When consensus hurts: experts' advice and electoral support *

Antoni-Italo de Moragas[†]

This version: January 2026

Abstract

When voters rely on expert advice to form opinions about policy, greater expert consensus should—in standard models—be more persuasive. I show this logic can fail. If voters cannot observe whether individual experts have vested interests, and these interests are correlated (so that when one expert is biased, others likely are), then unanimous expert endorsement can be *less* persuasive than a simple majority. The intuition is that consensus may arise either from shared information (informative) or from shared bias (uninformative), while disagreement among experts reveals they are not coordinating on private interests. I characterize when this non-monotonicity arises and show it persists when experts behave strategically. The results identify conditions under which voter skepticism toward expert consensus reflects rational inference rather than irrationality or bias.

Keywords: Voting; Experts; Consensus.

JEL Codes: D72, D83.

*I am grateful to my supervisors Andrea Mattozzi and David K. Levine for their help and support. I'm particularly indebted to Santiago Sánchez-Pages for his enthusiasm at the first stages of the project. I would also like to thank Steven Callander, Agustín Casas, Hülya Eraslan, Helios Herrera, Matías Iaryczower, Aniol Llorente-Saguer, Vincent Maurin, Ken Shotts, Oriol Tejada, Bauke Visser as well as seminar participants at SAET, Spring Meeting of Young Economists, ASSET, IOEA, NICEP and SAEe for many valuable suggestions.I acknowledge the support of PGC2018-099415-B-100 MICINN/FEDER/UE.

[†]Department of Economics, CUNEF Universidad. Email: antoni.demoragas@cunef.edu

1 Introduction

Citizens often lack the expertise to evaluate complex policy issues, from trade agreements to pandemic responses. A natural solution—both theoretically and in practice—is to rely on expert advice. Standard models of Bayesian learning suggest that as more experts endorse a policy, voters should become more confident it is correct (a logic dating to Condorcet's Jury Theorem). Yet recent events challenge this prediction: overwhelming expert consensus on Brexit's economic costs, the dangers of Trump's trade policies, and the reality of climate change has often failed to sway public opinion—and may have even provoked backlash. More recently, widespread vaccine hesitancy persisted despite near-unanimous expert endorsement of COVID-19 vaccines. After the Brexit vote, Paul Johnson, director of the Institute for Fiscal Studies, remarked, "It is clear that economists' warnings were not understood or believed by many. So we economists need to be asking ourselves why that was the case, why our near-unanimity did not cut through."

This paper provides a rational explanation for this puzzle. I develop a model in which experts observe informative signals about the correct policy but may have private interests favoring one alternative. The key innovation is that expert biases are *correlated*: in some decisions, experts as a class have shared interests that may not align with voters. When voters cannot observe whether a given decision triggers these shared interests, they face an inference problem. Expert consensus can arise from two sources: (a) all experts received the same informative signal, or (b) the decision is one where experts share biased interests. Unanimous endorsement increases the likelihood of both scenarios, but only (a) is informative about policy quality.

The main result is that the relationship between the number of experts endorsing an alternative and voter support can be non-monotonic. In particular, an alternative endorsed by a narrow majority of experts may receive *more* electoral support than one backed unanimously. Intuitively, when one expert dissents, voters can rule out the "shared bias" scenario—the dissenter proves that experts are responding to information, not coordinating on their private interests.

How does this result depend on voter information? I extend the model to allow voters to acquire independent, albeit imperfect, information in addition to expert advice, and I show that the non-monotonic relationship still holds. Furthermore, when voters are moderately confident in their own information, an alternative can only win a majority if most, but not all, experts endorse it. Interestingly, when voters are overly confident in their information, this non-monotonicity can persist even when it would be optimal for citizens to always follow the majority of expert advice.

The paper proceeds as follows. Section 2 reviews related literature. Section 3 presents the model. Section 4 derives the main results on non-monotonicity of voter support in expert consensus. Section 5 extends the model to informed citizens and analyzes overconfidence. Section 6 considers several extensions: larger expert panels with partial correlation, symmetric bias, and strategic experts. Section 7 concludes.

2 Related Literature

The puzzle that citizens may be more persuaded by experts who express some disagreement has been empirically documented by several studies. Sapienza and Zingales (2013) compare responses to policy questions between experts from the University of Chicago Booth Economic Expert Panel and ordinary U.S. citizens. They find an average difference of 35 percentage points in agreement between the two groups, and this gap is *larger* when there is strong expert consensus. Moreover, informing citizens about expert consensus makes little difference: the belief that stock prices are hard to predict decreases only from 55% to 42% when citizens learn that all experts agree. Johnston and Ballard (2016) find moderate shifts in public opinion when citizens are informed of economists' consensus, but interestingly, responsiveness is greater when the consensus is attributed to a generic sample rather than to economists specifically. Darmofal (2005) shows that citizens are more likely to disagree with experts when political elites they support challenge the expert opinion. The present paper provides a theoretical explanation for why such behavior may be rational.

This paper contributes to the literature on information aggregation in elections, which dates to de Condorcet (1785). The Condorcet Jury Theorem establishes that majority voting aggregates information efficiently when voters have common interests and receive independent signals. Feddersen and Pesendorfer (1997) extend this framework to strategic voting, showing that voters condition on being pivotal—the "swing voter's curse." Feddersen and Pesendorfer (1998) demonstrate that unanimous jury verdicts can be inferior to majority rule because unanimity requirements distort strategic behavior. My paper shares the theme that unanimity can be problematic, but through a different channel: voters rationally discount unanimous expert advice because it may reflect correlated bias rather than correlated information. Martinelli (2006) studies costly information acquisition in elections, showing that rational voters may remain uninformed. In my model, the issue is not whether voters acquire information, but how they interpret expert consensus.

The paper also relates to the cheap talk literature on information transmission between informed experts and uninformed decision makers. The seminal work by Crawford and Sobel (1982) has been extended to multiple experts by Gilligan and Krehbiel (1989), Austen-Smith (1993), Battaglini (2002), Wolinsky (2002), Krishna and Morgan (2004), and Gerardi et al. (2009). These studies focus on how competition among experts can improve information transmission. My paper differs in three respects: (i) experts are non-strategic in the baseline model, (ii) there is an electoral aggregation stage, and (iii) the key friction is unobserved correlation in expert biases, not preference misalignment per se. Recent work on strategic communication includes Kamenica and Gentzkow (2011) on Bayesian persuasion and Kartik et al. (2021) on strategic advisors, both of which study how senders optimally design information disclosure.

The mechanism here also differs from both classic cheap talk and the credibility literature. In cheap talk models (Crawford and Sobel, 1982; Krishna and Morgan, 2004), experts strategically garble information, and the receiver must filter. In reputation models (Sobel, 1985; Morris, 2001), experts may distort advice to signal their type or protect their reputation; Ottaviani and Sørensen (2006) show that reputational concerns can lead experts to herd on conventional forecasts. In contrast, I assume experts naively report their preferred action—the strategic friction arises entirely from voters' inference problem when expert biases are unobserved and potentially correlated. This distinction is important: in reputation models, the distortion comes from experts' strategic behavior, while here it comes from voters' rational skepticism about correlated interests.

The model's structure—where expert consensus may arise from shared bias—is related to the literature on media bias and herding. Gentzkow and Shapiro (2006) and Mullainathan and Shleifer (2005) study how media outlets may slant news to match audience priors or advertiser interests, creating correlated bias across sources. The classic models of herding and information cascades by Banerjee (1992) and Bikhchandani et al. (1992) show how rational agents may ignore private information and follow predecessors, leading to fragile consensus. In my model, expert consensus can similarly arise from coordination on private interests rather than aggregation of information, though the mechanism is bias correlation rather than sequential observation.

Finally, this paper contributes to the literature on disagreement and belief formation. Several theories explain persistent disagreement: confirmation bias (Rabin and Schrag, 1999), overconfidence (Ortoleva and Snowberg, 2015), and correlation neglect (Levy and Razin, 2015). The closest paper is Cheng and Hsiaw (2021), which shares two features with this paper: experts are non-strategic, and citizens face uncertainty about both the correct policy and expert credibility. However, the mechanisms differ substantially. In their model, voters are uncertain about expert *quality* (whether experts are informed); in mine, voters are uncertain about expert *bias correlation* (whether the decision triggers shared interests). This distinction generates different predictions: in their framework, more experts always increase persuasion, while in mine, the relationship can be non-monotonic.

3 The Model

The model has three components: a continuum of voters who must choose between a reform and the status quo, three experts who observe signals about which policy is correct, and two types of decisions (high-conflict and low-conflict) that determine whether experts have shared biases. Voters observe expert recommendations but cannot directly observe expert signals or biases. The key friction is that when a decision is high-conflict, experts are likely to share a common interest, creating correlation in their advice that is uninformative about policy quality.

An electorate composed of a continuum of voters has to choose between passing a reform $R \in (0,1)$ or keeping status quo $S = -R$. We will assume a simple majority rule: the reform can only pass if it obtains the majority of votes [1]. Each voter i has a preference parameter towards the status quo v_i distributed[2] according to a uniform distribution on $[-1,1]$. In addition to this parameter, the utility of a voter i also depends on the realization of a state of the world $\omega \in \{-1,1\}$ and both states are equally likely.[3] More precisely, the utility of a voter i when a decision $d \in \{S,R\}$ is taken in a state of the world ω is given by:

$$u_i(\omega, d) = -(\omega + v_i - d)^2 \tag{1}$$

Voters do not observe the state of the world but they receive the advices of n informed experts. Each expert j receives an individual private signal $s_j \in \{-1,1\}$ of the state of the world. More precisely $Pr(s_j = \omega|\omega) = q$. In addition to this signal, each expert j can have some vested interest on status quo. When an expert j has an interest over the status quo[4] we will say that the expert is biased and $\beta_j = -1$; otherwise, we will say that the expert is neutral and $\beta_j = 0$. Some decisions are more likely to affect the interests of experts than

[1] When both alternatives obtain the same number of votes we will assume that status quo is implemented.
[2] As it will become clear later, this heterogeneity is only needed to guarantee that the share of votes for the reform take values in $(0,1)$.
[3] Fixing one state more likely would reduce tractability but would not affect qualitatively the outcome.
[4] The case where experts can be biased towards both alternatives is discussed in one of the extensions.

others. With probability $p_H \in [0,1]$ the decision is high-conflict ($b = H$)—meaning experts' shared interests are at stake—and the probability that each expert is biased is $\sigma_H \in [0,1]$. With probability $1 - p_H$ the decision is low-conflict ($b = L$) and the probability that each expert is biased is $\sigma_L \in [0, \sigma_H]$. Notice that when $\sigma_L = \sigma_H$, biases are fully uncorrelated and when $\sigma_L = 0$ and $\sigma_H = 1$ biases are fully correlated. Voters do not observe the type of the decision.

Given that the correlation among experts' bias is a key feature of our modelling assumptions, it is important to clarify its interpretation. The rationale for assuming this correlation is that experts are not representative of the whole population (they work in particular economic sectors such as universities, think tanks or institutions) and they can have particular interests that differ with those of the rest of the society. The importance of these interests are captured by the type of the decision: when the decision is high-conflict, the decision to be taken has large impact on these interests and it is more likely that experts will advise according to them, whereas when the decision is low-conflict, we may think that the relative importance of these interests are lower and experts will be more likely to advise according to the interests of society.[5]

The utility of an expert j with private interest β_j is:

$$u_j(\omega, d) = -(\omega + \beta_j - d)^2 \tag{2}$$

Notice that the optimal decision of a biased expert is S regardless of the state of the world and the optimal decision of a neutral expert is ω. We will assume that each expert j provides an advice $a_j \in \{S, R\}$ and, for simplicity, I assume this advice coincides with the expert's optimal decision.[6] That is, I assume experts are *naive* in the sense that they report their bliss point rather than strategically manipulating voters. Thus, the advice a_j of an

[5]Instead of assuming that biases are correlated we could assume that the precision of the signals they receive are correlated. For example, experts could be informed or uninformed and, when decisions are easy, the probability of being informed is higher than when they are difficult. Informed experts advise according to the information they have and uninformed ones simply advise according to their bias.

[6]I discuss what happens when experts are strategic in the extensions.

expert j is:

$$a_j(\beta_j, s_j) = \begin{cases} S & \text{if } \beta_j = -1 \\ s_j & \text{if } \beta_j = 0 \end{cases} \tag{3}$$

We will assume that $n = 3$, which is the simplest environment that allows us to distinguish between majority and consensus in expert's advice. Since I assume a continuum of sincere voters, each voter chooses the alternative that he expects to provide higher utility to himself. The timing of the model is the following:

1. State of the world ω and the type of the decision b are realized.
2. Experts receive signals s_j of the state of the world and observe their bias β_j.
3. Experts give advices a_j to voters.
4. Elections are held and each voter elects the alternative that maximizes his utility.

Importantly, voters observe the advice vector (a_1, a_2, a_3) but not the signals (s_1, s_2, s_3) or the biases $(\beta_1, \beta_2, \beta_3)$.

4 Results

First of all we will prove that, despite the heterogeneity of voters' preferences, in each state of the world the optimal decision of all voters coincides:

Lemma 1 *Let $d_i^*(\omega)$ be the optimal decision of a voter i. For all i,*

$$d_i^*(\omega) = \begin{cases} S & \textit{if } \omega = -1 \\ R & \textit{if } \omega = 1 \end{cases} \tag{4}$$

The intuition is that the salience of the common component of citizens' preferences overweights the private one. Given that all citizens share the same optimal decision in each state of the world, we will say that a decision d is correct when $d = d^*$ and wrong otherwise. Notice that despite all voters prefer a correct decision with respect to a wrong one, given that voters do not know the realization of the state ω and they have different individual parameters, they will vote different alternatives. In particular, a voter with preference parameter $v_k = -1$ will always vote for status quo whereas a voter with preference parameter $v_l = 1$ will always vote for the reform.

The only relevant information a voter i has before casting his vote is the number of experts advising each alternative. Let Π_k be the perceived probability that the state of the world is status quo conditional on k experts supporting it. Given that voters are sincere, a voter i votes for status quo if and only if $E[U_i(s)|(k, n-k)] \geq E[u_i(r)|(k, n-k)]$ and this happens if and only if:

$$v_i \leq 2\Pi_k - 1 \tag{5}$$

.

Lemma 2 *The share of votes for status quo is* $2\Pi_k - 1$*, which is increasing in* Π_k*.*

Not surprisingly, given that voters value taking correct decisions, an increase in the probability that the decision is correct increases the electoral support for that decision. In the next sections we will concentrate our attention on the relationship between the number of experts advising an alternative and the probability that this alternative is correct. In particular we will study what happens when $\sigma_L = \sigma_H$ (uncorrelated bias) and when $\sigma_L = 0$ and $\sigma_H = 1$ (fully correlated bias).

4.1 Uncorrelated Bias

When $\sigma_H = \sigma_L = \sigma$, the probability that an expert j advises status quo when $\omega = S$ is $\sigma + (1-\sigma)q$ and in state of the world R is $\sigma + (1-\sigma)(1-q)$. Therefore,

$$\Pi_k = \frac{\pi(S, k)}{\pi(S, k) + \pi(R, k)} \tag{6}$$

where

$$\pi(S,k) = \binom{3}{k}(\sigma + (1-\sigma)q)^k(1-(\sigma+(1-\sigma)q))^{3-k} \tag{7}$$

$$\pi(R,k) = \binom{3}{k}(\sigma + (1-\sigma)(1-q))^k(1-(\sigma+(1-\sigma)(1-q)))^{3-k} \tag{8}$$

Lemma 3 *When experts' biases are uncorrelated, the share of votes for status quo is increasing in the number of experts advising it.*

Notice that this result applies both when experts are always neutral ($\sigma = 0$) or when they are biased with some probability ($\sigma > 0$). In both scenarios, the likelihood of a policy being correct increases on the number of experts advising for it. Increasing the probability of being biased ($\sigma > 0$) makes the advice of experts less informative but does not change the monotonicity.

4.2 Correlated Bias

This section introduces correlation among expert biases. For simplicity, we will study the fully correlation case such that $\sigma_L = 0$ and $\sigma_H = 1$. That is, all experts are neutral when $b = L$ and all experts are biased when $b = H$. Now, suppose that $k < 3$, that is, at least one expert advises the reform. This expert has to be necessarily neutral but if this expert is neutral, this means that all experts are neutral and voters know for certain that $b = L$ and the expression of Π_k is the one we derived when we studied the uncorrelated bias for $\sigma = 0$. In particular, when one expert endorses the reform and two experts endorse status quo, given that the prior probability is symmetric, the first two advices cancel out and the posterior probability that status quo is the correct decision is simply the precision q.

Thus, we can restrict our analysis to the case $k = 3$. There are different situations where all agents advise status quo, either all agents received a status quo signal independently of the type of the decision or the decision is high-conflict and all experts are biased. When the decision is low-conflict, the probability that the state of the world is S is just the ex-ante probability which we assumed it was $\frac{1}{2}$. Therefore the probability that status quo is correct when all experts advise it is:

$$\Pi_3 = \frac{(1-p_H)\pi(S,3) + \frac{p_H}{2}}{(1-p_H)(\pi(S,3)+\pi(R,3)) + p_H} \tag{9}$$

Clearly, when p_H gets close to 1, Π_3 gets close to $\frac{1}{2}$ and it can occur that $\Pi_3 < q = \Pi_2$. The next result formalizes this intuition:

Proposition 1 *When experts biases are correlated, there exists a $\hat{p} \in (0,1)$ such that:*

(i) When $p_H \leq \hat{p}$, the share of votes for status quo is increasing in the number k of experts advising it.

(ii) When $p_H > \hat{p}$, the share of votes for status quo is non-monotonic in the number k of experts advising it. In particular, it is increasing when $k < 3$ but the share of votes for status quo is higher when $k = 2$ than when $k = 3$.

Interestingly, when $p_H \geq \hat{p}$, there is a non-monotonic relationship between the number of experts supporting status quo and the probability that status quo is the correct decision. In particular, unanimity makes status quo less likely to be correct than a majority with one dissent expert. The intuition is that, when some decisions are high-conflict and others are low-conflict, that is when interests are correlated, consensus among experts is informative of the likelihood of the state of the world but also of the likelihood of the decision being high-conflict. When the probability that the decision is high-conflict is higher than $\hat{p}$, the second phenomenon dominates the first.

Corollary 1 *The threshold $\hat{p}$ is decreasing in q and* $\lim_{q\to 1} \hat{p} = 0$

The previous result means that, ceteris paribus, an increase in the quality of the signals of experts makes the non-monotonic result more likely because when the precision of the

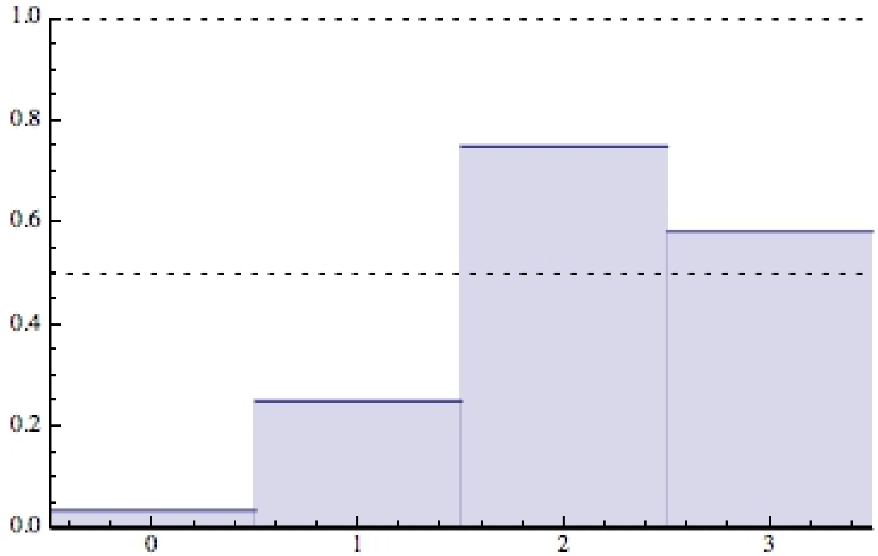

Figure 1: Probability that status quo is correct (Π_k) as a function of the number of experts endorsing it. Parameters: $n = 3$, $q = 0.85$, $p_H = 0.5$, fully correlated biases. The non-monotonicity is evident: $\Pi_2 > \Pi_3$, so a majority with one dissent is more informative than unanimity.

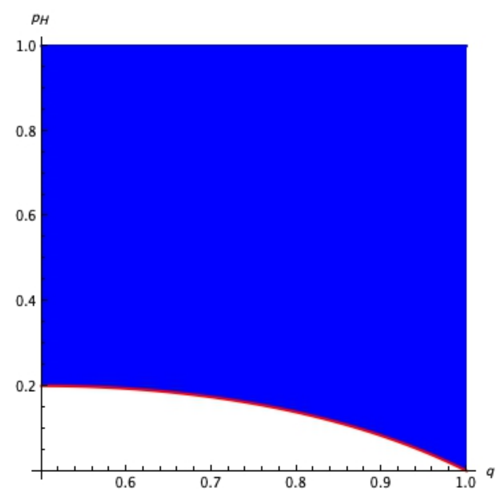

Figure 2: Parameter region where non-monotonicity arises ($p_H \geq \hat{p}$). The shaded area shows combinations of expert signal precision q and high-conflict probability p_H for which unanimous endorsement is less persuasive than a majority with one dissent. As $q \to 1$, the threshold $\hat{p} \to 0$.

signal increases, the posterior of status quo being correct increases both for consensus and for simple majority but that increase is always larger for the latter.

Moreover, when experts tend to be perfectly informed, the non monotonic result holds for any positive probability of a high-conflict decision.

5 Informed citizens and overconfidence

So far, experts were the only source of voter information and even when consensus was less informative than a simple majority, a majority of voters would still follow the advice of experts. Therefore, the non-monotonicity result had no effect for elections with majority vote. In this section we will assume that voters have access to other information to illustrate why some reforms that wouldn't pass when a simple majority of experts advise against them, can pass when all experts advise against them. Assume that, in addition to observing experts' advices, all citizens get a common signal $\theta_c \in \{-1, 1\}$ of the state of the world with precision $q_c = Pr(\omega = \theta_c|\theta_c) > \frac{1}{2}$. We allow citizens to be overconfident about the precision of their signal, that is, instead of thinking that their precision is q_c, they think that their precision is $q'_c \geq q_c$. Nevertheless, they are aware that $q'_c < q$ (which implies $q_c < q$), that is, not only their signal is less informative than the signal of an expert but they are aware of it. Let $\pi_k(\theta_c)$ denote the probability that status quo is the correct decision when citizens receive a signal θ_c and k experts advise it and let $\pi'_k(\theta_c)$ be the subjective probability of citizens.

We will concentrate on the case where biases are correlated since this is the one that can hold the non-monotonicity result. In particular, we will derive the conditions such that citizens only follow experts advice when there is a simple majority and follow their signal otherwise despite the fact that following expert's consensus would increase their welfare.

Lemma 4 *When most experts advise the reform, the reform is implemented independently of the prior of voters.*

When most experts advise the reform, given that citizens know that experts are better informed than them and they can't be biased towards the reform, citizens think that the reform is more likely to be correct and a majority of citizens vote for the reform. Thus, the interesting case is what happens when most experts advise status quo.

Proposition 2 *When most experts advise status quo,*

1. *if* $\theta_c = -1$*, status quo is implemented.*
2. *if* $\theta_c = 1$ *and and one expert advises the reform, status quo is implemented.*
3. *if* $\theta_c = 1$ *and all experts advise status quo, there exists a* $\underline{p} \in (0, 1)$ *such that status quo is implemented if and only if* $p_H \leq \underline{p}$*.*

We have just proven that the reform can be implemented when all experts advise status quo but, given the correlation of experts' biases, it could be the case that voters maximize the probability of taking correct decisions and, therefore, we shouldn't worry when voters ignore experts' consensus. However, in the following corollary we show that this is not always the case.

Corollary 2 *When* $q_c' > q_c$*, there exists a* $\bar{p} \in (\underline{p}, 1)$ *such that if* $p_H \in (\underline{p}, \bar{p})$ *the reform is implemented when all experts advise status quo and status quo is more likely to be correct than the reform.*

This result shows that when voters are poorly informed but they they are very overconfident about their information, voters ignore expert's consensus when they would be better off by following their advice. The intuition is that overconfident voters do not aggregate correctly the advice of experts and overestimate the probability that all experts are biased when experts' unanimous advice does not coincide with voters' prior. When there is no consensus, even overconfident voters follow experts' majoritarian advice because disagreement among experts proves that they are not biased and, despite being overconfident, voters think that experts are still better informed than them.

6 Extensions

6.1 Larger Expert Panels and Partial Correlation

We have already shown that the relationship between the number of experts advising status quo and the probability that status quo is correct can be non-monotonic. In particular we have shown that it can decrease when we move from two experts advising status quo to all experts advising status quo. A reasonable concern is that when the number of experts n is arbitrarily large, the non-monotonic result only holds for the very extreme case of unanimity. In this section we will assume that the bias is not fully correlated, that is $\sigma_L < \sigma_H$ and we will show that non-monotonicity is not only a property of unanimity but, more generally, it can also be a property of less demanding majorities.

When biases are not fully correlated, even in high-conflict decisions it can happen that some expert advises Reform. Given that having some expert advising Reform is not fully informative about the type of the decision anymore, all Π_k change with respect to the neutral experts benchmark (recall that when bias was fully correlated, the only probability that changed was Π_n). In order to derive Π_k, we need to compute $Pr(\omega, b, k)$, that is the joint probability that the state of the world is ω, the type of the decision is b and k experts advise status quo. Conditional on state of the world ω and type of the decision b, the probability that an expert j advises status quo is:

$$\alpha(\omega, b) = \begin{cases} \sigma_b + (1 - \sigma_b)q & \text{if } \omega = S \\ \sigma_b + (1 - \sigma_b)(1 - q) & \text{if } \omega = R \end{cases} \tag{10}$$

Thus the joint probability that k experts advise status quo and state of the world is ω and type of the decision is b

$$Pr(\omega, b.k) = \frac{p_b}{2} Pr(k|\omega, b) = \frac{p_b}{2} \binom{n}{k} (\alpha(\omega, b))^k (1 - \alpha(\omega, b))^{n-k} \tag{11}$$

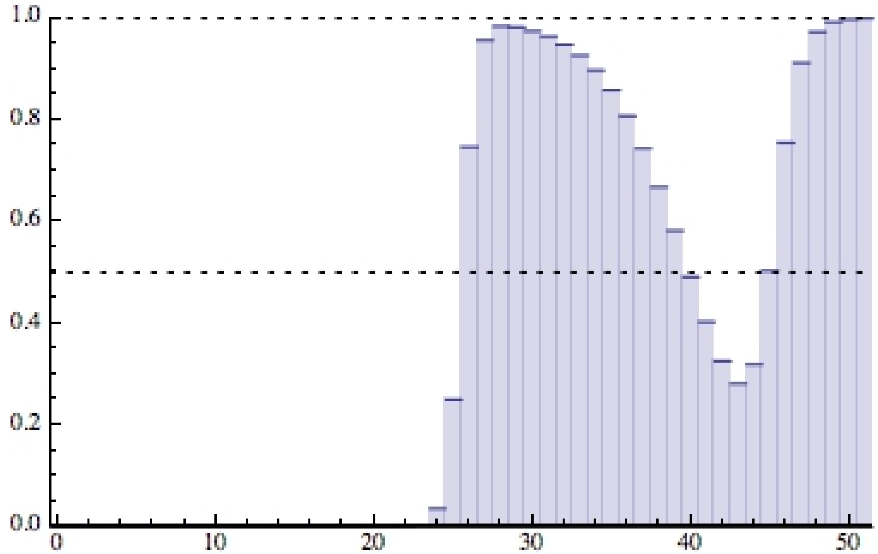

Figure 3: Probability that status quo is correct (Π_k) with a larger expert panel and partial correlation. Parameters: $n = 51$, $q = 0.75$, $p_H = 0.5$, $\sigma_H = 0.75$, $\sigma_L = 0$. Non-monotonicity extends beyond unanimity to a range of supermajorities.

Once we have computed $Pr(\omega, b.k)$, we apply Bayesian updating and we obtain the probability that the state of the world is status quo conditional on the number of experts advising it:

$$\Pi_k = \frac{p_H Pr(S,H,k) + (1-p_H)Pr(S,L,k)}{p_H Pr(S,H,k) + (1-p_H)Pr(S,L,k) + p_H Pr(R,H,k) + (1-p_H)Pr(R,L,k)} \quad (12)$$

In figure 2 we can observe what can happen when biases are partially correlated. In particular we see that the non-monotonicity is not only a problem of the extreme case $k = n$ but it is a property that can appear for all $k > \frac{n+1}{2}$. The intuition is the following. If k is close to unanimity, it means that there is a huge fraction of biased experts ($b = H$) and neutral agents receive status quo signals ($\omega = S$). If k is slightly above the simple majority threshold, experts are neutral ($b = L$) and they receive status quo signals $\omega = S$. If k is under the majority, experts are neutral $b = L$ and they receive Reform signals. Finally, when k is above small majority and under unanimity, there is a huge fraction of biased experts

$b = H$ and neutral experts receive Reform signals ($\omega = R$).

When q is larger, that is, when neutral decision makers receive more precise signals, there is a compression of the distribution towards the right that can potentially erase the sink between simple majority and consensus.

6.2 Symmetric bias

In all the article we have assumed that experts could only be biased towards one of the options that we have labeled as status quo. The underlying idea is that experts, as an intellectual elite, may have some common interests that voters can ex-ante identify. In this section we relax this assumption allow experts to be biased to both directions.

That is we will extend the model such that the bias β_j of expert j can take values $\{-1, 0, 1\}$ and the type of the decision can also take values $S, 0, R$. Conditional on the type of the decision S (R), the probability that an expert has bias $\beta_j = S$ ($\beta_j = R$) is σ_S (σ_R) and the probability that he has bias 0 is $1 - \sigma_S$ $(1 - \sigma_S)$. When $b = 0$ all experts are neutral. When the bias of experts can go on both directions, the non-monotonicity can also apply to the situations where $k < \frac{n}{2}$. The reason is that, when the bias was only towards status quo, consensus advice for Reform could only happen when all experts received a signal for Reform. This is not the case anymore if experts can also be biased towards Reform. Figure 4 shows this non-monotonicity when biases are fully correlated.

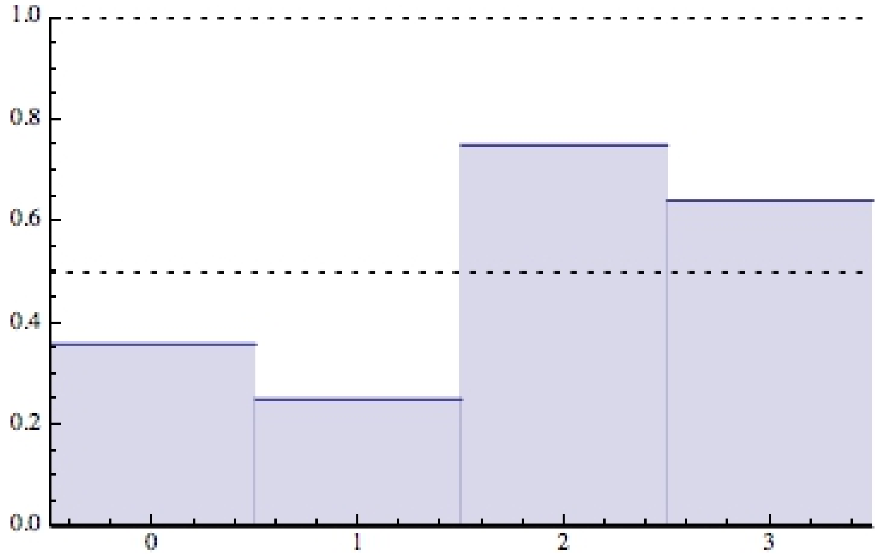

Figure 4: Probability that status quo is correct (Π_k) with symmetric bias (experts can be biased toward either alternative). Parameters: $n = 3$, $q = 0.75$, $p_H = 0.5$, fully correlated biases. Non-monotonicity now appears for both high and low values of k.

6.3 Strategic experts

One of the critical assumptions of the model is that experts are not strategic and they simply advise the reform they prefer without anticipating the electoral effect of their advice. We have seen that ,under this assumption, there can exist an equilibrium such that the relationship between the number of experts endorsing a policy and the votes for that policy is non-monotonic. Does this equilibrium survive when experts are strategic?

Let's suppose that experts' biases are correlated and $p_H > \hat{p}$, that is, when experts are not strategic, the share of votes for status quo is non-monotonic in the number of experts advising it. This can't be an equilibrium outcome because strategic biased experts have incentives to deviate. Given that biases are fully correlated, if an expert is biased, he knows that all other experts are biased too. The biased expert knows that the other two experts are already endorsing status quo, and he has incentives to endorse the reform because the share of votes for status quo increases when the number of experts endorsing it goes from 3 to 2. Does the non-monotonic result vanish when we consider strategic experts? Not necessarily.

We will restrict our analysis to symmetric equilibria (i.e. equilibria such that all experts play the same strategy given their signal and their vested interest) and we will only allow biased experts to be strategic. If there exists a symmetric equilibrium where biased experts use mixed strategies, it has to be that they are indifferent between advising status quo or advising the reform and this indifference condition requires a non-monotonic relationship between the number of experts endorsing status quo and support for status quo. Suppose that the relationship was monotonic and increasing, then an extra endorsement can only increase the support for status quo and this violates the indifference condition.

Regarding the existence, notice that if biased experts advise status quo with probability q, that is, they mimic the behaviour of neutral experts when the state of the world is -1, they have incentives to advise always the status quo. On the contrary, as we have discussed before, if biased experts always endorse status quo, they have incentives to endorse the reform. By continuity there exists a $\hat{q} \in (q, 1)$ such that, in equilibrium, biased experts endorse the status quo with probability $\hat{q}$. The following proposition formalizes the previous reasoning:

Proposition 3 *When experts biases are correlated and $\hat{p} < p_H$, there exists an equilibrium in which biased experts endorse status quo with probability $x > q$ and the share of votes for status quo is non-monotonic on the number of experts endorsing it.*

7 Conclusion

This paper shows that voters may rationally discount unanimous expert advice. When expert biases are correlated—so that a decision affecting one expert's interests likely affects others'—unanimous endorsement can be *less* persuasive than a narrow majority. The key mechanism is that consensus can arise from two sources: shared information (informative) or shared bias (uninformative). A dissenting expert rules out the latter scenario, signaling that experts are responding to evidence rather than coordinating on private interests.

This result has several policy implications. First, it provides a rationale for ensuring diverse perspectives in expert panels. While independent regulatory agencies typically strive for consensus, our analysis suggests that unanimous recommendations may be viewed with suspicion by a skeptical public. Publicly acknowledging minority views—even when the majority position is correct—may paradoxically increase credibility. Second, transparency about potential conflicts of interest becomes crucial: if voters could directly observe whether experts face conflicts, the inference problem would disappear. Third, the results suggest caution about "consensus messaging" strategies that emphasize near-universal expert agreement, as in climate communication campaigns. Such messaging may backfire precisely when voters suspect shared expert interests.

Several limitations warrant discussion. First, I assume voters cannot observe any characteristics that would allow them to distinguish biased from unbiased experts. If some credentialing or disclosure mechanisms reveal vested interests, the inference problem changes. Second, the model treats the pool of experts as exogenous. Endogenizing expert participation—for instance, allowing biased experts to strategically choose whether to advise—could generate additional predictions. Third, while I show non-monotonicity persists with strategic experts (Section 6.3), a richer model of strategic communication would allow experts to optimally design their messages anticipating voter skepticism.

Several extensions merit future research. Repeated interaction between voters and expert communities could allow voters to learn about the correlation structure of expert bias over time. Extending the model to multiple policy issues would permit studying how voter trust evolves as they observe expert track records across different domains. Empirically, the model suggests that survey experiments varying the degree of expert consensus could identify the non-monotonicity predicted here. Finally, while this paper focuses on rational voters, the interaction between the mechanisms identified here and behavioral factors such as confirmation bias or motivated reasoning remains an open question.

References

D. Austen-Smith. Interested experts and policy advice: Multiple referrals under open rule. *Games and Economic Behavior*, 5(1):3–43, 1993.

A. V. Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, 107 (3):797–817, 1992.

M. Battaglini. Multiple referrals and multidimensional cheap talk. *Econometrica*, 70(4): 1379–1401, 2002.

S. Bikhchandani, D. Hirshleifer, and I. Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100(5):992–1026, 1992.

H. Cheng and A. Hsiaw. Distrust in experts and the origins of disagreement. *Journal of Economic Theory*, page 105401, 2021.

V. P. Crawford and J. Sobel. Strategic information transmission. *Econometrica*, 50(6): 1431–1451, 1982.

D. Darmofal. Elite cues and citizen disagreement with expert opinion. *Political Research Quarterly*, 58(3):381–395, 2005.

M. de Condorcet. Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. 1785.

T. Feddersen and W. Pesendorfer. Voting behavior and information aggregation in elections with private information. *Econometrica*, 65(5):1029–1058, 1997.

T. Feddersen and W. Pesendorfer. Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting. *American Political Science Review*, 92(1):23–35, 1998.

M. Gentzkow and J. M. Shapiro. Media bias and reputation. *Journal of Political Economy*, 114(2):280–316, 2006.

D. Gerardi, R. McLean, and A. Postlewaite. Aggregation of expert opinions. *Games and Economic Behavior*, 65(2):339–371, 2009.

T. W. Gilligan and K. Krehbiel. Asymmetric information and legislative rules with a heterogeneous committee. *American Journal of Political Science*, 33(2):459–490, 1989.

C. D. Johnston and A. O. Ballard. Economists and public opinion: Expert consensus and economic policy judgments. *The Journal of Politics*, 78(2):443–456, 2016.

E. Kamenica and M. Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6): 2590–2615, 2011.

N. Kartik, F. X. Lee, and W. Suen. Opinion leaders and the design of collective wisdom. *Journal of Political Economy*, 129(9):2504–2545, 2021.

V. Krishna and J. Morgan. The art of conversation: eliciting information from experts through multi-stage communication. *Journal of Economic Theory*, 117(2):147–179, 2004.

G. Levy and R. Razin. Correlation neglect, voting behavior, and information aggregation. *The American Economic Review*, 105(4):1634–1645, 2015.

C. Martinelli. Would rational voters acquire costly information? *Journal of Economic Theory*, 129(1):225–251, 2006.

S. Morris. Political correctness. *Journal of Political Economy*, 109(2):231–265, 2001.

S. Mullainathan and A. Shleifer. The market for news. *American Economic Review*, 95(4): 1031–1053, 2005.

P. Ortoleva and E. Snowberg. Overconfidence in political behavior. *The American Economic Review*, 105(2):504–535, 2015.

M. Ottaviani and P. N. Sørensen. Reputational cheap talk. *The RAND Journal of Economics*, 37(1):155–175, 2006.

M. Rabin and J. L. Schrag. First impressions matter: A model of confirmatory bias. *The Quarterly Journal of Economics*, 114(1):37–82, 1999.

P. Sapienza and L. Zingales. Economic experts versus average americans. *The American Economic Review*, 103(3):636–642, 2013.

J. Sobel. A theory of credibility. *The Review of Economic Studies*, 52(4):557–573, 1985.

A. Wolinsky. Eliciting information from multiple experts. *Games and Economic Behavior*, 41(1):141–160, 2002.

Appendix: Proofs

Proof of Lemma 1. When $\omega = -1$, a citizen i prefers S to R if and only if $u_i(-1, S) > u_i(-1, R)$. And,

$$\begin{aligned} u_i(-1,S) - u_i(-1,R) &= -(-1+v_i-S)^2 + (-1+v_i-R)^2 \\ &= -(-1+v_i-S)^2 + (-1+v_i+S)^2 \\ &= -((-1+v_i)^2 - 2(-1+v_i)S + S^2) + ((-1+v_i)^2 + 2(-1+v_i)S + S^2) \\ &= 4(-1+v_i)S > 0 \end{aligned}$$

Analogously when $\omega = 1$. ■

Proof of Lemma 2. Given Π_k, the indifferent voter between status quo and the reform is $2\Pi_k - 1$. Given that voter's preferences follow a uniform distribution, this is also the share of voters for status quo and is increasing in Π_k. ■

Proof of Lemma 3.

$$\Pi_k < \Pi_{k+1} \leftrightarrow \frac{\pi(S,k)}{\pi(S,k+1)} < \frac{\pi(R,k)}{\pi(R,k+1)} \tag{13}$$

Now we just have to plug the expressions of $\pi(\omega, k)$:

$$\frac{\binom{n}{k}(\sigma+(1-\sigma)q)^k(1-(\sigma+(1-\sigma)q))^{n-k}}{\binom{n}{k+1}(\sigma+(1-\sigma)q)^{k+1}(1-(\sigma+(1-\sigma)q))^{n-k-1}} < \frac{\binom{n}{k}(\sigma+(1-\sigma)(1-q))^k(1-(\sigma+(1-\sigma)(1-q)))^{n-k}}{\binom{n}{k+1}(\sigma+(1-\sigma)(1-q))^{k+1}(1-(\sigma+(1-\sigma)(1-q)))^{n-k-1}} \leftrightarrow \tag{14}$$

$$\frac{(\sigma+(1-\sigma)q)^k(1-(\sigma+(1-\sigma)q))^{n-k}}{(\sigma+(1-\sigma)q)^{k+1}(1-(\sigma+(1-\sigma)q))^{n-k-1}} < \frac{(\sigma+(1-\sigma)(1-q))^k(1-(\sigma+(1-\sigma)(1-q)))^{n-k}}{(\sigma+(1-\sigma)(1-q))^{k+1}(1-(\sigma+(1-\sigma)(1-q)))^{n-k-1}} \leftrightarrow \tag{15}$$

$$\frac{(1-(\sigma+(1-\sigma)q))}{(\sigma+(1-\sigma)q)} < \frac{(1-(\sigma+(1-\sigma)(1-q)))}{(\sigma+(1-\sigma)(1-q))} \leftrightarrow \tag{16}$$

$$\frac{(1-\sigma)(1-q)}{(\sigma+(1-\sigma)q)} < \frac{(1-\sigma)q)}{(\sigma+(1-\sigma)(1-q))} \leftrightarrow \tag{17}$$

$$(1-q)(\sigma+(1-\sigma)(1-q)) < q(\sigma+(1-\sigma)q) \leftrightarrow \tag{18}$$

$$\sigma(1-2q) < (1-\sigma)(q^2-(1-q)^2) \tag{19}$$

Which from $\sigma \in [0,1]$ and $q \in (\frac{1}{2}, 1)$ always holds because the LHS is negative and the RHS is positive. ■

Proof of Proposition 1.

(i) When $k < n-1$, the proof is analogous to the proof of Lemma 3 for $\sigma = 0$.

(ii) When $k = n$,

$$\Pi_{n-1} < \Pi_n \leftrightarrow \tag{20}$$

$$\Pi_{n-1} < \frac{(1-p_H)\pi(S,n) + \frac{p_H}{2}}{(1-p_H)(\pi(S,n)+\pi(R,n)) + p_H} \leftrightarrow \tag{21}$$

$$\Pi_{n-1}\left((1-p_H)(\pi(S,n)+\pi(R,n)) + p_H\right) < (1-p_H)\pi(S,n) + \frac{p_H}{2} \leftrightarrow \tag{22}$$

$$\Pi_{n-1}\left((\pi(S,n)+\pi(R,n)) + p_H\left(1-(\pi(S,n)+\pi(R,n))\right)\right) < \pi(S,n) + p_H\left(\frac{1}{2} - \pi(S,n)\right) \leftrightarrow \tag{23}$$

$$p_H < \hat{p} = \frac{\pi(S,n) - \Pi_{n-1}\left((\pi(S,n)+\pi(R,n))\right)}{\left(\Pi_{n-1}\left(1-(\pi(S,n)+\pi(R,n))\right) - \left(\frac{1}{2} - \pi(S,n)\right)\right)} \tag{24}$$

Finally we have to show that $\hat{p} \in [0,1]$. But notice that $\lim_{p\to 0}\Pi_n = \frac{\pi(S,n)}{\pi(S,n)+\pi(R,n)} > \Pi_{n-1}$ and $\lim_{p\to 1}\Pi_n = \frac{1}{2} < \Pi_{n-1}$. Therefore, $\hat{p} \in [0,1]$.

■

Proof of Corollary 1.

$$\Pi_{n-1} = q \tag{25}$$

$$\Pi_n = \frac{p_H + (1-p_H)q^3}{1 + p_H - 3(1-p_H)(q-q^2)} \tag{26}$$

And $\Pi_n \leq \Pi_{n-1}$ if and only if

$$\frac{p_H + (1-p_H)q^3}{1 + p_H - 3(1-p_H)(q-q^2)} \leq q \tag{27}$$

Rearranging,

$$\hat{p} := 1 - \frac{1}{1+(1-q)q} \leq p_H \tag{28}$$

And $\hat{p}$ is decreasing in q and $\lim_{q \to 1} \hat{p} = 0$. ■

Proof of Lemma 4.

Suppose that three experts advise the reform. Given that experts only advise the reform if they are unbiased and they received a signal in favour of the reform, this means that all three experts received a signal in favour of the reform.

$$\pi'_0(s) = \frac{q'_c(1-q)^3}{q'_c(1-q)^3 + (1-q'_c)q^3} \tag{29}$$

And $\pi'_0(s) < \frac{1}{2}$ if and only if:

$$q'_c(1-q)^3 < (1-q'_c)q^3 \tag{30}$$

$$q'_c(1-q)^3 < (1-q'_c)q^3 \tag{31}$$

From $q'_c < q$, we have that $1 - q'_c > 1 - q$ and, from both inequalities we have $\frac{q'_c}{1-q'_c} < \frac{q}{1-q}$ and, therefore $q'_c(1-q) < q(1-q'_c)$. Now, we multiply both sides by $(1-q)^2$ and we get $q'_c(1-q)^3 < q(1-q)^2(1-q'_c)$ and the RHS is smaller than $q^3(1-q'_c)$. Thus, $\pi'_0(s) < \frac{1}{2}$ and, from the previous lemma, $\pi'_0(r) < \pi'_0(s) \leq \frac{1}{2}$.

If only two experts advise the reform, we have that $\pi'_1(s) < \frac{1}{2}$ if and only if $q'_c(1-q) < (1-q'_c)q$ which always holds. ■

Proof of Proposition 2. The first two statements are trivial and follow a proof similar to the one of the previous lemma. Regarding the third one,

$$\pi'_3(r) = \frac{(1-q'_c)(p_H + (1-p_H)q^3)}{(1-q'_c)(p_H + (1-p_H)q^3) + q'_c(p_H + (1-p_H)(1-q)^3)} \tag{32}$$

And status quo is implemented if and only if $\pi'_3(r) \geq \frac{1}{2}$ which happens if and only if:

$$(1-q'_c)(p_H + (1-p_H)q^3) \geq q'_c(p_H + (1-p_H)(1-q_c)^3) \tag{33}$$

$$p_H \leq \underline{p} = \frac{(1-q'_c)q^3 - q'_c(1-q)^3}{(1-(1-q)^3)q'_c - (1-q^3)(1-q'_c)} \tag{34}$$

Finally we have to prove that $\underline{p} \in (0,1)$.

When $p_H = 0$,

$$\pi_3'(r) = \frac{(1-q_c')q^3}{(1-q_c')q^3 + q_c'(1-q)^3} \tag{35}$$

And $\pi_3'(r) > \frac{1}{2}$ if and only if:

$$(1-q_c')q^3 > q_c'(1-q)^3 \tag{36}$$

And, from the proof of the previous proposition, this is always satisfied.

When $p_H = 1$,

$$\pi_3'(r) = \frac{(1-q_c')}{(1-q_c') + q_c'} \tag{37}$$

And $\pi'3(r) < \frac{1}{2}$ if and only if $q_c' > \frac{1}{2}$ which always holds.

Therefore, $\underline{p} \in (0,1)$.

■

Proof of Corollary 2. The proof is analogous to the proof of the previous proposition. ■

Proof of Proposition 3.

We know Proposition 1 that when $\hat{p} < p_H$, $x = 1$ can't be an equilibrium because an individual expert has incentives to deviate and sustain the reform because it will move support from Π_3 to Π_2. Hence it can't be an equilibrium.

Now, let's suppose that $x = q$, we will prove that this can't be an equilibrium either because an individual expert would have incentives to deviate and sustain status quo. Suppose all biased experts are supporting status quo with probability q. In this situation,

$$\pi(S,k) = \binom{3}{k}(p_H \cdot q + (1-p_H)q)^k(1-(p_H \cdot q + (1-p_H)q))^{3-k} = \binom{3}{k}q^k(1-q)^{3-k} \tag{38}$$

$$\pi(R,k) = \binom{3}{k}(p_H \cdot q + (1-p_H)(1-q))^k(1-(p_H \cdot q + (1-p_H)(1-q)))^{3-k} \tag{39}$$

Now we'll prove that for all k, $\Pi_k < \Pi_{k+1}$

$$\Pi_k < \Pi_{k+1} \leftrightarrow \frac{\pi(S,k)}{\pi(S,k+1)} < \frac{\pi(R,k)}{\pi(R,k+1)} \tag{40}$$

We plug in the expression we computed before:

$$\frac{q^k(1-q)^{3-k}}{q^{k+1}(1-q)^{3-(k+1)}} < \frac{(p_H \cdot q + (1-p_H)(1-q))^k(1-(p_H \cdot q + (1-p_H)(1-q)))^{3-k}}{(p_H \cdot q + (1-p_H)(1-q))^{k+1}(1-(p_H \cdot q + (1-p_H)(1-q)))^{3-(k+1)}} \tag{41}$$

Simplifying,

$$\frac{1-q}{q} < \frac{(1-(p_H \cdot q + (1-p_H)(1-q)))}{(p_H \cdot q + (1-p_H)(1-q))} \tag{42}$$

$$\frac{1}{q} < \frac{1}{(p_H \cdot q + (1-p_H)(1-q))} \tag{43}$$

$$p_H \cdot q + (1-p_H)(1-q) < q \tag{44}$$

$$1 < 2q \tag{45}$$

Which always holds.

Now, given that Π_k is increasing in k for all k, an individual expert has incentives to advise status quo.

Finally, by continuity, there exists an $x \in (q, 1)$ such that an expert is indifferent to support status quo or the reform given that the other two biased experts are endorsing status quo with probability x. Moreover, for this x, Pi_k can't be monotonic in k because if it was, the expert would not be indifferent. ■