Over-Caution of Large Committees of Experts*

Rune Midjord, Tomás Rodríguez Barraquer, and Justin Valasek†

May 24, 2015

Abstract

In this paper, we consider a committee of experts that decides 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 payoff (negative) if they vote to pass a drug that proves to be fatal for some users. Under the standard voting model, we show that information is aggregated in large committees provided disesteem payoffs are not overly large. However, we go on to document an empirically-relevant discontinuity in the standard model: if an arbitrarily large number of signals does not perfectly reflect the state of the world then, no matter how small the disesteem payoffs are, information aggregation fails in large committees and the committee rejects the innovation almost surely, providing an explanation for over-caution in committees. Finally, we extend the model to include a disesteem payoff when an agent votes to reject a beneficial innovation, and provide closed form expressions capturing the limit behavior of the committee for generic combinations of the two kinds of disesteem payoffs.

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

*We thank David Austen-Smith, 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, Public University of Navarra, the University of the Basque Country, the Center for Rationality at Hebrew University, University of Bern, IGIER-Bocconi, and the AEA Annual Meeting. 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, Universitat Autònoma de Barcelona, and WZB Berlin respectively. Contact: Rune Midjord, Oester Farigmagsgade 5, 1353 Copenhagen. Telefon: +4535323059. Email: 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. When decisions are made by groups, however, each individual's ability to influence the final decision is diluted, which can lead to a magnification of individual biases. Specifically, if committee members face idiosyncratic payoffs tied to their vote, such as an expressive motive or a preference for voting for the winning option, then information aggregation can fail in large committees (see Callander (2008) and Morgan and Várdy (2012)): In these settings, individuals prefer that the committee chooses the right option, but have an incentive to *vote* for one of the options *independently* of which option is right. Therefore, in a large committee, where the probability of being pivotal is low, individuals rationally vote to maximize their idiosyncratic payoffs, rather than to accurately aggregate information. In contrast to previous research, the question we ask here is, does a committee effectively aggregate information when, in addition to caring about making the right decision, committee members also care about individually voting for the right decision (an idiosyncratic payoff that is *dependent* on which option is right)?

A particularly relevant example is the advisory committees under the United States Food and Drug Administration (FDA), which are called upon to decide whether or not to approve a new 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 an additional negative (disesteem) payoff if they personally voted to approve the drug. 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 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.[1]

Contrary to the intuition that idiosyncratic biases will dominate, we find that in the standard

[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; 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. 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 convict a suspect who later turns out to be innocent.

model of information aggregation, the introduction of disesteem payoffs (provided they are not overly-large) does not lead to a failure of information aggregation in large committees. The reason is that although the probability of being pivotal approaches zero as the committee becomes large, the probability that the committee, conditional on voting optimally, takes a wrong decision also approaches zero at the same rate as the probability of being pivotal.

While this result is interesting in itself, we find that it is not robust to certain real-world considerations. The standard approach in voting models assumes that the information held by individuals is generated by a true state of the world, so that the aggregation of information held by an arbitrarily large group of individuals reveals the state with arbitrary precision. We analyze information aggregation under disesteem payoffs in a proposed alternative to the standard model, where each expert's information is generated by a technology which itself may be incorrect. Under this view, where even the collective knowledge contained in a very large number of signals has some probability of being wrong, we reveal a discontinuity in the standard model in the presence of idiosyncratic payoffs. That is, we show that no matter how small disesteem payoffs are, and no matter how small the probability that the collective knowledge is wrong, a large enough committee will always reject the innovation regardless of the information held by its members. This result undermines the idea formally captured by the generalizations of the Condorcet Jury Theorem to strategic voting, according to which a robust way of improving collective decision making is by increasing committee sizes.

Our analysis highlights that even committees of experts whose idiosyncratic payoffs depend on which option is correct 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.

Turning to FDA's advisory panels, we find that larger 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.[2] Intuitively, the logic proposed above suggests these observations might be due to the collective nature of the decision-making process magnifying the caution of individual committee members.

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 world, which is revealed only if the innovation is approved.[3] 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. Under the standard state-of-the-world model (SoW), each individual's signal is a noisy signal of the state of the world (ω). Alternatively, 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 (SoA). 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.[4] The standard (SoW) model can be seen as a special case of our model in which the state of the art is a perfect reflection of the true state of the world.

We show that for each set of values of the exogenous parameters there is essentially a unique equilibrium. As exemplified by Callander (2008), voting games with idiosyncratic payoffs may have multiple non-trivial equilibria, and this possibility may limit the significance of comparative static analysis. We establish that in our model, non-trivial-equilibria are unique by applying the recent results of Quah and Strulovici (2012). The application of these results in order to explore uniqueness in voting problems with non-standard payoffs in general seems

[2]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.

[3]In our FDA example, new information on harmful side-effects of a drug (not contained in clinical trials) is only later generated if the drug is made generally available.

[4]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.

quite promising.[5] This is important since in the absence of uniqueness the insights that can be obtained via comparative statics only apply locally and are thus of limited interest.

Besides our main result, we characterize this unique equilibrium and study the comparative statics. We also show that in the special case of the standard (SoW) model, the Condorcet Jury Theorem continues to hold as long as K is small enough. Of interest, yet unsurprisingly, when 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. Next, 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 speed of dilution of the disesteem payoffs for the main results to hold.

Lastly, we extend the model to include the possibility that other voting errors are punished. For example, committee members who incorrectly vote to reject a good drug may face internal social or professional sanctions from their committee peers. Specifically, we consider the case where agents receive a negative payoff, k_1, for voting for an innovation that is approved and is shown to be bad (equivalent to K in the main analysis), and a negative payoff, k_2, for voting against an innovation that is approved and is shown to be good.

We find that limit behavior as n goes to infinity depends only on the ratio of the two disesteem payoffs; similar to the main analysis, k_1 and k_2, no matter how small, generically distort the committee decision away from perfect information aggregation in the limit. For k_1/k_2 large enough, the result of the main analysis holds, and the committee always rejects in the limit. Similarly, for k_1/k_2 small enough, the only limit equilibria of the model are for the committee to always accept or always reject. For intermediate values of k_1/k_2 the committee decision cannot perfectly match the state of the art in the limit since, conditional on perfect information aggregation, all agents prefer to vote to accept to avoid k_2. Instead, the committee always accepts the drug when the state of the art is accept *and* accepts with positive probability when the state of the art is reject, increasing the exposure to k_1 and making agents with a signal of r indifferent between voting to accept and voting to reject.

[5] As we note in our analysis, standard techniques for establishing the single crossing property in these kinds of models do not apply, since the functions involved, along with their first and second derivatives are non-monotonic.

We provide closed form expressions capturing the limit behavior of the committee in the intermediate case.

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, provides comparative statics, discusses robustness to dilution, and extends the model to also consider disesteem payoffs due to type II voting errors. All proofs are relegated to Appendix A. Appendix B contains a detailed discussion of the empirical analysis noted in footnote 2. 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.[8]

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 by the committee. Visser and Swank show that reputation concerns make the a priori unconventional decision more attractive and lead 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. One difference with respect to our model is that in Visser and Swank the additional reputational (idiosyncratic) payoffs do not directly depend on the state (the true value of the project is never revealed).

[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).

[8] In another branch of the literature the committee members have no concern for the aggregate decision and care only about voting (or giving recommendations) to maximize the belief that the "market" holds about their level of competence i.e. the precision of their private signals. See e.g. Ottaviani and Sorensen (2001) and Levy (2007).

Callander (2008) analyzes idiosyncratic payoffs in elections under simple majority rule when voters wish for the better candidate to be elected, but also to personally vote for the winner. The 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*.[9] 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.[10] If K is small these payoffs represent a seemingly small departure from a pure common values situation, 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.

[9] We consider any q-rule with a fixed q, such as the majority rule used in FDA committees. This excludes decision rules such as the unanimity rule, where $q = (n-1)/n$; for an analysis of the case of unanimity and communication, see a working version of the paper, available online at http://mwpweb.eu/JustinValasek/.

[10] 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.

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 ($\alpha = p(\omega = A|t = r), 0 \leq \alpha < \frac{1}{2}$), and let β denote the probability that the state of the art is wrong when it indicates that the innovation should be accepted ($\beta = p(\omega = R|t = a), 0 \leq \beta < \frac{1}{2}$). 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 ($p(s_i = t) = 1 - \varepsilon$, $\varepsilon < \frac{1}{2}$), and with probability ε it differs with respect to the state of the art ($p(s_i \neq t|t) = \varepsilon$).[11] The standard model, where signals are generated directly by the state of the world (SoW model) corresponds to the case of $\alpha = \beta = 0$.

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 $s_i = r$. Throughout the analysis we rely on the concept of Bayesian Nash equilibrium and focus on symmetric strategies only; that

[11]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.

is, conditioning on signals, all members use the same decision rule.[12] 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:

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

where p_σ denotes the probability of the event given other agents play strategies $\sigma = (\sigma(a), \sigma(r))$.

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:*

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

3 Analysis

We first characterize the equilibria of the model and then present the main results 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 $R_{s_i}(n, \sigma)$ is nonpositive:[13]

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

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 $K = 0$, $\sigma(a) = \sigma(r) = 1$ is not an equilibrium.[14] As is also the case with $K = 0$, non-babbling equilibria often involve mixed strategies.

[12] Focusing on strategies which just depend on players' types (in this case their signals), is a standard practice in Bayesian games. The justification in our case is that our players are ex-ante identical in every aspect but their labels and it is unappealing to let behavioral differences depend purely on payoff irrelevant characteristics (such as the labels). Restricting attention to symmetric strategies is specifically common in the voting literature when voting is simultaneous; see for example Palfrey and Rosenthal (1985) and Feddersen and Pesendorfer (1997).

[13] 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.

[14] 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 $s_i = a$).[15]

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 $R_r(n, (1, \sigma(r))) = 0$.[16]

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 crossing at the point of discontinuity in case $z = 1$, which corresponds to the equilibrium $\sigma = (1, 0)$. We can now present the proposition characterizing the non-babbling equilibrium:

[15] 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 $s_i = r$.

[16] 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.

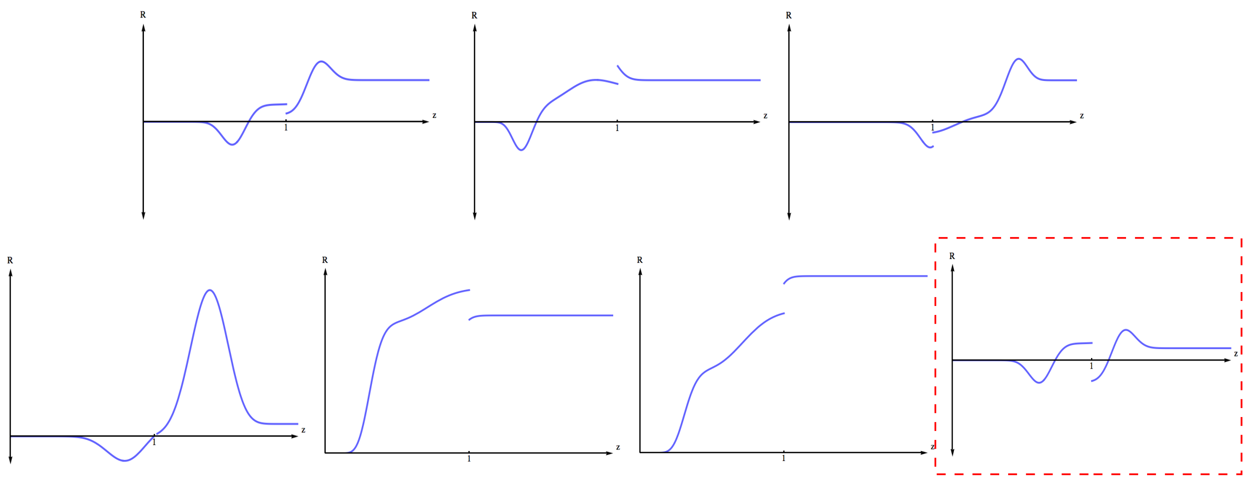

Figure 1: The results of Quah and Strulovici (2012) show that $R(n, z)$ can only conform with one of these seven stylized cases; in addition, the only case that admits multiple equilibria, illustrated by the framed diagram on the bottom right, is ruled out by Lemma 1.

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 $\frac{\partial R(n,z^*)}{\partial z} > 0$.

As exemplified by Callander (2008), in voting games with idiosyncratic payoffs there are often multiple equilibria, and this multiplicity limits the significance of comparative static analysis. It turns out that in our model non-trivial-equilibria are unique, and we establish this to be the case by applying the recent results of Quah and Strulovici (2012). We discuss the main idea of the proof in what follows, relegating the details to the Appendix.

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 two properties of R that hold when $K > 0$ and which exclude the possibility of multiplicity. These are: (*Property 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. 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. Establishing that R has the single crossing property in each of these two segments using standard techniques is not possible, since the signs of its first and second derivatives change frequently. We can however express R as a linear combination of functions which are always non-negative or non-positive and which can be readily seen to satisfy the single crossing property. The main result of Quah and Strulovici (2012) establishes conditions under which linear combinations of single-crossing functions are themselves single crossing, and these conditions are met by each of the continuous segments of R. So this means that R must take one of the stylized shapes shown in the 7 panels of Figure 1. (*Property 2*) If $R_r(n,(1,0)) \leq 0$ then the jump at the discontinuity is positive ($R_a(n,(1,0)) < R_r(n,(1,0))$), which follows from Lemma 1. By (1) the only way in which there can be an equilibrium at $z \geq 1$ is if $R_r(n,(1,0)) \leq 0$, but then it follows that $R_a(n,(1,0)) < R_r(n,(1,0)) \leq 0$, so by (2) in this case there is no equilibrium $z \in (0,1)$. This argument therefore excludes the possibility illustrated by the bottom right panel of Figure 1.

Next, we turn to the comparative statics of the non-babbling equilibrium with respect to K. 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 Claim.

Claim **1 (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.

Claim 1 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 established using the results of Quah and Strulovici (2012), any crossing can only take place from negative to positive and therefore this upward shift causes the new crossing to take place at $z' < z^K_{n,q}$ (unless the crossing happens exactly at the discontinuity ($z = 1$)). Generically, a small enough change of K at an equilibrium $z^* = 1$ preserves this as the unique equilibrium of the game.[18]

The straightforward intuition for Claim 1 is that K indexes the conflict of interest among committee members. 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

[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] There is a simple argument, that essentially provides the same result as Claim 1 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.

only way of doing this is by lowering z.[19]

3.1 Large Committees

In this section, we analyze the consequences of disesteem payoffs for large n. In order to compare, we first characterize large committee outcomes for $K = 0$ ($G^0_{n,q}$).

PROPOSITION **2 (No disesteem payoffs: Information aggregation)**
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 results. In particular, the following proposition shows that for positive yet sufficiently small K, in the special case in which the state of the art is a perfect reflection of the truth (SoW), committees make the correct decision with probability 1 as $n \to \infty$.

PROPOSITION **3 (Disesteem payoffs, SoW: Conditional information aggregation)**
Assume that $\frac{W}{C} \in \left(\frac{\varepsilon^2(1-p_A)}{(1-\varepsilon)^2 p_A}, \frac{(1-p_A)}{p_A}\right)$ and

$$K < \frac{p_A W \left(\frac{(1-\varepsilon)^2}{\varepsilon}\right) - C\varepsilon(1 - p_A)}{\varepsilon(1 - p_A)\frac{1-\varepsilon}{1-2\varepsilon}} \tag{1}$$

If $q = \frac{1}{2}$, then $\sigma(a) = 1$, $\sigma(r) = 0$ is an equilibrium for all sufficiently large n. In particular, this implies that the Condorcet Theorem holds for all K satisfying the inequality above.[20]

The intuition for this result (proved in the appendix) is that, with $\sigma(a) = 1$, $\sigma(r) = 0$, the probability that the committee makes the wrong decision converges to 0 at the same rate as the probability that a given agent is pivotal, and therefore the ratio of the the two probabilities, $p(piv|\omega = A)/p(X = A|\omega = R)$, approaches a strictly positive constant. Therefore, when K is sufficiently small relative to W, the benefit of voting to accept given

[19] It is important to note that this is a property of the equilibrium and not a global property of R.

[20] Note that this is a sufficient, but not necessary, condition. In particular, this proposition specifies conditions under which voting is truthful ($\sigma(a) = 1$, $\sigma(r) = 0$) given a majority rule and K positive, a stronger condition than is needed for the Condorcet Theorem to hold. Also, if the disesteem payoff is "diluted" as n grows, then the condition on K is not restrictive. We discuss this, and the robustness of the following SoA result to dilution, in section 3.3.

a signal of accept outweighs the exposure to the disesteem payoff in large committees, and truthful voting is supported in equilibrium.

In stark contrast to Proposition 3, we show that under the SoA model, 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, for any $\beta, \alpha > 0$ the committee converges to always rejecting the innovation. This is true independently of how precise is the the state of the art (that is, independently of how small α and β are), as long as it is not perfectly precise (which is the special SoW case analyzed in Proposition 3.

Proposition **4 (Disesteem payoffs, SoA: No information aggregation)**
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 4 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,

$$W p_\sigma(piv_i, \omega = A|s_i) - C p_\sigma(piv_i, \omega = R|s_i) \geq K p_\sigma(X = a, \omega = R|s_i) \qquad (1')$$

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 **2 (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 4 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, Propositions 2 and 3 show 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 (state of the world). In contrast, Proposition 4 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.

Comparing the results of Propositions 3 and 4, the propositions show that: under the SoW model, information aggregation can be sustained with disesteem payoffs since both the probability of being pivotal and the probability the committee wrongly accepts the innovation approach zero; under the SoA model, however, the probability the committee wrongly accepts the innovation is bounded away from zero whenever $p(X = A)$ is bounded away from zero. Therefore, the mechanism that sustains information aggregation in the SoW model is absent in the SoA model. Moreover, we demonstrate that this difference is particularly stark since it holds in the limit as the SoA model approaches the SoW model (α and β approach zero). This exposes a discontinuity in the standard model, where only a marginal deviation away from the SoW assumption changes equilibrium behavior from truthful to babbling.

3.2 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 **3 (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:

- $\frac{\partial z^*}{\partial W} > 0$, $\frac{\partial z^*}{\partial C} < 0$.
- $\frac{\partial z^*}{\partial p_A} > 0$, $\frac{\partial z^*}{\partial \alpha} > 0$ *and* $\frac{\partial z^*}{\partial \beta} < 0$.
- 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

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

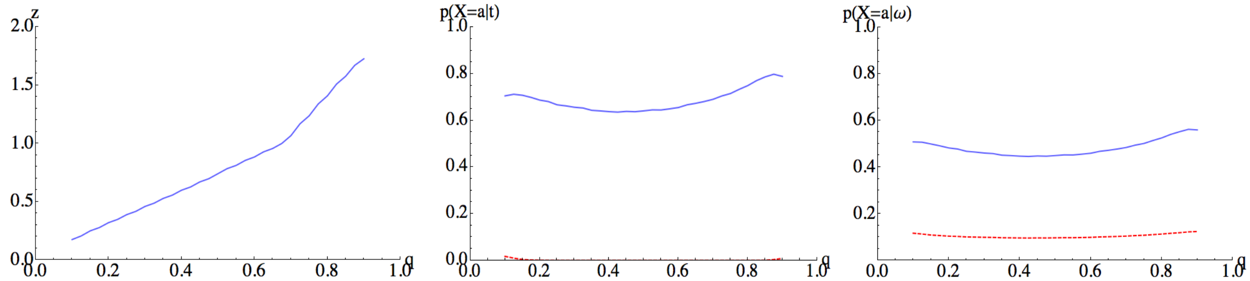

Figure 2: Monotonicity of z and non-monotonicity of $p(X = a|\omega)$ and $p(X = a|t)$ in q. Parameters: $\varepsilon = 0.3$, $W = 3$, $C = 4$, $K = 1/3$, $p_A = 0.5$, $\alpha = 0.4$, $\beta = 0.3$, $n = 101$. Left graph: z is monotonic in q, yet $p(X = A|t)$ (middle graph) $p(X = a|\omega)$ (right graph) 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 $p(X = a|\omega = A)$.

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 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

[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.

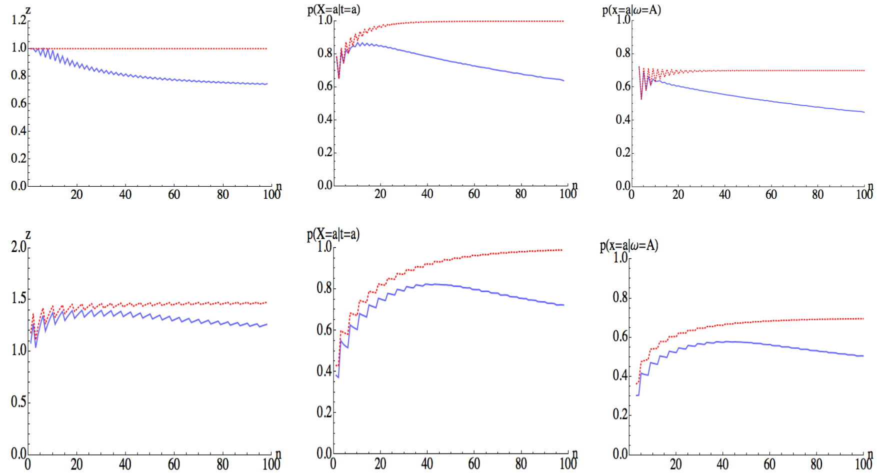

Figure 3: Non-monotonicity of $\sigma(a)$, $p(X = a|t = a)$ and $p(X = a|\omega = A)$ in n. The jaggedness of the figures is due to the discreteness of the problem (we are interested in the "low frequency variation"). Parameters: $\varepsilon = 0.3$, $W = 3$, $C = 4$, $p_A = 0.5$, $\alpha = 0.4$, $\beta = 0.3$ 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 ($K = 0$).

relative salience of disesteem payoffs–which accrue regardless of whether he is pivotal or not–increases. Proposition 4 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 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.

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

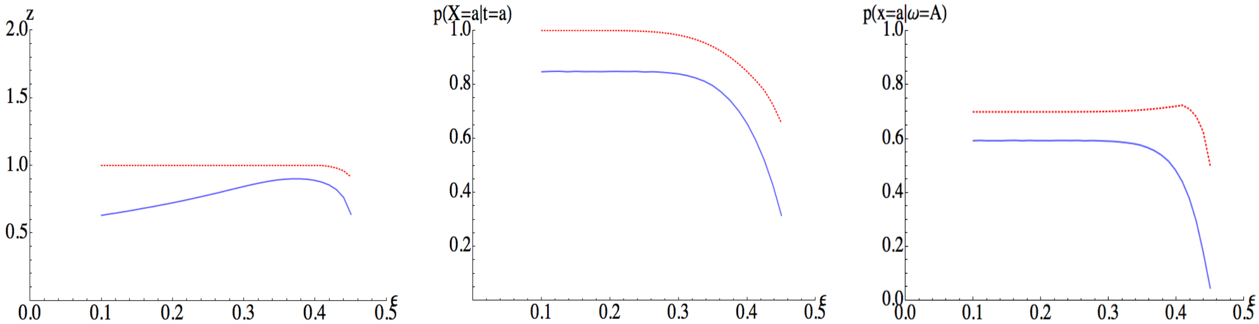

Figure 4: Non-monotonicity of z, $p(X = a|t = a)$, and $p(X = a|\omega = A)$ in ε. Parameters: $n = 25$, $W = 3$, $C = 4$, $p_A = 0.5$, $\alpha = 0.4$, $\beta = 0.3$. Note that z is non-monotonic. Despite the initial rise in z, $p(X = a|t = a)$ is weakly decreasing throughout, and $p(X = a|\omega = A)$ 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 ($K = 0$).

3.3 Dilution of Disesteem Payoffs

Lastly, we discuss the extent to which our result is robust to dilution of disesteem payoffs.[24] 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 results 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 3.1. 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 $\lim_{n\to\infty} \sqrt{n} g_n(n) \to \infty$.

Note that the games (G_n^K) of Section 3.1 are a special case of this formulation, as $f(n) = K$ being constant in n certainly has the required property ($\lim_{n\to\infty} \sqrt{n}K \to \infty$). This definition also accommodates other interesting cases; for example, let $f(n, \{v_i\}_{i=1}^n) = \frac{K}{log(n)}$.

First, note that Proposition 3 extends to the case of dilution:

[24] A discussion of deliberation and the robustness of our main results to information pooling can be found in a working version of the paper, available online at http://mwpweb.eu/JustinValasek/.

Corollary 4

Assume that $\frac{W}{C} \in \left(\frac{\varepsilon^2(1-p_A)}{(1-\varepsilon)^2 p_A}, \frac{(1-p_A)}{p_A}\right)$. *If* $\lim_{n\to\infty} f(n, \{v_i\}_{i=1}^n) = 0$ *and* $q = \frac{1}{2}$, *then* $\sigma(a) = 1$, $\sigma(r) = 0$ *is an equilibrium for all sufficiently large* n.

Trivially, the condition on K satisfied for large n if $\lim_{n\to\infty} f(n, \{v_i\}_{i=1}^n) = 0$.

Next, we show that Proposition 4 extends to dilution of the form outlined above:

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 4, and follows by simply dividing both sides of $(1'')$ by g_n.

3.4 Extension: Additional Disesteem Payoffs

Due to the one-sided revelation of the state of the world, errors are only discovered when the committee approves the innovation. Above, we analyzed the effect of a disesteem payoff relating to a committee member voting to pass an innovation that is ultimately approved by the committee and revealed to be bad. This focus is motivated by the potential for negative media exposure faced by committee members who vote to approve a drug that turns out to have harmful side-effects, following the perception that committee decisions are only scrutinized when they go wrong. However, we cannot exclude the possibility that errors revealed after the committee's approval are punished through other channels. For example, committee members who incorrectly vote to reject a drug which is ultimately approved by the committee and turns out to be good may face internal social or professional sanctions from their committee peers. Furthermore, with other applications in mind, studying the behavior of committees allowing for this richer incentive structure is interesting in its own right.

In this subsection, we incorporate payoffs related to both errors into our framework by distinguishing between the disesteem payoff relating to a type I error, labeled k_1 (K in the analysis above), and a disesteem payoff relating to a type II error, labeled k_2, and characterize the behavior of large committees under this extended framework. Specifically, the payoff structure is modeled as follows:

$$U(v_i, X, \omega) = \begin{cases} 0 & \text{if } X = r \\ W & \text{if } X = a,\ \omega = A,\ v_i = a \\ W - k_2 & \text{if } X = a,\ \omega = A,\ v_i = r \\ -C & \text{if } X = a,\ \omega = R,\ v_i = r \\ -(C + k_1) & \text{if } X = a,\ \omega = R,\ v_i = a \end{cases}$$

where $W, C, k_1, k_2 > 0$.

The limit behavior of the committee as $n \to \infty$ when $\mathbf{k} = (k_1, k_2)$ can be generically characterized using the ratio of k_1/k_2.[25] That is, the absolute magnitude of the disesteem payoffs are irrelevant: no matter how small, the results of the model in the limit are determined by the relative magnitude of k_1 and k_2. Also, as shown by the following propositions, perfect information aggregation in the limit can only occur when $k_1/k_2 = (1-\beta)/\beta$; for any value of k_1/k_2 other than this point, the committee decision, in the limit, is always biased towards either rejecting or accepting.

The following expressions will be helpful in characterizing the limit equilibria with k_1, k_2.

$$\begin{aligned} L_b &= \frac{(1-\beta)p(t=a|s_i=r) + \alpha p(t=r|s_i=r)}{(1-\alpha)p(t=r|s_i=r) + \beta p(t=a|s_i=r)} \\ Z' &= \left(\frac{k_2(1-\beta) - k_1\beta}{k_1(1-\alpha) - k_2\alpha}\right) \frac{p(t=a|s_i=r)}{p(t=r|s_i=r)} \\ Z_1^n(\sigma) &= p_\sigma(X = A|t = a) \\ Z_2^n(\sigma) &= p_\sigma(X = A|t = r) \end{aligned}$$

Proposition **6**

Assume $q \in (\varepsilon, 1-\varepsilon)$.[26] *Define* σ^n *as an equilibrium of* $G^{\mathbf{k}}_{n,q}$. *Given the specified range of* $\frac{k_1}{k_2}$, *there exists a sequence of equilibria* $\{\sigma^n\}$ *such that as* $n \to \infty$ *the corresponding sequence of* $\{(Z_1^n(\sigma^n), Z_2^n(\sigma^n))\}$ *converges to:*

(i) For all values of k_1/k_2 : $(0, 0)$.

(ii) For $k_1/k_2 \in [L_b, (1-\beta)/\beta)$: $(1, Z')$.[27]

(iii) For $k_1/k_2 < L_b$: $(1, 1)$.

[25] Besides a finite number of values of k_1/k_2, such as $k_1/k_2 = (1-\beta)/\beta$ where the asymptotic committee's behavior does depend on the actual magnitudes of k_1 and k_2, not just on the value of the ratio.

[26] The results for $q \notin (\varepsilon, 1-\varepsilon)$ are largely analogous and are detailed in the Appendix.

[27] Note that $k_1/k_2 = (1-\beta)/\beta$ is not explicitly covered by this proposition; however, we address this case in detail in the Appendix following the proof of the proposition. At this value of k_1/k_2, the game is analogous to the SoW model analyzed in Section 3.1, which implies that for certain parameter ranges, there exists an equilibrium sequence of $\{(Z_1^n(\sigma^n), Z_2^n(\sigma^n))\}$ that converges to $(1, 0)$ (information aggregation).

Proposition 6 constitutes a proof of existence of limit equilibria (all formal proofs for this section are given in the Appendix, A.1). (i) follows directly from the existence a babbling equilibrium of the form $\sigma^n(a) = \sigma^n(r) = 0$. Note that this babbling equilibrium persists with the introduction of k_2 since information regarding the state of the world is not revealed when the innovation is rejected, which implies that neither disesteem payoff is incurred. (ii) shows that for an intermediate range of k_1/k_2, it is a limit equilibrium for the committee to accept the proposal when $t = a$, and to accept the proposal with a strictly positive probability when $t = r$. Finally (iii) shows that for k_1/k_2 small enough, it is an equilibrium for the committee to accept the innovation under both states of the art.

The following proposition establishes the uniqueness of the non-babbling limit equilibrium.

Proposition **7**

Assume $q \in (\varepsilon, 1-\varepsilon)$. Given the stated range of k_1/k_2, for any $\delta > 0$ there exists N, such that for all $n > N$, for all equilibria σ^n of $G^{\mathbf{k}}_{n,q}$, the corresponding $(Z^n_1(\sigma^n), Z^n_2(\sigma^n))$ are within a δ-neighborhood of:

(i) For $k_1/k_2 > (1-\beta)/\beta : (0,0)$.

(ii) For $k_1/k_2 \in [L_b, (1-\beta)/\beta) : (0,0)$ or $(1, Z')$.

(iii) For $k_1/k_2 < L_b$, $k_1/k_2 \neq \alpha/(1-\alpha) : (0,0)$ or $(1,1)$.[28]

First, (i) demonstrates that the main result of the previous analysis is robust to the introduction of k_2 as long as k_1 is large relative to k_2: for k_1/k_2 greater than $(1-\beta)/\beta$, a large committee essentially always reject the innovation.[29] Again, it is only the relative size of the disesteem payoffs that matter, so it may be the case that rejecting all innovations with arbitrarily high probability (as the size of the committee increases) is the only equilibrium even for very small values of k_1. To be more precise, how large k_1/k_2 needs to be depends on accuracy of the state of the art conditional on recommending approval; given both k_1 and k_2 strictly positive, the range over which rejecting the innovation is the unique limit outcome increases as the state of the art becomes less accurate.

Similarly, (iii) demonstrates that if k_1 is small enough relative to k_2, a large committee essentially always accepts the innovation (outside of the babbling equilibrium). The intuition for this result is quite similar to (i): for k_2 large relative to k_1, as long as there is a positive

[28] When $k_1/k_2 = \alpha/(1-\alpha)$, $(1,1)$ is the unique limit equilibrium of the form $(1, Z_2)$. However, as we discuss following the formal proof, we cannot exclude the existence of equilibria of the form $(0, Z_2)$.

[29] Note that for $k_1, k_2 > 0$, Propositions 6 and 7 also generally characterize the limiting behavior of the committee in the standard State of the World information aggregation model ($\alpha = \beta = 0$). The difference between the SoW and SoA models is further emphasized by Proposition 7; with $k_1, k_2 > 0$ and $\alpha = \beta = 0$ (SoW model), in the non-babbling equilibrium with n large, $Z^n_1(\sigma^n)$ is close to one no matter how large k_1 is relative to k_2.

probability of the committee passing the proposal, agents strictly prefer to vote for a to avoid k_2.

Case (ii) is less straightforward than the other two. For values of $(k_1/k_2) < (1-\beta)/\beta$, conditional on the committee decision matching the state of the art, voting to reject is not a best response for an agent with an r signal. Therefore, perfect information aggregation cannot be supported as an equilibrium in the limit; instead, agents with a signal of r mix between voting to accept and reject at a ratio such that $Z_1^n(\sigma^n) = 1$ and $Z_2^n(\sigma^n)$ is greater than zero, and their exposure to k_1 and k_2 is precisely balanced. The value of $Z_2^n(\sigma^n)$ that makes agents with a signal of r indifferent between voting to reject and accept, Z', increases continuously from 0 to 1 as k_1/k_2 decreases from $(1-\beta)/\beta$ to L_b.

Finally, note that Proposition 7 implies a discontinuity in the limit value of $p_{\sigma^n}(X = a|t = a) = Z_1^n(\sigma^n)$, since $\lim_{n\to\infty} Z_1^n(\sigma^n) = 0$ for $(k_1/k_2)^+ \to (1-\beta)/\beta$ and $\lim_{n\to\infty} Z_1^n(\sigma^n) = 1$ for $(k_1/k_2)^- \to (1-\beta)/\beta$. This jump occurs since for all $(k_1/k_2) < (1-\beta)/\beta$ a unique non-babbling limit equilibrium exists with $\sigma(a) = 1$, which implies $Z_1^\infty = 1$, while for $(k_1/k_2) > (1-\beta)/\beta$ the committee always rejects. To see why there is no limit equilibrium where agents with a signal of a mix between voting to reject and voting to accept and $Z_1^\infty \in (0,1)$, $Z_2^\infty = 0$, note that for n very large, the willingness to vote to reject for an agent with $s_i = a$ is strictly positive for $(k_1/k_2) < (1-\beta)/\beta$ as long as the committee votes to accept only when the state of the art is accept ($Z_1^n > 0$, $Z_2^n = 0$). Therefore, $Z_1^\infty \in (0,1)$, $Z_2^\infty = 0$ cannot be a limit equilibrium since it requires $\sigma_a < 1$, $\sigma_r = 1$, but the willingness to reject for agents with $s_i = a$ is strictly negative for all $Z_1^\infty \in (0,1)$, $Z_2^\infty = 0$.[30]

4 Conclusion

In this paper, we detail the effect of 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. Our paper also shows that the predictions of models of information aggregation can be sensitive to the standard assumption that experts' signals are independently distributed conditional on the state of the world. This distinction is empirically relevant, since it is unlikely that the decision that aggregates all current knowledge perfectly identifies the true state of the world; that is, due to imperfect evidence, even the "best" decision might be wrong ex post. Additionally,

[30]Put differently, in the limit game where the probability of being pivotal is equal to zero, given $(k_1/k_2) = (1-\beta)/\beta + h$ for $h > 0$, the willingness to vote to reject for $s_i = a$ is strictly positive when $Z_1^\infty \in (0,1)$, $Z_2^\infty = 0$. However, given $(k_1/k_2) = (1-\beta)/\beta - h$ for $h > 0$, the willingness to vote to reject for $s_i = a$ is strictly *negative* when $Z_1^\infty \in (0,1)$, $Z_2^\infty = 0$.

the state of the art model in this paper implies a particular correlation structure between experts' signals, and the general implications of such 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, an interesting area for future study is 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:

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

Unless otherwise stated we establish the following results, by analyzing the slightly more general game with the following payoff structure:[31]

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

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 :

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

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:

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

[31] 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

$R_a(n,\sigma) < R_r(n,\sigma) \leq 0$, 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 $q < \frac{n-1}{n}$,[32] 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:

$$G(y) : D_1 p_z(piv_i|t = r) + D_2 p_z\left(\frac{|H_i|}{n} > q|t = a\right) + D_3 p_z\left(\frac{|H_i|}{n} > q|t = r\right) - D_4 p_z(piv_i|t = a).$$ [33]

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:*

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

Note that $G(z)$ has at most one solution if and only if it satisfies (SCP).[34] $G(z)$ is a non-

[32] 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.

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

[34] 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 $(0,1)$.[35] 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 $\sum_{i=1}^{M} \alpha_i f_i$, 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:

[35]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 $(1-\varepsilon)$.[36] More generally, we will show that the hazard ratio evaluated at k:

$\dfrac{\binom{m}{k}\mu_a^k(1-\mu_a)^{m-k}}{\sum\limits_{j=k+1}^{m}\binom{m}{j}\mu_a^j(1-\mu_a)^{m-j}}$, 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:

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

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 $\sigma(a)=1$, $\sigma(r)=0$

[36] 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[37] 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)).[38]

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.

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

[38]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.

(1) $-p_z(piv_i|t=r)$ and $p_z(piv_i|t=a)$. 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 $\varepsilon < \frac{1}{2}$.

(2) $p(piv_i|t=a)$ and $-p\left(|H_i| > m|t=a\right)$. 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 $\mu_a \in (0,1)$.[39]

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

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

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

[39]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.

(3) $p(piv_i|t=a)$ and $-p\left(\frac{|H_i|}{n}>q|t=r\right)$. 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.[40]

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).[41] Assume that it is of form (1).[42] By Lemma 5 we

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

[41]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 $R_r(m',(0,1)) \geq 0$, 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.

[42]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:

$\frac{\partial z^*(p_A, \alpha, \beta, \varepsilon, q, n, K, W, C)}{\partial \theta}(>)(=)(<)0$ *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: From the analysis section, we know that truthful voting is an equilibrium when $K > 0$ if and only if:

$$\frac{\varepsilon p_A W p(piv_i|\omega = A) - (1-\varepsilon)(1-p_A)Cp(piv_i|\omega = R)}{(1-\varepsilon)(1-p_A)p(X = a|\omega = R)} \leq K$$
$$\leq \frac{(1-\varepsilon)p_A W p(piv_i|\omega = A) - \varepsilon(1-p_A)Cp(piv_i|\omega = R)}{\varepsilon(1-p_A)p(X = a|\omega = R)}$$

We begin the proof by showing that the RHS of the inequality is bounded below by expression 1 in Proposition 3.

Let $\mu_R = \sigma(a)pr(s_i = a|\omega = R) + \sigma(r)(1 - pr(s_i = a|\omega = R)) = \sigma(a)\varepsilon + \sigma(r)(1-\varepsilon)$ and $\mu_A = \sigma(a)(1-\varepsilon) + \sigma(r)\varepsilon$.
Note that the probability of incorrectly accepting the innovation, $p(X = a|\omega = R)$, is equal to:

$$\begin{aligned} &= \sum_{k=\lfloor nq \rfloor}^{n-1} \binom{n-1}{k} \mu_R^k (1-\mu_R)^{n-1-k} \\ &= \binom{n-1}{\lfloor nq \rfloor} \mu_R^{\lfloor nq \rfloor} (1-\mu_R)^{n-1-\lfloor nq \rfloor} \left[\sum_{k=0}^{n-1-qn} \prod_{m=1}^{k} \left(\frac{n-m-\lfloor nq \rfloor}{\lfloor nq \rfloor + m} \right) \left(\frac{\mu_R}{1-\mu_R} \right)^k \right] \end{aligned}$$

Now, looking at the term in brackets:

$$\begin{aligned} \left[1 + \sum_{k=1}^{n-1-qn} \prod_{m=1}^{k} \left(\frac{n(1-q)-m}{nq+m} \right) \left(\frac{\mu_R}{1-\mu_R} \right)^k \right] &< \left[1 + \sum_{k=1}^{n-1-qn} \prod_{m=1}^{k} \left(\frac{n(1-q)}{nq} \right) \left(\frac{\mu_R}{1-\mu_R} \right)^k \right] \\ &= \left[\sum_{k=0}^{n-1-qn} \left(\frac{(1-q)\mu_R}{q(1-\mu_R)} \right)^k \right] \end{aligned}$$

Returning to the main expression, the above equation shows that:

$$
\begin{aligned}
& \frac{(1-\varepsilon)p_A W p(piv_i|\omega=A) - \varepsilon(1-p_A)Cp(piv_i|\omega=R)}{\varepsilon(1-p_A)p(\frac{\lfloor H_i \rfloor}{n} > \frac{nq-1}{n}|\omega=R)} \\
> \; & \frac{(1-\varepsilon)p_A W \binom{n-1}{\lfloor nq \rfloor}\mu_A^{\lfloor nq \rfloor}(1-\mu_A)^{n-1-\lfloor nq \rfloor} - \varepsilon(1-p_A)C\binom{n-1}{\lfloor nq \rfloor}\mu_R^{\lfloor nq \rfloor}(1-\mu_R)^{n-1-\lfloor nq \rfloor}}{\varepsilon(1-p_A)\binom{n-1}{\lfloor nq \rfloor}\mu_R^{\lfloor nq \rfloor}(1-\mu_R)^{n-1-\lfloor nq \rfloor}\left(\sum_{k=0}^{n-1-qn}\left(\frac{(1-q)\mu_R}{q(1-\mu_R)}\right)^k\right)} \\
= \; & \frac{(1-\varepsilon)p_A W(\mu_A/\mu_R)^{\lfloor nq \rfloor}((1-\mu_A)/(1-\mu_R))^{n-1-\lfloor nq \rfloor}}{\varepsilon(1-p_A)\left(\sum_{k=0}^{n-1-qn}\left(\frac{(1-q)\mu_R}{q(1-\mu_R)}\right)^k\right)} - \frac{C}{\left(\sum_{k=0}^{n-1-qn}\left(\frac{(1-q)\mu_R}{q(1-\mu_R)}\right)^k\right)} \\
= \; & \frac{(1-\varepsilon)p_A W\left((\mu_A/\mu_R)^q((1-\mu_A)/(1-\mu_R))^{1-q}\right)^n\left(\frac{1-\mu_R}{1-\mu_A}\right)}{\varepsilon(1-p_A)\left(\sum_{k=0}^{n-1-qn}\left(\frac{(1-q)\mu_R}{q(1-\mu_R)}\right)^k\right)} - \frac{C}{\left(\sum_{k=0}^{n-1-qn}\left(\frac{(1-q)\mu_R}{q(1-\mu_R)}\right)^k\right)}
\end{aligned}
$$

Next, we proceed by construction. Assume $\sigma(a)=1$ and $\sigma(r)=0$, and $q=\frac{1}{2}$. Then:

$$
\left((\mu_A/\mu_R)^q((1-\mu_A)/(1-\mu_R))^{1-q}\right)^n = 1 \tag{2}
$$

since $\mu_A=(1-\varepsilon)$ and $\mu_R=\varepsilon$:

$$
\lim_{n\to\infty}\left(\sum_{k=0}^{n-1-qn}\left(\frac{(1-q)\mu_R}{q(1-\mu_R)}\right)^k\right) = \lim_{n\to\infty}\left(\sum_{k=0}^{n-1-qn}\left(\frac{\varepsilon}{(1-\varepsilon)}\right)^k\right) = \frac{1-\varepsilon}{1-2\varepsilon} \tag{3}
$$

Equations 2 and 3, taken together with above inequality, shows that:

$$
\begin{aligned}
& \lim_{n\to\infty}\frac{(1-\varepsilon)p_A W p(piv_i|\omega=A) - \varepsilon(1-p_A)Cp(piv_i|\omega=R)}{\varepsilon(1-p_A)p(\frac{\lfloor H_i \rfloor}{n} > \frac{nq-1}{n}|\omega=R)} \\
> \; & \frac{p_A W\left(\frac{(1-\varepsilon)^2}{\varepsilon}\right)}{\varepsilon(1-p_A)\frac{1-\varepsilon}{1-2\varepsilon}} - \frac{C}{\frac{1-\varepsilon}{1-2\varepsilon}} \\
= \; & \frac{p_A W\left(\frac{(1-\varepsilon)^2}{\varepsilon}\right) - C\varepsilon(1-p_A)}{\varepsilon(1-p_A)\frac{1-\varepsilon}{1-2\varepsilon}}
\end{aligned}
$$

which shows that the RHS of the inequality is bounded below by the expression in 1.

Now note that the numerator in the LHS of the inequality is given by:
$(p_A W - (1-p_A)C)\binom{n-1}{\lfloor nq \rfloor}(\varepsilon(1-\varepsilon))^{\frac{n}{2}}$,
which is strictly smaller than 0 as long as $\frac{p_A W}{(1-p_A)C} < 1$. So the above shows that truthful voting is an equilibrium when $q=\frac{1}{2}$ and $K>0$.

When $K = 0$ the analogous inequalities are:
$$\frac{\varepsilon p_A W p(piv_i|\omega = A) - (1-\varepsilon)(1-p_A)Cp(piv_i|\omega = R)}{(1-\varepsilon)(1-p_A) + \varepsilon_A} \leq 0$$
$$\leq \frac{(1-\varepsilon)p_A W p(piv_i|\omega = A) - \varepsilon(1-p_A)Cp(piv_i|\omega = R)}{(1-\varepsilon)p_A + \varepsilon(1-p_A)}$$
As shown above, the numerator of the LHS negative as long as $\frac{p_A W}{(1-p_A)C} < 1$. The numerator od the RHS is given by:
$$\left(\frac{1-\varepsilon}{\varepsilon} p_A W - \frac{\varepsilon}{1-\varepsilon}(1-p_A)C\right)\binom{n-1}{\lfloor nq \rfloor}(\varepsilon(1-\varepsilon))^{\frac{n}{2}}$$
which is positive as long as $\frac{\varepsilon^2}{(1-\varepsilon)^2} < \frac{p_A W}{(1-p_A)C}$. Note that this inequality also guarantees that our bound for K in the proposition:
$$\frac{p_A W \left(\frac{(1-\varepsilon)^2}{\varepsilon}\right) - C\varepsilon(1-p_A)}{\varepsilon(1-p_A)\frac{1-\varepsilon}{1-2\varepsilon}}$$
is positive. ∎

Proof of Proposition 4:

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

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:

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

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:

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

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* $\lim_{n\to\infty} f(n) \to 0$.

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 $\left(\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n} = 1\right)$ 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 4)*

Proof of Corollary 2: 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 2)*

Proof of Corollary 3:

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

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

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 $R_a(m', (\sigma^*(a), 0)) \leq 0$.[43] 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 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 3)*

A.1: Proofs for Section 3.4

For the proofs in this section, it will be useful for us to use the following function of $Z = (Z_1, Z_2)$:

$$\begin{aligned}\mathbf{R}_{s_i}(Z) =& Z_1\big(k_1 pr(\omega = R|t = a) - k_2 pr(\omega = A|t = a)\big)pr(t = a|s_i) + \\ & Z_2\big(k_1 pr(\omega = R|t = r) - k_2 pr(\omega = A|t = r)\big)pr(t = r|s_i),\end{aligned}$$

since, as the following Lemma shows, the willingness to vote to reject when observing signal s_i is close to $\mathbf{R}_{s_i}(Z)$ for n large.

Lemma **8**
For any $h > 0$, there exists N such that for all $n > N$, the willingness to vote to reject $R_{s_i}(n, \sigma)$ is within h of $\mathbf{R}_{s_i}(Z_1^n(\sigma), Z_2^n(\sigma))$, where $Z_1^n(\sigma) = p_{\sigma,n}(X = A|t = a)$ and $Z_2^n(\sigma) = p_{\sigma,n}(X = A|t = r)$.[44]

Proof: For n finite, the willingness to vote to reject can be rewritten as follows:

$$\begin{aligned}R_{s_i}(n, \sigma) =& - Wp(piv_i, \omega = A|s_i) + Cp(piv_i, \omega = R|s_i) + k_1 p(piv_i, \omega = R|s_i) \\ & + k_1 p(X = a, \neg piv_i, \omega = R|s_i) - k_2 p(X = a, \neg piv_i, \omega = A|s_i),\end{aligned}$$

where $\neg piv_i$ indicates the event that i is not pivotal. This in turn can be written as:

$$F'_{s_i}(piv_i, \sigma, n) + k_1 p(X = a, \neg piv_i, \omega = R|s_i) - k_2 p(X = a, \neg piv_i, \omega = A|s_i),$$

[43] 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.

[44] For notational ease, when there is no risk of confusion we do not explicitly denote the dependence of the probability terms on n.

where $F'_{s_i}(piv_i, \sigma, n)$ gathers the terms that involve the event that i is pivotal.

With some additional algebra, the above expression can be shown to equal:

$$F'_{s_i}(piv_i, \sigma, n)+[k_1p(\omega = R|t = a) - k_2p(\omega = A|t = a)]p(t = a|s_i)p(X = a, \neg piv_i|t = a)+ \\ [k_1p(\omega = R|t = r) - k_2p(\omega = A|t = r)]p(t = r|s_i)p(X = a, \neg piv_i|t = r).$$

Since $p(X = a, \neg piv_i|t = a) = Z_1^n(\sigma) - p(X = a, piv_i|t = a)$ and $p(X = a, \neg piv_i|t = r) = Z_2^n(\sigma) - p(X = a, piv_i|t = r)$, we can substitute these terms into the above equation and rearrange to get:

$$R_{s_i}(n, \sigma) = Z_1^n(\sigma)[k_1\beta - k_2(1-\beta)]p(t = a|s_i)+Z_2^n(\sigma)[k_1(1-\alpha)-k_2\alpha]p(t = r|s_i)+F_{s_i}(piv_i, \sigma, n),$$

where $F_{s_i}(piv_i, \sigma, n)$ captures all payoffs associated with the pivotal event. Note that since the probability of being pivotal approaches zero uniformly with respect to σ (see Lemma 7), there exists a function $m_{s_i}(n)$ converging to 0 as $n \to \infty$ such that for all σ we have $|F_{s_i}(piv_i, \sigma, n)| < m_{s_i}(n)$. Let $m(n) = max\{m_a(n), m_r(n)\}$ and consider N such that $m(n) < h$ for all $n > N$. Then we have that for all $n > N$:

$$\frac{k_1}{k_2} = |R_{s_i}(n, \sigma) - \mathbf{R}_{s_i}(Z_1^n(\sigma), Z_2^n(\sigma))| < h$$

∎

Next, we prove an analogous result to Lemma 1, showing that equilibrium strategies can only take a certain form:

Lemma **9**
If $k_1/k_2 < (1 - \beta)/\beta$, $k_1/k_2 \neq \alpha/(1 - \alpha)$, then for any $\delta > 0$ there exists N such that for all $n > N$, if there is an equilibrium of $G^{\mathbf{k}}_{n,q}$ such that either $Z_1^n(\sigma^n) > \delta$ or $Z_2^n(\sigma^n) > \delta$, then $\sigma_n(a) = 1$ and $\sigma_n(r) \leq 1$ or $\sigma_n(a) \geq 0$ and $\sigma_n(r) = 0$.

Proof of Lemma 9:
First, consider the case of $k_1/k_2 \in (\alpha/(1 - \alpha), (1 - \beta)/\beta)$. Take $h > 0$ small enough such that the following inequality holds:

$$\begin{aligned}&\mathbf{R}_r(Z_1^n(\sigma^n), Z_2^n(\sigma^n)) - \mathbf{R}_a(Z_1^n(\sigma^n), Z_2^n(\sigma^n)) \\ &=Z_1^n(\sigma^n)\big[\big(k_1\beta - k_2(1 - \beta)\big)\big(p(t = a|s_i = r) - p(t = a|s_i = a)\big)\big] \\ &\quad + Z_2^n(\sigma^n)\big[\big(k_1(1 - \alpha) - k_2\alpha\big)\big(p(t = r|s_i = r) - p(t = r|s_i = a)\big)\big] > 2h,\end{aligned}$$

where h is well-defined since both terms within brackets are strictly positive for $k_1/k_2 \in (\alpha/(1 - \alpha), (1 - \beta)/\beta)$, and either $Z_1^n(\sigma^n)$ or $Z_2^n(\sigma^n)$ is strictly greater than δ.

Take N large enough such that Lemma 8 holds for h, then we have that for $n > N$:

$$\mathbf{R}_r(Z_1^n(\sigma^n), Z_2^n(\sigma^n)) - \mathbf{R}_a(Z_1^n(\sigma^n), Z_2^n(\sigma^n)) > 2h \Rightarrow R_r(n,\sigma) - R_a(n,\sigma) > 0$$

So any equilibrium of $G_{n,q}^{\mathbf{k}}$ such that either $Z_1^n(\sigma^n) > \delta$ or $Z_2^n(\sigma^n) > \delta$ must be of the form $\sigma_n(a) = 1$ and $\sigma_n(r) \leq 1$ or $\sigma_n(a) \geq 0$ and $\sigma_n(r) = 0$.

Next, we consider the case of $k_1/k_2 < \alpha/(1-\alpha)$. When $k_1/k_2 < \alpha/(1-\alpha)$, then $\mathbf{R}_{s_i}(Z) < 0$ for all $Z \neq (0,0)$. Therefore, by Lemma 8 for n large enough the only possible equilibria are $\sigma = (0,0), (1,1)$. ∎

Lemma 10 allows us to further narrow the candidate non-babbling equilibria in the case of $\varepsilon < q < 1-\varepsilon$.

LEMMA **10**
If $\varepsilon < q < 1-\varepsilon$ and if $k_1/k_2 < (1-\beta)/\beta$, $k_1/k_2 \neq \alpha/(1-\alpha)$, then for any $\delta > 0$, there exists n^* such that for all $n > n^*$, if there is an equilibrium of $G_{n,q}^{\mathbf{k}}$ such that either $Z_1^n(\sigma^n) > \delta$ or $Z_2^n(\sigma^n) > \delta$, then $\sigma_n(a) = 1$ and $\sigma_n(r) \leq 1$.

Proof of Lemma 10:
By contradiction, assume that for all N, there is $n > N$ such that there exists an equilibrium, $\hat{\sigma}^n$, such that $\hat{\sigma}^n(r) = 0$ and $\hat{\sigma}^n(a) \geq 0$, and $Z_1^n(\hat{\sigma}^n) > \delta$ or $Z_2^n(\hat{\sigma}^n) > \delta$. Consider N_1 large enough so that Lemma 8 holds for h and so that we can apply Lemma 9 for all $n > N_1$. Given $\sigma^n(r) = 0$ and $\sigma^n(a) \geq 0$, then there exists $N > N_1$ such that for all $n > N$, $Z_2^n(\sigma^n) < \delta$ and $Z_2^n(\sigma^n)\big(k_1(1-\alpha) - k_2\alpha\big)p(t=r|s_i=r) < -\delta\big(k_1\beta - k_2(1-\beta)\big)p(t=a|s_i=r) - h$ for some small h. This implies:

$$\begin{aligned}\mathbf{R}_r(Z_1^n(\hat{\sigma}^n), Z_2^n(\hat{\sigma}^n)) + h &= Z_1^n(\hat{\sigma}^n)\big(k_1\beta - k_2(1-\beta)\big)p(t=a|s_i=r) \\ &\quad + Z_2^n(\hat{\sigma}^n)\big(k_1(1-\alpha) - k_2\alpha\big)p(t=r|s_i=r) + h < 0,\end{aligned}$$

since if $Z_2^n(\hat{\sigma}^n) < \delta$ it must be the case that $Z_1^n(\hat{\sigma}^n) > \delta$.

However, Lemma 8 implies

$$R_r(n,\hat{\sigma}) < \mathbf{R}_r(Z_1^n(\hat{\sigma}^n), Z_2^n(\hat{\sigma}^n)) + h < 0,$$

which contradicts $\hat{\sigma}^n(r) = 0$. ∎

Proof of Proposition 6:
Note that (i) follows trivially from the existence of a babbling equilibrium at $\sigma_a = \sigma_b = 0$. Also, we consider the cases of $\frac{k_1}{k_2} = \frac{\alpha}{1-\alpha}$, $\frac{k_1}{k_2} = \frac{1-\beta}{\beta}$ at the end of the proof, since these cases requires special treatment.

For $\frac{k_1}{k_2} < \frac{1-\beta}{\beta}$, $\frac{k_1}{k_2} \neq \frac{\alpha}{1-\alpha}$, Lemma 9 allows us to constrain the analysis to $\sigma(r) = 0, 0 \leq \sigma(a) \leq 1$, and $0 \leq \sigma(r) \leq 1, \sigma(a) = 1$, allowing us to define equilibria of the game using the same willingness to vote to reject function used for the main analysis:

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

Where $z = \sigma_a + \sigma_r$.[45] Also, since either $Z_1^\infty(z) \in [0, 1]$, $Z_2^\infty(z) = 0$, or $Z_1^\infty(z) = 1$, $Z_2^\infty(z) \in [0, 1]$, it will be useful to define $\mathbf{R}(\bar{Z})$ as a function of $\bar{Z} = Z_1 + Z_2$ over this subset of (Z_1, Z_2). Formally:

$$\mathbf{R}(\bar{Z}) = \begin{cases} \mathbf{R}(Z_1, 0) & \text{if } Z_2 = 0 \\ \mathbf{R}(1, Z_2)) & \text{if } Z_1 = 1 \end{cases}$$

From this expression we derive the bounds of L_b and $(1-\beta)/\beta$, as well as Z'. Specifically, L_b is the value of k_1/k_2 that solves $\mathbf{R}(2) = 0$, implying that $\mathbf{R}(2) < 0$ for values of $k_1/k_2 < L_b$. Similarly, $(1-\beta)/\beta$ solves $\mathbf{R}(0) = 0$, and Z' solves $\mathbf{R}(1+Z') = 0$ for $k_1/k_2 \in (L_b, (1-\beta)/\beta)$.

Moreover, this notation allows us to prove the following lemma:[46]

LEMMA **11**
(i) Given (Z_1^*, Z_2^*) such that $\bar{Z}^* \in (0, 1)$ and $\mathbf{R}(\bar{Z}^*) = 0$, there exists a sequence of equilibria $\{\sigma^n\}$ such that as $n \to \infty$ the corresponding sequence of $\{(Z_1^n(\sigma^n), Z_2^n(\sigma^n))\}$ converges to (Z_1^*, Z_2^*) if $\mathbf{R}(\bar{Z}^* - h) < 0$ and $\mathbf{R}(\bar{Z}^* + h) > 0$ for all $h \in (0, \bar{h})$ for some $\bar{h} > 0$.
(ii) If $\mathbf{R}(2) \leq 0$, there exists a sequence of equilibria $\{\sigma^n\}$ such that as $n \to \infty$ the corresponding sequence of $\{(Z_1^n(\sigma^n), Z_2^n(\sigma^n))\}$ converges to $Z = (1, 1)$.

Proof of Lemma 11: We use Lemma 8 and the following expression,

$$R_{s_i}(n, \sigma) = Z_1^n(\sigma)[k_1\beta - k_2(1-\beta)]p(t = a|s_i) + Z_2^n(\sigma)[k_1(1-\alpha) - k_2\alpha]p(t = r|s_i) + F_{s_i}(piv_i, \sigma, n),$$

[45] What is crucial here, is that when characterizing equilibria in mixed strategies we only need to verify the indifference of the agent between accepting and rejecting, after observing the signal under which he mixes. Lemma 9 then implies the optimality of his behavior under the other signal when n is large enough.

[46] We are interested in the roots of $\mathbf{R}(\bar{Z})$ to the extent that they correspond to limit points of acceptance probabilities in sequences of equilibria of the finite games. The fact that Z that do not have this shape are not candidate limit point stems from the informativeness of the signals with respect to the state of the art ($\varepsilon < \frac{1}{2}$). This assumption along with the shape of the equilibria, imply that from the perspective of an observer the probability μ_a^n that a randomly chosen agent votes to accept conditional on $t = a$ is always strictly higher than conditional on $t = r$ (μ_r^n) -as long as $\sigma^n(a)$ and $\sigma^n(r)$ are not both 0 or both 1-. By the law of large numbers, $p_{\sigma^n}(X = A|t) \in (0, 1)$ converges to an interior point of $[0, 1]$ only if $\mu_t^n \to q$, but it can't be the case that both μ_r^n and μ_a^n are converging to an interior point of $[0, 1]$.

to prove (i) by first proving that for n large, there exists an equilibrium within an h-neighborhood of (Z_1^*, Z_2^*) for each n, and then proving that a sequence of equilibria within the h-neighborhood of (Z_1^*, Z_2^*) converges to (Z_1^*, Z_2^*).

Take a sequence of $\{z_1^n\}$, such that $(Z_1^n(z_1^n), Z_2^n(z_1^n))$ converges to $(\bar{Z}^* + h)$, and $\{z_2^n\}$, such that $(Z_1^n(z_2^n), Z_2^n(z_2^n))$ converges to $(\bar{Z}^* - h)$. Such sequences exist since Z_1, Z_2 are continuous in z. Moreover, for all $n > N^+$, for some N^+, $R(n, z_1^n) > 0$ by Lemma 8. Similarly, for all $n > N^-$, for some N^-, $R(n, z_2^n) < 0$. This implies that for each value of n greater than N, where $N = \max\{N^+, N^-\}$, there exists an equilibrium value z^*, such that $R(n, z^*) = 0$, since $R(n, z_2^n) < 0 < R(n, z_1^n)$ and $R(n, z)$ is continuous in z. This proves the existence of a sequence of equilibria within a h-neighborhood of (Z_1^*, Z_2^*) for $n > N$.

Take a sequence of equilibria, z^{n*}, such that the corresponding sequence of $\{Z_1^n(z^{n*}), Z_2^n(z^{n*})\}$ is within a h-neighborhood of (Z_1^*, Z_2^*) for $n > N$, where h is small enough so that the neighborhood does not include $(0, 0)$. By construction $R(n, z^{*,n}) = 0$ all along the sequence, and furthermore $F_{s_i}(piv_i, \sigma, n) \to 0$. Therefore, $Z_1^n(z^n)[k_2\beta - k_1(1-\beta)]p(t = a|s_i) + Z_1^n(z^n)[k_2\alpha - k_1(1-\alpha)]p(t = r|s_i)$ must converge to 0. This can only happen if (1) $(Z_1^n(z^{n*}), Z_2^n(z^{n*}))$ are converging to (Z_1^*, Z_2^*); (2) $(Z_1^n(z^{n*}), Z_2^n(z^{n*}))$ are converging to $(0, 0)$; or (3) $(Z_1^n(z^{n*}), Z_2^n(z^{n*}))$ is alternating between a neighborhood of (Z_1^*, Z_2^*) and a neighborhood of $(0, 0)$. However, since $(Z_1^n(z^{n*}), Z_2^n(z^{n*}))$ is bounded from $(0, 0)$ for $n > N$ by construction, (2) and (3) are impossible.

The proof of (ii) follows a similar logic. First, if $\mathbf{R}(2) = 0$, then a sequence of equilibria such that the corresponding $\{(Z_1^n(z^{n*}), Z_2^n(z^{n*}))\}$ converge to $(1, 1)$ exists by the same argument as above since $\mathbf{R}(\bar{Z}^* - h) < 0$ for all h.

For $\mathbf{R}(2) < 0$, take $h > 0$, but small enough that $\mathbf{R}(2) + h < 0$. Next, take N large enough such that $|F(piv_i, z, n)| < h$ for all $n > N$. This implies that for $n > N$, $z = 2$ is an equilibrium since:

$$R_{s_i}(n, 2) = \mathbf{R}_{s_i}(2) + F(piv_i, z, n) < \mathbf{R}(2) + h < 0.$$

This in turn implies that a sequence of equilibria exist such that $(Z_1^n(z^{n*}), Z_2^n(z^{n*})) = (1, 1)$ is an equilibrium for all $n > N$. ⋄

The proof of (ii) and (iii) now follow directly from Lemma 11: For (ii), $\mathbf{R}(1 + Z') = 0$ and $\mathbf{R}(1 + Z' - h) < 0 < \mathbf{R}(1 + Z' + h)$ for all $h \in (0, \bar{h})$ for $\bar{h}$ small. For (iii), $\mathbf{R}(2) \leq 0$. For $\frac{k_1}{k_2} = \frac{\alpha}{1-\alpha}$, Lemma 11 does not apply; however, since $\mathbf{R}_{s_i}(1, 1) < 0$ it follows that an equilibrium with $(Z_1^n(\sigma^n), Z_2^n(\sigma^n))$ close to $(1, 1)$ exists for n finite but large. ▮

The case of $\frac{k_1}{k_2} = \frac{1-\beta}{\beta}$:
Here we show that when $\frac{k_1}{k_2} = \frac{1-\beta}{\beta}$, information aggregation can be supported in the limit

for certain parameter ranges. Notice that for this value of k_1/k_2, $\mathbf{R}_{s_i}(Z_1, Z_2)$ is equal to:

$$\mathbf{R}_{s_i}(Z_1, Z_2) = Z_2\big(k_1(1-\alpha) - k_2\alpha\big)p(t = r|s_i),$$

which suggests that information may be aggregated in the limit, since information aggregation implies $\lim_{n\to\infty} Z_2^n = 0$. In fact, we show below that the case of $\frac{k_1}{k_2} = \frac{1-\beta}{\beta}$ is analogous to the SoW model, which implies that information aggregation in reached in the limit for the comparable parameters detailed in Proposition 3.

PROPOSITION **8**
When $\frac{k_1}{k_2} = \frac{k_1}{k_2} = \frac{1-\beta}{\beta}$, *then there exist parameter values such that* $\{(Z_1^n(\sigma^n), Z_2^n(\sigma^n))\}$ *converges to* $(1, 0)$.

Proof: First, we show that the case of $\frac{k_1}{k_2} = \frac{1-\beta}{\beta}$ is analogous to the SoW model. Under parameters W, C, k_1, k_2, ε, p_A, α and β, when $\frac{k_1}{k_2} = \frac{1-\beta}{\beta}$ we have that for each signal s_i:

$$\begin{aligned} R_{s_i}(n,\sigma) = & Z_2^{\sigma^n}\big(k_1(1-\alpha) - k_2\alpha\big)p(t = r|s_i) \\ & - p(piv_i|t = a)\big((W - k_2)(1-\beta) - C\beta\big)p(t = a|s_i) \\ & + p(piv_i|t = r)\big(C(1-\alpha) - (W - k_2)\alpha\big)p(t = r|s_i) \end{aligned}$$

On the other hand under parameters W', C', k_1', k_2', α', β', $\varepsilon' = \varepsilon$, p_A', when $\beta' = 0, \alpha' = 0$ and $k_2' = 0$ (the state of the world model analyzed in Proposition 3) we have that for each signal s_i:

$$R'_{s_i}(n,\sigma) = Z_2^{\sigma^n} k_1' p'(t = r|s_i) - p(piv_i|t = a)W'p'(t = a|s_i) + p(piv_i|t = r)C'p'(t = r|s_i)$$

So let $k_1' = k_1(1-\alpha) - k_2\alpha$, $W' = (W - k_2)(1-\beta) - C\beta$, $C' = C(1-\alpha) - (W - k_2)\alpha$ and $p_A' = \frac{p_A - \alpha}{(1-\beta)-\alpha}$. Then $p'(t = r|s_i) = p(t = r|s_i)$ (and therefore also $p'(t = a|s_i) = p(t = a|s_i)$) , and therefore $R'_{s_i}(n,\sigma) = R_{s_i}(n,\sigma)$ for all n and σ.[47] Therefore, when $\frac{k_1}{k_2} = \frac{1-\beta}{\beta}$, Proposition 3 applies directly to W', C', p_A', k_1', p_A' and ε'.

Next, we show that the parameter set for which Proposition 3 applies is non-empty. Note that if the parameters W, C, k_1, ε, and p_A meet the two conditions of Proposition 3, then for sufficiently small (but positive) α and β and k_2, they will be met by W' C', p_A'. and k_1', p_A' and ε'. The reason is that as $\alpha \to 0$, $\beta \to 0$ and $k_2 \to 0$, the parameters in the second problem converge to those of the first problem, and so do the corresponding boundaries of the open sets that define the two conditions.[48]

[47] $Z_2^{\sigma^n}$, $p(piv_i|t = a)$ and $p(piv_i|t = r)$ are exactly the same in both expressions for all σ and n because $\varepsilon' = \varepsilon$.

[48] Concretely, pick W, C and k_1 which meet the conditions of proposition 3, given p_A and ε. Then set α, β and k_2 low enough so that the conditions are met by W', C', k_1 and p_A'. If at this stage $\frac{k_1}{k_2}$ is larger than $\frac{1-\beta}{\beta}$ then just decrease β further in order to restore equality, and this can only make the conditions slacker. Similarly, if at this stage $\frac{k_1}{k_2}$ is smaller than $\frac{1-\beta}{\beta}$ then just decrease k_2 further in order to restore equality.

Applying Proposition 3 we therefore have that when $q = \frac{1}{2}$ truth-telling is an equilibrium of the second problem for all sufficiently large n (and therefore of the first one, as they are exactly the same). Furthermore, since $\varepsilon < \frac{1}{2} < 1 - \varepsilon$ we also have that in the limit $Z_1^\infty = 1$ and $Z_2^\infty = 0$. That is, the committee aggregates information perfectly in the sense that for sufficiently large n it approximates the state of the art with arbitrarily precision. ▮

Proof of Proposition 7:
We proceed by contradiction, assuming that there exists δ such that for all N, there is $n > N$ such that there exists an equilibrium $\hat{\sigma}^n$ such that either $Z_1^n(\hat{\sigma}^n)$ or $Z_2^n(\hat{\sigma}^n)$ is more than δ away from the points specified in the proposition. The details of the proof differ between the individual cases. However, for each case, we make use of the fact that by Lemma 8, given σ^n, the willingness to vote to reject, $R_{s_i}(n, \sigma^n)$, will be close to:

$$\mathbf{R}_{s_i}(Z_1^n(\sigma^n), Z_2^n(\sigma^n)) = Z_1^n(\sigma^n)\big(k_1\beta - k_2(1-\beta)\big)p(t = a|s_i) + Z_2^n(\sigma^n)\big(k_1(1-\alpha) - k_2\alpha\big)p(t = r|s_i),$$

and proceed by showing that $\hat{\sigma}^n$ cannot be an equilibrium for n large.

Case (i) $\left(\frac{k_1}{k_2} > \frac{1-\beta}{\beta}\right)$*:* Here we consider $\hat{\sigma}^n$ such that either $Z_1^n(\hat{\sigma}^n)$ or $Z_2^n(\hat{\sigma}^n)$ is more than δ away from 0. Take $h > 0$ such that $min\{\delta(k_1\beta - k_2(1-\beta))p(t = a|s_i), \delta(k_1(1-\alpha) - k_2\alpha)p(t = a|s_i)\} > h$ for each signal $s_i = a$ and $s_i = r$ (such an h exists since both expressions within the brackets are strictly positive given $k_1/k_2 > (1-\beta)/\beta > \alpha/(1-\alpha)$).

By assumption, there exists $\hat{\sigma}^n$ such that either $Z_1^n(\hat{\sigma}^n) > \delta$ or $Z_2^n(\hat{\sigma}^n) > \delta$ and, by Lemma 8, such that $R_{s_i}(n, \hat{\sigma}^n)$ is within h of $\mathbf{R}_{s_i}(Z_1^n(\hat{\sigma}^n), Z_2^n(\hat{\sigma}^n))$. Therefore:

$$\begin{aligned} R_{s_i}(n, \hat{\sigma}^n) >& Z_1^n(\hat{\sigma}^n)\big(k_1\beta - k_2(1-\beta)\big)p(t = a|s_i = r) + Z_2^n(\hat{\sigma}^n)\big(k_1(1-\alpha) - k_2\alpha\big)p(t = r|s_i) - h \\ >& \min\{\delta(k_1\beta - k_2(1-\beta))p(t = a|s_i), \delta(k_1(1-\alpha) - k_2\alpha)p(t = a|s_i)\} - h > 0. \end{aligned}$$

So it follows that the unique best response of an agent is to vote to reject under both signals, and thereby $\hat{\sigma}^n$ with corresponding $Z_1^n(\hat{\sigma}^n)$ cannot be an equilibrium of $G_{n,q}^{\mathbf{k}}$.

Case (ii) $\frac{k_1}{k_2} \in \left[L_b, \frac{1-\beta}{\beta}\right)$*:* Here we consider $\hat{\sigma}^n$ such that $(Z_1^n(\hat{\sigma}^n), Z_2^n(\hat{\sigma}^n))$ is more than δ away from $(0, 0)$ or $(1, Z')$, in either dimension.

First notice that within the stated range of k_1/k_2, the unique interior crossing of $\mathbf{R}(\bar{Z})$ is at $(1 + Z')$, which implies that for some small enough h, there exists $w > 0$ such that if $Z_1^n(\sigma^n) > 1 - w$ and $|Z_2^n(\sigma^n) - Z'| > \delta$ then $|\mathbf{R}_r(\infty, (Z_1^n(\sigma^n), Z_2^n(\sigma^n)))| > h$. Take h small enough so that this relationship holds, and so that $\mathbf{R}_r(\infty, (1, 1)) > h$ and $\mathbf{R}_r(\infty, (1, 0)) < -h$ (these two expressions are strictly positive and negative (respectively) for the range of k_1/k_2 considered).

Next, pick N large enough so that (1) Lemma 10 holds for δ, (2) Lemma 8 holds for h, (3) $Z_1^n(\hat{\sigma}^n) > \max\{1-\delta, 1-w\}$, and (4) $R_r(n,(1,0)))$ is within h of $\mathbf{R}_r(\infty,(1,0))$ (which follows from the fact that $q \in (\varepsilon, 1-\varepsilon)$ implies $\lim_{n\to\infty}(Z_1^n(1,0), Z_2^n(1,0)) = (1,0)$). It follows that for the conjectured equilibrium $\hat{\sigma}^n$, $R_a(n,\hat{\sigma}^n) \neq 0$ and $R_r(n,\hat{\sigma}^n) \neq 0$ since $\mathbf{R}_{s_i}((Z_1^n(\hat{\sigma}^n), Z_2^n(\hat{\sigma}^n)))$ are at least h away from 0. So both $\hat{\sigma}^n(a)$ and $\hat{\sigma}^n(r)$ must be extreme points of $[0,1]$, which implies that the only candidates for $\hat{\sigma}^n$ are $(0,0)$, $(1,0)$ and $(1,1)$ (pure strategies).

However, $\hat{\sigma}^n$ cannot equal $(0,0)$, as this implies $(Z_1^n(\hat{\sigma}^n), Z_2^n(\hat{\sigma}^n)) = (0,0)$. Nor can $\hat{\sigma}^n$ equal $(1,1)$, as this requires $R_a(n,\sigma^n) < 0$ and $R_r(n,\sigma^n) < 0$ and implies $(Z_1^n(\hat{\sigma}^n), Z_2^n(\hat{\sigma}^n)) = (1,1)$, which cannot be true since $\mathbf{R}_r(\infty,(1,1)) > h$ implies $R_r(n,(1,1)) > 0$. Finally, $\hat{\sigma}^n$ cannot equal $(1,0)$, since this requires $R_a(n,\sigma^n) < 0$ and $R_r(n,\sigma^n) > 0$, which cannot be true since $\mathbf{R}_r(\infty,(1,0)) < -h$ and $R_r(n,(1,0)))$ is within h of $\mathbf{R}_r(\infty,(1,0))$, implying $R_r(n,(1,0))) < 0$. This chain of arguments contradicts that σ^n is an equilibrium of $G^{\mathbf{k}}_{n,q}$.

Case (iii) $\frac{k_1}{k_2} \in \big(0, L_b\big)$*:* Here we consider $\hat{\sigma}^n$ such that either $Z_1^n(\hat{\sigma}^n)$ or $Z_2^n(\hat{\sigma}^n)$ is more than δ away from 1 and 0.

Note that since $\mathbf{R}(\bar{Z}) < 0$ for all $\bar{Z} \in (0,1]$, $R_a(n,\hat{\sigma}^n) \neq 0$ and $R_r(n,\hat{\sigma}^n) \neq 0$ for large enough n by the same argument as in Case (ii). Therefore, $\hat{\sigma}^n$ must be a pure strategy equilibrium. However, as above, $\hat{\sigma}^n$ cannot equal $(0,0)$ or $(1,1)$ since this implies $(Z_1^n(\hat{\sigma}^n), Z_2^n(\hat{\sigma}^n))$ equals $(0,0)$ or $(1,1)$ (respectively). Also, $\hat{\sigma}^n$ cannot equal $(1,0)$ for n large enough since, as above, $\mathbf{R}_r(\infty,(1,0)) < -h$.

Finally, note that when $k_1/k_2 = \alpha/(1-\alpha)$, the same argument as above demonstrates that all equilibria with $Z_1^n(\sigma_n) >> 0$ must be within δ of $(1,1)$ for n large. However, since $\mathbf{R}(0, Z_2) = 0$ at this point, Lemma's 9 and 10 do not apply, and we cannot exclude the existence of equilibria close to $(0, Z_2)$. ∎

Propositions 6 and 7 for $q \notin (\varepsilon, 1-\varepsilon)$

Here we detail how the results of Propositions 6 and 7 extend to the case of $q \notin (\varepsilon, 1-\varepsilon)$. First, note that the statement and the proof of Proposition 6 also apply to $q \geq (1-\varepsilon)$. Also, the proof of Proposition 7 is only slightly different for $q \geq (1-\varepsilon)$, since the proof of (i) applies to all values of q, and for *(ii)* and *(iii)* the argument that for n large enough an equilibrium outside the specified range must involve pure strategy equilibria also applies, and it is straightforward to show that $\sigma^n = (1,0)$ is not an equilibrium for n large. Below, we state and prove the analogous propositions for $q \leq \epsilon$.

For the case of $q \leq \epsilon$, the following definitions will be helpful:

$$\hat{L}_b = \frac{(1-\beta)p(t=a|s_i=a) + \alpha p(t=r|s_i=a)}{(1-\alpha)p(t=r|s_i=a) + \beta p(t=a|s_i=a)}$$

$$\hat{Z}' = \frac{k_1(1-\beta) - k_2\beta}{k_2\alpha - k_1(1-\alpha)} \frac{p(t=a|s_i=a)}{p(t=r|s_i=a)}$$

Case 1 ($q < \epsilon$): For $q < \epsilon$, we provide the analogous result to Proposition 6 (the first paragraph of the results are unchanged, and are not restated here). The logic of the proofs below are very similar to the proofs of Propositions 6 and 7; therefore, we do not provide the same level of detail.

Proposition 6' ($q < \epsilon$):
(i) For all values of k_1/k_2, $\{(Z_1^n(\sigma^n), Z_2^n(\sigma^n))\} \to (0,0)$.
(ii) For $k_1/k_2 \in [\hat{L}_b, (1-\beta)/\beta)$, $\{(Z_1^n(\sigma^n), Z_2^n(\sigma^n))\} \to (1, \hat{Z}')$.
(iii) For $k_1/k_2 < \hat{L}_b$, $\{(Z_1^n(\sigma^n), Z_2^n(\sigma^n))\} \to (1,1)$.

Proof: For $q < \epsilon$, $\sigma_a = 1$ implies that $Z_1 = Z_2 = 1$ for $n = \infty$, which in turn implies that interior limit values of (Z_1, Z_2) can only be achieved in the range $z < 1$. That is, to find analogous limit equilibria to those in Proposition 6, we must look over the range of z where agents with a signal of accept are mixing between voting to accept and reject. Given $\hat{L}_b$ and $\hat{Z}'$, the proof of Proposition 6' (ii) and (iii) follow directly from Lemma 11 as above. ∎
The extension of Proposition 7 is similar.

Proposition 7' ($q < \epsilon$):
(i) For $k_1/k_2 > (1-\beta)/\beta : (0,0)$.
(ii) For $k_1/k_2 \in (L_b, (1-\beta)/\beta) : (0,0)$ or $(1, \hat{Z}')$.
(iii) For $k_1/k_2 \leq L_b$, $k_1/k_2 \neq \alpha/(1-\alpha)$: $(0,0)$ or $(1,1)$.

Proof: Note that, as in the proof of Proposition 7, for n large enough an equilibrium outside the specified range must involve pure strategy equilibria. (Note that for $q < \epsilon$, the comparable result to Lemma 10 specifies that all non-babbling equilibria be of the form $\sigma^n = (\sigma(a), 0)$ with $\sigma(a) \in (0,1]$ for large n; otherwise the argument is analogous.) For (ii), however, $\sigma = (1,0), (1,1)$ cannot be equilibria for n large, since $\{(Z_1^n(\sigma^n), Z_2^n(\sigma^n))\} \to (1,1)$. This argument also proves (iii) since $\{(Z_1^n(\sigma^n), Z_2^n(\sigma^n))\} \to (1,1)$ or $(0,0)$ for all pure strategy equilibria. ∎

Case 2 ($q = \epsilon$): The case of $q = \epsilon$ requires special treatment since interior values of (Z_1, Z_2) can be achieved with $\sigma(a) = 1$ and $\sigma(r) = 0$, which introduces some ambiguity regarding the limit equilibria in the interior case (ii).

Proposition 6" ($q = \epsilon$):
(i) For all values of k_1/k_2, $\{(Z_1^n(\sigma^n), Z_2^n(\sigma^n))\} \to (0,0)$.
(ii) For $k_1/k_2 \in (\hat{L}_b, (1-\beta)/\beta)$, for $n > N$ and any $h > 0$ an equilibrium σ^n exists such that $(Z_1^n(\sigma^n), Z_2^n(\sigma^n))$ are within an h-neighborhood of the range $(1, (Z', \hat{Z}'))$.
(iii) For $k_1/k_2 < \hat{L}_b$, $\{(Z_1^n(\sigma^n), Z_2^n(\sigma^n))\} \to (1,1)$.

Proof: The case of $q = \epsilon$ is complicated by the fact that $\mathbf{R}_a(\infty, (1 + Z_2)) < 0$ for $Z_2 \in (Z', \hat{Z}')$, while $\mathbf{R}_r(\infty, (1 + Z_2)) > 0$ for $Z_2 \in (Z', \hat{Z}')$. This implies that a convergent sequence of equilibria that is bounded from $Z = (0, 0)$ does not necessarily exist. However, since $\mathbf{R}_r(\infty, (1 + Z' - h)) < 0$ and $\mathbf{R}_a(\infty, (1 + \hat{Z}' + h)) > 0$, for some $n > N$ a sequence z_1 converging to $\bar{Z} = 1 + Z' - h$ will have $R(n, z_1^n) < 0$, and a sequence z_2 converging to $\bar{Z} = 1 + \hat{Z}' + h$ will have $R(n, z_2^n) > 0$. By continuity of R in z, for $n > N$ an equilibrium must therefore exist for some z^* with $(Z_1^{z^*}, Z_2^{z^*})$ in an h-neighborhood of the range $(1, (Z', \hat{Z}'))$. ∎

Proposition 7" ($q = \epsilon$):
(i) For $k_1/k_2 > (1 - \beta)/\beta : (0, 0)$.
(ii) For $k_1/k_2 \in (L_b, (1 - \beta)/\beta) : (0, 0)$ *or* $(1, (Z', \hat{Z}'))$.
(iii) For $k_1/k_2 \leq L_b$, $k_1/k_2 \neq \alpha/(1 - \alpha)$*:* $(0, 0)$ *or* $(1, 1)$.

Proof: Again, for n large enough an equilibrium outside the specified range must involve pure strategy equilibria. For (ii), $\sigma = (1, 1)$ cannot be equilibria for n large, since $\{(Z_1^n(\sigma^n), Z_2^n(\sigma^n))\} \to (1, 1)$. For $\sigma = (1, 0)$, the limit value of $\{(Z_1^n(\sigma^n), Z_2^n(\sigma^n))\}$ is not unique; however, in both (ii) and (iii), for n large enough $\{(Z_1^n(\sigma^n), Z_2^n(\sigma^n))\}$ outside the specified range cannot be an equilibrium, since either $R_r(n, (1, 0)) < 0$ or $R_a(n, (1, 0)) > 0$. ∎

Appendix B: The Case of the FDA

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."[49]

We collect data from FDA committee meetings held between January 2008 and August 2013.[50] 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 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

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

[50]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.

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 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

	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.

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.

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.