Discussion Paper 1234

Public Debate Among Experts

by

Ezra Friedman

November 13, 1998

http://www.kellogg.nwu.edu/research/math

Public Debate Among Experts

Preliminary Draft

Ezra Friedman[1]

Center for Mathematical Studies of Economics And Management Sciences
Northwestern University

November 13, 1998

Abstract

This paper presents a model of public debate in which experts attempt to influence public policy by making recommendations about controversial issues. However the decision to become an expert is taken to be endogenous, and consequently depends on the potential expert's bias. Under certain conditions there exist multiple equilibria. There may be one uninformative equilibrium in which only agents with strong biases are likely to become experts, and as a result the public gives experts little credibility. However, there might also exist another more informative equilibrium, in which more moderates function as experts, and the public places more weight on their reports. In the most informative equilibrium, increasing the heterogeneity of the public or decreasing the number of potential experts leads to an improvement in public information.

[1] Math Center, Northwestern University, Room 383, Leverone Hall, 2001 Sheridan Rd, Evanston, IL 60208. email: e-friedman@nwu.edu. I would like to thank Abhijit Banerjee for many involved and informative discussions regarding this work. I would also like to thank Bengt Holmstrom, Daron Acemoglu, Jean Tirole, Michael Kremer, David Spector, and seminar participants at the Math Center for useful advice regarding this paper and its precursors. Any errors are my own.

Section 1. Introduction

Information is seldom distributed uniformly through society. Because information is often costly to obtain and analyze it is inefficient for each member of society to independently research all issues of concern to them. However, sharing conclusions based on information tends to be very cheap and easy . Thus we might expect to see specialization in the gathering of information. There is a clear role for 'experts' to gather information on an issue, and then to analyze the information and share their conclusions with the public. People might be motivated to gather information and become an expert by the promise of direct rewards, as in the case of journalists or professional researchers. On the other hand, if the expert cares about public action, the incentive to gather information could come from the influence the expert has over the actions of the public. This is likely to occur in political debates in democracies, where the expert might be motivated by a desire for influence over public policy.

If the expert is motivated by a desire to influence the public, but the interests of the public and the expert do not exactly coincide, the expert might have something to gain by distorting the information she provides. For example an expert with a personal dislike of weapons might exaggerate the arguments for gun control. Recognizing that experts may have incentive to distort information, the public is unlikely to place complete faith in the recommendation of any expert.

Although experts who enter the public debate often provide facts and logical arguments to support their position, they generally have the ability to manipulate these facts and arguments to serve their own private interest. The fact that an expert can provide some evidence to back up her stated opinion does not imply that her recommendation is not based upon her biases, the expert may be suppressing compelling evidence that goes against her stated position. Thus experts have the ability to misuse their expertise. They can use their prestige as an expert to manipulate the public into taking an action advantageous to the expert, but against the public interest. The entire concept of expert opinion is based on the tenet that experts cannot easily transmit all of the information or analysis which leads them to hold a certain position, so at least some of the experts basis for a recommendation is unobservable. Because experts have the ability to manipulate information, the public will take into account its beliefs about the motivation of experts when interpreting the report of any expert.

This paper is concerned with how the ability of experts to misrepresent evidence affects the public debate, and how this ability interacts with potential experts' choice of whether to become involved in the debate. Specifically, the potential experts with the strongest vested interest in the issue, and the most to gain by manipulating the public are most likely to report on the issue. These experts, who will be referred to as extremists, are likely to base their report on their vested interest rather than on the information they have gathered. Thus their recommendations are not necessarily likely to be reflective of the true state of the world or the optimal policy. If the only experts that are likely to enter the debate are those with the greatest stake in the issue, rational citizens will not place much weight on the recommendations of any expert. As a result, moderate potential experts, those without a strong bias and with less incentive to enter the debate, will be unlikely to report on the issue. This can be thought of as an adverse selection effect, in which the very fact that someone wishes to join debate causes the public to question their motivation and lend little credence to their recommendation.

In this paper, I present a model where potential experts choose whether or not to spend resources learning the truth about an issue and make a non-verifiable recommendation on the issue to the public. The experts' primary incentive to report on the issue is a desire to influence the actions or beliefs of the public in choosing between two opposing policies. However, experts may also be motivated by intellectual curiosity. Some of these experts, referred to as moderates, do not have as strong a vested interest in either policy and always prefer the socially optimal policy. Other experts, referred to as extremists, have strong vested interest in one policy and prefer that policy whether or not it is socially optimal.

I find that under some conditions, there is an equilibrium in which only extremists are likely to enter the public debate, and as result the public learns little from the experts. I refer to this as the 'uninformative' equilibrium. However, there may also be another equilibrium in which almost all potential experts, including moderates, report on a particular issue. Consequently the public will be more responsive to any reports about this issue, and will be able to learn more from the experts. As a result many potential experts, both moderate and extremist will find it worthwhile to enter the debate about this issue. In addition, I show that if there exists both an informative equilibrium and an uninformative equilibrium there must be an intermediate equilibrium

Although this paper presents a static model, it might be reasonable to suppose that the public's expectations about experts' behavior are formed based on their past behavior. Thus the existence of multiple equilibria can be interpreted as suggesting that the past tone of the debate can affect the present tone of the debate. Issues that have in the past been discussed by moderates, whose positions are responsive to new information, are likely to be continue to be debated rationally and draw more moderates with new information into the fold, whereas issues that have been debated only by extremists are likely to continue to be debated only by extremists.

Decreasing the number of potential experts increases the likelihood that an informative equilibrium exists and increases the informativeness of this equilibrium. The intuition behind this result is that in the informative equilibrium all extremist potential experts already are experts, thus decreasing the number of potential experts will decrease the number of extremist experts in the debate. Decreasing the number of extremist experts can be thought of as decreasing the noise in the aggregate signal, making any moderate experts who enter more informative and more influential. Another way of looking at it is that decreasing the number of potential experts makes it more likely that any potential expert will choose to become an expert. This can be though of as exogenizing the choice of who becomes an expert and diminishing the adverse selection effect.

This result suggests that an arbitrary limit on who is able to participate in the debate might have positive effects on the quality of public information. In addition I show that limiting who is able to take part in the debate need not be done formally, it can instead take the form of a social norm. If the public believes that anyone outside of an arbitrary class who enters the public debate is likely to be an extremist, potential experts outside of that class will be given little credence, and will be unlikely to enter the debate.

Through the model, I find that increasing the heterogeneity of the population will generally lead to better public information in the informative equilibrium. This occurs because, in the informative equilibrium, the collected recommendations of the experts are likely to strongly indicate one state of the world, and as a result unbiased members of the public are likely to have already been convinced to act in the socially optimal fashion. The only members of society who are likely to be influenced by additional information are those who have a bias against the action that the majority of the experts are recommending. In a heterogeneous society more people will be biased and a marginal expert will be more influential, as a result more marginal experts will enter and the public will be better informed. In other words, once there is a preponderance of evidence in favor of one point of view, debate might stop in a homogenous society, but in a heterogeneous society there will be skeptics who remain to be convinced, and the debate will continue until the evidence is more conclusive.

Another result is that increasing the importance of intellectual curiosity or decreasing the degree to which extremists are extreme both increase the informativeness of the least informative equilibrium and might even block its existence. This occurs because both changes increase the relative likelihood that a moderate expert reports even when her report is unlikely to be influential, diminishing the strength of the adverse selection effect.. As a result, the public will be less likely to dismiss an expert as extremist when there are few experts entering.

Although the model assumes that the experts learn the true state of the world for certain, the results of the model are generally robust to relaxing the assumption. If this assumption is relaxed, The "Swing Voter's Curse" described by Feddersen and Pesendorfer (1996) becomes relevant, in that a moderate expert is more likely to be influential when she is wrong about the state of the world. I show that the effect of this "curse" will be that some moderately biased experts abstain from making recommendations after receiving their signal. However this does not lead to substantive changes in the results of the model, in particular the comparative statics results remain the same, and multiple equilibria are still possible.

Although the model presented in this paper is not a model of formal voting, the aggregation of information in this paper is similar to the aggregation of information through voting. Consequently the paper is related to the recent literature on aggregation of information through voting, including such papers as Myerson(1994) and Feddersen and Pesendorfer(1997). This paper shares more with Piketty(1994), which points out that the people who have the most interest in voting on an issue may be the people whose vote is least responsive to information. Allowing a market for votes may preclude information aggregation if it allows the extremists, whose votes convey no information, to purchase all the votes of the moderates, whose votes do convey information.

This model is more accurately described as a sender-receiver game in the line of Crawford and Sobel(1982), but with multiple senders and receivers. Krishna and Morgan (98) construct a model where there are two senders with known biases, and ask when the receiver would wish to consult only one sender and when he would wish to consult both. Dewatripont and Tirole (1996) explores the consequences of intentionally giving researchers vested interests, in order to give them incentive to discover information. Banerjee and Somanathan (1997) looks at issues closer to the concerns of this paper in that it models the transmission of information where the receiver questions the sender's motivation in speaking.

The model used in this paper is similar in structure to a model presented by Susanne Lohmann in a series of papers concerning political signaling [1994,1995]. In Lohmann's models, members of the public, who are heterogeneous in interest, receive private information and are able to take costly actions to signal this information to a policy maker. The results in this paper that the public learns only from moderates echoes her finding that the actions of the most extreme agents do not depend on their information, and that the only information comes from the actions of the moderates. However the model presented in this paper differs from hers in two key aspects. The first major difference is that actors in her model receive information *before* deciding whether or to take the costly action. As a result, the central question in her model is who decides to incur a cost and act after receiving the information, rather than the question of who decides to pay for the information. Consequently, the public information in Lohmann's model is generated by agents whose choice of whether or not to act is contingent on the information they receive. The second, and perhaps more important difference between the models is that in Lohmann's models, the distribution of interests is known, although the interest of any one particular agent is not. Because the exact number of extremists is known, it is possible to simply subtract the actions of the extremists from the aggregate actions and the presence of extremists does not add any noise or detract from the information imparted by the moderates. These two differences in formulation lead to major differences in results, most notably she does not obtain a result of multiple equilibria.

The remainder of the paper is organized as follows. Section 2 contains a formal presentation of the model , and section 3 outlines how the various actors will behave. The first major result I present is a sufficient condition for the existence of the uninformative equilibrium, which is presented in section 4. This is followed by sufficient conditions for the informative and intermediate equilibria in section 5. In section 6, I present the comparative statics results, notably that limiting the number of experts can have beneficial effects on public information, and this limitation can be achieved informally, through social norms. Section 7 demonstrates that the results of the model would not change drastically if the experts were imperfectly, rather than perfectly informed.

I conclude by trying to relate the model presented in this paper to actual debates encountered in the real world. I make an argument that in some public debates the static model considered in this paper is more appropriate than a dynamic model. Furthermore I argue that newspapers and other media outlets are often not effective in filtering out the recommendations of extremist experts from the public debate. because they print the opinions of all experts so as to counter any possible charges of bias. Finally I offer some possible examples of society increasing the quality of public information by the formation of a norm which excludes some people from the debate.

Section 2: The Model

The participants in the model consist of a general public of I individuals indexed by i and J potential experts indexed by j. In addition, nature determines the state of the world, denoted ω, which can take on two values, $\omega_H = 1$ or $\omega_L = -1$. Both the general public and the potential experts are assumed to place equal prior probabilities on either state of the world.

Each potential expert simultaneously chooses whether or not to do costly research to learn the true state of the world. After doing research, an expert makes a public statement claiming the state of the world is either ω_H or ω_L. The public is able to observe whether the expert has actually engaged in research, but the truth of the claim the expert makes is not verifiable. This represents a situation where a citizen can tell by the expert's statement that the expert is well informed and has "done her homework", but the citizen does not have enough independent information to judge the validity of the expert's claim. After the experts make their claims, each citizen sees the claims of all experts[2], and based on these claims, he chooses a social action x_i which can take on a value $x_H = 1$ or $x_L = -1$.

In any model of this type, absent further restrictions, there will exist a babbling equilibrium, an equilibrium in which there is no generally agreed upon correspondence between the experts' recommendations and the state of the world. That is to say experts are not influential not because they are not trusted, but because they are not understood. In this paper I assume a natural language, that is I assume that a recommendation by an expert always reflects the action preferred by the expert. Since an expert is always (weakly) more likely to prefer an action when it is socially optimal this assumption presents no contradictions.

The experts care mostly about the social actions taken by the public, but they also derive some intellectual satisfaction simply from doing the research and becoming experts. They have an objective function given by:

$$f_j(\mathbf{x},\omega) = \frac{1}{I}\sum_{i=1}^{I} (s_j+\omega)x^i + (b_j - c)d_j \qquad (1)$$

[2] The assumption that the citizen sees all recommendations is not vital to the central arguments of the paper. Although the quantitative results of the paper would change if it was assumed that citizens only saw a limited number of recommendations, or if the likelihood that a citizen saw a particular recommendation was a declining function of the number experts, the results of multiple equilibria, and the desirability of limiting the potential expert pool would still hold.

Where the variable s_j represents an expert's vested interest in the social action of the public and s_j is drawn independently for each potential expert according to the distribution function G(•). Because there are a finite number of experts, the realized distribution of interest is not precisely determined by G(•), but is subject to some degree of sampling error. The cost to researching the issue is denoted c, and d_j is an indicator variable representing whether the potential expert chooses to become an expert. The intellectual satisfaction derived from entering the debate is denoted b_j, and is assumed to vary independently of s_j with a distribution given by the function F(•). Experts are anonymous, in that the public is unable to directly observe the vested interest or the intellectual curiosity of any expert.

The experts can be divided into three types according to their interest. Any expert for whom $|s_j| < 1$ will be referred to as a moderate, and the action she prefers the public to take will depend on the state of the world. Any expert for whom $s_j < -1$ will always prefer that the public takes action x_L regardless of the state of the world, and will be called a leftist extremist, likewise if $s_j > 1$, the expert always prefers x_H and is a rightist extremist. Because a rightist extremist always prefers that citizens take action x_H, she will always claim that the state of the world is ω_H, likewise a leftist extremist will always claim the state of the world is ω_L. On the other hand, a moderate expert prefers that the public takes action x_H in state ω_H and action x_L in state ω_L. Thus she will always truthfully report the state of the world.

In order to simplify the calculation and expression of some of the results, we will impose very specific forms on the distribution functions.

F1: $f(b) = de^{-b/d}$

G1: $$G(s) = \begin{cases} 0 & s \leq -S \\ \frac{\gamma}{2} & -S < s \leq 0 \\ 1-\frac{\gamma}{2} & 0 < s \leq S \\ 1 & S \leq s \end{cases}$$

Note that **F1** implies that $F(b) = 1-e^{-b/d}$. The general results of the model are not critically sensitive to the exact form of the distribution over intellectual curiosity, but do depend on the thickness of the tails of this distribution. For example variations in d, which can be thought of as the importance of intellectual curiosity, will affect the properties of the equilibria, especially the uninformative equilibrium. Under **G1**, s_j is assumed to be drawn from three possible values -S, 0, and S, thus if experts are differentiated by interest, there are only three types. Again this atomistic distribution is chosen for simplicity, very similar results would obtain if the experts were assumed to be more smoothly distributed along interests. What is important is the relative mass of extremists, denoted here by γ and the degree to which the extremists are extreme, denoted here by S.

The public citizens, like the experts, have an interest in the social action of all citizens, including themselves. Their Von-Neumann Morgenstern objective function is:

$$w_i(\mathbf{x},\omega) = \frac{1}{I}\sum_{k=1}^{I}(z_i+\omega)x_k \qquad (2)$$

The distribution of z_i, the citizen's interests, is given by the cumulative probability distribution H(•) defined over the interval [-1,1] with probability density denoted h(•). The probability density is assumed to be symmetric, in that h(z)= h(-z). This implies that H(z) = 1-H(-z). Lastly, it is assumed that h(z) is bounded above by $\bar{h}$ and is continuous and weakly monotone decreasing in |z|. In other words the interests of the citizens are concentrated in the center, but there is a limit to how concentrated they are. We should note that the restriction of citizen's interests to the interval [-1,1] is not a substantive restriction of the model. Any agent with vested interest outside of this interval would prefer the same action regardless of his beliefs about the state of the world and would not be influenced by any expert opinion, and thus would be of no interest to the model. Likewise, the assumption that agents care about the other actions of other agents is not of central importance to the results, rather it is incorporated to maintain the flavor of the model as a model of debate over public policy.

Section 3: The Actions and Beliefs of the Public

Having set up the model, our next step is to describe how the beliefs and actions of the public respond to the recommendations of the experts. This enables us to calculate the value to an expert of making a report, so that we can derive the actions of the potential experts as a best response to public beliefs and actions. These preliminary results set the groundwork for the characterization of various equilibria in further sections of the paper.

Because the reports of moderate experts reflect the state of the world, but the reports of extremists do not, the degree to which the public is influenced by an expert depends on the probability the public assigns to that expert being a moderate. The public has no prior information about the interests of any particular experts, so all experts' reports are treated equally. Thus a citizen's posterior beliefs about the state of the world are a function of the aggregate report of the experts and the probability he places on any expert who enters the debate being a moderate. We will denote the number of experts reporting ω_H as v_H, the number reporting ω_L as v_L and define $v = v_H - v_L$. Let m be the probability that a citizen places on an expert being a moderate. We obtain the following result concerning the public's posterior beliefs:

Lemma 1: *The posterior probability a citizen places on state ω_H is given by:*

$$p_H = \frac{(1+m)^v}{(1+m)^v + (1-m)^v} \tag{3}$$

Proof: See Appendix

Note that according to the lemma the posterior beliefs that the public forms depend only on the *absolute* difference between the number of experts recommending the respective actions rather than the relative difference. Given the public's beliefs about the experts' reliability, the beliefs will be the same if only one expert reports and recommends x_L or if 51 experts recommend x_L and 50 recommend x_H. Furthermore if $(1+m)^v$ is close to 1, suggesting that the recommendations of the experts have not been very informative in aggregate, as will occur when m or v is small, the expression for p_H can be approximated by $\frac{1+mv}{2}$. In this case, the probability a citizen places on one state of the world will be linear in v, the vote of experts. Thus, when the aggregate recommendation is not very informative, the marginal effect of an additional recommendation on public beliefs is proportional to m and does not depend on the recommendations of the other experts.

The optimal action of a citizen is a function of his beliefs about the state of the world and his vested interest. Because the citizens act simultaneously, they cannot influence the actions of other citizens, thus they choose an action considering only its direct effect on their utility[3]. Citizen i chooses x_i such that.

$$x_i = \operatorname{argmax}_{x_i \in \{-1,1\}} E(z_i+\omega)x_i \tag{4}$$

Let p_H denote the posterior probability on state of the world ω_H induced by the experts' reports. A citizen with interest z_i will prefer to take action x_L only if $p_H(z_i+1) + (1-p_H)(z_i-1) \leq -p_H(z_i+1) - (1-p_H)(z_i-1)$, and will otherwise choose x_H. Note that increasing

[3] If we were to assume that readers cared only about their own actions so that $w_i(x_i,\omega) = (z_i+\omega)x_i$ identical results would obtain.

z_i increases the right side of the inequality, and decreases the left side, so there exists some $z^*(p_H)$ such that a citizen will choose x_L only if $z_i \leq z^*(p_H)$. The inequality reduces to $z_i+2p_H-1 > 0$, implying:

$$z^*(p_H) = 1-2p_H \tag{5}$$

Thus when the posterior beliefs are p_H, the likelihood that any given citizen chooses x_L is $H(z^*(p_H))$. Thus if an expert's report changes the posterior probability from p_H to p_H', the mass of citizens who change their actions is $H(z^*((p_H))-H(z^*(p_H'))$

As noted above a potential expert's motivation to enter the debate consists of intellectual interest, and desire for influence over policy. The value that an expert places on her influence over policy is the product of the likelihood that her recommendation changes a citizen's action and the benefit she derives when a citizen's action actually does change. The likelihood that an expert's recommendation changes a citizen's action depends on the interests of the public, the recommendation of the other experts and the likelihood that the citizen places on the expert being a moderate. The benefit the expert derives from actually changing a citizen's action depends on the state of the world and the expert's vested interest. We will relegate the exact derivation of the value of influence to the appendix and present the results. For the sake of compactness we define $q = \frac{1+m}{1-m}$, and **r** as a two dimensional vector $\{r_E, r_M\}$ which refers to the likelihood that an extremist or moderate potential expert chooses to enter the debate. We define $\beta(s_j, m, \mathbf{r})$ as the value an expert with interest s_j places on her influence over policy, given that public beliefs are m, and the actions of the other experts are summarized by the vector **r**.

Proposition 2: *The value an expert places on her influence over policy is given as follows:*

$$\beta(s_j,m,\mathbf{r}) = 2\sum_{v=-J}^{J} \pi(v|\omega_H,\mathbf{r})\ [H(1-\frac{2}{1+q^{-v}}) - H(1-\frac{2}{1+q^{-(v+1)}})] \qquad |s_j| \leq 1 \tag{6}$$

$$\beta(s_j,m,\mathbf{r}) = (|s_j|+1)\sum_{v=-J}^{J} \pi(v|\omega_H,\mathbf{r})\ [H(1-\frac{2}{1+q^{-v}}) - H(1-\frac{2}{1+q^{-(v+1)}})] +$$

$$(|s_j|-1)\sum_{v=-J}^{J} \pi(v|\omega_H,\mathbf{r})\ [H(1-\frac{2}{1+q^{-(v-1)}}) - H(1-\frac{2}{1+q^{-v}})] \qquad |s_j| \geq 1 \tag{7}$$

The first term on the right hand side of equation (7) represents the value of influence to a moderate when the state of the world is in line with her bias, and the second term denotes the value of influence when her bias is against the state of the world. There is only one term on the right hand side of equation (6) since a moderate always recommends the socially optimal action and is as likely to be pivotal whether the state of the world is in line with or opposed to her bias. Note that by equation (6) that if $|s_j| \leq 1$ then β, the value of influence, does not vary with s_j, but by equation (7) if $s_j > 1$, β increases at least proportionally with $|s_j|$. That is to say, all moderates value their influence equally regardless of the strength of their bias, whereas the value moderates place on their influence increases in the strength of their bias. The intuition for this is simply that a moderate will always recommend the policy appropriate for the true state of the world. Increasing the strength of the moderate's bias will increase her reward to being pivotal when the state of the world is in line with her bias, but will decrease the benefit to being pivotal when her bias is against the true state of the world, and will have no net effect[4]. On the other hands extremists always make a recommendation according to their bias, and strengthening their bias will increase the amount they have to gain in both states of the world.

Careful comparison of expressions 6 and 7 reveals that expression 7 is always greater, so compared to moderates, extremists will always place greater value on the influence derived from becoming an expert. In fact, we can more precisely express the relationship between a moderate's value and that of an extremist .

Corollary: $\dfrac{\beta(s_j,m,r)}{\beta(1,m,r)} > |s_j|$, *furthermore,* $\lim_{q \to 1} \dfrac{\beta(s_j,m,r)}{\beta(1,m,r)} = |s_j|$

Proof: See Appendix.

The corollary holds because an extremist expert is more likely to be pivotal when her recommendation is contrary to the state of the world, that is to say that the sum in the second term of equation 6 is greater than the sum in the first term. Since increasing her bias leads to a more than proportionate increase in the benefit when she is pivotal and her

[4] This result depends on the assumption of a linear loss rule, i.e. the utility can be thought of as distance from the actual policy to the expert's preferred policy. If the loss function were convex, the rewards to being an expert would be increasing in strength of bias for moderates, and the difference between the rewards to the extremist and the rewards to a moderate would be exaggerated. If the loss function were concave moderates could have greater rewards to being an expert than extremists.

recommendation is against the state of the world, increasing her bias leads to a more than proportionate increase her value of influence. It is worth noting that as q approaches 1, an extremist becomes nearly as likely to be influential in both states of the world, and the value of her influence is nearly proportional with her bias.

Because the influence of a recommendation depends on the strategies of the other experts, the value that experts place on this influence naturally depends on the experts' beliefs about the other experts' strategies. However we are searching for equilibria in which both the experts' beliefs and the public's beliefs are correct and correspond to the actual strategies of the experts. We will not miss any equilibria if we only consider situations where the experts' beliefs about the experts' actions (**r**) are consistent with the public's beliefs (m). It should be pointed out that m is not a complete description of experts' beliefs; the experts care not only about how likely it is that other experts are moderates, but also about how many other experts are debating the issue. However we can impose the partial equilibrium condition that the potential experts' beliefs about how many extremists become experts coincides with the actual proportion of extremists who find it optimal to become experts. In this case we find that there is only one possible set of expert beliefs that are consistent within public beliefs. We define **r**(m) as the expert beliefs which are consistent with public beliefs m[5].

The likelihood that an actual expert is a moderate is naturally a function of how likely it is that a moderate finds it optimal to enter compared to the likelihood that an extremist finds it optimal. Any potential expert for whom $\beta(s_j,m) > c-b_j$ will wish to enter the debate while any potential expert for whom $\beta(s_j,m) < c-b_j$ will choose not to. The actual likelihood that a potential expert with interest s finds it optimal to become an expert and enter the debate is given by:

$$\mathbf{r}^*(s) = 1 - F(\beta(s,m)).$$

Note that the function $\mathbf{r}^*(s)$ determines the vector $\hat{\mathbf{r}}$, which represents the actual likelihood that a potential expert of any given type enters the debate[6]. Of the potential

---

[5] $\mathbf{r}(m)=\{y(m), y(m)\frac{\gamma m}{(1-\gamma)(1-m)}\}$ and $y(m) = \pi(c-b_j < \beta(s_j,m) \mid |s_j|<1)$. There is a unique y(m) that satisfies this for any m, because $\beta(s_j,m)$ is decreasing in y(m) but y(m) is increasing in $\beta(s_j,m)$.

[6] $\hat{r} = \{\hat{r}_E, \hat{r}_M\}$ where $\hat{r}_E = \dfrac{\int_{-\infty}^{-1} r^*(s)g(s)\,ds + \int_{1}^{\infty} r^*(s)g(s)\,ds}{G(-1)+(1-G(1))}$, $\hat{r}_M = \int_{-1}^{1} \frac{r^*(s)g(s)}{G(1)-G(-1)}\,ds$

experts who enter the debate, we refer to the proportion who are moderate as m*, and m* is given by:

$$m^*(m) = \frac{\pi(\beta(s_j, m) < c - b_j \text{ and } |s_j| < 1)}{\pi(\beta(s_j, m) < c - b_j)} = \frac{\int_{-1}^{1} \mathbf{r}^*(s) g(s)\, ds}{\int_{-\infty}^{\infty} \mathbf{r}^*(s) g(s)\, ds} \tag{10}$$

Since $\beta(s_j, m)$ is weakly increasing in $|s_j|$, extremists always value their influence more than moderates, and extremists are always at least as likely to enter as moderates. Consequently, regardless of the public's beliefs $m^*(m) \leq 1-\gamma$.

We are now ready to introduce our definition of an informational equilibrium:

Definition: *An informational equilibrium is a set of public beliefs m and expert actions* ***r*** *such that the experts actions are optimal given public beliefs, and public beliefs correctly describe expert actions.*

Because m*(m) describes the experts' optimal actions given public beliefs m, it is easy to see that an equilibrium will occur if and only if m*(m)=m. The existence of at least one informational equilibrium is guaranteed because m*(m) is continuous in m, and $m^*(m) \in [0,1-\gamma]$. It is clear that there will be a fixed point in the interval $[0,1-\gamma]$ which represents an informational equilibrium. Having established the existence of an informational equilibrium, our next task is to describe the possible candidate equilibria, and to develop an intuition for when a particular equilibrium will exist.

Section 4: Uninformative Equilibrium

This section examines the conditions necessary for the existence of an equilibrium in which public information is poor and characterizes that equilibrium. Under the assumptions of our model, if the public believes that *only* extremists will become experts, experts will have no influence over the public, and potential experts will become experts only if motivated by intellectual interest. Since moderates would be as likely as extremists to enter if motivated by intellectual interest, an actual expert would be as likely to be a moderate as any potential expert, and public beliefs that no moderates enter must be wrong. Hence there can be no equilibrium where experts are completely uninformative.

On the other hand there could be an equilibrium where the public believes extremists are much more likely to participate than moderates, and consequently places only a little faith in the recommendations of experts. This equilibrium is especially likely to exist if intellectual interest is not very important. As soon as the public places some probability on an expert being moderate (m>0), experts gain some influence over the public, and by proposition 2, the value of influence is greater for extremists than for moderates. As a result extremists will be more likely to enter. As m increases so does $\beta(\bullet,m)$ and by proposition 2, the magnitude of the difference in the value of influence between moderates and extremists increases. If the probability distribution over intellectual interest falls off sharply as one moves away from 0, a relatively small difference in value of influence might lead to a large proportionate increase in the likelihood of a report. Thus an extremist may be many times more likely to enter than a moderate, leading to an equilibrium where the recommendations of experts are of a very little value.

We can directly evaluate the proportion of moderate experts entering as a function of public beliefs under our assumptions about the distribution functions. Under **G1** and **F1**, Expression (10) reduces to:

$$m^*(m) = \frac{1-\gamma}{1-\gamma+\gamma e^{(\beta(S,m)-\beta(1,m))/d}}. \tag{11}$$

It is easy to see that m* is decreasing in the difference between $\beta(S,m)$ and $\beta(1,m)$ as long as both are lower than c. As long as the aggregate recommendation is not expected to be very informative, $\beta(S,m) - \beta(1,m)$ is an increasing function of m, so m*(m) is decreasing in m as long as $\beta(1,m) < c$. Thus there will be only one equilibrium where $\beta(1,m) < c$. We will refer to this equilibrium as the uninformative equilibrium and define m_0 as the value of m associated with this equilibrium. This equilibrium is referred to as the uninformative equilibrium because it will always be less informative than the other possible equilibria.

Once $\beta(m) = c$, any further increase in m and $\beta(m)$ will lead to an increase in m*(m) because all the extremists will have already entered, and increasing the value of influence will only lead more moderates to enter. The minimum possible value for m*(m) thus occurs when $\beta_S(m) = c$ and is defined as $\underline{m}$, approximately given by:

$$\underline{m} = \frac{1-\gamma}{1-\gamma+\gamma e^{(c-(1/S)c)/d}}. \tag{12}$$

Note that by the Corollary to Proposition 2, the approximation in (12) is a maximum value for $\underline{m}$. If we define m° such that $\beta(S,m^\circ) = c$, we know that $m^*(m^\circ) = \underline{m}$. Since m*(0)

$=1-\gamma$ and m^* is continuous, if $\underline{m}^* < m^\circ$ there must be an $m_0 \in (0, m^\circ)$ such that $m^*(m_0) = m_0$. This m_0 represents an uninformative equilibrium, which is illustrated in figure 1.We are able to derive sufficient conditions for $\underline{m}^* < m^\circ$, and state the following proposition:

Proposition 3: *Under* ***G1*** *and* ***F1****, If* $S\bar{h}\left(1+\frac{\gamma}{1-\gamma}e^{\frac{(S-1)c}{Sd}}\right)^{-1} < c$ *there is an 'uninformative' equilibrium where not all extremists enter.*[7]

Proof: See Appendix.

[7] For a general distribution F(•) the parallel condition is $\bar{h}\left(1+\frac{\gamma}{1-\gamma}(1-F(\frac{S-1}{S}\bar{c}))^{-1}\right)^{-1} < 1$. This is much less intuitive, but one can see it is likely to hold if F(•) is concentrated close to zero. Unlike under F1, for general distributions F(•), it is possible that several 'uninformative' equilibria occur.

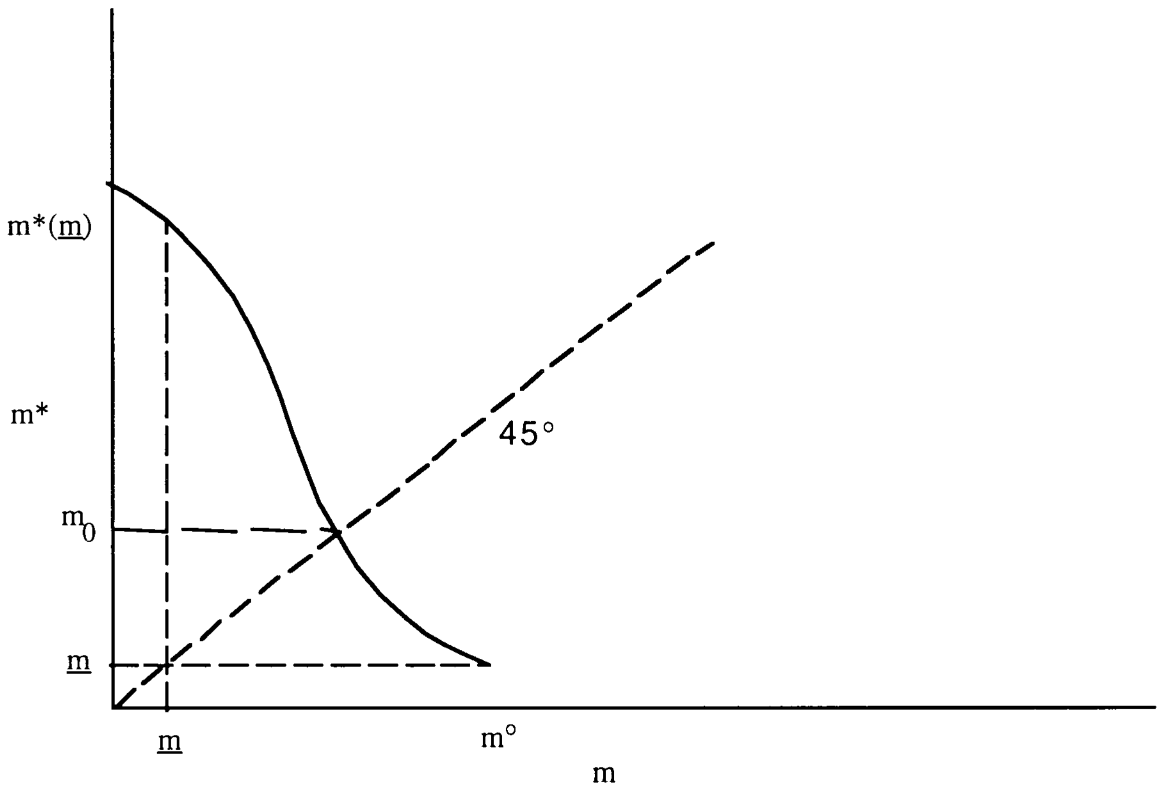

Figure 1: If the condition in proposition 3 is satisfied, $m^*(\underline{m}^*) > \underline{m}^*$ and $m^*(m^\circ) < m^\circ$ ensuring that there is an equilibrium at m_0.

The left side of the condition in proposition 3 is decreasing in c, decreasing in γ and increasing in d, thus as long is cost is high enough, there are sufficiently many extremists, or intellectual satisfaction is sufficiently unimportant, the uninformative equilibrium will exist. It can also be shown that if intellectual satisfaction is not very important, the left side of the condition is decreasing in S, so if the extremists are more extreme the uninformative equilibrium is more likely to exist.

If the extremists are not very extreme, (S is close to 1), or intellectual curiosity is not very concentrated (d is large), or there are very few extremists (γ is close to 0) then $\underline{m}$ could be greater than m°. In this case it would never be rational to believe that moderates are very unlikely to enter, and as a consequence experts would always be somewhat influential and extremists would always find it worthwhile to enter. The equilibrium in such an economy would occur when so many moderates entered that the public was almost certain to have strong beliefs about the state of the world and any additional expert was unlikely to be influential. This is the "most informative" equilibrium, and is discussed in the next section.

Whenever the uninformative equilibrium does exist, its informativeness will be a function of m_0, the equilibrium likelihood that an expert is moderate, as well as the number of experts at that equilibrium. By expression (11) it can be seen that m^* as a function of m_0 is decreasing in d, the expectation (and variance) of the intellectual interest, implying that the uninformative equilibrium will be less informative if experts are less motivated by intellectual interest. Also as intellectual interest, d, approaches zero, m_0, the equilibrium likelihood that an expert is moderate, also goes to zero. This shows that the equilibrium becomes arbitrarily close to being completely uninformative as intellectual interest becomes completely unimportant. The intuition behind this result is that as intellectual interest becomes less important, an individual's decision to become an expert relies more on how much they value influence, and it becomes less likely that any significant number of moderates will be experts unless the value of influence is so great that all extremists become experts. By expression 11, $m^*(m)$ is also decreasing in S, the strength of the extremists' biases. An increase in the extremism of extremists means that the difference in value of influence between moderates and extremists will be large, so that few moderates will enter relative to the number of extremists. Consequently if either d is small or S is large, there will be an equilibrium where there are very few moderates who enter and there is very poor public information.

Finally, it should be pointed out that for some parameter values the "uninformative" equilibrium could be quite informative. If there are very many potential experts, even if any one expert is very likely to be an extremist, the sheer number of signals could lead to a fairly reliable public signal. As long as the public is able to observe and process every recommendation, a large number of low value recommendations can add up to a very informative signal. In this case, under extreme circumstances there is a possibility of more than one uninformative equilibrium. This could occur if information is already very good so increasing m leads to an increase in the expectation that information will be very good. This, in turn leads to a decrease in the likelihood that an agent will be pivotal and hence a decrease in $\beta(\bullet,m)$.[8] However $m^*(m)$ is decreasing in $\beta(\bullet,m)$, so it will be increasing in m, and may cross above the 45° line, indicating a second uninformative equilibrium.

We have now established sufficient conditions for the existence of an uninformative equilibrium, and given a little intuition as to how informative this equilibrium will be. The existence of an uninformative equilibrium does not rule out the existence of more informative equilibria, and in the next section we will discuss the more informative equilibria.

Section 5: Informative equilibria

When an uninformative equilibrium does exist, there is still a possibility that there will be two more informative equilibria which occur when all extremists enter. We can conceive of an informative equilibrium where the public thinks that any expert is fairly likely to be a moderate. As a result the benefit to becoming an expert is high, and many potential experts both moderate and extremist find it optimal to enter. In this case the number of experts who enter is not limited by mistrust of experts, but by a crowding out among experts. There may be so many moderate experts that the aggregate recommendation is almost certain to strongly indicate one state of the world, at this point the recommendation of an additional expert is not likely to change the actions of many citizens, so few additional experts are likely to enter. There also may be an 'intermediate' equilibrium, where the public places just enough weight on the opinion of experts to convince some moderates to join and justify the weight placed on experts. The remainder of this sections more rigorously describes these equilibria and puts forth conditions sufficient for their existence.

If the cost to entering is not too high there will be some public beliefs which lead to values of influence great enough so that all extremists wish to enter. Since all extremists are already entering, further increasing the value of influence will now increase m* by increasing the number of moderates, and m*, the likelihood that an expert is a moderate will be given by:

$$m^*(m) = \left(1+\frac{\gamma}{1-\gamma}e^{(c-\beta(1,m))/d}\right)^{-1} \tag{13}$$

[8] It should be noted that unless S is much greater than 1, an increase in m implies a decrease in the number of moderates, making it unlikely that public information improves, and unlikely that there will be more uninformative equilibria.

We find that is convenient to take a slightly different approach to finding the informative equilibrium than we did for the uninformative equilibria. Rather than starting with public beliefs and asking whether those beliefs are consistent with experts optimal behavior, we will look at value of influence implied by certain public beliefs, and ask whether it matches the value of influence that implies those beliefs. We define $\beta^*(m)$ as the moderate's value of influence necessary so that if potential experts are allowed to choose whether to enter, and all extremists choose to enter, the likelihood that an expert is moderate is m. We can easily derive $\beta^*(m)$ by inverting expression (13) and we obtain:

$$\beta^*(m) = c - d \ln \left(\frac{1-\gamma}{\gamma}\frac{1-m}{m}\right) \tag{14}$$

It is easy to see that $\beta^*(m)$ is increasing in m. Since m is bounded above at $1-\gamma$, at which point all potential experts enter, $\beta^*(m)$ is bounded above by $\beta^*(1-\gamma) = c$.

Recall that $\beta(1,m)$ is the value a moderate expert places on her influence when public beliefs are m. An equilibrium occurs when $\beta(1,m) = \beta^*(m)$[9], so that the experts' expected actions are indeed optimal. If there is some m' such that $\beta(1,m') > c$, the value of influence is so great at m' that all potential experts would wish to enter. Thus, if $\beta(1,1-\gamma) < c$, as in figure 2, the graphs of $\beta^*(m)$ and $\beta(1,m)$ must cross so there is an m_2 such that $m' < m_2 < 1-\gamma$. This m_2 corresponds to the public beliefs in the informative equilibrium where not all moderates enter. If $S\beta(1,1-\gamma) > c$ there is an equilibrium at $m=1-\gamma$ where all experts enter. Thus if there is some m' such that $\beta(1,m') > c$, there exists an informative equilibrium.

Although the explicit expression for $\beta(1,m)$ is neither compact nor transparent and thus not shown in the body of the paper, it is straightforward to solve computationally. Furthermore it is possible to compactly express a lower bound for the maximum value of $\beta(1,m)$, and thus state sufficient conditions for the existence of an informative equilibrium.

Proposition 4: *If* $c \le \frac{.316}{\sqrt{J\gamma}}$ *and* $1-\gamma > \frac{1}{\sqrt{\gamma J}}$ *there is an informative equilibrium where at least* $\sqrt{J\gamma}$ *moderates become experts, and an unbiased citizen has at least a 84% chance of choosing his optimal action.*

Proof: See Appendix

---

[9] Whenever $\beta(S,m) > c$, If $\beta^*(n) = \beta(1,m)$ then $n=m^*(m)$, thus $\beta^*(m) = \beta(1,m)$ implies that $m^*(m) = m$, so the definitions of equilibrium are equivalent. Note that if $\beta(1,1-\gamma) > c$ this implies that there is an equilibrium where all potential experts enter.

The proof of proposition 4 involves a large amount of algebra, but is not particularly difficult conceptually. Essentially the proof establishes the minimum possible value for $\beta(1,m)$ as a function of m and evaluates this minimum at $\tilde{m} = \frac{1}{\sqrt{\gamma J}}$. At this point the expectation of the experts' vote conditional on the state of the world is one standard deviation away from 0, and it becomes likely that the aggregate recommendation of the experts is already informative, so further increasing m begins to decrease the value of influence, $\beta(1,m)$. If $\beta(1,\tilde{m}) >$ c there must be an informative equilibrium where $m > \tilde{m}$, so the conditional expectation of the experts' vote must be more than one standard deviation away from zero, and the likelihood that the experts' vote reflects the state of the world must be greater than 84%. Thus there is a lower bound to the quality of information in the informative equilibrium.

Having established the existence of an informative equilibrium, we now show that by continuity there will be an intermediate equilibrium whenever there is also an uninformative equilibrium. If an uninformative equilibrium exists then we know from Proposition 3 that $\beta(1,m^\circ) \leq \frac{c}{S}$. But we know that $\beta^*(m^\circ) > \frac{c}{S}$. Now if proposition 4 is satisfied then there is some m' where $\beta(1,m') > \beta^*(m')$. In this case, the graphs of $\beta(1,m)$ and $\beta^*(m)$ must intersect at some m_1 where $m^\circ < m_1 < m'$. This represents public beliefs in the intermediate equilibrium, where the quality of public information and number of moderates entering is between the uninformative and informative equilibrium. The intermediate and informative equilibria are illustrated in Figure 2

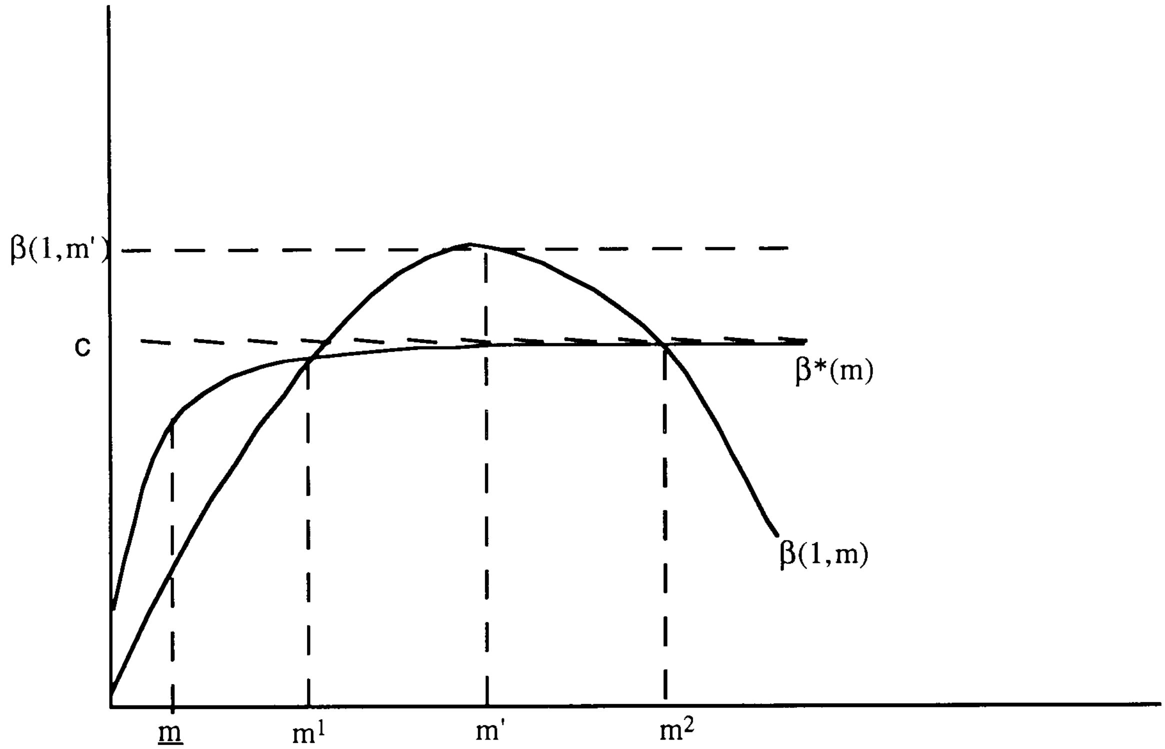

Figure 2: The informative equilibria

If we take the point of view that citizens form beliefs based on the past behavior of the experts we might conclude that the intermediate equilibrium is unstable, whereas the most informative and least informative equilibria are both stable. At m_1 the graph of β crosses β^* from below, so if m is decreased slightly below m', β will be lower than β^*, indicating that m* will be lower than m. This would lead to a further decrease in m which would lead to a decrease in m*, and so forth, until the uninformative equilibrium is reached. Similarly, a small increase in m would lead to β being greater than β^*, leading to increases in m* and m until the informative equilibrium is reached. On the other hand at the informative equilibrium β crosses β^* from above. A slight increase in m will lead to m*(m) decreasing below m, and the public will adjust their beliefs back to the equilibrium values.

Because the conditions in proposition 3 can always be satisfied as intellectual curiosity becomes irrelevant, (i.e. $d \to 0$), we will have multiple equilibria whenever intellectual curiosity is sufficiently unimportant. This finding of multiple equilibria suggests that the public expectations concerning the tone of a debate can actually determine the tone of the debate. If the public expects the debate to be purely ideological, or based on interests, rational public opinion will not be greatly affected by the experts' recommendations and only the most extreme potential experts are likely to become experts. On the other hand if the public expects moderate experts to participate in the debate, experts will have much more influence, and moderate experts might indeed wish to join the debate.

Section 6: Comparative Statics

Having established sufficient conditions for the existence of the informative equilibria, we are now able to show how the properties of these equilibria depend on the parameters of the model. Our first result is that as long as not all potential extremists enter, decreasing the absolute number of extremists improves information in the informative equilibrium, even if the relative number of extremists is increased. Since decreasing the absolute number of extremists improves information, decreasing the number of potential experts while holding the distribution of potential experts constant will increase the quality of public information, so an arbitrary restriction on expertise might be beneficial. We will also show that increasing the heterogeneity of the population increases the quality of public information in the most informative equilibrium.

As a preliminary step to obtaining the comparative statics results it is helpful to establish that as long as intellectual curiosity is not a major factor, both the intermediate and the informative equilibria will occur at a point where $\beta(1,m)$ is very close to c. The remaining results of this section will make use of this fact, and rely on intellectual interest being relatively minor. An examination of expression 14 shows that if d is very small, for any $m^*>>0$, $\beta^*(m)$ is very close to c. Thus in equilibrium , the value an expert places on her influence must be close to c. Consequently, rather than looking for points where $\beta(1,m) = \beta^*(m)$, we can just look for points where $\beta(1,m) = c$, and know that an equilibrium will be very close by. In general $\beta(m)$ is increasing at m_1, the intermediate equilibrium, but decreasing at m_2, the informative equilibrium. Thus anything that tends to increase $\beta(1,m)$ or decrease c will increase m_2 and decrease m_1, making the informative equilibrium more informative and the intermediate equilibrium less informative.

We now present our result that increasing the number of extremists will lead to a decrease in the quality of public information available at the informative equilibrium.

Proposition 5: *If m is small and γJ is large, whenever all moderates do not choose to become experts, increasing the absolute number of extremists decreases the quality of* *public information in the informative equilibrium. (If $\gamma J' > \gamma J$ and $m_2' < 1-\gamma$, $m_2' < m_2$ and $\pi(v'>0 \mid \omega_H') < \pi(v>0 \mid \omega_H)$)*

Proof: See Appendix

The intuition behind proposition 5 can be illustrated by the following thought exercise. Compare a debate where there are J experts, each with a probability of m of being moderate, to a debate with J' > J experts with an m' probability of being moderate. Now suppose that m' < m , and m' is such that the debates are equally informative. In this case an additional expert's signal in the first debate will be of higher quality, and since the debates are equally informative, an additional expert will have more influence in the first debate. Hence if the second debate is at the informative equilibrium, where the marginal moderate expert is just indifferent between entering and not entering, a moderate expert in the smaller debate will strictly wish to enter, showing that the equilibrium in the smaller debate will be more informative. The quality of public information does not depend directly on γ, the proportion of potential experts who are extremist, but rather depends on γJ, the *absolute* number of extremists. Thus it can be seen that if the distribution of potential experts is held constant, decreasing J, the total number of potential experts, leads to an improvement in public information.

The above results suggest that exogenously reducing the number of potential experts can have positive effects on the informational content of public debate even if there is no special screen or test aimed at filtering out extremists. Given the existence of an informative equilibrium, reducing the number of potential experts improves information as long as all moderate potential experts do not choose to become experts. Once the number of potential experts is small enough so that all potential experts choose to become experts, reducing the size of the debate no longer improves the expected quality of an individual expert's signal, but it reduces the number of signals, therefore worsening the quality of public information. Thus the optimum size of the potential expert pool will be such that $\beta(1,1-\gamma) = c$. That is when all potential experts become experts and the last moderate expert is just indifferent between becoming an expert or not.

Shrinking the potential expert pool can be thought of as reducing the adverse selection effect by partially exogenizing the decision to become an expert and making it more likely that any actual expert will be moderate. It can also be thought of as reducing the free riding problem among moderates, in which a moderate potential expert thinks "in order for me to have many influence, the public must think that there is an appreciable chance that any expert is a moderate. But there are so many experts that if there is an appreciable chance that an expert is a moderate, public information will be very good, and I will be unlikely to be influential"

Reducing the size of the potential expert pool need not be done by a formal or legal restriction. If the conditions for the existence of an uninformative equilibrium are met, then there can exist a mixed equilibrium. In this mixed equilibrium experts who satisfy an arbitrary observable criterion are thought somewhat likely to be moderates, while experts who do not satisfy the criterion are thought to be extremists. In other words the informative equilibrium holds for some class of potential experts but the uninformative equilibrium holds for the other class. As long as society is able to coordinate on a criterion, this will be an equilibrium. The moderate potential experts who satisfy the criterion will find it worthwhile to enter the debate since they will be trusted. On the other hand those who do not satisfy the criterion will have little incentive to enter the debate, knowing they are likely to be ignored. The public is justified in ignoring those who do not meet the criterion, but nevertheless become experts, since these experts are likely to be motivated by extremism.

We will now show that if the aggregate recommendation of the experts is sufficiently informative, increasing the heterogeneity of the public increases the value of influence ($\beta(1,m)$) and thus makes the informative equilibrium more informative. If m is small expression 6 can be approximated by:[10]

$$\beta(1,m) = \int_{-1}^{1} (\pi(v(z)|\omega_L) + \pi(v(z)|\omega_H))\, h(z)\, dz \qquad (8)$$

Where v(z) is the aggregate recommendation necessary so that a citizen with interest z is indifferent between the two possible actions.

Expression (8) can be interpreted as showing that holding m constant, the value of influence will be greatest when there are likely to be many citizens who are close to indifferent between the two actions. If there are relatively few moderates, the unconditional probability of seeing an aggregate recommendation of v(z),which is given by $\pi(v(z)|\omega_L)+ \pi(v(z)|\omega_H)$, will be greatest when z and v(z) are zero. At this point the value of influence is maximized when citizens are homogenous and h(0) is high. As the number of moderates increases it becomes less likely that the aggregate recommendation, v, is close to zero, and more likely that v represents the true state of the world. The unconditional probability $\pi(v(z)|\omega_L)+ \pi(v(z)|\omega_H)$ becomes bimodal with a trough at zero and peaks at M and -M,

[10] This is shown as an intermediate step in the proof of proposition 4.

where M is the number of moderate experts in the debate. Now the influence of experts will be greater when h(z(M)) is large. However, since M is considerable, z(M) will be close to 1, indicating that only citizens with a strong bias are likely to be indifferent. Increasing the heterogeneity of the public thus increases the proportion of citizens who are likely be swayed by an additional moderate's recommendation. This in turn increases the marginal moderate expert's benefit to entering, so the informative equilibrium occurs with more moderate experts and the quality of public information is higher. It should be pointed out that even though the quality of public *information* is improved, the likelihood that the socially optimal *action* is taken is not necessarily increased. Because the interests of the public citizens differ more, if the quality of public information is held constant, a citizen is more likely to be guided by his bias rather than by the aggregate recommendations when choosing an action. While the effect of increasing heterogeneity on information is clear, its effect on social welfare is ambiguous.

Section 7: Less Informative Signals

The assumption that experts who do research learn the true state of the world with certainty is admittedly a very strong assumption, and it would be disturbing if our results critically depended upon it. We first show that if all experts receive the same imperfect signal about the sate of the world are results do not change at all. As long as experts all receive the same imperfect signal about the state of the world, we can reparameterize the distributions so that there is an equivalent model where experts learn the true state of the world with certainty.[11] It is only necessary to define intermediate states of the world which correspond to the signals received by the experts, and compute expected utilities from either action conditional on the experts having received that signal.

However if the experts receive *independent* signals, the behavior of the potential experts is subtly changed, yet all of the major qualitative results of the paper still hold.

---

[11] Suppose that the signal received by the experts represented the true state of the world with probability $a>.5$. Let $\omega' \in \{-\frac{a-.5}{.5}, \frac{a-.5}{.5}\}$ refer to the signal received by the experts.

The experts objective function can be written $f_j(x,\omega') = \frac{1}{I} \sum_I (s_j+\omega')x_i$ - cr, and any expert for whom $|s_j| > \frac{a-.5}{.5}$ is an extremist. We can rewrite the citizens' utility function in a similar manner, and will find that the results are the same as if the experts learned the true state of the world, but the distribution functions were given by $G'(\bullet)$ and $H'(\bullet)$ so that $G'(s) = G(\frac{a-.5}{.5}s)$ and $H'(z) = H(\frac{a-.5}{.5}s)$.

Suppose that rather than learning the true state of the world an expert received an independent signal $\sigma \in \{\sigma_H, \sigma_L\}$, and that the likelihood that an expert received signal σ_H was a in state ω_H and $1\text{-}a$ in state ω_L. If there are many moderate experts, the "Swing Voter's Curse" described in Feddersen and Pesendorfer(1996) becomes relevant. When a moderate with some bias or a borderline extremist receives a signal contrary to her bias, she may strictly prefer to abstain.

If the aggregate recommendation of the experts favors one state, a new recommendation that contradicts the aggregate information will influence more citizens than a recommendation that agrees. If there are many moderate experts the aggregate recommendation is likely to reflect the true state of the world, so on average an expert is more likely to be pivotal if she reports a false signal than if she reports a true signal. This makes it less attractive for all potential experts to make recommendations because the likelihood that an expert 's recommendation is appropriate to the state of the world conditional on it being pivotal is less than the unconditional likelihood that it is appropriate.

Because the likelihood that the aggregate recommendation reflects the state of the world is increasing in m, the strength of the 'Swing Voter's Curse' is also increasing in m. However the following proposition shows that, conditional on being pivotal, the recommendation of an expert who truthfully reports her signal is always more likely to be appropriate than not.

Proposition 6: *If an expert receives signal σ_H and truthfully reports her signal, the probability she places on state ω_H is denoted p_H and $p_H > .5$.*

Proof: See Appendix.

It follows from proposition 6 that a completely unbiased expert will always wish to truthfully report her signal. On the other hand if an expert is somewhat biased and receives a signal against her bias she might wish to abstain. Consider an expert for whom $s_j = 2a\text{-}1$ who receives signal σ_L. Conditional on her being pivotal when she recommends x_L, the probability of being in state ω_L is less than a, so she would prefer that citizens take action x_H, and would not wish to recommend x_L. However if she recommends x_H, the probability of being in state ω_L when she is pivotal is more than a and she would prefer that citizens take action x_L. Since neither recommendation benefits the expert she will wish to abstain. Note that if she received signal σ_H, she would wish to recommend x_H, since even conditional on her being pivotal, the likelihood of being in state ω_H would be greater than half. We can refer to this type of expert as a semi-moderate[12], she reports truthfully if she receives a signal indicative of her preferred state and abstains otherwise.

Despite the added complications, the flavor of the results do not change if experts receive only an informative signal rather than learning the true state of the world. Save for the effect of the swing voters curse, the results are qualitatively identical. The graph of the payoffs to entering the debate may be flatter and more stretched out, but will have the same shape as when experts learn the true state of the world for certain. If the cost of becoming an expert is low enough, but not too low, the same multiple equilibria result will obtain.

Section 8: Extensions and Conclusion

We have shown that if experts are motivated by a desire to influence the public, the potential experts who have the most extreme interest in an issue will be most eager to influence the public's beliefs on this issue. Unfortunately, the recommendations of these extremist experts are unlikely to be influenced by the state of the world and are not informative. In short, the potential experts who are most likely to make a recommendation on an issue are precisely those whose recommendations society is least interested in. Because of this, if it is sufficiently costly to become an expert, there might be an equilibrium where only the most extreme are likely to become experts and enter the debate, and the public will pay little attention to the debate.

Reducing the number of extremist potential experts makes it more likely that the informative equilibrium will exist. If society arbitrarily limits the number of people who could potentially enter the debate, it will also limit the number of extremists. By doing this society may be able to insure against the uninformative equilibrium, or ensure the existence of the informative equilibrium. The limit would not need to be a formal limit: if there were social norms that only experts belonging to an arbitrary but observable class would be trusted, there could be an equilibrium where few potential experts outside of that class would enter the debate. Any expert outside of that class who entered the debate would be thought likely to be extremist, and would not affect the influence of the experts within that class. A social norm where members of a large arbitrary class of experts are ignored on a particular issue could develop to prevent extremists from taking over the debate. That is to say, it could be more efficient for society to arbitrarily choose whom to listen to, rather than to listen to all of those who wish to speak.

[12] The term semi-moderate is apt because the combined recommendations of a leftist and a rightist semi-moderate will be exactly the same as that of one moderate.

The lack of real world examples where society's norms of who to listen to are completely arbitrary should not be taken as a weakness in the theory. If society has a weak signal of an expert's reliability or bias, it would make more sense to coordinate on this signal. For example it is not unusual for a group of Nobel prize winners to make a pronouncement on some socially relevant issue, perhaps writing on something that is not directly related to their prize. These Nobel prize winners might not be the most informed of all possible people about the subject they are writing about, but the fact that they are Nobel prize winners reassures the public that they are not being motivated to participate by their own extremist biases. By the same intuition, we might expect editorials to be taken more seriously than letters to the editor, because the pool of those who can write letters to the editor is less restricted than those chosen to be editorialists. Furthermore when reading a letter to the editor, we are likely to put more weight on a letter from someone with a serious professional concern with the issue than a layperson. The involvement of entertainers in politics may be a case where a truly arbitrary class has disproportionate influence. Even though being an entertainer is not generally seen as a signal of expertise, entertainers are able to influence public opinion. If the public expects entertainers to enter public debates, the public is less likely to think that the entertainer is motivated by extremism. Thus the public will place more weight on the pronouncements of a celebrity than a layperson, even if they appear equally well informed.

This paper has focused on a static model of an individual debate, and assumed that experts were in a way anonymous, that is that the public could not distinguish moderates from extremists, except on the basis of their recommendation. This can be justified in that in many public debates the participants are known principally for their participation in that debate. Although experts gain a reputation for their position in this one debate, it might be very difficult to know what their biases were before they entered the debate. For example it may be well known that an activist strongly favors gun control, and this activist might be thought extreme on that issue, but his reputation as an extremist arises only from his position on gun control. It is very difficult to divine a person's motivation simply by her actions, it might be impossible to distinguish if someone's strong view ex-post was the result of her ex-ante extremism, or the result of a strong signal. In other words, it cannot really be said he had a reputation before the debate.

In situations where experts were not anonymous, any moderate expert might wish to send some signal about her bias to the public. Of course if an expert's bias was perfectly known there would be no informational problem, the public would simply disregard all the recommendations of those known to be extremist. If the public has a noisy estimation of an expert's bias, the public will place more weight on a recommendation by an expert suspected of a bias against her recommendation, rather than for her recommendation. This explains why participants in a debate will always be quick to point out factors that would lead them to have a bias against the position they are espousing[13].

Concern over reputation could conceivably aggravate the multiple equilibria findings. If a potential expert's bias over several issues is correlated, there will be an additional cost to entering a debate about a particular issue when the uninformative equilibrium holds. Because anyone who enters this debate is thought likely to be an extremist, by entering this debate they will diminish their credibility in further debates. Therefore, the people who would be most likely to enter this debate would be those are most extreme on the first issue, so extreme that they are willing to trade away future influence on future issues. Furthermore people might have direct preferences over reputations, for example moderates might not wish to be seen as extremists out of fear of being alienated from friends and colleagues. Extremists on the other hand might wear their extremism as a badge of pride. If there were a direct cost to a moderate arising from being thought of as an extremist, she might be even less likely to speak under the uninformative equilibrium.

Finally it is worth saying a word about mediation of debate. Even if it is inefficient for one citizen to do the research necessary to determine the bias of an expert who participates in a debate, it might seem worthwhile for the public to appoint gatekeepers, such as newspapers or media outlets to do this research and to exclude clear extremists from public debate. However this only raises the question of how to monitor the gatekeepers. If the gatekeepers have a bias, these gatekeepers might exclude moderates whose recommendations are against their biases, and might include extremists whose biases agree with their own. Thus the mediated debate might be less informative or reliable than the unmediated debate.

In conclusion, this paper shows that public debate among anonymous experts poses some special problems for information aggregation. Problems of both an adverse selection and a moral hazard nature arise: those who are most drawn to provide information may be those whose information is least trustworthy, and all providers of information might have incentive to manipulate the information they provide. Society can deal with the adverse selection problem by exogenizing the choice of entering the debate, that is society can set up norms by which it chooses who to listen to, rather than listening to whomever chooses to speak.

[13] This intuition is similar to that of Cukierman and Tomassi (1997)

References

Banerjee, A. and R. Somanathan. 1997. "A Simple Model of Voice", mimeo, MIT

Crawford, V. and J. Sobel. 1982, "Strategic Information Transmission" *Econometrica,* **50**, 1431-1451

Cukierman, A and M. Tommasi. 1997. "When Does it Take a Nixon to Go to China?", mimeo

Dewatripont, M and J. Tirole. 1995. "Advocates", mimeo, IDEI and MIT

Feddersen, T and W. Pesendorfer. 1996. "The Swing Voter's Curse", *American Economic Review* **86**(3), 408-424

Feddersen, T and W. Pesendorfer. 1997. "Voting Behavior and Information Aggregation in Elections with Private Information", *Econometrica* **65**(5), 1029-1059

Krishna, V. and J. Morgan. 1998, "A Model of Expertise", mimeo, Penn State and Princeton University

Lohmann, Susanne. 1994. "Information Aggregation Through Costly Political Action". *American Economic Review* **84**(3), 518-30

Lohmann, Susanne. 1995. "A Signaling Model of Competitive Political Pressures". *Economics & Politics* **7**(3), 181-206

Myerson, Roger B. 1994. "Extended Poisson Games and the Condorcet Jury Theorem", Northwestern University Math Center (CMS-EMS) Discussion Paper: 1103.

Piketty, T. 1994. "Information Aggregation Through Voting and Vote Trading", Mimeo MIT

Appendix

Proof of Lemma 1

If the reader thinks that the probability that a columnist is moderate is m, and places an equal probability on the columnist being a left or right extremist, the likelihood that the newspaper truly reports the state of the world is $m+\frac{1-m}{2}$. The likelihood that v_L columnists report the left state of the world and v_R columnists report the right state, given that the world is in the left state is given by $\frac{(v_L+v_R)!}{v_L!v_R!}(.5+\frac{m}{2})^{v_L}(.5-\frac{m}{2})^{v_R}$. The likelihood that the same report occurs given the right state of the world is

$$\frac{(v_L+v_R)!}{v_L!v_R!}(.5-\frac{m}{2})^{v_L}(.5+\frac{m}{2})^{v_R}.$$

Define $v= v_R-v_L$

By Bayes' Rule, if p_p is the prior probability placed on $\omega=1$:

$p_1 = \pi(\omega=1|v) =$

$$\frac{p_p\frac{(v_L+v_R)!}{v_L!v_R!}(.5-\frac{m}{2})^{v_L}(.5+\frac{m}{2})^{v_R}}{p_p\frac{(v_L+v_R)!}{v_L!v_R!}(.5-\frac{m}{2})^{v_L}(.5+\frac{m}{2})^{v_R}+(1-p_p)\frac{(v_L+v_R)!}{v_L!v_R!}(.5+\frac{m}{2})^{v_L}(.5-\frac{m}{2})^{v_R}}$$

Dividing through by $\frac{(v_L+v_R)!}{v_L!v_R!}(.5+\frac{m}{2})^{v_L}(.5-\frac{m}{2})^{v_L}$ gives us

$$\frac{p_p\ (.5+\frac{m}{2})^{v}}{p_p(.5+\frac{m}{2})^{v}+(1-p_p)(.5-\frac{m}{2})^{v}}$$

By assumption $p_p = .5$

$$\text{So } p_1 = \frac{(.5+\frac{m}{2})^{v}}{(.5+\frac{m}{2})^{v}+(.5-\frac{m}{2})^{v}}=\frac{(1+m)^{v}}{(1+m)^{v}+(1-m)^{v}}$$ ♦

Proof of Proposition 2

Any columnist for whom $|s_j| = 1$ is a moderate and will always make a recommendation based on the state of the world. Her gain from being pivotal in state ω_H is: $(s_j+ 1) -(-s_j - 1) = 2+2s_j$. Her gain from being pivotal in state ω_L is $(-s_j+1) - (s_j -1) = 2-2s_j$. Since her actions are symmetric, she is equally likely to be pivotal in both states so her expected gain conditional on being pivotal is 2. If the state of the world is ω_H and the vote of the other columnists is v, she will always vote right, so the new vote will be v+1. The likelihood that any one reader changes her action due to the columnists vote is the likelihood that his interests lie between z*(v) and z*(v+1). Substituting in from expression (5), this is given by: $H(1-\frac{2}{1+q^{-v}}) - H(1-\frac{2}{1+q^{-(v+1)}})$. Thus the likelihood that she is pivotal conditional on being in state ω_H is $\sum_{v=-J}^{J} \pi(v|\omega_H,r)\,(H(1-\frac{2}{1+q^{-v}}) - H(1-\frac{2}{1+q^{-(v+1)}}))$

Since her behavior is symmetric, the likelihood that she is pivotal is the same in both states, so her expected benefit is her benefit conditional on being pivotal multiplied by the likelihood that she is pivotal, and is given by:

$$\beta(s_j,m,r)=2\sum_{v=-J}^{J} \pi(v|\omega_H,r)\,(H(1-\frac{2}{1+q^{-v}}) - H(1-\frac{2}{1+q^{-(v+1)}})) \qquad |s_j| \le 1 \qquad (6)$$

A rightist extremist with $s_j > 1$ will always recommend x_H, if the state of the world is ω_H her benefit conditional on being pivotal is $2s_j+2$, we have already established the likelihood of a columnist who recommends x_H being pivotal in state of the world ω_H. Thus conditional on ω_H the expected benefit for an extremist is:

$$(2s_j+2)\sum_{v=-J}^{J} \pi(v|\omega_H,r)\,(H(1-\frac{2}{1+q^{-v}}) - H(1-\frac{2}{1+q^{-(v+1)}}))$$

If the state of the world is ω_L, a rightist extremist will still recommend x_H and see a benefit of $2s_j -2$ if she is pivotal. Conditional on a vote of -v, the likelihood that she is pivotal is $H(1-\frac{2}{1+q^{v}}) - H(1-\frac{2}{1+q^{v-1}})$, by symmetry $\pi(-v|\omega_L) = \pi(v|\omega_H)$ and $H(1-\frac{2}{1+q^{v}}) - H(1-\frac{2}{1+q^{v-1}})$ $= H(1-\frac{2}{1+q^{-(v-1)}}) - H(1-\frac{2}{1+q^{-v}})$ thus the benefit to an extremist in state ω_L is

$$(2s_j-2)\sum_{v=-J}^{J} \pi(v|\omega_H,r)\,(G(1-\frac{2}{1+q^{-(v-1)}}) - G(1-\frac{2}{1+q^{-v}}))$$

Since the prior probability placed on either state is .5

$$\beta(1,s_j,m) = (s_j+1)\sum_{v=-J}^{J} \pi(v|\omega_H,r)\,(H(1-\frac{2}{1+q^{-v}}) - H(1-\frac{2}{1+q^{-(v+1)}})) \;+$$

$$(s_j-1)\sum_{v=-J}^{J}\pi(v|\omega_H,r)\left(H\left(1-\frac{2}{1+q^{-(v-1)}}\right)-H\left(1-\frac{2}{1+q^{-v}}\right)\right) \qquad |s_j|\geq 1 \qquad (7)$$

Proof of Corollary to Proposition 2

Define $\psi_1 = \sum_{v=-J}^{J}\pi(v|\omega_H,r)\left(H\left(1-\frac{2}{1+q^{-v}}\right)-H\left(1-\frac{2}{1+q^{-(v+1)}}\right)\right)$ and

$$\psi_2 = \sum_{v=-J}^{J}\pi(v|\omega_H,r)\left(H\left(1-\frac{2}{1+q^{-(v-1)}}\right)-H\left(1-\frac{2}{1+q^{-v}}\right)\right)$$

Claim $\psi_2 > \psi_1$:
Proof
Define

$$\xi(v) = H\left(1-\frac{2}{1+q^{-v}}\right)-H\left(1-\frac{2}{1+q^{-(v+1)}}\right) \qquad (2.1)$$

If v = 0, $\xi(v) < \xi(v-1)$. and if v =1 $\xi(v) > \xi(v-1)$. Furthermore by the symmetry of h(•),
$\xi(v) = \xi(-v-1)$. (2.2)

$$\psi_1 = \pi(0|\omega_H,r)\,\xi(0) + \sum_{v=1}^{J}\pi(v|\omega_H,r)\xi(v) + \sum_{v=-J}^{-1}\pi(v|\omega_H,r)\,\xi(v) \qquad (2.3)$$

Define the three terms in the sum as a_1^0, a_1^+ and a_1^- respectively

$$\psi_2 = \pi(0|\omega_H,r)\,\xi(v-1) + \sum_{v=1}^{J}\pi(v|\omega_H,r)\,\xi(v-1) + \sum_{v=-J}^{-1}\pi(v|\omega_H,r)\xi(v-1) \qquad (2.4)$$

Likewise, define the three terms in the sum as ψ_2^0, ψ_2^+ and ψ_2^-. One can see that $a_1^0 = \psi_2^0$,

Since $\pi(v|\omega_H) = \pi(-v|\omega_L)$ We can rewrite

$$\psi_2^- = \sum_{v=1}^{J}\pi(v|\omega_L,r)\xi(v) \qquad (2.5)$$

and ψ_1^- as $\sum_{v=1}^{J}\pi(v|\omega_L,r)\xi(v-1)$. (2.6)

So $\psi_2^+ + \psi_2^- = \sum_{v=1}^{J}(\pi(v|\omega_H,r)\,\xi(v-1)+\pi(v|\omega_L,r)\xi(v))$ and

$\psi_1^+ + \psi_1^- = \sum_{v=1}^{J}(\pi(v|\omega_H,r)\,\xi(v)+\pi(v|\omega_L,r)\xi(v-1))$. Since for any $v>0$ $\pi(v|\omega_H) > \pi(v|\omega_L)$ and $\xi(v-1) > \xi(v)$, $\psi_2^+ + \psi_2^- > \psi_1^+ + \psi_1^-$ Thus $\psi_2 > \psi_1$

By (6) and (7), $\dfrac{\beta(s_j,m,r)}{\beta(1,m,r)} = \dfrac{(|s_j|+1)\psi_1 + (|s_j|-1)\psi_2}{2\psi_1}$ since $\psi_2 > \psi_1$

$\dfrac{\beta(s_j,m,r)}{\beta(1,m,r)} > \dfrac{(|s_j|+1)\psi_1 + (|s_j|-1)\psi_1}{2\psi_1} = |s_j|$ so the first part of the proposition is proven.

By the continuity of $h(\bullet)$ as $q \to 1$, $H(1-\frac{2}{1+q^{-v}}) - H(1-\frac{2}{1+q^{-(v+1)}}) \to$

$2h(1-\frac{2}{1+q^{-v}})((1+q^{-v})^{-1} - (1+q^{-(v+1)})^{-1}) = 2h(1-\frac{2}{1+q^{-v}}) \ln q \frac{q^{-v}}{(1+q^{-v})^2} h(\bullet)$ so as $q \to 1$,

$\dfrac{\xi(v-1)}{\xi(v)} \to 1$, and $a_2 \to a_1$. Thus $\dfrac{\beta(s_j,m,r)}{\beta(1,m,r)} \to |s_j|$ ♦

Proof of Proposition 3:

Consider the maximum possible influence of a columnist when m = m̲. Since the interests of readers is concentrated in the center, a columnist's influence is greatest when the vote of the other columnists is zero.

So holding m constant the maximum likelihood that a columnist is pivotal is $H(1-\frac{2}{1+q^0}) - H(1-\frac{2}{1+q^{-1}})$. Since $h(\bullet)$ is bounded by $\bar{h}$ this must be less than $\bar{h}\,(0-(1-\frac{2}{1+q^{-1}}))$. Recall that $q = \frac{1+m}{1-m}$, so the expression becomes $-\bar{h}(1-\frac{2}{1+\frac{1-m}{1+m}}) =$

$$-\bar{h}\left(\frac{\frac{1-m}{1+m}-1}{1+\frac{1-m}{1+m}}\right) = -\bar{h}\left(\frac{\frac{-2m}{1+m}}{1+\frac{1-m}{1+m}}\right) = \bar{h}m$$

Thus for any $m \le \underline{m}$ the likelihood that a columnist is pivotal must be less than $\bar{h}\underline{m}^*$ and $\beta(S,\underline{m}) < S\bar{h}\underline{m}$. We can rewrite (12) as $\underline{m}^* = \left(1+\frac{\gamma}{1-\gamma}e^{\frac{(S-1)c}{Sd}}\right)^{-1}$. Thus if $S\left(1+\frac{\gamma}{1-\gamma}e^{\frac{(S-1)c}{Sd}}\right)^{-1} < c$, $\beta(S,\underline{m}) < c$. By (11) if $\beta(S,\underline{m}) < c$ then $m^*(m) > \underline{m}$, for any $m \le \underline{m}$.

Suppose there is some $m' \in [0,1-\gamma]$ s.t. $\beta^*(S,m') = c$, then $m^*(m') = \underline{m}$. By above $m' > \underline{m}$, so by the continuity of $m^*(m)$ $\exists\, m_0 \in (\underline{m}, m')$ s.t. $m^*(m_0) = m_0$, and this represents the uninformative equilibrium.

Suppose $\beta^*(S,m) < c\ \forall\ m$, in this case, no matter what the public beliefs are, some extremists will not wish to enter. Thus the only possible equilibrium will be the uninformative equilibrium ♦

Proof of Proposition 4

By the continuity of h(•) if q is small so that 1 vote is not very important we can approximate the likelihood that a columnist is pivotal given a vote v as $h(z^*(v))\frac{dz^*}{dv}$. Thus the unconditional likelihood a columnist is pivotal is written Λ and is approximately given by:

$$\Lambda = .5\int_{-J}^{J} \pi(v|\omega_H)\frac{dz^*}{dv}\, h(z^*(v))\, dv + .5\int_{-J}^{J} \pi(v|\omega_L)\frac{dz^*}{dv}\, h(z^*(v))\, dv \tag{4.1}$$

$$\Lambda = .5\int_{-J}^{J} (\pi(v|\omega_H) + \pi(v|\omega_L))\frac{dz^*}{dv}\, h(z^*(v))\, dv \tag{4.2}$$

Note that if we integrate over interest rather than vote we obtain

$$\Lambda = .5\int_{-1}^{1} (\pi(v(z)|\omega_H) + \pi(v(z)|\omega_L))\, h(z)\, dz \tag{4.3}$$

Recall that $z^*(v) = 1-\frac{2}{1+q^{-v}}$. Differentiating and substituting in we obtain:

$$\Lambda = .5\int_{-J}^{J} (\pi(v|\omega_L)+\pi(v|\omega_H))\frac{\ln q}{q^{v}+2+q^{-v}}\, f(z^*(v))\, dv \tag{4.4}$$

recall $\frac{\pi(v|\theta_H)}{\pi(v|\theta_L)} = q^{v}$, so $\frac{1}{q^v+2+q^{-v}} = \frac{1}{\frac{\pi(v|\theta_H)}{\pi(v|\theta_L)}+2+\frac{\pi(v|\theta_L)}{\pi(v|\theta_H)}} = \frac{\pi(v|\theta_H)\pi(v|\theta_L)}{(\pi(v|\theta_H)+\pi(v|\theta_L))^2}$

Thus

$$\Lambda = .5 \ln q \int_{-J}^{J} \frac{\pi(v|\theta_H)\pi(v|\theta_L)}{(\pi(v|\theta_H)+\pi(v|\theta_L))} f(z^*(v))\, dv \tag{4.5}$$

If there are many columnists, M of whom are moderates and γJ of whom are extremists the conditional distribution $\pi(v|\omega_H)$ is approximated by a normal distribution with mean M and variance γJ. So if $M = \frac{1}{\sqrt{\gamma J}}$ the unconditional distribution $\pi(v|\omega_H)+\pi(v|\omega_L)$ peaks at 0. By equation 4.3, and our assumption that h(•) peaks (weakly) at 0, we see that over admissible distribution functions Λ is minimized when h(•) is completely flat and $h(z) = \frac{1}{2}$

Hence if $m \leq \frac{1}{\sqrt{\gamma J}}$

$$\Lambda \geq .5 \ln q\left(\int_{-J}^{J} \frac{\pi(v|\theta_H)\pi(v|\theta_L)}{(\pi(v|\theta_H)+\pi(v|\theta_L))} \frac{1}{2}\, dv\right) \tag{4.6}$$

$$\Lambda \geq .5 \ln q \int_{-J}^{0} \frac{\pi(v|\theta_H)}{\left(1+\frac{\pi(v|\theta_H)}{\pi(v|\theta_L)}\right)} \frac{1}{2}\, dv + \int_{0}^{J} \frac{\pi(v|\theta_L)}{\left(1+\frac{\pi(v|\theta_L)}{\pi(v|\theta_H)}\right)} \frac{1}{2}\, dv \tag{4.7}$$

$$\Lambda \geq .5 \ln q \left(\int_{-J}^{0} \frac{\pi(v|\theta_H)}{2} \frac{1}{2}\, dv + \int_{0}^{J} \frac{\pi(v|\theta_L)}{2} \frac{1}{2}\, dv\right) \tag{4.8}$$

By the fact that $m \leq \frac{1}{\sqrt{\gamma J}}$, The expectation of v conditional on ω_H is less than the standard deviation. Since v is has an approximately normal distribution:

$.5\int_{-J}^{0} \frac{\pi(v|\theta_H)}{2}\, dv \;\; > .5\frac{\Phi(-1)}{2} \approx .072$ where Φ is the cumulative standard normal distribution function.

Thus $\beta(1,m) = 2\Lambda > \;.316 \ln q$, Set $m = \frac{1}{\sqrt{\gamma J}}$, Then $\ln q = \;\approx \frac{2}{\sqrt{\gamma J}}$, So

$$\beta(1, \frac{1}{\sqrt{\gamma J}}) > .316 \frac{1}{\sqrt{\gamma J}} \tag{4.9}$$

Hence if $c < .316 \frac{1}{\sqrt{\gamma J}}$ when there are γJ extremist columnists, there is an equilibrium where at least $\sqrt{\gamma J}$ moderates become columnists. Since the variance of v is γJ and the mean conditional on ω_H is $M > \sqrt{\gamma J}$, The likelihood that v<0 and an unbiased reader takes action x_H conditional on ω_H is less than $\Phi(-1) \approx .158$ ♦

Proof of Proposition 5:

Recall that m_2 is defined s.t. $\beta(1,m_2) \approx c$

Let $\theta(m)$ refer to the scenario where are expected to be γJ extremists and $\frac{m\gamma J}{1-m}$ moderates and likewise let $\theta'(m)$) refer to the scenario where are expected to be $\gamma' J'$ extremists and $\frac{m\gamma' J'}{1-m}$ moderates.

$$\bar{\beta}(m_2) = 2\sum_{v=-J}^{J} \pi(v|\omega_H,\theta(m_2)) \left(H(1-\frac{2}{1+q^{-v}}) - H(1-\frac{2}{1+q^{-(v+1)}})\right) = c$$

We can rewrite this as an integral of the step function v*(z), v*(z) represents the minimum expert vote under which a citizen with interest z prefers to vote right.

$\bar{\beta}(m_2) = 2\int_{-1}^{1} h(z)\, \pi(v^*(z)|\omega_H,\theta(m_2))\, dz$, Where v*(z) is defined as the unique integer such that $1-\frac{2}{1+q^{-(v^*+1)}} \leq z < 1-\frac{2}{1+q^{-v^*}}$. Note that $v^*(z) = -v^*(-z)-1$

So the integer becomes

$$\bar{\beta}(m_2) = 2\int_{0}^{1} h(z)[\pi(v^*(z)|\omega_H,\theta(m_2)) + \pi(-v^*(z)-1|\omega_H,\theta(m_2))]dz$$

If m is small and γJ is large

$\pi(v|\omega_H,\gamma J')$ can be approximated by:

$$\pi(v|\omega_H,\gamma J') = \frac{1}{\sqrt{2\pi\gamma J}} e^{-(M-v)^2/2\gamma J}$$

Suppose $M' = M_2\sqrt{\frac{\gamma' J'}{\gamma J}}$ and define $m' = \dfrac{\frac{m_2 \gamma J}{1-m_2}\sqrt{\frac{\gamma' J'}{\gamma J}}}{\frac{m_2 \gamma J}{1-m_2}\sqrt{\frac{\gamma' J'}{\gamma J}}+\gamma' J'} = \dfrac{\frac{m_2}{1-m_2}}{\frac{m_2}{1-m_2}+\sqrt{\frac{\gamma' J'}{\gamma J}}}$ so that if there are M' moderates the chance that a given columnist is moderate is m'.

Then $\pi(v^*(z),\omega_H,\theta(m_2)) = \pi(v^*(z),\omega_H,\theta'(m'))\sqrt{\frac{\gamma' J'}{\gamma J}}$

Since $\sqrt{\frac{\gamma' J'}{\gamma J}}>1$, $\beta^*(m',\theta') \leq \beta^*(m_2,\theta)$. Therefore at m' the payoffs to becoming a columnist are at less than the equilibrium value, so if there are $\gamma J'$ extremists the informative equilibrium occurs at $m_2' < m' < m_2$

Furthermore if we define $\Phi(v|\theta)$ as the cumulative probability function over v, we find that $\Phi(v|\theta(m_2)) = \Phi(v|\theta'(m'))$. Which is to say that when there are $\gamma J'$ extremists and m=m' the public is as well informed as when there are γJ extremists and $m=m_2$. Since if there are $\gamma J'$ experts the informative equilibrium occurs at $m_2' < m'$, the public will be less well informed at the equilibrium. ♦

Proof of Proposition 6:

If a columnist receives signal σ_H, conditional on being pivotal $\pi(\omega_H) > \pi(\omega_L)$:
Recall the definitions of ψ_1 and ψ_2 from the proof of proposition 2.
The likelihood that a columnist recommending x_H is pivotal conditional on state ω_H is ψ_1 and conditional on ω_L it is ψ_2. By Bayes law if p_0 is the prior and p_H is the posterior probability a columnist places on ω_H

$$p_H = \frac{p_0\psi_1}{p_0\psi_1+(1-p)\psi_2}$$

Recall from the proof of Proposition 2

$\psi_1 = \psi_1^+ + \psi_1^- + \psi_1^0$

and $\psi_2 = \psi_{2'}^+ + \psi_2^- + \psi_2^0$

Where

$$\psi_1^+ = \sum_{v=1}^{J}\pi(v|\omega_H)\,\xi(v),\ \psi_2^+ = \sum_{v=1}^{J}\pi(v|\omega_H)\,\xi(v-1),\ \psi_1^- = \sum_{v=1}^{J}\pi(v|\omega_L)\xi(v-1).$$

$$\psi_2^- = \sum_{v=1}^{J} \pi(v|\omega_L)\xi(v), \quad \psi_1^0 = \psi_2^0 = \pi(0)\,\xi(0)$$

We can separate the first term of ψ_2^+ and obtain:

$$\psi_2^+ = \sum_{v=1}^{J} \pi(v+1|\omega_H)\,\xi(v) + \pi(1|\omega_H)\,\xi(0), \text{ likewise}$$

$$\psi_1^- = \sum_{v=1}^{J} \pi(v+1|\omega_L)\xi(v) + \pi(-1|\omega_H)\xi(0).$$

$$\text{Thus } \frac{\psi_2}{\psi_1} = \frac{\sum_{v=0}^{J} (\pi(v+1|\omega_H)+\pi(v|\omega_L))\xi(v)}{\sum_{v=0}^{J} (\pi(v|\omega_H)+\pi(v+1|\omega_L))\xi(v)}$$

$$\text{Define } \phi_2(N) = \sum_{v=0}^{N} (\pi(v+1|\omega_H)+\pi(-v|\omega_H))$$

$$\phi_1(N) = \sum_{v=0}^{N} (\pi(v|\omega_H)+\pi(-v-1|\omega_H))$$

$$\text{Claim } \frac{\phi_2(N)}{\phi_1(N)} < q \text{ for any N}$$

Proof: Suppose N is odd

$$\phi_2(N) = \sum_{v=1}^{N+1/2} (\pi(2v|\omega_H)+\pi(-2(v-1)|\omega_H)) + \sum_{v=0}^{N-1/2} (\pi(2v+1|\omega_H)+\pi(-(2v+1)|\omega_H))$$

$$\phi_1(N) = \sum_{v=1}^{N+1/2} (\pi(2(v-1)|\omega_H)+\pi(-2(v)|\omega_H)) + \sum_{v=0}^{N-1/2} (\pi(2v+1|\omega_H)+\pi(-(2v+1)|\omega_H))$$

Suppose the actual number of columnists who enter is J*

$$\text{By the Bernoulli distribution } \frac{\pi(2(v+1)|\omega_H,J^*)}{\pi(2(v)|\omega_H,J^*)} = q\frac{J^*-v}{J^*+v+2} = \alpha q$$

So $\dfrac{\pi(2v|\omega_H,J^*)+\pi(-2(v-1)|\omega_H)}{\pi(2(v-1)|\omega_H,J^*)+\pi(-2v|\omega_H)} = \dfrac{\alpha q\pi(2(v-1)|\omega_H,J^*)+ \pi(-2(v-1)|\omega_H,J^*)}{\pi(2(v-1)|\omega_H,J^*)+(\alpha/q)\pi(-2(v-1)|\omega_H,J^*)}$

Since $\pi(2(v-1)|\omega_H,J^*) > \pi(-2(v-1)|\omega_H,J^*)$, and $q> 1$,one can see that

$\alpha q\pi(2(v-1)|\omega_H,J^*)+ \pi(-2(v-1)|\omega_H,J^*)< q\pi(2(v-1)|\omega_H,J^*)+\alpha\pi(-2(v-1)|\omega_H,J^*)$

So $\dfrac{\pi(2v|\omega_H,J^*)+\pi(-2(v-1)|\omega_H)}{\pi(2(v-1)|\omega_H,J^*)+\pi(-2v|\omega_H)} < q$

Thus $\dfrac{\sum_{v=1}^{N+1/2}(\pi(2v|\omega_H)+\pi(-2(v-1)|\omega_H))}{\sum_{v=1}^{N+1/2}(\pi(2(v-1)|\omega_H)+\pi(-2(v)|\omega_H))} < q$ and

$$\frac{\sum_{v=1}^{N+1/2}(\pi(2v|\omega_H)+\pi(-2(v-1)|\omega_H)) + \sum_{v=0}^{N-1/2}(\pi(2v+1|\omega_H)+\pi(-(2v+1)|\omega_H))}{\sum_{v=1}^{N+1/2}(\pi(2(v-1)|\omega_H)+\pi(-2(v)|\omega_H)) + \sum_{v=0}^{N-1/2}(\pi(2v+1|\omega_H)+\pi(-(2v+1)|\omega_H))} < q\ ,$$

so $\dfrac{\phi_2}{\phi_1} < q$ if N is odd. Now suppose N is even

Then $\phi_2(N) = \sum_{v=1}^{N/2}(\pi(2v+1|\omega_H)+\pi(-2v+1|\omega_H)) +\pi(1|\omega_H)+ \sum_{v=0}^{N}\pi(2v-N|\omega_H)$

and $\phi_1(N) = \sum_{v=1}^{N/2}(\pi(2v-1|\omega_H)+\pi(-2v-1|\omega_H)) +\pi(-1|\omega_H)+ \sum_{v=0}^{N}\pi(2v-N|\omega_H)$

Recall from above that $\dfrac{\pi(2v|\omega_H,J^*)+\pi(-2(v-1)|\omega_H)}{\pi(2(v-1)|\omega_H,J^*)+\pi(-2v|\omega_H)} < q$

So $\dfrac{\sum_{v=1}^{N/2}(\pi(2v+1|\omega_H)+\pi(-2v+1|\omega_H))}{\sum_{v=1}^{N/2}(\pi(2v-1|\omega_H)+\pi(-2v-1|\omega_H))} < q$, since $\dfrac{\pi(1|\omega_H)}{\pi(-1|\omega_H)} = q$

$$\frac{\sum_{v=1}^{N/2}(\pi(2v+1|\omega_H)+\pi(-2v+1|\omega_H)) +\pi(1|\omega_H)+ \sum_{v=0}^{N}\pi(2v-N|\omega_H)}{\sum_{v=1}^{N/2}(\pi(2v-1|\omega_H)+\pi(-2v-1|\omega_H)) +\pi(-1|\omega_H)+ \sum_{v=0}^{N}\pi(2v-N|\omega_H)} < q$$

Thus $\frac{\phi_2(N)}{\phi_1(N)} < q$ and the claim is proved.

Define $\lambda(v) = \xi(v) - \xi(v+1)$, Note that $\psi_2 = \phi_2(J)\xi(J)+ \sum_0^{J-1}\phi_2(v)\lambda(v)$ and $\psi_1 = \phi_1(J)\xi(J)+ \sum_0^{J-1}\phi_1(v)\lambda(v)$. Since $\frac{\phi_2(v)}{\phi_1(v)} < q$ for any v,

$$\frac{\psi_2}{\psi_1} = \frac{\phi_2(J)\xi(J)+ \sum_0^{J-1}\phi_2(v)\lambda(v)}{\phi_1(J)\xi(J)+ \sum_0^{J-1}\phi_1(v)\lambda(v)} < q$$

and $\frac{\psi_2}{\psi_1} < q$. Since $\frac{\psi_2}{\psi_1} < q$, $p_H > \frac{p_0}{p_0+(1-p_0)\frac{1+m(2a-1)}{1-m(2a-1)}}$. If the agent receives signal σ_H,

$p_0 = a$ so $p_H > \frac{a}{a+(1-a)\frac{1+m(2a-1)}{1-m(2a-1)}}$. Since $a > .5$ and $m < 1$,

$a - ma(2a-1) > 1+m(2a-1) - a - ma(2a-1)$. So $(1-a)\frac{1+m(2a-1)}{1-m(2a-1)} < a$, and $p_H > .5$ ♦

The following is demonstration that if J is large we would expect that $\frac{\psi_2}{\psi_1} < q$

$$\frac{\sum_{v=1}^{J/2}(\pi(2v+1|\omega_H)+\pi(2v|\omega_L))\xi(2v)+\sum_{v=1}^{J/2}(\pi(2v|\omega_H)+\pi(2v-1|\omega_L))\xi(2v-1)+\pi(0)\,\xi(0)+\pi(1|\omega_H)\,\xi(0)}{\sum_{v=1}^{J/2}(\pi(2v|\omega_H)+\pi(2v+1|\omega_L))\xi(2v)+\sum_{v=1}^{J/2}(\pi(2v-1|\omega_H)+\pi(2v|\omega_L))\xi(2v-1)+\pi(0)\,\xi(0)+\pi(1|\omega_L)\,\xi(0)}$$

Define $q = \frac{1+m(2a-1)}{1-m(2a-1)}$, As before $\frac{1}{1+q^{-1}}$ represents the likelihood that an individual columnist's vote reflects the state of the world.

Note by the Bernoulli equation that $\frac{\pi(v+2|\omega_H)}{\pi(v|\omega_H)} = \alpha^2 q$, $\frac{\pi(v+2|\omega_L)}{\pi(v|\omega_L)} = \frac{\alpha^2}{q}$ where if $v \geq 0$, $\alpha<1$.

We smooth this so $\frac{\pi(v+1|\omega_H)}{\pi(v|\omega_H)} = \alpha\sqrt{q}$ and $\frac{\pi(v+1|\omega_L)}{\pi(v|\omega_L)} = \frac{\alpha}{\sqrt{q}}$ $\frac{\pi(v|\omega_H)\alpha\sqrt{q}+\pi(v|\omega_L)}{\pi(v|\omega_H)+\pi(v|\omega_L)\frac{\alpha}{\sqrt{q}}} < \sqrt{q}$. Since $\alpha<1$ and $\pi(v|\omega_H) > \pi(v|\omega_L)$,

$\pi(v|\omega_H)\alpha\sqrt{q}+\pi(v|\omega_L) < \pi(v|\omega_H)\sqrt{q}+\alpha\pi(v|\omega_L)$. Thus

$$\frac{\sum_{v=1}^{J}(\pi(v+1|\omega_H)+\pi(v|\omega_L))\xi(v)}{\sum_{v=1}^{J}(\pi(v|\omega_H)+\pi(v+1|\omega_L))\xi(v)} < \sqrt{q}$$

Furthermore $\frac{\pi(1|\omega_H)}{\pi(1|\omega_L)} = q$. So $\frac{\pi(0)\,\xi(0)+\pi(1|\omega_H)\,\xi(0)}{\pi(0)\,\xi(0)+\pi(1|\omega_L)\,\xi(0)} < q$ and

$\frac{\psi_2}{\psi_1} < q$. In fact we can approximate $\frac{\pi(0)\ \xi(0) + \pi(1|\omega_H)\ \xi(0)}{\pi(0)\ \xi(0) + \pi(1|\omega_L)\ \xi(0)} = \frac{1+\sqrt{q}}{1+\frac{1}{\sqrt{q}}}$

which is equal to $\sqrt{q}$ if q is close to 1, so if m is small and q is close to 1

$\frac{\psi_2}{\psi_1} < \sqrt{q}$

In any case $p_H = \frac{p_0\psi_1}{p_0\psi_1+(1-p)\psi_2}$ and $q = \frac{1+m(2a-1)}{1-m(2a-1)}$. Since $\frac{\psi_2}{\psi_1} < q$

$p_H > \frac{p_0}{p_0+(1-p_0)\frac{1+m(2a-1)}{1-m(2a-1)}}$. If the agent receives signal σ_H, $p_0 = a$ so

$p_H > \frac{a}{a+(1-a)\frac{1+m(2a-1)}{1-m(2a-1)}}$. since a>.5 and m< 1,

a -ma(2a-1) > 1+m(2a-1) -a-ma(2a-1). So $(1-a)\frac{1+m(2a-1)}{1-m(2a-1)} < a$, and $p_H > .5$ ♦