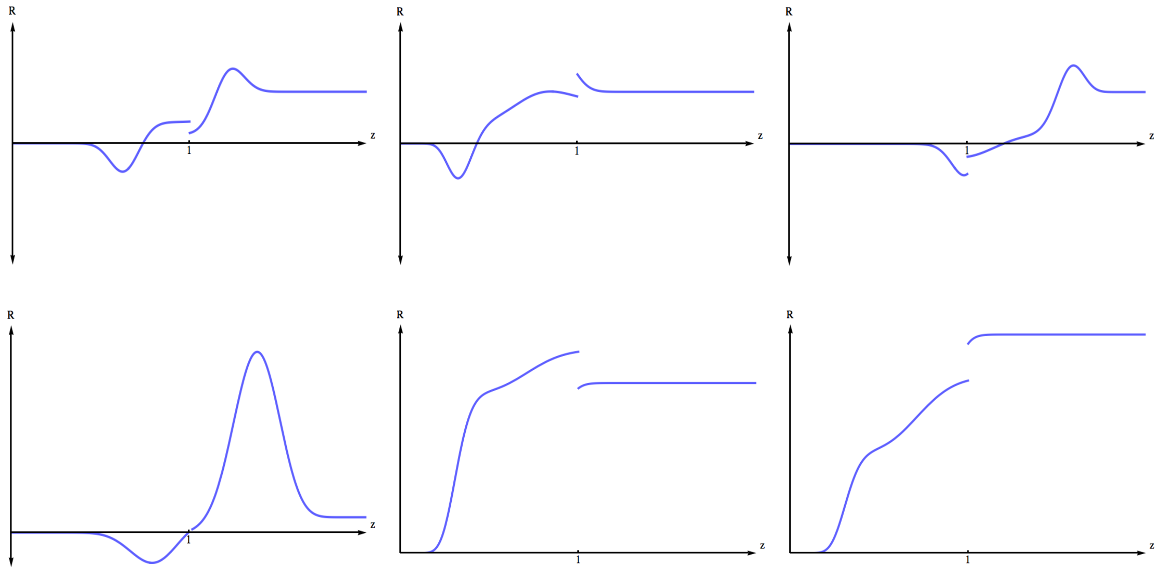

Figure 1: From top left, clockwise (a) and (b) equilibria of the form $\sigma(r) = 0, \sigma(a) > 1$; (c) an equilibrium of the form $\sigma(r) > 0, \sigma(a) = 1$; (d) equilibria of the form $\sigma(r) = 0, \sigma(a) = 1$; (e) and (f) no equilibrium.

Proposition **1 (Equilibrium Characterization)**

(1) If a non-babbling equilibrium z^* exists, it is unique.

(2) If $G^K_{n,q}$ has a non babbling equilibrium, then so does $G^K_{n,q'}$ for any $q' > q$.

(3) If an equilibrium $z^* \neq 1$ exists then $\frac{\partial R(n,z^*)}{\partial z} > 0$.

The difficulty in characterizing the set of roots of $R(n, z)$ (and thereby the equilibria of the game), stems from the fact that the function is non-monotonic, and discontinuous at $z = 1$. However, there are (3) main properties of R that hold when $K > 0$, which taken together give uniqueness. These are: (1) In each of the two continuous segments ($z \in (0, 1]$ and $z \in (1, 2)$) R has the single crossing property in z.[16] This implies both that R has at most one root in each of these two segments, and also that the crossing of the z axis must be from negative to positive. (2) If $R_r(n, 1) \leq 0$ then the jump at the discontinuity is positive, which follows from Lemma 1. (3) $R(n, 2) > 0$, since if everyone votes a, it is strictly a best response to deviate to voting r (for all values of $q < \frac{n-1}{n}$).

These three properties place $R(n, z)$ in one of the six classes shown in Figure 1, in each of which either a non-babbling equilibrium does not exist, or it exists and is unique. First, if $R(n, z)$ has a crossing in $(0, 1]$, then while property (2) tells us that the jump at the

[16]This result follows from applying the main result of Quah and Strulovici (2012); see the formal proof in Appendix A for the details of Quah and Strulovici's result, and for its application to our problem.

discontinuity can either be negative (Figure 1 (a)) or positive (as in Figure 1 (b)), the second continuous segment must start being positive, which by property (1) implies that there is no crossing in $(1, 2)$, since any crossing of the z axis must be from negative to positive.

Second, if R has no roots in $(0, 1]$, and is always negative in $(0, 1]$, then by property (2), the jump at the discontinuity must be positive. If the second continuous segment starts being negative, then there must be a crossing in $(1, 2)$ (Figure 1 (c)), since by property (3), $R(n, 2)$ is strictly positive. Furthermore, by property (1) the crossing must be unique. If the second segment starts being positive, then we have an equilibrium at $z = 1$ (Figure 1 (d)), and no crossing in $(1, 2)$.

Finally, if R has no roots in $(0, 1]$, and is always positive in $(0, 1]$ then, once more, the second segment must start being positive, which implies that R has no roots in $(1, 2)$ and therefore no roots at all (Figure 1 (e) and (f)).

It is worthwhile to note two main qualitative differences between the cases $K = 0$ and $K > 0$ regarding the willingness to reject function R: First, when $K = 0$ the jump at the discontinuity must always be positive, which excludes the cases shown in Figure 1 (a) and (e). Second, when $K = 0$, property (3) does not hold, and in particular $R(n, 2) = 0$.

We denote the unique non-babbling equilibrium of $G^K_{n,q}$ (if it exists) by $\sigma^K_{n,q}$ and its one dimensional representation by $z^K_{n,q} = \sigma^K_{n,q}(a) + \sigma^K_{n,q}(r)$.[17] Throughout what follows, we alternate between the $\sigma^K_{n,q}$ and $z^K_{n,q}$ based on convenience. Proposition 1 allows us to characterize the effect of increasing K, which is captured by the following Corollary.

COROLLARY **2 (Comparative statics: K)**
If $z^* \neq 1$ then $\frac{\partial z^*_{n,q}}{\partial K} > 0$. It then follows that $p_{z^K_{n,q}}(X = a)$ and in particular, both $p_{z^K_{n,q}}(X = a|t = a)$ and $p_{z^K_{n,q}}(X = a|t = r)$ are decreasing in K.

Corollary 2 follows immediately from observing in equation (1) that as long as $z = \sigma(a)+\sigma(r)$ is positive, increasing K simply shifts R upwards. As shown in each of the cases depicted in Figure 1 in which the non-babbling equilibrium exists,[18] this upwards shift causes the new crossing to take place at $z' < z^K_{n,q}$. Generically, a small enough change of K at an equilibrium $z^* = 1$ preserves this as the unique equilibrium of the game.[19]

[17] Given $z^K_{n,q}$, the actual values of $\sigma^K_{n,q}(a)$ and $\sigma^K_{n,q}(r)$ can be recovered as $\sigma^K_{n,q}(a) = z^K_{n,q}$ and $\sigma^K_{n,q}(r) = 0$ if $z^K_{n,q} \leq 1$ and $\sigma^K_{n,q}(a) = 1$ and $\sigma^K_{n,q}(r) = z^K_{n,q} - 1$ if $z^K_{n,q} > 1$.

[18] With the exception of the one corresponding to $z^*_{n,q} = 1$.

[19] There is a simple argument, that essentially provides the same result as Corollary 2 without requiring Proposition 1 (single crossing). Let $z^{**}_{n,q}$ denote the maximum crossing (in case there are many). Since we know that when all other agents vote to accept, any agent i finds it strictly optimal to reject $(R(n, 2) > 0)$, this last crossing at $z^{**}_{n,q}$ must be from bottom to top (with the exception of a possible tangency). It follows that if $\frac{\partial z^{**}_{n,q}}{\partial K}$ exists it must be positive.

The straightforward intuition for Corollary 2 is that K indexes the conflict of interest among committee members. For any $q < \frac{n-1}{n}$ (all decision rules excluding unanimity), a positive K implies that if the members believe that the innovation should be approved, any given expert i would rather have the rest of the committee to approve it, and hedge against the disesteem payoff by setting $v_i = r$. The motive for avoiding potential disesteem is increasing in K. Therefore, to sustain positive approval rates at higher values of K it is necessary that all committee members are relatively more pivotal, and/or for $P(X = a|\omega = R)$ to decrease. Part (3) of Proposition 1 implies that in equilibrium the only way of doing this is by lowering z.[20]

The second part of corollary 2 tells us that a higher K makes committees more cautious. Whether the expected outcome is better or worse from a social standpoint depends on the particular social welfare function. In applications of the model such as the FDA approval process, the social welfare function might be by and large independent of the payoffs of committee members. In general, given the imperfection of the state of the art ($\alpha, \beta > 0$), there is an unavoidable tradeoff between the probability of approving bad drugs $p(X = a|\omega = R)$ and the probability of rejecting good drugs $p(X = r|\omega = A)$, making any welfare judgment conditional on society's valuation of the various outcomes.

Large Committees

While the welfare implications of positive K for small committees depend on the particular social welfare function, the consequences of K are stark for sufficiently large n. In order to compare, we first characterize large committee outcomes for $K = 0$ ($G^0_{n,q}$).

Proposition **2 (Convergence to the state of the art)**
When $K = 0$ and committee members act according to the non-babbling equilibrium, the decision of the committee converges almost surely to the state of the art for all $q \in (0, 1)$ as n approaches infinity.

Proposition 2 states the analogous result to Feddersen and Pessendorfer's (1998) Proposition 3 (proved in the supplementary appendix as Corollary 2): in the absence of disesteem payoffs, regardless of q, decisions by large committees almost surely converge to the state of the art.

This gives us an appropriate benchmark for our main result: as n grows, the behavior of the committee given any equilibrium of $G^K_{n,q}$ converges to its behavior under the babbling strategy. That is, the committee converges to always rejecting the innovation.

[20] It is important to note that this is a property of the equilibrium and not a global property of R, as follows from inspecting the sketches in Figure 1.

Proposition **3 (The probability of acceptance converges to** 0 **as** $n \to \infty$**)**
Let $K > 0$ *and consider the sequence of games* $G^K_{n,q}$ *and any sequence of symmetric strategy profiles* σ^n*, such that for each* n*,* σ^n *is an equilibrium of* $G^K_{n,q}$*. We let* $p_{\sigma^n}(X = a)$ *denote the probability that the committee accepts the innovation in game* $G^K_{n,q}$*, playing according to* σ^n*. Then,* $p_{\sigma^n}(X = a) \to 0$ *as* $n \to \infty$*. That is, for all* $\delta > 0$*, there exists* n_δ *such that for all* $n > n_\delta$*,* $p_{\sigma^n}(X = a) < \delta$*.*

The proof of Proposition 3 has two parts, which can be illustrated by reference to the RHS and LHS of the following rearrangement of equation (1), representing expert i's willingness to vote to accept the innovation upon receiving signal s_i, when all other members play according to σ,

$$Wp_\sigma(piv_i, \omega = A|s_i) - Cp_\sigma(piv_i, \omega = R|s_i) \geq Kp_\sigma(X = a, \omega = R|s_i) \tag{1'}$$

First, we show that under any q-rule, LHS converges to zero as n approaches infinity. Next we show that, due to the state of the art layer, the RHS, while decreasing under some $\{\sigma_a, \sigma_r\}$, is always strictly bounded away from zero. Intuitively, as the size of the committee grows, the probability of influencing the committee decision, and hence of obtaining W rather than $-C$, approaches zero. The probability the negative disesteem payoff realizes, however, is bounded away from zero.

Corollary **3 (Behavior)**
Let $K > 0$*. There exists* n^* *such that for all* $n > n^*$*,* $\sigma^n(a) < \frac{q}{1-\varepsilon}$*, where* σ^n *is any symmetric equilibrium of* $G^K_{n,q}$*.*

Proposition 3 and its corollary implies a striking difference in the equilibrium behavior of committees of sufficiently large size with respect to their behavior with no disesteem payoffs no matter how small these disesteem payoffs are. In particular, Proposition 2 shows that the unique non-babbling equilibrium of $G^0_{n,q}$ converges to the decision of a single agent with representative preferences and direct access to the state of the art. In contrast, Proposition 3 tells us that no matter how small, when $K > 0$, and for sufficiently large n, the committee essentially always rejects the innovation, implying that it will wrongly reject the innovation with high probability.

It is worth noting precisely how Proposition 3 depends on the state of the art view of expertise. If the signals of the experts were independent conditioning on the state of the world, as is frequently assumed in the literature on information aggregation, then under any strategy σ such that $\sigma(r) < \frac{q-\varepsilon}{1-\varepsilon}$ the RHS of (1′) would converge to 0 at the same rate as the LHS. This allows for the *possibility* of an equilibrium where the committee votes optimally, since under optimal voting, the probability of the event that the committee wrongly approves the innovation becomes negligible as n grows. However, our analysis of the state of the art view of expertise shows that if there is even a minuscule probability that the state of the art does not correctly identify the state of the world, then the incentive for the committee members to protect themselves against disesteem payoffs dominates, and large committees reject the innovation almost surely.

3.1 Comparative Statics

In this section, we characterize the marginal effect of changes in the exogenous parameters on the non-babbling equilibrium z^*. The following is a Corollary of Proposition 1:

COROLLARY **4 (Signs)**
For all parameter values such that $z^*(p_A, \alpha, \beta, \varepsilon, q, n, K, W, C)$ exists and is different from 1, we have:

- $\dfrac{\partial z^*}{\partial W} > 0$, $\dfrac{\partial z^*}{\partial C} < 0$.
- $\dfrac{\partial z^*}{\partial p_A} > 0$, $\dfrac{\partial z^*}{\partial \alpha} > 0$ *and* $\dfrac{\partial z^*}{\partial \beta} < 0$.
- z^* *is weakly increasing in* q.[21]

In order to establish the first two sets of results we rely on the characterization of the willingness to reject R, summarized in Figure 1. Note that the effects of $\alpha = p(\omega = A|t = r)$ and $\beta = p(\omega = R|t = a)$ have opposite signs; a higher α and a lower β both map onto a greater likelihood of $\omega = A$, which shifts the expert's willingness to reject, R, down. Since R is increasing in z at all equilibria, the unique non-babbling equilibrium under a higher α (or lower β) must occur at a higher z.

In terms of the effect of q on z^*, we use the result of Quah and Strulovici (2012) to show that the negative of the willingness to reject, $-R$, has the single crossing property in $\lfloor nq \rfloor$. Thus, it follows that in any non-babbling equilibrium, the willingness of any committee member i to vote to *accept* the innovation is increasing in the decision threshold $\lfloor nq \rfloor$. This result also has an intuitive explanation: Fixing the behavior of all other agents, an increase of $\lfloor nq \rfloor$ from m to m', has two effects. First, it makes the committee less likely to accept, and thus reduces i's exposure to the disesteem payoffs. Second, conditional on being pivotal, i infers that the other agents have received a greater number of a signals under m' than under m,[22] and thus assigns a higher probability on $\omega = A$. Since both these effects lower the agent's

[21] q affects z^* through $\lfloor nq \rfloor$ and therefore z^* is discontinuous and not differentiable in q.

[22] Formally, under m', the distribution of the number of a signals conditional on i being pivotal first order stochastically dominates the same distribution under m.

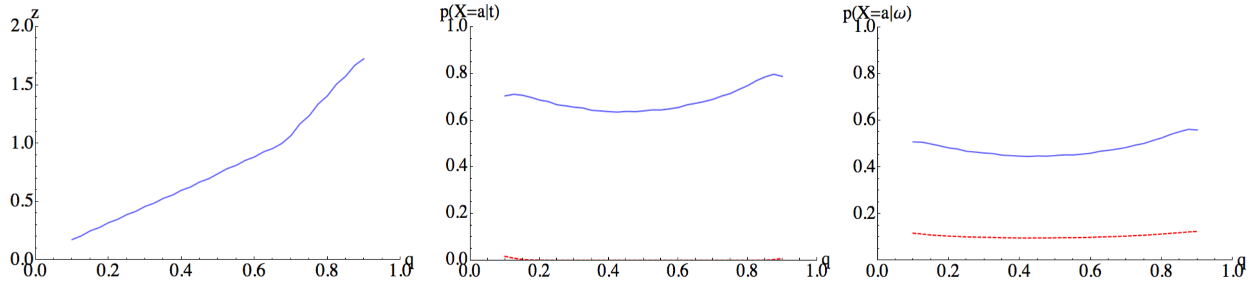

Figure 2: Monotonicity of z and non-monotonicity of $p(X = a|\omega)$ and $p(X = a|t)$ in q. Parameters: $\varepsilon = 0.3$, $W = 3$, $C = 4$, $K = 1/3$, $p_A = 0.5$, $\alpha = 0.4$, $\beta = 0.3$, $n = 101$. Left graph: z is monotonic in q, yet $p(X = A|t)$ (middle graph) $p(X = a|\omega)$ (right graph) are not. The dashed lines represent $p(X = a|t = r)$ and $p(X = a|\omega = R)$ and the continuous lines $p(X = a|t = a)$ and $p(X = a|\omega = A)$.

willingness to reject the unique non-babbling equilibrium under a higher q must occur at a (weakly) higher z.

Note, however, that the overall effect of an increase in q on the probability that the committee accepts the innovation depends on whether the increase in z is high enough to outweigh the increase in the decision threshold. In general, the relation between q and the probability of acceptance is non-monotonic and, somewhat surprisingly, as shown in Figure 2, a higher value of q may imply a higher acceptance probability $p(X = a)$.

The comparative statics with respect to n and ϵ are non-monotonic, and therefore cannot be generally classified by sign. However, these non-monotonicities represent interesting cases that we explore further. Fixing the behavior of all members other than i, increasing n has two effects: (1) Fixing the fraction of a signals received by other experts, i's confidence on his inference on the state of the world increases. Therefore, conditional on i being pivotal, voting for a becomes less 'risky.' (2) The probability of i being pivotal decreases, and therefore so does the importance of his payoffs that condition on being pivotal. Thus, the relative salience of disesteem payoffs–which accrue regardless of whether he is pivotal or not–increases. Proposition 3 shows that for large enough increases in n, (2) always predominates. However, for small increases in n this may not be the case, as seen in Figure 3.

The case of ε is also interesting. On the one hand, a smaller ε implies that any expert's signal is more likely to reflect the state of the art, and indirectly the state of the world. From this perspective, any given member i becomes more willing to vote for a upon receiving an a signal. On the other hand, a lower ε implies that all else equal, i has a better prediction of how the other experts will vote. In particular, under a smaller ε, holding the strategy

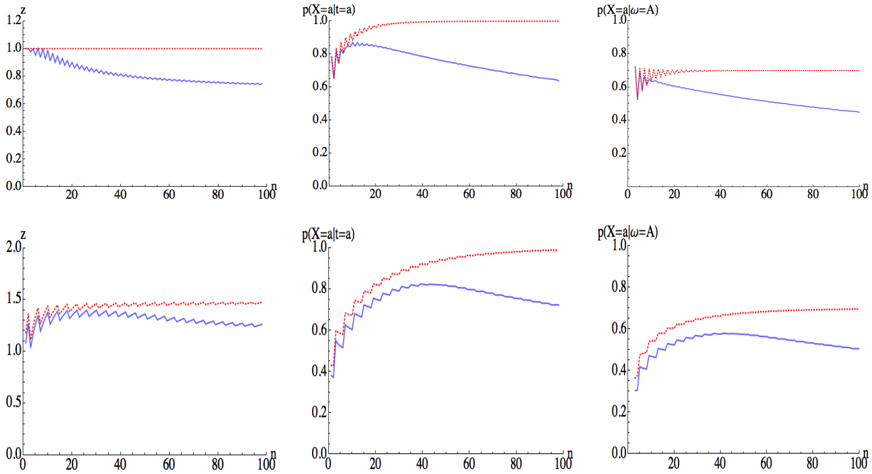

Figure 3: Non-monotonicity of $\sigma(a)$, $p(X = a|t = a)$ and $p(X = a|\omega = A)$ in n. The jaggedness of the figures is due to the discreteness of the problem (we are interested in the "low frequency variation"). Parameters: $\varepsilon = 0.3$, $W = 3$, $C = 4$, $p_A = 0.5$, $\alpha = 0.4$, $\beta = 0.3$ Top Figures: $q = 0.5$, z is weakly monotonic, yet $p(X = a|t = a)$ and $p(X = a|\omega = a)$ is non-monotonic. Bottom Figures: $q = 0.75$, None of z, $p(X = a|t = a)$ and $p(X = a|\omega = A)$ are monotonic. The smooth lines represent the case $K = \frac{1}{3}$. As a benchmark, the dotted lines represent the situation with no disesteem payoff ($K = 0$).

used by other experts constant, upon receiving an a signal i is more confident that other members will vote a.[23] Therefore, conditional on receiving signal a expert i is less likely to be pivotal, and has a smaller incentive for vote a than with the higher ε. These competing effects can result in non-monotonicity, which can be seen in the example shown in Figure 4.

4 Extensions

We first consider the effect of information pooling of the experts' signals on the collective action problem introduced by disesteem payoffs. Since committees of experts most often discuss prior to voting, the committee members can then share their private signals with the other members of the committee. Second, we characterize the extent to which Proposition 3 is robust to dilution of disesteem payoffs. It is reasonable that the size of the disesteem payoff is smaller in a larger committee, since more individuals share the blame for approving a bad innovation. The main result still obtains, however, as long as the speed of dilution is "slow enough."[24]

[23]Note that equilibria are always ordered, in the sense that $\sigma(a) > \sigma(r)$.

[24]In a different vein, a possible extension of our model is to include disesteem payoffs that are realized when the committee rejects the innovation and esteem payoffs that accrues to members who supported a successful

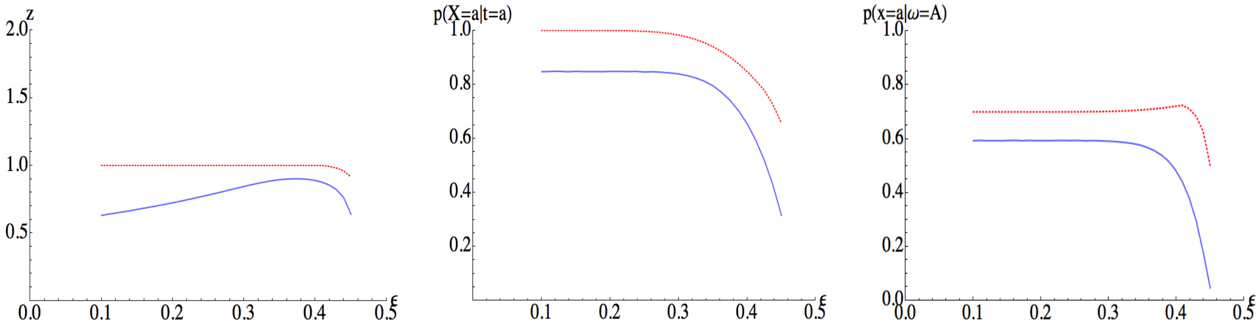

Figure 4: Non-monotonicity of z, $p(X = a|t = a)$, and $p(X = a|\omega = A)$ in ε. Parameters: $n = 25$, $W = 3$, $C = 4$, $p_A = 0.5$, $\alpha = 0.4$, $\beta = 0.3$. Note that z is non-monotonic. Despite the initial rise in z, $p(X = a|t = a)$ is weakly decreasing throughout, and $p(X = a|\omega = A)$ is non-monotonic. The smooth lines represent the case of $K = \frac{1}{3}$. As a benchmark, the dotted lines represent the situation with no disesteem payoff ($K = 0$).

4.1 Information Pooling

Suppose each member reveals his private signal to the committee prior to voting, which implies that the precision of the information available to individual committee members is increasing in n. Since each expert now has the same information set, $\{s_1, ..., s_n\} \equiv \{s_i\}^n$, symmetric strategies imply that every member chooses to approve or reject the innovation with the same probability, which we denote by $\sigma(s_1, s_2, ..., s_n)$.

Corollary **5 (Robustness to Increasing Precision)**
Let $K > 0$ and let each committee member observe the full set of signals $\{s_i\}^n$. Following the notation of Proposition 3, for any q-rule, $p_{\sigma^n}(X = a) \to 0$ as $n \to \infty$. That is, for all $\delta > 0$, there exists n_δ such that for all $n > n_\delta$, $p_{\sigma^n}(X = a) < \delta$.

Corollary 5 follows directly from Proposition 3 since that result holds under perfect signals ($\varepsilon = 0$). More specifically, the structure of each expert's decision rule is unchanged, with the only difference being that the probabilities condition on the full set of signals:

$$Kp_\sigma(X = a, \omega = R|\{s_i\}^n) - Wp_\sigma(piv_i, \omega = A|\{s_i\}^n) + Cp_\sigma(piv_i, \omega = R|\{s_i\}^n) \leq 0 \qquad (1'')$$

Corollary 5 extends Proposition 3 and shows that under pooling of private signals, large committees will reject innovations almost surely.

In contrast to Corollary 5, without disesteem payoffs ($K = 0$) and under information pooling, there always exists a *single-agent efficient* equilibrium. We call an equilibrium single-agent

innovation. Notice that Proposition 3 is robust to such "idiosyncratic" payoffs if they are sufficiently close to zero, which follows from the fact that the RHS of $(1')$ is bounded away from zero for any equilibrium where the innovation is accepted with positive probability.

efficient if for every realization of signals $(s_1, s_2, ..., s_n)$, the committee approves the innovation if, and only if, it would be approved by a single-agent committee whose only member has access to $(s_1, s_2, ..., s_n)$.[25] This disparity with respect to $K > 0$ stems from the fact that with disesteem payoffs and any q-rule different from unanimity ($q < \frac{n-1}{n}$), payoffs are heterogeneous since it is only the experts who vote to accept that are exposed to disesteem payoffs. Hence, even though the experts agree on the optimal committee outcome, the collective action problem remains and each expert will face an incentive to free-ride and vote to reject.

Given Corollary 5, the result of the following proposition is remarkable: under unanimity and $K > 0$ a single-agent efficient equilibrium exists, and this is true also if the experts have the option to misrepresent their signals. In appendix A, we present a formal model with an initial stage of costless communication.

Proposition 4 **(Efficiency under Unanimity)**
Let $K = 0$ and consider any $q < 1$. Under information pooling there exists a single-agent efficient equilibrium, even if experts are able to mis-represent their signals. This is only true for $K > 0$ when $\frac{n-1}{n} \leq q < 1$ (unanimity).

The intuition for the proof of Proposition 4 is simple: Assuming all experts observe the full set of signals and, if approval is single-agent efficient, then no agent has an incentive to deviate. This is true when $K = 0$ because the payoff of each committee member coincides, for all possible outcomes, with that of the single agent. This is also the case under unanimity and $K > 0$, since each expert's vote is pivotal when the committee accepts and thus the expert faces exactly the same decision problem as the representative single agent. Furthermore, by mis-representing his signal, an expert is only able to alter the final outcome in the same way as he would by correctly presenting his signal and modifying his vote (see also Coughlan (2000)).

Proposition 4 shows that costless communication can lead to efficiency in the presence of disesteem payoffs, *but only if it is paired with a unanimity rule.* This result supports Coughlan (2000) argument for why certain committees, such as juries, are better off using a unanimity rule, despite its disadvantages (see Feddersen and Pesendorfer (1998)).

Our analysis of information aggregation under information pooling yields the following two insights. First, it shows that, to the extent that it provides committee members with more information, information sharing cannot overcome the collective action problem in committees. Second, it highlights the fact that in our setting, the conflict of interest among committee members depends on the decision rule.

[25] This representative agent receives a payoff of 0 if the innovation is rejected, W if it is accepted and $\omega = A$ and $-(C + K)$ if he accepts it and $\omega = R$.

4.2 Dilution of Disesteem Payoffs

Lastly, we discuss the extent to which our result is robust to dilution of disesteem payoffs. It is reasonable that the size of the disesteem payoff is smaller in a larger committee, since more individuals share the blame for approving a bad innovation. The main result still obtains, however, as long as the speed of dilution is "slow enough." Consider the following variation of the sequence of games (G_n^K) analyzed in Section 3. We let the game $(G_n^{f(n)})$ be just as (G_n^K) with the exception of the disesteem component of the payoffs, which we define in a slightly more general way. In particular let the payoff function be given by:

$$U(X, v_i, \omega) = \begin{cases} 0 & \text{if } X = r \\ W & \text{if } X = a,\ \omega = A \\ -C & \text{if } X = a,\ \omega = R,\ v_i = r \\ -(C + f_n(n, \{v_i\}_{i=1}^n)) & \text{if } X = a,\ \omega = R,\ v_i = a \end{cases}$$

Where for each n, $f(n, \{v_i\}_{i=1}^n) > 0$ and is bounded from below by some deterministic function g_n of n, $g_n : \mathbb{N}^+ \to \mathbb{R}$ such that the sequence $g_n(n)$ converges to 0 at a lower speed than $\frac{1}{\sqrt{n}}$. That is, for all n, $f_n(n, \{v_i\}_{i=1}^n) \geq g_n(n)$, where $\lim\limits_{n\to\infty} \sqrt{n} g_n(n) \to \infty$.

Note that the games (G_n^K) of Section 3 are a special case of this formulation, as $f(n) = K$ being constant in n certainly has the required property ($\lim\limits_{n\to\infty} \sqrt{n}K \to \infty$). This definition also accommodates other interesting cases; for example, let $f(n, \{v_i\}_{i=1}^n) = \frac{K}{log(n)}$.

PROPOSITION 5
Let (f_n) be a sequence of functions satisfying the properties discussed above and consider the sequence of games $G_{n,q}^{f_n}$ and any sequence of symmetric strategy profiles σ^n, such that for each n, σ^n is an equilibrium of $G_{n,q}^{f_n}$. We let $p_{\sigma^n}(X = a)$ denote the probability that the committee accepts the innovation in game $G_{n,q}^{f_n}$, playing according to σ^n. Then, $p_{\sigma^n}(X = a) \to 0$ as $n \to \infty$. That is, for all $\delta > 0$, there exists n_δ such that for all $n > n_\delta$, $p_{\sigma^n}(X = a) < \delta$.

The proof of Proposition 5 is analogous to the proof of Proposition 3, and follows by simply dividing both sides of $(1'')$ by g_n.

5 Conclusion

In this paper, we detail the effect of idiosyncratic disesteem payoffs on information aggregation in committees. We show that under the "state of the art" model of expertise, disesteem payoffs lead large committees to be over-cautious and reject new innovations as individual committee members seek to save face and avoid being blamed for a bad decision. If committee members communicate prior to voting, then over-caution can be addressed by a unanimity rule: if all committee members are responsible for approval, then communication will be effective and the committee will choose the option that maximizes their ex-ante aggregate payoff.

Our paper suggests multiple areas for further study. First, our paper shows that the predictions of models of information aggregation in committees can be sensitive to the standard assumption that experts' signals are generated by the state of the world. It is unlikely that the state of the art decision perfectly identifies the true state of the world; that is, due to imperfect evidence, even the "best" decision might be wrong ex post. If experts' signals are generated by an imperfect state of the art, then information aggregation no longer results in fully accurate decisions. Additionally, the state of the art model implies a particular correlation structure between experts' signals, and the general implications of this correlation warrant further study.

Second, our paper shows that idiosyncratic payoffs can affect information aggregation even when they reinforce common payoffs. Specifically, idiosyncratic payoffs can distort decisions when they introduce asymmetry in payoffs. This asymmetry need not be large; we show here that even a marginal deviation from common payoffs can distort outcomes in large committees. Asymmetry can occur either due to informational asymmetry, e.g. when information regarding the adequacy of a drug is only revealed when the drug is passed, or if the saliency of individual votes vary with the committee outcome. One particularly relevant environment is a political setting, where idiosyncratic payoffs can be interpreted as changes in reelection probabilities. Voting records of politicians are heavily scrutinized in US legislatures, and the saliency of a particular representative's vote might condition on the legislative outcome. Therefore, a particularly interesting area for future study might be the effect of idiosyncratic payoffs on information aggregation in legislatures.

Appendix A: Proofs

We begin by noting that our game is equivalent to a single layer game (that does not make reference to the state of the world ω, but just to the state of the art t) with the following payoff function:

$$U(X, v_i, \omega) = \begin{cases} 0 & \text{if } X = r \\ Wp(\omega = A|t = a) - Cp(\omega = R|t = a) & \text{if } X = a,\, t = a,\, v_i = r \\ -(Cp(\omega = R|t = r) - Wp(\omega = A|t = r)) & \text{if } X = a,\, t = r,\, v_i = r \\ Wp(\omega = A|t = a) - (C + K)p(\omega = R|t = a) & \text{if } X = a,\, t = a,\, v_i = a \\ -((C + K)p(\omega = R|t = r) - Wp(\omega = A|t = r)) & \text{if } X = a,\, t = r,\, v_i = a \end{cases}$$

Unless otherwise stated we establish the following results, by analyzing the slightly more general game with the following payoff structure:[26]

$$U(X, v_i, \omega) = \begin{cases} 0 & \text{if } X = r \\ W' & \text{if } X = a,\, t = a,\, v_i = r \\ -C' & \text{if } X = a,\, t = r,\, v_i = r \\ W' - K_1 & \text{if } X = a,\, t = a,\, v_i = a \\ -C' - K_2 & \text{if } X = a,\, t = r,\, v_i = a \end{cases}$$

Our game is a special case of this second one. However this second structure is strictly more general. For instance, in our game we would always have $K_1 = Kp(\omega = R|t = a)$, $K_2 = Kp(\omega = R|t = r)$ which implies $K_1 < K_2$ since $\beta < 1 - \alpha$. Denote the set of all agents $j \neq i$, such that $v_j = a$, by H_i, and let piv_i, denote the event $|H_i| = \lfloor nq \rfloor$. Expert i finds it optimal to set $v_i = a$ upon receiving signal s_i, when all other agents are using strategy σ if, and only if, $R_{s_i}(p_A, \alpha, \beta, \varepsilon, q, n, K_1, K_2, W, C, \sigma)$, abbreviated $R_{s_i}(n, \sigma)$, is nonpositive :

$$\begin{aligned} R_{s_i}(n, \sigma) = {} & K_1 p_\sigma \left(\frac{|H_i| + 1}{n} > q, t = a|s_i \right) + K_2 p_\sigma \left(\frac{|H_i| + 1}{n} > q, t = r|s_i \right) \\ & - W' p_\sigma(piv_i, t = a|s_i) + C' p_\sigma(piv_i, t = r|s_i) \leq 0 \end{aligned}$$

Proof of Lemma 1:

Assume $R_r(n, \sigma) \leq 0$, and that at least one of $\sigma(r)$ or $\sigma(a)$ is strictly positive. Then:

$$-W' p_\sigma(piv_i|t = a) p_a \varepsilon + C' p_\sigma(piv_i|t = r)(1 - p_a)(1 - \varepsilon) +$$

[26] We are abusing notation slightly, as we are just referring to the structure of the payoff function. The coincidence of W and C in the representations above and below does not mean that they are equal.

$$K_1 p_\sigma \left(\frac{|H_i|+1}{n} > q|t=a\right) p_a\varepsilon + K_2 p_\sigma \left(\frac{|H_i|+1}{n} > q|t=r\right)(1-p_a)(1-\varepsilon) \leq 0$$

$$\equiv p_a\varepsilon\left(-W'p_\sigma(piv_i|t=a) + K_1 p_\sigma\left(\frac{|H_i|+1}{n} > q|t=a\right)\right) \leq$$

$$(1-p_a)(1-\varepsilon)\left(-C'p_\sigma(piv_i|t=r) - K_2 p_\sigma\left(\frac{|H_i|+1}{n} > q|t=r\right)\right)$$

$$\Rightarrow p_a\varepsilon\left(-W'p_\sigma(piv_i|t=a) + K_1 p_\sigma\left(\frac{|H_i|+1}{n} > q|t=a\right)\right) < 0$$

Since $\left(-C'p_\sigma(piv_i|t=r) - K_2 p_\sigma\left(\frac{|H_i|+1}{n} > q|t=r\right)\right) < 0$, given that since by assumption at least one of $\sigma(a)$ and $\sigma(r)$ is positive, and therefore $p\left(\frac{|H_i|+1}{n} > q|t=r\right) > 0$.

Now,

$$R_a(n,\sigma) =$$

$$(1-\varepsilon)p_a\left(-W'p_\sigma(piv_i|t=a) + K_1 p_\sigma\left(\frac{|H_i|+1}{n} > q|t=a\right)\right) -$$

$$\varepsilon(1-p_a)\left(-C'p_\sigma(piv_i|t=r) - K_2 p_\sigma\left(\frac{|H_i|+1}{n} > q|t=r\right)\right)$$

$$< \varepsilon p_a\left(-W'p_\sigma(piv_i|t=a) + K_1 p_\sigma\left(\frac{|H_i|+1}{n} > q|t=a\right)\right) -$$

$$(1-\varepsilon)(1-p_a)\left(-C'p_\sigma(piv_i|t=r) - K_2 p_\sigma\left(\frac{|H_i|+1}{n} > q|t=r\right)\right)$$

$$= R_r(n,\sigma)$$

The strict inequality follows from the facts that:
(1) $\left(-W'p_\sigma(piv_i|t=a) + K_1 p_\sigma\left(\frac{|H_i|+1}{n} > q|t=a\right)\right) < 0$ and
$\left(-C'p_\sigma(piv_i|t=r) - K_2 p_\sigma\left(\frac{|H_i|+1}{n} > q|t=r\right)\right) < 0$ (proved above)
(2) $\varepsilon < \frac{1}{2}$ and therefore $(1-\varepsilon) > \varepsilon$. ▮*(Lemma 1)*

Proof of Corollary 1:
The proof of Corollary 1 proceeds by showing that if an equilibrium $\sigma(a), \sigma(r)$ is not babbling ($\sigma(a) = \sigma(r) = 0$) then it necessarily must be in one of the two categories (1) $\sigma(r) = 0$, $\sigma(a) > 0$, or (2) $\sigma(r) > 0$, $\sigma(a) = 1$. So suppose the equilibrium is not a babbling equilibrium. There are two possibilities: either $\sigma(r) = 0$ or $\sigma(r) > 0$. If $\sigma(r) = 0$, then we are done, since by assumption the equilibrium is not babbling, and therefore it must be the case that $\sigma(a) > 0$, in which case the equilibrium is in category (1). So assume $\sigma(r) > 0$. Then for a player to be best responding it must be the case that $R_r(n,\sigma) \leq 0$, as otherwise he would find it strictly better to set $v_i = r$ (contradicting $\sigma(r) > 0$). By Lemma 1 this implies

$R_a(n,\sigma) < R_r(n,\sigma) \le 0$, and therefore the expert finds it strictly optimal to set $v_i = a$. It must therefore be the case that $\sigma(a) = 1$ and the equilibrium is in category (2). ▮*(Corollary 1)*

Proof of Proposition 1:

We begin by showing part (1). The proof establishes that $R(n,z)$ has the single crossing property in $z \in (0,2]$. Given that $R_r(n,(1,1)) > 0$ for all $q < \frac{n-1}{n}$,[27] it follows that any crossing must actually take place in $(0,2)$, so it suffices to show that $R(n,z)$ has the single crossing property in $z \in (0,2)$. We proceed as follows: (A) We first show that $R_a(n,(z,0))$ has the single crossing property for $z \in (0,1]$ and $R_r(n,(1,z-1))$ has the single crossing property for $z \in [1,2)$. We do so by relying on the main result of Quah and Strulovici (2012) which provides sufficient and necessary conditions for non-negative sums of functions having the single crossing property to also have the single crossing property. And then (B) we use Lemma 1 to argue that if $R(n,z)$ has a crossing in $(0,1)$ then it cannot have one in $[1,2)$.

(A) $R_a(n,(z,0))$ has the single crossing property for $z \in (0,1]$ and, $R_r(n,(1,z-1))$ for $z \in [1,2)$.

Note that for $z \in (0,1]$, $R_a(n,(z,0))$, is just

$$R_a(n,(z,0)) = Kp_z(X = a, \omega = R|a) - Wp_z(piv_i, \omega = A|a) + Cp_z(piv_i, \omega = R|a).$$

Which is a special case of the general form:

$$G(y) : D_1 p_z(piv_i|t = r) + D_2 p_z\left(\frac{|H_i|}{n} > q|t = a\right) + D_3 p_z\left(\frac{|H_i|}{n} > q|t = r\right) - D_4 p_z(piv_i|t = a).$$ [28]

where D_1, D_2, D_3 and D_4 are nonnegative constants. The result will follow as a direct application of Lemma 1 in the appendix of Quah and Strulovici (2012). For convenience we reproduce the Lemma and the relevant definitions below (as they apply to our paper).

DEFINITION **2 (Quah and Strulovici (2012))** *Let S be partially ordered set. A function $f : S \to \mathbb{R}$ satisfies the single crossing property (SCP) if:*

- $f(s) \ge (>)0 \Longrightarrow f(s') \ge (>)0$ *whenever* $s' > s$.

Note that $G(z)$ has at most one solution if and only if it satisfies (SCP).[29] $G(z)$ is a non-

[27] Given that as long as the decision rule is not unanimity, if all other agents surely accept the innovation, any agent's unique best reply is to reject it.

[28] We suppress explicitly noting the dependence on n as throughout this section n is kept constant.

[29] Note $G(0) = 0$ so in principle it could satisfy (SCP) by being constant at 0 for all z or by remaining constant for an interval and then becoming positive. This possibility can be ruled out by verifying that there exist points arbitrarily close to 0 whose image under G is not 0.

negative linear combination of functions that satisfy (SCP) in $(0,1)$.[30] In their work, Quah and Strulovici provide necessary and sufficient conditions under which such linear combinations also satisfy (SCP).

Definition **3 (Quah and Strulovici (2012))** *A pair of functions $f : S \to \mathbb{R}$ and $g : S \to \mathbb{R}$ satisfy the signed ratio monotonicity property (SR) if:*

a) If $g(s) < 0$ and $f(s) > 0$ then $-\frac{g(s)}{f(s)} \geq -\frac{g(s')}{f(s')}$ when $s' > s$.

b) If $g(s) > 0$ and $f(s) < 0$ then $-\frac{f(s)}{g(s)} \geq -\frac{f(s')}{g(s')}$ when $s' > s$.

Lemma **2 (Quah and Strulovici (2012) (Lemma 1 in the Appendix))**
Let $\mathcal{F} = \{f_i\}_{1 \leq i \leq M}$ be a family of functions satisfying (SCP) such that any two members satisfy (SR). Then $\sum\limits_{i=1}^{M} \alpha_i f_i$, where $\alpha_i \geq 0$ for all i, satisfies (SCP).

Consider the family of functions (1) $p_y(piv_i|t=r)$, (2) $p_y\left(\frac{|H_i|}{n} > q|t=a\right)$, (3) $p_y\left(\frac{|H_i|}{n} > q|t=r\right)$ and (4) $-p_z(piv_i|t=a)$, and notice that they all satisfy (SCP) when $z \in (0,1]$. The first 3 are nonnegative, so any pair among them satisfies (SR). It therefore suffices to show that all the pairs formed by (4) and each of (1), (2) and (3) satisfy (SR).

Lemma **3**
All pairs in the family $\{p_z(piv_i|t=r), p_z\left(\frac{|H_i|}{n} > q|t=a\right), p_z\left(\frac{|H_i|}{n} > q|t=r\right), -p_z(piv_i|t=a)\}$ satisfy (SR) for $z \in (0,1]$.

Proof of Lemma 3:
As stated above, we just need to check the pairs involving $-p_z(piv_i|t=a)$, as all other pairs involving components with the same sign satisfy the condition vacuously.

<u>(1) $-p_z(piv_i|t=a)$ and $p_z(piv_i|t=r)$</u>. In this case, the condition is equivalent to $\frac{p_z(piv_i|t=a)}{p_z(piv_i|t=r)}$ being non-increasing in z.

$p_z(piv_i|t=a) = \binom{n-1}{\lfloor nq \rfloor} \mu_a^{\lfloor nq \rfloor}(1-\mu_a)^{n-1-\lfloor nq \rfloor}$ and

$p_z(piv_i|t=r) = \binom{n-1}{\lfloor nq \rfloor} \mu_r^{\lfloor nq \rfloor}(1-\mu_r)^{n-1-\lfloor nq \rfloor}$

where $\mu_a = (1-\varepsilon)z$ and $\mu_r = \varepsilon z$. Therefore:

[30] They do so trivially, as each of the four functions ((1) $p_z(piv_i|t=r)$, (2) $p_z\left(\frac{|H_i|}{n} > q|t=a\right)$, (3) $p_z\left(\frac{|H_i|}{n} > q|t=r\right)$ and (4) $-p_z(piv_i|t=a)$) are 0 when evaluated at $z=0$, and then either always positive (the first 3) or alway negative (the 4^{th}).

$$\frac{p_z(piv_i|t=a)}{p_z(piv_i|t=r)} = \left(\frac{1-\varepsilon}{\varepsilon}\right)^{\lfloor nq \rfloor} \left(\frac{1-(1-\varepsilon)z}{1-\varepsilon z}\right)^{n-1-\lfloor nq \rfloor}$$

This expression is non-increasing in z, if for all z, $z' \in (0,1]$ where $z < z'$ we have

$$\left(\frac{1-(1-\varepsilon)z}{1-\varepsilon z}\right) \geq \left(\frac{1-(1-\varepsilon)z'}{1-\varepsilon z'}\right)$$

which can be seen to be true whenever $\varepsilon \leq 0.5$.

(2) $-p_z(piv_i|t=a)$ and $p_z\left(\frac{|H_i|}{n} > q|t=a\right)$. This case amounts to showing that the hazard ratio of the binomial distribution evaluated at $\lfloor nq \rfloor$ is decreasing for all success probabilities between 0 and $(1-\varepsilon)$.[31] More generally, we will show that the hazard ratio evaluated at k:

$$\frac{\binom{m}{k}\mu_a^k(1-\mu_a)^{m-k}}{\sum_{j=k+1}^{m}\binom{m}{j}\mu_a^j(1-\mu_a)^{m-j}}$$, is decreasing in $\mu_a \in [0,1)$ for all m.

The hazard ratio is decreasing if, and only if, its multiplicative inverse is increasing, which is true since:

$$\frac{\sum_{j=k+1}^{m}\binom{m}{j}\mu_a^j(1-\mu_a)^{m-j}}{\binom{m}{k}\mu_a^k(1-\mu_a)^{m-k}} = \sum_{j=k+1}^{m}\frac{\binom{m}{j}}{\binom{m}{k}}\left(\frac{\mu_a}{1-\mu_a}\right)^{j-k}.$$

and $\frac{\mu_a}{1-\mu_a}$ is strictly increasing in $\mu_a \in [0,1)$ as required.

(3) $-p_z(piv_i|t=a)$ and $p_z\left(\frac{|H_i|}{n} > q|t=r\right)$. The analogous expression to the inverse hazard ratio in this case (as a function of z) is:

$$\sum_{j=k+1}^{m}\frac{\binom{m}{j}}{\binom{m}{k}}\left(\frac{(\varepsilon z)^j(1-\varepsilon z)^{m-j}}{((1-\varepsilon)z)^k(1-(1-\varepsilon)z)^{m-k}}\right).$$

The derivative of $\left(\frac{(\varepsilon z)^j(1-\varepsilon z)^{m-j}}{((1-\varepsilon)z)^k(1-(1-\varepsilon)z)^{m-k}}\right)$ w.r.t. z is:

$$\left(\frac{(\varepsilon z)^j(1-\varepsilon z)^{m-j-1}}{z((1-\varepsilon)z)^k(1-(1-\varepsilon)z)^{m-k+1}}\right)((j-k)+z((1-2\varepsilon)m-j(1-\varepsilon)+k\varepsilon))$$

The sign of this expression just depends on the sign of the linear function of z , $(j-k)+z((1-2\varepsilon)m-j(1-\varepsilon)+k\varepsilon)$ which can be straightforwardly verified to be always non-negative for $\varepsilon < 0.5$, and $k < j \leq m$. We therefore have that the sum above is nondecreasing in z and $\frac{p_z(piv_i|t=a)}{p_z\left(\frac{|H_i|}{n}>q|t=a\right)}$ is nonincreasing in z, as required. ∎ *(Lemma 3)*

We can therefore apply Quah and Strulovici's Lemma (Lemma 2 above) to conclude that $G(z)$ can have at most one other solution (other than $\sigma(a)=0$), in the interval $z \in [0,1]$. We end by noting that the "extreme" configuration $z=1$, corresponding to $\sigma(a)=1$, $\sigma(r)=0$

[31] These are the bounds for μ_a as z varies between 0 and 1.

requires just $G(1) \leq 0$ and not the more restrictive $G(1) = 0$. The definition of (SCP) also implies that any crossing must take place from below the x-axis. But this means that either $G(z)$ is negative throughout the range (with the exception of $G(0) = 0$), in which case $z = 1$ defines an equilibrium provided that $R_r(n, z) \geq 0$ (and is the only one), or it crosses the x-axis, but if this is the case then having $G(1) \leq 0$ would require a second crossing, which we have shown to be impossible.

We now verify the analogous steps for the case $z \in [1, 2)$
For $z \in (0, 1]$, $R_r(n, (1, z-1))$, is just

$R_r(n, (1, z-1)) = Kp_z(X = a, \omega = R|r) - Wp_z(piv_i, \omega = A|r) + Cp_z(piv_i, \omega = R|r)$.
Once more, it is a special case of the form:

$$M(z) : D'_1 p_z(piv_i|t = r) + D'_2 p_z\left(\frac{|H_i|}{n} > q|t = a\right) +$$
$$D'_3 p_z\left(\frac{|H_i|}{n} > q|t = r\right) - D'_4 p_z(piv_i|t = a) = 0.$$

for some nonnegative constants D'_1, D'_2, D'_3 and D'_4. However, z now belongs to $[1, 2)$. Since $M(z)$ and $G(z)$ have the same form, analogous arguments to those used in (1) (2) and (3) above apply, the main difference being that now $\mu_a = (1-\varepsilon) + \varepsilon(z-1)$ and $\mu_r = (1-\varepsilon)(z-1) + \varepsilon$. Or letting $y = z - 1$, $\mu_a = (1-\varepsilon) + \varepsilon y$ and $\mu_r = (1-\varepsilon)y + \varepsilon$, $y \in [0, 1)$.

Lemma 4
All pairs in the family $\{p_z(piv_i|t = r), p_z\left(\frac{|H_i|}{n} > q|t = a\right), p_z\left(\frac{|H_i|}{n} > q|t = r\right), -p_z(piv_i|t = a)\}$ *satisfy* (SR) *for* $z \in [1, 2)$.

Proof of Lemma 4:
As above we just need to check the pairs involving $-p_z(piv_i|t = a)$, as all other pairs, involving components with the same sign, satisfy the condition vacuously.
(1) $-p_z(piv_i|t = a)$ and $p_z(piv_i|t = r)$. In this case, the condition is equivalent to $\frac{p_z(piv_i|t=a)}{p_z(piv_i|t=r)}$ being non-increasing in z.

$$\frac{p_z(piv_i|t = a)}{p_z(piv_i|t = r)} = \left(\frac{1-\varepsilon}{\varepsilon}\right)^{\lfloor nq \rfloor - n + 1} \left(\frac{(1-\varepsilon) + \varepsilon y}{(1-\varepsilon)y + \varepsilon}\right)^{\lfloor nq \rfloor}$$

This expression is non-increasing in y, if for all y, $y' \in (0, 1]$ where $y < y'$ we have

$$\left(\frac{(1-\varepsilon) + \varepsilon y}{(1-\varepsilon)y + \varepsilon}\right) \geq \left(\frac{(1-\varepsilon) + \varepsilon y'}{(1-\varepsilon)y' + \varepsilon}\right)$$

which can be seen to be true whenever $\varepsilon \leq 0.5$.

(2) $-p_z(piv_i|t = a)$ and $p_z\left(\frac{|H_i|}{n} > q|t = a\right)$. The argument presented above (for $z \in (0, 1]$) just relied on $\mu_a \in [0, 1)$, which contains the full range of μ_a, $(1-\varepsilon, 1)$, for $z \in (1, 2)$, so it applies directly to this case.

(3) $-p_z(piv_i|t=a)$ and $p_z\left(\frac{|H_i|}{n} > q|t=r\right)$. The analogous expression[32] to the inverse hazard ratio in this case (as a function of z) is:

$$\sum_{j=k+1}^{m} \frac{\binom{m}{j}}{\binom{m}{k}} \left(\frac{((1-\varepsilon)y+\varepsilon)^j((1-\varepsilon)(1-y))^{m-j}}{((1-\varepsilon)+\varepsilon y)^k(\varepsilon(1-y))^{m-k}} \right).$$

The derivative of $\left(\frac{((1-\varepsilon)y+\varepsilon)^j((1-\varepsilon)(1-y))^{m-j}}{((1-\varepsilon)+\varepsilon y)^k(\varepsilon(1-y))^{m-k}} \right)$ w.r.t. y is:

$$\left(\frac{(1-\varepsilon)(y(1-\varepsilon)+\varepsilon)^{j-1}((1-y)(1-\varepsilon))^{m-j-1}}{(1-(1-y)\varepsilon)^{k+1}(\varepsilon(1-y))^{m-k}} \right)(j(1-\varepsilon)-\varepsilon k + y(j\varepsilon - k(1-\varepsilon)))$$

The sign of this expression just depends on the sign of the linear function of z, $j(1-\varepsilon) - \varepsilon k + z(j\varepsilon - k(1-\varepsilon))$ which can be straightforwardly verified to be always non-negative for $\varepsilon < 0.5$, and $k < j \leq m$. We therefore have that the sum above is nondecreasing in z and $\frac{p_z(piv_i|t=a)}{p_z\left(\frac{|H_i|}{n} > q|t=a\right)}$ is nonincreasing in z, as required. ▮ *(Lemma 4)*

(B) If $R(n,z)$ has a crossing in $z \in [1,2)$, the it does not have a crossing in $z \in (0,1)$.

If there is a crossing with $z \in [1,2)$ then $R_r(n,(1,0)) \leq 0$, as the crossing must be from below the $x-axis$. By Lemma 1, this implies $R_a(n,(1,0)) < R_r(n,(1,0)) \leq 0$, and therefore there can't be any crossing in $z \in (0,1)$ ▮*(Part (1) Proposition 1)*

So far we have shown that when a non-babbling equilibrium exists it is unique. We now go on to the proof of part (2) of Proposition 1. For that purpose we study the willingness to reject, as a function of $m = \lfloor nq \rfloor$ and denote it $R(m)$ ($R_a(m)$ and $R_r(m)$ when referring to the two continuous segments (as functions of σ)).[33]

Lemma **5**
$-R(m)$ has the single crossing property (as a function of $m = \lfloor nq \rfloor$), for $m \in \{0,1,2,...,n-1\}$, for all $z \in (0,2)$.

The proof shows that when $z \in (0,2)$, the family
$\{-p_z(piv_i|t=r), -p_z(|H_i| > m|t=a), -p_z(|H_i| > m|t=r), p_z(piv_i|t=a)\}$ satisfies (SR). As argued in the proof of Lemma 3, $R_a(m)$ and $R_r(m)$ are both nonnegative linear combinations of this family of functions (they only differ in the values of the coefficients in the linear combination). Lemma 2 from Quah and Strulovici (2012) then immediately leads to the result.

[32] As above, we let y=z-1, and therefore $y \in [0,1)$.

[33] Note that the dependence of $R_{s_i}(p_A, \alpha, \beta, \varepsilon, q, n, K, W, C, \sigma)$ on q is only through the *number of votes* required for acceptance, that is $\lfloor nq \rfloor$.

Proof of Lemma 5:

We only need to check the pairs involving $p_z(piv_i|t=a)$, as all other pairs, involving components with the same sign, satisfy the condition vacuously.

<u>(1) $-p_z(piv_i|t=r)$ and $p_z(piv_i|t=a)$</u> . In this case, the condition is equivalent to $\frac{p_z(piv_i|t=r)}{p_z(piv_i|t=a)}$ being non-increasing in m, or equivalently $\frac{p_z(piv_i|t=a)}{p_z(piv_i|t=r)}$ being non-decreasing in m.

(1a) When $z \in (0,1]$, $\dfrac{p_z(piv_i|t=a)}{p_z(piv_i|t=r)} = \left(\dfrac{1-\varepsilon}{\varepsilon}\right)^m \left(\dfrac{1-(1-\varepsilon)z}{1-\varepsilon z}\right)^{n-1-m}$

Note that

$$\left(\frac{1-\varepsilon}{\varepsilon}\right)^{m+1} \left(\frac{1-(1-\varepsilon)z}{1-\varepsilon z}\right)^{n-1-(m+1)} > \left(\frac{1-\varepsilon}{\varepsilon}\right)^{m} \left(\frac{1-(1-\varepsilon)z}{1-\varepsilon z}\right)^{n-1-m}$$

if and only if:

$\left(\dfrac{1-\varepsilon}{\varepsilon}\right) > \left(\dfrac{1-(1-\varepsilon)z}{1-\varepsilon z}\right)$. Which is true for all $\varepsilon < \frac{1}{2}$.

(1b) When $z \in (1,2)$, $\dfrac{p_z(piv_i|t=a)}{p_z(piv_i|t=r)} = \left(\dfrac{1-\varepsilon}{\varepsilon}\right)^{m-n+1} \left(\dfrac{(1-\varepsilon)+\varepsilon y}{(1-\varepsilon)y+\varepsilon}\right)^{m}$

where $y = z-1$. Note that,

$$\left(\frac{1-\varepsilon}{\varepsilon}\right)^{m-n+2} \left(\frac{(1-\varepsilon)+\varepsilon y}{(1-\varepsilon)y+\varepsilon}\right)^{m+1} > \left(\frac{1-\varepsilon}{\varepsilon}\right)^{m-n+1} \left(\frac{(1-\varepsilon)+\varepsilon y}{(1-\varepsilon)y+\varepsilon}\right)^{m}$$

if and only if:

$\left(\dfrac{1-\varepsilon}{\varepsilon}\right) > \left(\dfrac{(1-\varepsilon)y+\varepsilon}{(1-\varepsilon)+\varepsilon y}\right)$. Which is true for all $\varepsilon < \frac{1}{2}$.

<u>(2) $p(piv_i|t=a)$ and $-p\left(|H_i| > m|t=a\right)$</u>. This case amounts to showing that the hazard ratio of the binomial distribution is non-decreasing in $m \in \{0, ..., n-1\}$. That is:

$\dfrac{\binom{n-1}{m}\mu_a^m(1-\mu_a)^{n-1-m}}{\sum_{j=m+1}^{n-1}\binom{n-1}{j}\mu_a^j(1-\mu_a)^{n-1-j}}$, is non-decreasing in m, for all $\mu_a \in (0,1)$.[34]

Consider $m \in \{0, ..., n-2\}$ (so $m+1 \leq n-1$). Then we require:

$$\frac{\binom{n-1}{m}\mu_a^m(1-\mu_a)^{n-1-m}}{\sum_{j=m+1}^{n-1}\binom{n-1}{j}\mu_a^j(1-\mu_a)^{n-1-j}} \leq \frac{\binom{n-1}{m+1}\mu_a^{m+1}(1-\mu_a)^{n-1-(m+1)}}{\sum_{j=m+2}^{n-1}\binom{n-1}{j}\mu_a^j(1-\mu_a)^{n-1-j}}$$

$$\equiv \frac{m+1}{n-1-m} \leq \left(\frac{\mu_a}{1-\mu_a}\right) \frac{\sum_{j=m+1}^{n-1}\binom{n-1}{j}\mu_a^j(1-\mu_a)^{n-1-j}}{\sum_{j=m+2}^{n-1}\binom{n-1}{j}\mu_a^j(1-\mu_a)^{n-1-j}}$$

[34] Note that for $z \in (0,1]$, $\mu_a = (1-\varepsilon)z$ and for $z \in (1,2)$, $\mu_a = (1-\varepsilon)+\varepsilon(z-1)$. In either case $\mu_a \in (0,1)$.

But we can write,

$$\left(\frac{\mu_a}{1-\mu_a}\right)\frac{\sum\limits_{j=m+1}^{n-1}\binom{n-1}{j}\mu_a^j(1-\mu_a)^{n-1-j}}{\sum\limits_{j=m+2}^{n-1}\binom{n-1}{j}\mu_a^j(1-\mu_a)^{n-1-j}}=\frac{(1-\mu_a)\left(\sum\limits_{h=m+2}^{n-1}(\frac{h}{n-h})\binom{n-1}{h}\mu_a^h(1-\mu_a)^{n-1-h}+\frac{\mu_a^n}{1-\mu_a}\right)}{(1-\mu_a)\sum\limits_{j=m+2}^{n-1}\binom{n-1}{j}\mu_a^j(1-\mu_a)^{n-1-j}}$$

$$\geq\frac{\sum\limits_{h=m+2}^{n-1}(\frac{h}{n-h})\binom{n-1}{h}\mu_a^h(1-\mu_a)^{n-1-h}}{\sum\limits_{j=m+2}^{n-1}\binom{n-1}{j}\mu_a^j(1-\mu_a)^{n-1-j}}\geq\frac{\frac{m+1}{n-m-1}\sum\limits_{j=m+2}^{n-1}\binom{n-1}{j}\mu_a^j(1-\mu_a)^{n-1}}{\sum\limits_{j=m+2}^{n-1}\binom{n-1}{j}\mu_a^j(1-\mu_a)^{n-1-j}}=\frac{m+1}{n-m-1}$$

as required.

<u>(3) $p(piv_i|t=a)$ and $-p\left(\frac{|H_i|}{n}>q|t=r\right)$</u>. We need to verify:

$$\equiv\frac{m+1}{n-1-m}\leq\left(\frac{\mu_a}{1-\mu_a}\right)\frac{\sum\limits_{j=m+1}^{n-1}\binom{n-1}{j}\mu_r^{j+1}(1-\mu_r)^{n-1-j}}{\sum\limits_{j=m+2}^{n-1}\binom{n-1}{j}\mu_r^j(1-\mu_r)^{n-1-j}}$$

Using the same arguments as in (2) above we have:

$$\equiv\frac{m+1}{n-1-m}\leq\left(\frac{\mu_r}{1-\mu_r}\right)\frac{\sum\limits_{j=m+1}^{n-1}\binom{n-1}{j}\mu_r^{j+1}(1-\mu_r)^{n-1-j}}{\sum\limits_{j=m+2}^{n-1}\binom{n-1}{j}\mu_r^j(1-\mu_r)^{n-1-j}}$$

$$\leq\left(\frac{\mu_a}{1-\mu_a}\right)\frac{\sum\limits_{j=m+1}^{n-1}\binom{n-1}{j}\mu_r^{j+1}(1-\mu_r)^{n-1-j}}{\sum\limits_{j=m+2}^{n-1}\binom{n-1}{j}\mu_r^j(1-\mu_r)^{n-1-j}}$$

Since $\frac{\mu_a}{1-\mu_a}\geq\frac{\mu_r}{1-\mu_r}$, given that $\sigma(a)\leq\sigma(r)$ throughout our region of interest.[35]

We can therefore apply Lemma 2 (from Quah and Strulovici (2012)) to conclude that $R(m)$ has the single crossing property in $m=\lfloor nq\rfloor$, whenever $z\in(0,2)$. ▮ *(Lemma 5)*

Suppose that there exists a non-babbling equilibrium for some q. Let $q'>q$ and $m=\lfloor nq\rfloor$, $m'=\lfloor nq'\rfloor$. Evaluated at m we have that either (1) $R_a(m)=0$, or (2) $R_r(m)=0$ (depending on what kind of equilibrium we have).[36] Assume that it is of form (1).[37] By Lemma 5 we

[35]When $z\in(0,1]$, $\sigma(r)=0$ and $\sigma(a)>0$. When $z\in(1,2)$, $\sigma(r)<1$ and $\sigma(a)=1$.

[36]If the equilibrium occurs at $z=1$, and is of the form $R_a(m,(1,0))<0$ and $R_r(m,(0,1))>0$, then $R_a(m',(1,0))<0$. If $R_r(m',(0,1))\geq0$, then $z=1$ is also an equilibrium at m'. If $R_r(m',(0,1))<0$ then the last case considered in this paragraph applies.

[37]It will be readily seen that the argument applies to the other case.

know that $-R_a(m) = 0$ has the single crossing property in m, and therefore evaluated at $m' > m$, $R_a(m') \leq 0$.

If $R_a(m') = 0$ then we have an equilibrium, so assume $R_a(m') < 0$. Now lets fix m' and look at R_a as a continuous function of z, $R_a(m', (z, 0))$. If $R_a(m', (1, 0)) \geq 0$ then we have an equilibrium, since given that $R_a(m', (z, 0))$ is continuous in z, it must have crossed the z-axis in order to change sign. If at $z = 1$, $R_a(m', (1, 0)) < 0$, then either $R_r(m', (1, 0)) \geq 0$ (in which case we have an equilibrium at $z = 1$), or $R_r(m', (1, 0)) < 0$. In this case, there are two possibilities: either $R_r(m', (1, 1)) \leq 0$ (which can only be possible if $m' = n - 1$), in which case we have an equilibrium ($z = 2$); or $R_r(m', (1, 1)) > 0$. Then due to the continuity of $R_r(m', (1, z - 1))$ as a function of z, it must cross the z axis at some point in order to change signs, so we have an equilibrium. ▮ *(part (2), Proposition 1)*

To finish the proof of Proposition 1, we go on to part (3). Note that, excluding $z = 1$, $R(p_A, \alpha, \beta, \varepsilon, q, n, K, W, C, z)$ is continuously differentiable in all variables with the exception of n and q. So for any exogenous parameter θ different from n and q, and at all equilibria $z^* \neq 1$, we have that:

$$\frac{\partial z^*(p_A, \alpha, \beta, \varepsilon, q, n, K, W, C)}{\partial \theta} = -\frac{\frac{\partial R(p_A, \alpha, \beta, \varepsilon, q, n, K, W, C, \sigma^*)}{\partial \theta}}{\frac{\partial R(p_A, \alpha, \beta, \varepsilon, q, n, K, W, C, z*)}{\partial z}}$$

As shown in the proof of the uniqueness of the non-babbling equilibrium, as a function of z, $R(p_A, \alpha, \beta, \varepsilon, q, n, K, W, C, z)$ vanishes at most once, and when it does, the crossing is from negative to positive. Relying on the implicit function theorem, this implies that $\frac{\partial R(p_A, \alpha, \beta, \varepsilon, q, n, K, W, C, z^*)}{\partial z} > 0$ (where z^* is just shorthand for $z^*(p_A, \alpha, \beta, \varepsilon, q, n, K, W, C)$). The following lemma therefore immediately follows:

LEMMA **6**

For all p_A, α, β ε, K, W and C, such that $z^*(p_A, \alpha, \beta, \varepsilon, q, n, K, W, C)$ exists and does not equal one, we have:

$$\frac{\partial z^*(p_A, \alpha, \beta, \varepsilon, q, n, K, W, C)}{\partial \theta}(>)(=)(<)0 \textit{ if and only if}$$

$$\frac{\partial R(p_A, \alpha, \beta, \varepsilon, q, n, K, W, C, \sigma^*)}{\partial \theta}(<)(=)(>)0.$$

Part (3) of Proposition 1 follows from Lemma 6 and the single crossing property, which implies that any crossing is from below, and hence $\frac{\partial R(p_A, \alpha, \beta, \varepsilon, q, n, K, W, C, z^*)}{\partial \theta} > 0$ when z^* exists and is different from 1.

For the case in which $z^* = 1$, note that increasing K shifts both continuous branches of R upwards (R as a function of z). Thus, for a small enough increase $z^* = 1$ continues to be an equilibrium, otherwise the new equilibrium (if it exists), must necessarily be at $z < 1$.

Since R_r shifts up, there can't be any crossing with $z \in (1, 2]$. ▮ *(Part (3) Proposition 1)*

Proof of Proposition 3:

We prove the proposition by contradiction. That is, suppose that there exists a sequence of symmetric strategy profiles σ^n such that for each n, σ^n is an equilibrium of $G^K_{n,q}$ and $p_{\sigma^n}(X = a)$ does not converge to 0. This implies that there exists $\delta > 0$ such that for every m, there exists $n_m > m$ with $p_{\sigma^{n_m}}(X = a) > \delta$.

Let i be any expert. Then, by expression (1") i finds it optimal to set $v_i = a$ upon receiving signal s_i if and only if:

$$Wp_{\sigma_n}(piv_i, \omega = A|s_i) - Cp_{\sigma^n}(piv_i, \omega = R|s_i) \geq Kp_{\sigma^n}\left(\frac{|H_i| + 1}{n} > q, \omega = R|s_i\right) \qquad (1')$$

The argument is divided into two parts. First, we show that if $p_{\sigma^n}(X = a) > \delta$ then $Kp_{\sigma^n}\left(\frac{|H_i|+1}{n} > q, \omega = R|s_i\right) \geq K\delta min\{\beta p(t = a|s_i), (1-\alpha)p(t = r|s_i)\}$. Second, we show that the LHS of $(1')$ has an upper bound that is independent of σ_n and which converges to 0. Then putting the two together we arrive at a contradiction of the assumption that $p_{\sigma^n}(X = a)$ does not converge to 0.

<u>Part one: lower bound on the RHS</u>

Note that:

$$\begin{aligned}
Kp_{\sigma^n}\left(\frac{|H_i| + 1}{n} > q, \omega = R|s_i\right) &= Kp_{\sigma^n}\left(\frac{|H_i| + 1}{n} > q, \omega = R, t = a|s_i\right) \\
&\quad + Kp_{\sigma^n}\left(\frac{|H_i| + 1}{n} > q, \omega = R, t = r|s_i\right) \\
&= Kp_{\sigma^n}\left(\frac{|H_i| + 1}{n} > q, \omega = R|t = a\right)p(t = a|s_i) \\
&\quad + Kp_{\sigma^n}\left(\frac{|H_i| + 1}{n} > q, \omega = R|t = r\right)p(t = r|s_i)
\end{aligned}$$

where the second equality follows from the fact that conditional on t, s_i is independent of the state of the world ω and of the other committee members' signals.

Now note that $p_{\sigma^n}(X = a) = p_{\sigma^n}(X = a|t = a)p(t = a) + p_{\sigma^n}(X = a|t = r)(1 - p(t = a))$.

It must therefore be the case that at least one of (I) $p_{\sigma^n}(X = a|t = a) > \delta$ or (II) $p_{\sigma^n}(X = a|t = r) > \delta$ holds. First lets assume (I) holds, $p_{\sigma^n}(X = a|t = a) > \delta$.

Note that $p_{\sigma^n}(X = a|t = a) \leq p_{\sigma^n}\left(\frac{|H_i|+1}{n} > q|t = a\right)$ where $H_i = \{j \neq i : v_j = a\}$, and therefore (I) implies $p_{\sigma^n}\left(\frac{|H_i|+1}{n} > q|t = a\right) > \delta$ which in turn implies $p_{\sigma^n}(\frac{|H_i|+1}{n} > q, \omega = R|t = a) > \delta p(\omega = R|t = a)$, since:

$$p_{\sigma^n}\left(\frac{|H_i| + 1}{n} > q, \omega = R|t = a\right) = p_{\sigma^n}\left(\frac{|H_i| + 1}{n} > q|\omega = R, t = a\right)p(\omega = R|t = a)$$

$$= p_{\sigma^n}\left(\frac{|H_i|+1}{n} > q|t=a\right)p(\omega = R|t=a)$$

where the last equality follows from the fact that the voting behavior of the members only depends on their signals and these are independent from ω conditional on t. We can therefore conclude that:

$$Kp_{\sigma^n}\left(\frac{|H_i|+1}{n} > q, \omega = R|s_i\right) \geq K\delta p(\omega = R|t=a)p(t=a|s_i) = K\delta\beta p(t=a|s_i)$$

If (I) does not hold, then it must be the case that (II) holds, $p_{\sigma^n}(X = a|t=r) > \delta$, case in which we obtain $p_{\sigma^n}(\frac{|H_i|+1}{n} > q, \omega = R|t=r) > \delta p(\omega = R|t=r)$ and we can conclude:

$$Kp_{\sigma^n}\left(\frac{|H_i|+1}{n} > q, \omega = R|s_i\right) \geq K\delta p(\omega = R|t=r)p(t=r|s_i) = K\delta(1-\alpha)p(t=a|s_i).$$

Putting these two cases together it follows that it must be the case, as claimed, that:

$$Kp_{\sigma^n}\left(\tfrac{|H_i|+1}{n} > q, \omega = R|s_i\right) \geq K\delta min\{\beta p(t=a|s_i), (1-\alpha)p(t=r|s_i)\}$$

Part Two: The LHS has an upper bound wich converges to 0

Note that:

$$\begin{aligned} p_{\sigma^n}\left(piv_i, \omega = A|s_i\right) &= p_{\sigma^n}\left(piv_i, \omega = A, t=a|s_i\right) + p_{\sigma^n}\left(piv_i, \omega = A, t=r|s_i\right) \\ &= \ p_{\sigma^n}(piv_i|t=a, \omega = A, s_i)p(t=a, \omega = A|s_i) \\ &\quad + p_{\sigma^n}(piv_i|t=r, \omega = A, s_i)p(t=r, \omega = A|s_i) \\ &= \ p_{\sigma^n}(piv_i|t=a)p(t=a, \omega = A|s_i) \\ &\quad + p_{\sigma^n}(piv_i|t=r)p(t=r, \omega = A|s_i) \end{aligned}$$

where the second equality follows from Bayes' rule, and the third equality from the independence of signals (among them and from the state of the world), conditional on the state of the art. Given that there is an analogous expression for $p_{\sigma^n}(piv_i, \omega = R|s_i)$, it follows that the the LHS of $(1')$ is equal to :

$$C_1 p_{\sigma^n}(piv_i|t=a) + C_2 p_{\sigma^n}(piv_i|t=r) \qquad (*)$$

where C_1 and C_2 are constants that only depend on the exogenous parameters of the game other than n. In particular they do not depend on the strategy used by the agents. Now note that $p_{\sigma^n}(piv_i|t=a) = p_{\sigma^n}(|H_i| = \lfloor nq \rfloor |t=a)$ and $p_{\sigma^n}(piv_i|t=r) = p_{\sigma^n}(|H_i| = \lfloor nq \rfloor |t=r)$, where $|H_i| = |\{j \neq i : v_j = a\}|$. Letting $\mu_{a,n} = p(v_j = a|t=a) = (1-\varepsilon)\sigma^n(a) + \varepsilon\sigma^n(r)$ and $\mu_{r,n} = p_{\sigma^n}(v_j = a|t=r) = \varepsilon\sigma^n(a) + (1-\varepsilon)\sigma^n(r)$ and given the independence of the signals of different agents conditional on the state of the art we have:

$$p_{\sigma^n}(piv_i|t=a) = \binom{n-1}{\lfloor nq \rfloor}\mu_{a,n}^{\lfloor nq \rfloor}(1-\mu_{a,n})^{n-1-\lfloor nq \rfloor} \qquad \text{and}$$
$$p_{\sigma^n}(piv_i|t=r) = \binom{n-1}{\lfloor nq \rfloor}\mu_{r,n}^{\lfloor nq \rfloor}(1-\mu_{r,n})^{n-1-\lfloor nq \rfloor}$$

The fact that the LHS of (1″) is bounded above by an expression that is independent of σ^n and that this upper bound converges to 0, now follows from the above expressions and the following lemma.

Lemma **7 (Convergence of binomial points of mass)**
The set $\{\binom{n-1}{\lfloor nq \rfloor} p^{\lfloor nq \rfloor}(1-p)^{n-1-\lfloor nq \rfloor} : 0 \leq p \leq 1\}$ *is bounded above by a function* $f(n)$ *such that* $\lim\limits_{n\to\infty} f(n) \to 0$.

Proof of Lemma 7:
We prove the lemma by using Stirling's formula to establish an upper bound for the set $\{\binom{n-1}{\lfloor nq \rfloor} p^{\lfloor nq \rfloor}(1-p)^{n-1-\lfloor nq \rfloor} : 0 < p < 1\}$ and showing that this upper bound converges to 0.
By Stirling's formula $\left(\lim\limits_{n\to\infty} \frac{n!}{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n} = 1\right)$ we have that for any $\varepsilon > 0$ there exists n_1 such that if $n > n_1$ then:

$$\binom{n-1}{\lfloor nq \rfloor} p^{\lfloor nq \rfloor}(1-p)^{n-1-\lfloor nq \rfloor}$$
$$< (1-\varepsilon)\frac{\frac{(n-1)!}{\sqrt{2\pi(n-1)}\left(\frac{n-1}{e}\right)^{n-1}}}{\frac{\lfloor nq \rfloor!}{\sqrt{2\pi\lfloor nq \rfloor}\left(\frac{\lfloor nq \rfloor}{e}\right)^{\lfloor nq \rfloor}}\frac{(n-1-\lfloor nq \rfloor)!}{\sqrt{2\pi(n-1-\lfloor nq \rfloor)}\left(\frac{(n-1-\lfloor nq \rfloor)}{e}\right)^{(n-1-\lfloor nq \rfloor)}}} p^{\lfloor nq \rfloor}(1-p)^{n-1-\lfloor nq \rfloor}$$
$$= (1-\varepsilon)\left(\frac{n-1}{2\pi\lfloor nq \rfloor(n-1-\lfloor nq \rfloor)}\right)^{\frac{1}{2}}\left(\frac{(n-1)q}{\lfloor nq \rfloor}\right)^{\lfloor nq \rfloor}\left(\frac{(n-1)(1-q)}{n-1-\lfloor nq \rfloor}\right)^{n-1-\lfloor nq \rfloor}$$
$$\times\left(\frac{p}{q}\right)^{\lfloor nq \rfloor}\left(\frac{1-p}{1-q}\right)^{n-1-\lfloor nq \rfloor}$$

Note that $p^{\lfloor nq \rfloor}(1-p)^{n-1-\lfloor nq \rfloor}$ is strictly concave for sufficiently large n ($q < 1 - \frac{1}{n}$) and uniquely maximized at $p = \frac{\lfloor nq \rfloor}{n-1}$. At the maximum p^* we have:

$$\binom{n-1}{\lfloor nq \rfloor}(p^*)^{\lfloor nq \rfloor}(1-p^*)^{n-1-\lfloor nq \rfloor}$$
$$= (1-\varepsilon)\left(\frac{n-1}{2\pi\lfloor nq \rfloor(n-1-\lfloor nq \rfloor)}\right)^{\frac{1}{2}}\left(\frac{(n-1)q}{\lfloor nq \rfloor}\right)^{\lfloor nq \rfloor}\left(\frac{(n-1-nq+q)}{n-1-\lfloor nq \rfloor}\right)^{n-1-\lfloor nq \rfloor}$$
$$\times\left(\frac{nq}{(n-1)q}\right)^{\lfloor nq \rfloor}\left(\frac{n-1-nq}{n-1-nq+q}\right)^{n-1-\lfloor nq \rfloor}\left(\frac{\lfloor nq \rfloor}{nq}\right)^{\lfloor nq \rfloor}\left(\frac{n-1-\lfloor nq \rfloor}{n-1-nq}\right)^{n-1-\lfloor nq \rfloor}$$
$$= (1-\varepsilon)\left(\frac{n-1}{2\pi\lfloor nq \rfloor(n-1-\lfloor nq \rfloor)}\right)^{\frac{1}{2}}$$

We therefore have that for all $n > n_1$ and for all $p \in (0,1)$

$$\binom{n-1}{\lfloor nq \rfloor} p^{\lfloor nq \rfloor}(1-p)^{n-1-\lfloor nq \rfloor} < (1-\varepsilon)\left(\frac{n-1}{2\pi\lfloor nq \rfloor(n-1-\lfloor nq \rfloor)}\right)^{\frac{1}{2}}$$

Moreover $(1-\varepsilon)\left(\frac{n-1}{2\pi\lfloor nq \rfloor(n-1-\lfloor nq \rfloor)}\right)^{\frac{1}{2}}$ converges to 0 at rate $\frac{1}{\sqrt{n}}$. ∎*(Lemma 7)*

To end the proof let m be such that for all $n > m$, $(C_1+C_2)f(n) < \frac{K\delta min\{\beta p(t=a|s_i),(1-\alpha)p(t=r|s_i)\}}{2}$ and pick $n_m > m$ such that $p_{\sigma^{n_m}}(X=a) > \delta$ (which exists by the assumption that $p_{\sigma^n}(X =$

a) does not converge to 0). It follows that:

$$\begin{aligned} Wp_{\sigma^{n_m}}(piv_i, \omega = A|s_i) - Cp_{\sigma^{n_m}}(piv_i, \omega = R|s_i) &= C_1 p_{\sigma^{n_m}}(piv_i|t = a) + C_2 p_{\sigma^{n_m}}(piv_i|t = r) \\ &< \frac{K\delta min\{\beta p(t = a|s_i), (1-\alpha)p(t = r|s_i)\}}{2} \\ &< Kp_{\sigma^{n_m}}\left(\frac{|H_i|+1}{n} > q, \omega = R|s_i\right) \end{aligned}$$

So $(1')$ is violated. As i was arbitrary, this shows that every single expert strictly prefers to set $\sigma^{n_m}(s_i) = 0$. Moreover we can pick n large enough so that this is the case for both signals. For σ^{n_m} to be an equilibrium it must be the case that members are best responding and therefore $\sigma^{n_m}(a) = 0$ and $\sigma^{n_m}(r) = 0$, which contradicts $p_{\sigma^{n_m}}(X = a) > \delta$ which in turn contradicts the assumption that $p_{\sigma^{n_m}}(X = a)$ does not converge to 0 as $n \to \infty$. ▮*(Proposition 3)*

Proof of Corollary 3: The corollary holds trivially if beyond some point in the sequence the games have no non-babbling equilibria. So we assume this is not the case and focus on the maximal subsequence such that all along the games have non-babbling equilibria. Pick n_δ such that $p_{\sigma^n}(X = a) < \delta < \frac{p(t=a)}{8}$ for the unique non-babbling symmetric equilibrium σ^n of $G^K_{n,q}$. Pick $n^* > n_\delta$ large enough such that for all $n > n^*$:
$\sum_{m=\lfloor nq \rfloor}^{n} \binom{n}{m} q^m (1-q)^{n-m} > \frac{1}{4}$. Such n^* exists as this is the probability that the fraction of successes in n trials is greater or equal to q, where trials are independent and the probability of success of any one trial is q, and the binomial distribution can be approximated arbitrarily well (close to its mean) by the normal distribution which is symmetric. So in particular this probability converges to $\frac{1}{2}$.
Suppose the statement of the corollary is not true and pick $m > n^*$ such that σ^m is a symmetric equilibrium of $G^K_{m,q}$ and $\sigma^m(a) \geq \frac{q}{1-\varepsilon}$. Letting $\mu_a = (1-\varepsilon)\sigma(a) + \varepsilon\sigma(r)$, we have $\mu_a \geq q$ and therefore:

$$p_{\sigma^m}(X = a|t = a) = \sum_{m=\lfloor nq \rfloor}^{n} \binom{n}{m} \mu_a^m (1-\mu_a)^{n-m} \geq \sum_{m=\lfloor nq \rfloor}^{n} \binom{n}{m} q^m (1-q)^{n-m} > \frac{1}{4}$$

$\Rightarrow p_{\sigma^m}(X = a) = p_{\sigma^m}(X = a|t = a)p(t = a) + p_{\sigma^m}(X = a|t = r)p(t = r) > \frac{1}{4}p(t = a)$ a contradiction, as we picked $\delta < \frac{p(t=a)}{8}$. ▮*(Corollary 3)*

Proof of Corollary 4:
Fully writing R_a and R_r as a function of all the parameters of the model we have that:
$R_r(p_A, \alpha, \beta, \varepsilon, q, n, K, W, C, \sigma) =$
$-(W(1-\beta) - C\beta)p_{\sigma^*}(piv_i|t = a)p_A\varepsilon + (C(1-\alpha) - W\alpha)p_{\sigma^*}(piv_i|t = r)(1-p_A)(1-\varepsilon) +$

$K\beta p_{\sigma^*}\left(\frac{|H_i|+1}{n} > q|t=a\right)p_A\varepsilon + K(1-\alpha)p_{\sigma^*}\left(\frac{|H_i|+1}{n} > q|t=r\right)(1-p_A)(1-\varepsilon)$

and

$R_a(p_A,\alpha,\beta,\varepsilon,q,n,K,W,C,\sigma) =$
$-(W(1-\beta)-C\beta)p_{\sigma^*}(piv_i|t=a)p_A(1-\varepsilon) + (C(1-\alpha)-W\alpha)p_{\sigma^*}(piv_i|t=r)(1-p_A)\varepsilon +$
$K\beta p_{\sigma^*}\left(\frac{|H_i|+1}{n} > q|t=a\right)p_A(1-\varepsilon) + K(1-\alpha)p_{\sigma^*}\left(\frac{|H_i|+1}{n} > q|t=r\right)(1-p_A)\varepsilon$

(Ia) $\frac{\partial R_r(p_A,\alpha,\beta,\varepsilon,q,n,K,W,C,\sigma^*)}{\partial W} =$
$-(1-\beta)p_{\sigma^*}(piv_i|t=a)p_A\varepsilon - \alpha p_{\sigma^*}(piv_i|t=r)(1-p_A)(1-\varepsilon) < 0.$
and $\frac{\partial R_r(p_A,\alpha,\beta,\varepsilon,q,n,K,W,C,\sigma^*)}{\partial W} =$
$-(1-\beta)p_{\sigma^*}(piv_i|t=a)p_A(1-\varepsilon) - \alpha p_{\sigma^*}(piv_i|t=r)(1-p_A)\varepsilon < 0.$
By Lemma 6 we therefore have $\frac{\partial z^*(p_A,\alpha,\beta,\varepsilon,q,n,K,W,C)}{\partial W} > 0.$

(Ib) $\frac{\partial R_r(p_A,\alpha,\beta,\varepsilon,q,n,K,W,C,\sigma^*)}{\partial C} =$
$\beta p_{\sigma^*}(piv_i|t=a)p_A\varepsilon + (1-\alpha)p_{\sigma^*}(piv_i|t=r)(1-p_A)(1-\varepsilon) > 0.$
and $\frac{\partial R_a(p_A,\alpha,\beta,\varepsilon,q,n,K,W,C,\sigma^*)}{\partial C} =$
$\beta p_{\sigma^*}(piv_i|t=a)p_A(1-\varepsilon) + (1-\alpha)p_{\sigma^*}(piv_i|t=r)(1-p_A)\varepsilon > 0$. By Lemma 6 we therefore have $\frac{\partial z^*(p_A,\alpha,\beta,\varepsilon,q,n,K,W,C)}{\partial C} < 0.$

(IIa) $\frac{\partial R_r(p_A,\alpha,\beta,\varepsilon,q,n,K,W,C,\sigma^*)}{\partial p_A} =$
$-(W(1-\beta)-C\beta)p_{\sigma^*}(piv_i|t=a)\varepsilon - (C(1-\alpha)-W\alpha)p_{\sigma^*}(piv_i|t=r)(1-\varepsilon) +$
$K\beta p_{\sigma^*}\left(\frac{|H_i|+1}{n} > q|t=a\right)\varepsilon - K(1-\alpha)p_{\sigma^*}\left(\frac{|H_i|+1}{n} > q|t=r\right)(1-\varepsilon)$

The sign of which is the same as that of

$$\frac{-(W(1-\beta)-C\beta)p_{\sigma^*}(piv_i|t=a)\varepsilon + K\beta p_{\sigma^*}\left(\frac{|H_i|+1}{n} > q|t=a\right)\varepsilon}{(C(1-\alpha)-W\alpha)p_{\sigma^*}(piv_i|t=r)(1-\varepsilon) + K(1-\alpha)p_{\sigma^*}\left(\frac{|H_i|+1}{n} > q|t=r\right)(1-\varepsilon)} - 1$$

But $R_r(p_A,\alpha,\beta,\varepsilon,q,n,K,W,C,\sigma^*) = 0$ implies

$$\frac{-(W(1-\beta)-C\beta)p_{\sigma^*}(piv_i|t=a)\varepsilon + K\beta p_{\sigma^*}\left(\frac{|H_i|+1}{n} > q|t=a\right)\varepsilon}{(C(1-\alpha)-W\alpha)p_{\sigma^*}(piv_i|t=r)(1-\varepsilon) + K(1-\alpha)p_{\sigma^*}\left(\frac{|H_i|+1}{n} > q|t=r\right)(1-\varepsilon)} = -\frac{1-p_A}{p_A}$$

Similarly $\frac{\partial R_r(p_A,\alpha,\beta,\varepsilon,q,n,K,W,C,\sigma^*)}{\partial p_A} =$
$-(W(1-\beta)-C\beta)p_{\sigma^*}(piv_i|t=a)(1-\varepsilon) - (C(1-\alpha)-W\alpha)p_{\sigma^*}(piv_i|t=r)\varepsilon +$
$K\beta p_{\sigma^*}\left(\frac{|H_i|+1}{n} > q|t=a\right)(1-\varepsilon) - K(1-\alpha)p_{\sigma^*}\left(\frac{|H_i|+1}{n} > q|t=r\right)\varepsilon$

The sign of which is the same as that of

$$\frac{-(W(1-\beta)-C\beta)p_{\sigma^*}(piv_i|t=a)(1-\varepsilon)+K\beta p_{\sigma^*}\left(\frac{|H_i|+1}{n}>q|t=a\right)(1-\varepsilon)}{(C(1-\alpha)-W\alpha)p_{\sigma^*}(piv_i|t=r)\varepsilon+K(1-\alpha)p_{\sigma^*}\left(\frac{|H_i|+1}{n}>q|t=r\right)\varepsilon}-1$$

But $R_a(p_A,\alpha,\beta,\varepsilon,q,n,K,W,C,\sigma^*)=0$ implies

$$\frac{-(W(1-\beta)-C\beta)p_{\sigma^*}(piv_i|t=a)(1-\varepsilon)+K\beta p_{\sigma^*}\left(\frac{|H_i|+1}{n}>q|t=a\right)(1-\varepsilon)}{(C(1-\alpha)-W\alpha)p_{\sigma^*}(piv_i|t=r)\varepsilon+K(1-\alpha)p_{\sigma^*}\left(\frac{|H_i|+1}{n}>q|t=r\right)\varepsilon}=-\frac{1-p_A}{p_A}$$

By Lemma 6 the above imply that $\dfrac{\partial z^*(p_A,\alpha,\beta,\varepsilon,q,n,K,W,C)}{\partial p_A}>0.$

(IIb) $\dfrac{\partial R_r(p_A,\alpha,\beta,\varepsilon,q,n,K,W,C,\sigma^*)}{\partial\alpha}=$

$$-(W+C)p_{\sigma^*}(piv_i|t=r)(1-p_A)(1-\varepsilon)-Kp_{\sigma^*}\left(\frac{|H_i|+1}{n}>q|t=r\right)(1-\varepsilon)(1-p_A)<0$$

$\dfrac{\partial R_a(p_A,\alpha,\beta,\varepsilon,q,n,K,W,C,\sigma^*)}{\partial\alpha}=$

$$-(W+C)p_{\sigma^*}(piv_i|t=r)(1-p_A)\varepsilon-Kp_{\sigma^*}\left(\frac{|H_i|+1}{n}>q|t=r\right)\varepsilon(1-p_A)<0$$

So by Lemma 6, $\dfrac{\partial z^*(p_A,\alpha,\beta,\varepsilon,q,n,K,W,C)}{\partial\alpha}>0.$

(IIc) $\dfrac{\partial R_r(p_A,\alpha,\beta,\varepsilon,q,n,K,W,C,\sigma^*)}{\partial\beta}=$

$$(W+C)p_{\sigma^*}(piv_i|t=a)p_A)\varepsilon+Kp_{\sigma^*}\left(\frac{|H_i|+1}{n}>q|t=a\right)\varepsilon p_A>0$$

$\dfrac{\partial R_a(p_A,\alpha,\beta,\varepsilon,q,n,K,W,C,\sigma^*)}{\partial\beta}=$

$$(W+C)p_{\sigma^*}(piv_i|t=a)p_A)(1-\varepsilon)+Kp_{\sigma^*}\left(\frac{|H_i|+1}{n}>q|t=a\right)(1-\varepsilon)p_A>0$$

So by Lemma 6, $\dfrac{\partial z^*(p_A,\alpha,\beta,\varepsilon,q,n,K,W,C)}{\partial\beta}<0.$

(III) Let $q'>q$ and $m=\lfloor nq\rfloor$, $m'=\lfloor nq\rfloor$. Evaluated at m we have that either $R_a(m,(\sigma^*(a),0))=0$ or $R_r(m,(1,\sigma^*(r))=0$ (depending on what kind of equilibrium we have). Assume that it is of form (1) and therefore evaluated at $\sigma^*(r)=0$ and $\sigma^*(a)>0$, $R_a(m,(\sigma^*(a),0))=0$. By Lemma 5 we know that $-R_a(m,(\sigma(a),0))$ has the single crossing property in m, and therefore evaluated at $m'>m$, genericall $R_a(m',(\sigma^*(a),0))\leq 0.$[38] If $R_a(m',(\sigma^*(a),0))=0$ then we have an equilibrium, otherwise we fix fix m' and look at $R_a(m',(\sigma^*(a),0))$ as a

[38]Being a linear combination of different non-linear functions, with non-zero slopes at almost every point, R_a and R_r have non zero slopes at almost every point.

function of $\sigma(a)$. As $R_a(m', (\sigma^*(a), 0)) < 0$ and R_a has the single crossing property in $\sigma(a)$, the equilibrium (which exists by virtue of 5) must either involve $\sigma(a) > \sigma^*(a)$, or be of the form $\sigma(a) = 1$ and $\sigma(r) > 0$. ▮*(Corollary 4)*

Proof of Corollary 5:
In what follows, RHS and LHS are defined analogously to RHAS and LHS in the proof of Proposition 3.

Part One: Positive lower bound on RHS

$$RHS : \sum_S \left[K \sum_{t=a,r} p_{\sigma,m^{-i}}(X = a, \omega = R|t) p(t|\{s_i\}^n) \right] p_{m^{-i}}(\{s_i\}^n | s_i, m^{-i})$$

Take $lb^n = min\{K \sum_{t=a,r} p_{\sigma,m^{-i}}(X = a, \omega = R|t) p(t|\{s_i\}^n)\}$. For each n:

$$RHS \geq \sum_S (lb^n) p_{m^{-i}}(\{s_i\}^n | s_i, m^{-i}) = lb^n$$

By the proof of Proposition 3, $K \sum_{t=a,r} p_{\sigma,m^{-i}}(X = a, \omega = R|t) p(t|\{s_i\}^n)$ has a positive lower bound for each $\{s_i\}^n$, therefore lb^n, and by extension RHS, also has a positive lower bound.
Part Two: The LHS has an upper bound which converges to 0

$$LHS : \sum_S \left[W \sum_{t=a,r} p_{\sigma,m^{-i}}(piv_i|t) p(t, \omega = A|\{s_i\}^n) - \right. \tag{1}$$

$$\left. C \sum_{t=a,r} p_{\sigma,m^{-i}}(piv_i|t) p(t, \omega = R|\{s_i\}^n) \right] p_{m^{-i}}(\{s_i\}^n | s_i, m^{-i}) \tag{2}$$

Similarly to above, take $ub^n = max\{K \sum_{t=a,r} p_{\sigma,m^{-i}}(X = a, \omega = R|t) p(t|\{s_i\}^n)\}$. For each n:

$$LHS \leq \sum_S (ub^n) p_{m^{-i}}(\{s_i\}^n | s_i, m^{-i}) = ub^n$$

First, by the proof of Proposition 3 and Lemma 1, ub^n has an upper bound that converges to zero for each $\{s_i\}^n$ (that is, since Lemma 1 holds for any p, the fact that $\mu_{a,n}$ and $\mu_{r,n}$ condition on the set of messages does not affect the result). ▮ *(Corollary 5)*

Proof of Proposition 4: Consider an extended game with the following signal-pooling stage. After receiving their signals, experts simultaneously send public messages, $m_i \in \{r, a\}$. We denote the profile of messages by $m = (m_1, ..., m_n)$. A strategy for expert i is a speech function $\theta_i(s_i) : \{a, r\} \to [0, 1]$ indicating the probability that expert i sends message a given signal s_i and a voting rule $\nu_i(s_i, m) : \{a, r\} \times \{a, r\}^n \to [0, 1]$. As in the rest of the paper, we focus on symmetric strategies, such that $\theta_i = \theta$ and $\nu_i = \nu$ for all i. We denote the single-agent decision rule by $d(s_1, s_2, ..., s_n) \in \{a, r\}$, representing his choice of whether to accept or reject the proposal after observing $(s_1, s_2, ..., s_n)$.

We show that $\nu(s_i, (m_1, ..., m_{i-1}, m_i, m_{i+1}, ..., m_n)) = d(m_1, ..., m_{i-1}, s_i, m_{i+1}, ..., m_n)$ (second stage strategy), and $\theta(a) = 1$ and $\theta(r) = 0$ (first stage strategy), is a sequential equilibrium of the game when $(K = 0, q < 1)$ or $(K > 0, q = \frac{n-1}{n})$. We proceed by construction; assume that all experts other than i play according to these strategies.

Second stage:

Suppose expert i finds himself at information set $(s_i, (m_1, ..., m_i, ..., m_n))$. Then given the first stage strategy of the other experts, he believes that the true vector of signals is $(m_1, ..., s_i, ..., m_n)$. If $K = 0$ and $\frac{1}{n} < q < \frac{n-1}{n}$ then, given the other experts' second stage strategies, i's vote does not alter the outcome. Therefore, he finds it weakly optimal to vote according to $d(m_1, ..., s_i, ..., m_n)$.

Assume instead that $K \geq 0$ and $q = \frac{n-1}{n}$ (unanimity). If $d(m_1, ..., m_i, ..., m_n) = 1$ then, given the other experts' second stage strategy, i's vote is pivotal. His preferred vote is therefore $v_i = a$, since he faces the exact same payoffs as a single agent representative who observes $(m_1, ..., s_i, ..., m_n)$. When $d(m_1, ..., m_i, ..., m_n) = 0$, i's vote is never pivotal, and it is weakly optimal for him to set $v_i = r$.[39]

In comparison, if $K > 0$ and $q < \frac{n-1}{n}$ then the proposed strategy configuration is not an equilibrium: when $d(m_1, ..., m_i, ..., m_n) = 1$ expert i can profitably deviate, without affecting the committee's decision, by keeping his first stage strategy constant and altering his second stage strategy to set $\nu(s_i, (m_1, ..., m_i, ..., m_n)) = 0$.

Signal sharing stage:

Given the other experts' first stage strategy, all the second stage information sets (from i's perspective) are of the form $(s_i, (s_1, ..., m_i, ..., s_n))$. Note that for any vector of signals $(s_1, ..., s_i, ..., s_n)$, i prefers the outcome $X = d(s_1, ..., s_i, ..., s_n)$. Given the second-stage strategies, $X = d(s_1, ..., s_i, ..., s_n)$ if i sets $m_i = s_i$. If, however, i sets $m_i' \neq s_i$, then with positive probability $X = d(s_1, ..., m_i', ..., s_n) \neq d(s_1, ..., s_i, ..., s_n)$. He thus finds it optimal to reveal his true signal.

It follows that $\nu(s_i, (m_1, ..., m_i, ..., m_n)) = d(m_1, ..., s_i, ..., m_n)$, and $\theta(a) = 1$ and $\theta(r) = 0$ is a sequential equilibrium of the game. On the equilibrium path, agents correctly reveal their information, and the outcome is $d(s_1, ..., s_i, ..., s_n)$. ▮*(Proposition 4)*

[39] Analogous arguments follow for the cases with $K = 0$ and $q \leq \frac{1}{n}$ or $\frac{n-1}{n} \leq q < 1$.

Appendix B: Empirical Work

In this appendix, we utilize the rich set of data on decision-making in FDA boards to investigate whether there is correlation between the size of the committee and the rate of rejection of new drug applications. We find a weak negative relation between committee size and the proportion of approval votes out of the total number of votes cast. This finding could be explained by the mechanism we present in the paper, and the theoretical result that the approval rate is vanishing for sufficiently large committees.

In the United States, the Food and Drug Administration (FDA) must approve or reject new drugs by means of an assessment procedure called a "new drug application" (similarly a "biologic license application" for biologic products and "premarket approval" for medical devices). In most instances, the FDA has the option to refer a matter of drug approval to an expert committee for consideration. The members of the panel will then discuss scientific issues based on the studies provided by the sponsor company and then independently and simultaneously vote on approval; i.e. whether the benefits of the drug outweigh risks. As noted in the FDA's guidelines for voting procedures: "Since all members vote on the same question, the results help FDA gauge a committee's collective view on complex, multi-faceted issues."[40]

We collect data from FDA committee meetings held between January 2008 and August 2013.[41] The data comes from official meeting minutes (or 24 hour summary documents) downloaded via http://www.fda.gov. We only consider records from meetings that discuss drug/device/blood-product applications (NDA, sNDA, BLA, sBLA, PMA, sPMA) and where the approval question is posed in a single question. We have voting data on the approve/disapprove question from 174 FDA meetings across 21 different topical committees. In four cases, the FDA convened a joint meeting between two panels and in all these cases the Drug Safety and Risk Management Committee was part of the session.

For each meeting, the source reports the number of voting members present. This number varies between 3 and 26 in our sample and the average committee size is 13.14 members. The committee size varies for different reasons. First, the official number of permanent members vary across the topical committees; e.g. the Arthritis Drugs Committee has 11 permanent members, whereas the Dermatologic and Ophtalmic Drugs Committee has 15 permanent members. However, the actual number of permanent members is typically lower

[40]Guidance for FDA Advisory Committee Members and FDA Staff: Voting Procedures for Advisory Committee Meetings. August 2008.

[41]Prior to the FDA Amendments Act of 2007 the voting was sequential. Throughout the second half 2007, voting by "a show of hands" was replaced by a mechanical device whereby each member votes independently (Urfalino and Costa (2013)).

	Model 1	Std. Error	**Model 2**	Std. Error
Constant	0.807^{***}	(0.095)	0.46	(0.166)
# of voting members	-0.013^{*}	(0.007)	-0.009	(0.009)
Committee fixed effects	$-$		$+$	
Mean fraction of y votes	0.642		0.642	
R^2	0.0191		0.235	
Adjusted R^2	0.0134		0.129	
Num. Obs.	174		174	

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 1: Models 1 and 2. The dependent variable is the fraction of yes votes. The standard errors are heteroskedasticity robust.

due to many vacancies. Second, members often cancel on the meetings (meeting attendance and cancellations are stated in the official meeting minutes). Finally, the FDA invites a number of temporary voting members (including one patient representative) who are hand picked specialists or serve on other advisory committees. The average proportion of invited members out of the total number of voting members is 0.6.

Table 1 reports the results from an OLS regression of the fraction of acceptance votes in a session on the total number of voting members (Model 1). The table also reports the proportion of yes-votes (in favor of approval) out of the total number of votes. In the regressions we ignore abstentions, which are few and mostly due to declarations of conflict of interest. As reported in the Table the partial correlation associated to the number of voting members is negative with a p-value of 0.0819. We also ran a logit model of a binary variable taking a value of one if a simple majority of the committee members approved and zero otherwise. The results from this regression are similar to the OLS regression: the coefficient of the size variable is negative and the p-value is 0.111.

Some of the variation in size is due to variation in the number of permanent members across different topical committees, which raises the concern that the negative effect of size found in the 'naive' OLS regression is due to systematic differences in the medical products sent to the individual committees. For example, if the products generated in the area of Dermatologic and Ophtalmic Drugs are more likely to be "bad" (in terms of tour model, a lower p_A) than in the area of "Arthritis Drugs," then the negative correlation could be driven by the fact that the Dermatologic and Ophtalmic Drugs Committee has 15 permanent members whereas the Arthritis Drugs Committee has only 11. To explore this possibility, in Model 2 (also reported in Table 1), we include 20 dummies to account for committee fixed effects in the

	Coefficient	Std. Error
Constant	0.539^{***}	(0.112)
Fraction of invited members	0.131	(0.174)
R^2	0.004	
Adjusted R^2	-0.003	
Num. Obs.	144	

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 2: The dependent variable is the fraction of yes votes. Standard errors are robust to heteroskedasticity.

second regression. In the four cases where the meeting is joint between two committees, we assign the meeting to the Drug Safety and Risk Management Committee. We find that most of the committee dummies are significant, and while the sign on the "size" variable remains negative and of similar size (-0.00900 with fixed effects versus -0.0125), the significance of the coefficient drops, as reflected in the higher p-value, 0.2856.

Another concern is that the variation in committee size is endogenous, since the FDA invites additional, temporary, members to participate in the approval decision of individual medical products. This could explain the finding of a negative coefficient on committee size if, for example, temporary members are more likely to be added for 'difficult' decisions that have a higher downside risk (or in the terms of our model, a larger C). In order to study this possibility, we regress the proportion of yes votes out of total votes on the proportion of invited temporary voting members. We report the result in Table 2. For this specific regression we only have 140 observations, as for most meetings of PMA-committees there was no information available on the number of invited members. If endogenous variation in committee size is behind the negative relationship we find in the 'naive' regression, we would expect the proportion of invited members to be negatively correlated with the proportion of yes votes. However, we find that the sign of the "proportion of invited member" coefficient is not statistically significant (p-value=0.454), and is actually positive.

Lastly, we address a separate issue. The majority decision of an FDA board is not binding, and the final decision rests on FDA's division director. Therefore, in a legal sense, the decision of the committee is purely advisory. However, there is evidence of a norm for following the majority decision of the expert committee and the chairman usually has the task of breaking eventual voting ties. In our sample, 90 percent of the final FDA decisions follow the recommendation of the committee. However, the non-binding nature of committee decisions raises the following possibility: if the FDA is aware of a bias towards rejection in larger committees, they may try to counteract this bias by over-ruling close rejection outcomes in larger committees. Due to the small number of final decisions that go against the majority decision, we are not able explore this hypothesis statistically. Out of 174 committee meetings, we have the final FDA decision in 161 instances (some applications are still awaiting an answer from the division director) and out of these the FDA overturned 16 committee decisions. The committees recommended approval 117 times and the FDA overturned 11 of these applications (9.4 percent) and the average size of the "overturned" panels is 13.4. Further, the committees rejected 44 applications and the FDA overturned 5 of these recommendations (11.6 percent) and the average size of these five boards is 14.4.

References

Austen-Smith, David and Jeffrey Banks (1996), "Information Aggregation, Rationality, and the Condorcet Jury Theorem." *American Political Science Review*, 90, 34–45.

Bergstrom, Theodore, Lawrence Blume, and Hal Varian (1986), "On the Private Provision of Public Goods." *Journal of Public Economics*, 29, 25–49.

Brennan, Geoffrey and Philip Pettit (2004), *The Economy of Esteem*. An Essay on Civil and Political Society, Oxford University Press.

Callander, Steven (2008), "Majority rule when voters like to win." *Games and Economic Behavior*, 64, 393–420.

Coughlan, Peter (2000), "In Defense of Unanimous Jury Verdicts: Mistrials, Communication, and Strategic Voting." *American Political Science Review*, 94, 375–393.

Ellingsen, Tore and Magnus Johannesson (2008), "Pride and Prejudice: The Human Side of Incentive Theory." *The American Economic Review*, 98, 990–1008.

Feddersen, Timothy and Wolfgang Pesendorfer (1997), "Voting Behavior and Information Aggregation in Elections With Private Information." *Econometrica*, 65, 1029–1058.

Feddersen, Timothy and Wolfgang Pesendorfer (1998), "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting." *The American Political Science Review*, 92, 23–35.

Gerling, Kerstin, Hans Peter Grüner, Alexandra Kiel, and Elisabeth Schulte (2005), "Information acquisition and decision making in committees: A survey." *European Journal of Political Economy*, 21, 563–597.

Levy, Gilat (2007), "Decision Making in Committees: Transparency, Reputation, and Voting Rules." *The American Economic Review*, 97, 150–168.

Li, Hao (2001), "A Theory of Conservatism." *Journal of Political Economy*, 109, 617–636.

Li, Hao and Wing Suen (2009), "Viewpoint: Decision-making in committees." *Canadian Journal of Economics/Revue canadienne d'économique*, 42, 359–392.

McLennan, Andrew (1998), "Consequences of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Agents." *The American Political Science Review*, 92, 413–418.

Morgan, John and Felix Várdy (2012), "Mixed Motives and the Optimal Size of Voting Bodies." *Journal of Political Economy*, 120, 986–1026.

Olson, Mancur (1965), *The Logic of Collective Action*. Public Goods and the Theory of Groups., Harvard University Press.

Ottaviani, Marco and Peter Sorensen (2001), "Information aggregation in debate: who should speak first?" *Journal of Public Economics*, 81, 393–421.

Palfrey, Thomas and Howard Rosenthal (1985), "Voter Participation and Strategic Uncertainty." *The American Political Science Review*, 79, 62–78.

Peleg, Bezalel and Shmuel Zamir (2012), "Extending the condorcet jury theorem to a general dependent jury." *Social Choice and Welfare*, 39, 91–125.

Quah, John and Bruno Strulovici (2012), "Aggregating the Single Crossing Property." *Econometrica*, 80, 2333–2348.

Samuelson, Paul A (1954), "The Pure Theory of Public Expenditure." *The Review of Economics and Statistics*.

Urfalino, Philippe and Pascaline Costa (2013), "Secret-public voting in FDA advisory committees." *Cespra Working Paper*.

Visser, Bauke and Otto H Swank (2007), "On Committees of Experts." *The Quarterly Journal of Economics*, 122, 337–372.