Failing to correctly aggregate signals

Michele Piccione

London School of Economics

Ariel Rubinstein

School of Economics, Tel Aviv University
and Department of Economics, New York University

August 31, 2024

ABSTRACT: Individuals fail to recognize that two weak positive signals of a rare event together constitute strong evidence for the event. We demonstrate the dramatic effect of replacing Bayesian aggregation of signals with common aggregation procedures in a simple model of opinion formation and in a model of strategic voting.

KEYWORDS: signal aggregation, opinion formation, voting game
JEL classification: C72, D70, D80

We are grateful to Ayala Arad whose comments triggered this project, to Noga Alon and Michael Richter who suggested a proof for one of our results and to Aron Tobias for his insightful and generous comments. We also thank Sergiu Hart for his suggestions. Our special thanks go to Theodore Turocy, Anwar Ruff and Shayne Trotman for their generous assistance in organizing the surveys.

1. Introduction

We ask the reader to consider the following question:

> The proportion of newborns with a specific genetic trait is 1%. Two conditionally independent screening tests, A and B, are used to identify this trait in all newborns. However, the tests are not precise. Specifically, it has been found that:
> 70% of the newborns who are found to be positive according to test A have the trait.
> 20% of the newborns who are found to be positive according to test B have the trait.
> Suppose that a newborn is found to be positive according to both tests. What is your estimate of the probability that this newborn has the trait?

Notice that the properties of the two tests are described differently than they would be in standard economic theory. Usually, the tests would be described by the chances of a positive or negative result conditional on whether or not the newborn has the trait. In contrast, the question reports the likelihood of the trait given a positive result in either test. This type of description is widely used in real life. In medicine, for example, the results of a test are often assessed in terms of their *positive predictive value* (PPV) which is defined by the NIH as: "The likelihood that an individual with a positive test result truly has the particular gene and/or disease in question."[1] Similarly, the effectiveness of an alarm system is often measured by the *false alarm ratio* (FAR), which is "the number of false alarms per the total number of warnings or alarms in a given study or situation".[2] And just as importantly, this is the way we relate to a multitude of signals in daily life, such as the chances of snow when predicted by a forecaster, or the likelihood that someone is at the door when the dog barks.

In the case of one test, the above question is trivial: neither the prior nor the distribution of the signal given the states of the world are of any use. In the case of two tests, on the other hand, Bayesian updating is more complex and does depend on the prior.

[1] https://www.cancer.gov/publications/dictionaries/genetics-dictionary/def/positive-predictive-value

[2] https://www.statisticshowto.com/false-alarm-ratio-definition/

Suppose that $s \in (0,1)$ is the frequency of the trait and that a proportion $\phi_k \in [s,1)$ of the individuals who tested positive according to test k ($k=1,2$) indeed have the trait. Denote by p_k and n_k the probabilities of a positive result in test k for a newborn with and without the trait, respectively. By standard Bayesian updating:

$$\frac{\phi_k}{1-\phi_k} = \frac{s p_k}{(1-s)n_k}.$$

Denote by π the probability that a newborn who is found positive on both tests has the trait. Then:

$$\frac{\pi}{1-\pi} = \frac{s p_1 p_2}{(1-s)n_1 n_2} = \left(\frac{1-s}{s}\right)\frac{\phi_1\phi_2}{(1-\phi_1)(1-\phi_2)}.$$

Notice that: (i) π depends on s, (ii) if $\phi_1 = s$ then test 1 can be ignored and $\pi = \phi_2$, and (iii) given ϕ_1 and ϕ_2, π is close to 1 when s is very small. In our scenario, the probability that the newborn has the trait is $\pi = 0.983$!

Thus, Bayesian analysis leads to an unintuitive conclusion: when the trait is rare, two positive results indicate a high probability of the trait even if neither test is especially accurate. This can be significant in a variety of contexts: In medicine, the positive results of two conditionally independent tests for a rare disease might be concerning even though each result is not particularly alarming on its own. In legal proceedings, two conditionally independent clues about the guilt of a suspect, each unpersuasive on its own, may be strong evidence when considered jointly.[3]

We conjecture that most individuals would fail to compute the correct probability and would themselves be surprised by the correct answer, as were many of our colleagues when presented with the above question. The observation that in a variety of circumstances individuals fail to update beliefs via Bayes' rule is certainly not new. Introspection and a large psychological literature cast doubt on the hypothesis that Bayesian updating procedures resemble those used in real life, even when the information needed to form Bayesian posteriors is available and easily interpreted. In particular, much attention has been devoted to the "base-rate fallacy", identified in Kahneman and Tversky (1973) and elucidated in Bar-Hillel (1980). For a survey of the literature on errors in probabilistic reasoning, see Benjamin (2019) and for a survey of decision-making heuristics, see Gigerenzer and Gaissmaier (2011). In our context, the most relevant analysis is to be

[3] For a discussion of aggregation in law, see Porat and Posner (2012).

found in Budescu and Yu (2006) who present evidence that individuals tend to aggregate information from multiple sources by averaging.

Our own evidence, based on surveys reported in Section 6, confirms this. It also highlights some other intuitive non-Bayesian methods of signal aggregation. In particular, we find that in the above question individuals tend to use rules which lead to the answers 14%, 20%, 45%, 70% or 76%. These correspond to five simple formulae:

M: $\phi_1\phi_2$

Min: $min\{\phi_1, \phi_2\}$

Avg: $(\phi_1 + \phi_2)/2$

Max: $max\{\phi_1, \phi_2\}$

M^c: $1-(1-\phi_1)(1-\phi_2)$

Of these rules, only M^c is qualitatively consistent with the conclusion that both tests contribute positively to the posterior probability and thus the answer should exceed the larger of ϕ_1 and ϕ_2. Yet, even M^c fails to yield the conclusion that when the prior is small two moderately positive signals indicate a high probability of an event.

Since the above procedures for non-Bayesian belief aggregation are fairly evident and intuitive, it is worthwhile examining the implications of replacing Bayesian updating with some of these rules. We investigate two models whose conclusions are dramatically different from those of the standard analysis.

The first model is a simple setup of opinion formation in a society. There are two possible states of nature. An "opinion" is a belief about the state of nature. A "newcomer" to the society receives a private signal about the state as well as the opinion of one sampled individual in the population. We assume that the newcomer forms his belief by averaging the probabilities of the state estimated using these two signals. A distribution of opinions is stable if a newcomer's final belief has the same distribution as the beliefs of members of the society. We characterize the unique stable distribution and show that it cannot fully reveal the state of nature, in contrast to the case in which all agents would use Bayesian reasoning.

The second model is a special case of the voting game in Duggan and Martinelli (2001). A group of jurors is to decide whether to convict or acquit a defendant. Conviction is required to be unanimous. Every juror receives a private assessment of the probability that the defendant is guilty. As is standard, in this type of analysis, a juror aggregates his private belief with a fictitious, equilibrium signal, that is, the event that all other jurors voted to convict (and thus he is pivotal). In Duggan and Martinelli (2001) agents aggregate signals in a Bayesian way while, in our case, if the number of agents is large enough, then the correct decision is made with probability close to 1. We show that:

(1) If all the jurors use the Avg procedure and the number of jurors is large almost all defendants are acquitted.

(2) If all the jurors use the Max procedure and the number of jurors is large, then all guilty defendants are convicted, as well as some proportion of the innocent ones. Remarkably, the level of welfare converges to that achieved when the jurors never convict.

(3) While a "large" Bayesian jury can approximate the first-best, other types of "large" non-Bayesian juries achieve the same level of welfare in the limit as a jury that never convicts, even though the jurors use different equilibrium strategies.

2. Aggregation of Opinions

In this section, we consider a simple model of aggregation of opinions. There are two equally likely states of nature, a and b. There is a population of agents, each with his own belief about the state which we define always as the probability of state a. Denote by F_a the distribution of beliefs for state a in the population when the true state is a and by F_b the distribution of beliefs for state a (!) in the population when the true state is b.

A "new" agent receives a "personal" private signal that declares the state of nature. The signal is incorrect with probability $0 < \epsilon < \frac{1}{2}$. The agent is aware that his signal may be incorrect and thus his belief (again, for state a) based on his personal signal alone is a random variable which at state a (b) has the value $1-\epsilon$ with probability $1-\epsilon$ (ϵ) and ϵ with probability ϵ ($1-\epsilon$). Put differently, let B be a Bernoulli random variable that has value equal to 1 with probability $1-\epsilon$ and equal to 0 with probability ϵ. Then, the belief of an agent given his personal signal is distributed as the random variables $Y_a = (1-2\epsilon)B+\epsilon$ in state a and $Y_b = (1-\epsilon)-(1-2\epsilon)B$ in state b.

A new agent also samples one belief from the population of agents. We will refer to the outcome of this sampling process as the "sample signal". For $s = a, b$, denote by X_s a random variable, conditional independently of Y_s, which is distributed according to F_s. In contrast to the personal signal, the distribution of the sample signal will be endogenous, as will be shown below.

Given a state, the aggregation of the personal signal and the sample signal by the new agent, whatever the method used, yields a belief that is itself a random variable. We will look for distributions of beliefs in the population F_s, $s = a, b$, such that the new agent's beliefs, given the personal signal and the sample signal, are also distributed according to F_s. We refer to such distributions as *stable* and to (F_a, F_b) as a *stable pair*.

It is easily verified that if all agents are Bayesian, then a pair of distributions that reveals the true state is stable. That is, there is a stable pair of distributions such that according to F_a the entire population believes that the probability that the state is a is 1 and according to F_b the entire population believes that the probability that the state is a is 0.

The above conclusion does not hold if agents update their beliefs using the *Avg* procedure. One notable feature of the *Avg* procedure – and also for the *Max* procedure – is that the outcome of the aggregation in a state s is a random variable which depends on F_s alone but not on the distribution given the other state, as in Bayesian updating. Thus, we can study the stability of F_a and F_b independently.

Now assume that the agents update their beliefs by averaging the two estimates obtained from the personal signal and the sample signal. Stability is now equivalent to finding for each state s a random variable X_s, independent of Y_s, such that $(X_s + Y_s)/2$ is distributed as X_s. A solution to this mathematical problem was suggested to us by the mathematician Noga Alon and modified by Michael Richter:

Lemma: Let Y be a bounded random variable. There is then an independent random variable X such that X is distributed as $\lambda Y + (1-\lambda)X$.

Proof: Let $(Y^t)_{t=0,1,2,...}$ be a sequence of independent random variables distributed as Y. Let $X = \sum_{t=0} \lambda(1-\lambda)^t Y^t$. Then,

$$\lambda Y + (1-\lambda)X = \lambda Y + \sum_{t=0} \lambda(1-\lambda)^{t+1} Y^t \sim_F X$$

Claim A *If all agents use the Avg procedure then the stable pair of distributions is unique.*

Proof. For existence, consider the random variable Y_s defined earlier. By the above claim, there is a random variable X_s such that $(X_s + Y_s)/2$ is distributed as X_s.

For uniqueness, it suffices to show that the values of a stable distribution F_a are uniquely determined on a dense set (by symmetry it will follow that F_b is also unique).

First note that for a stable F_a, the following equation must be satisfied for any y:

$$F_a(y) = \epsilon F_a(2y - \epsilon) + (1-\epsilon) F_a(2y - 1 + \epsilon).$$

We now show that $F_a(\epsilon) = 0$ and $F_a(1-\epsilon) = 1$. Let y be the smallest x such that $F_a(x) = F_a(\epsilon)$. Since $2y - \epsilon \leq \epsilon$, then $F_a(2y - \epsilon) \leq F_a(y)$. Furthermore, since $2y - 1 + \epsilon < y$ if $F_a(\epsilon) > 0$ then $F_a(2y - 1 + \epsilon) < F_a(\epsilon)$ and thus the above equation cannot hold. Hence, $F_a(\epsilon) = 0$. Similarly, let $\hat{y}$ be the lowest upper bound to the set $\{x : F(x) = F(1-\epsilon)\}$. If this set contains only $1-\epsilon$, it is easily verified that the above equation holds only if $F_a(1-\epsilon) = 1$. Otherwise, choose $y < \hat{y}$ arbitrarily close to $\hat{y}$. Since $2y - \epsilon > \hat{y}$ and $2y - 1 + \epsilon > \hat{y}$, for y sufficiently close to $\hat{y}$, the above equality holds only if $F_a(y) = 1$.

Define the set $D_k = \{d_{n,k}\}_{n=0,1,...,2^k}$ where $d_{n,k} = \epsilon + n\frac{1-2\epsilon}{2^k}$. The set D_0 includes $d_{0,0} = \epsilon$ and $d_{1,0} = 1-\epsilon$. The set D_{k+1} includes D_k and all the half points between neighboring points in D_k. Let $D = \cup_{k=1,2,...} D_k$. Obviously, D is dense in $[\epsilon, 1-\epsilon]$.

The average belief is below a point $x \leq 1/2$ if and only if the belief from the personal signal is ϵ and from the sample signal at most $2x - \epsilon$. Thus, $F_a(x) = \epsilon F_a(2x - \epsilon)$. It follows that the value of F_a at $d_{n,k} \leq \frac{1}{2}$ is determined uniquely by $\epsilon F_a(2d_{n,k} - \epsilon)$ where, by definition, $2d_{n,k} - \epsilon = d_{n,k-1}$. Similarly, the average belief is below a point $x \geq 1/2$ if and only if (i) the belief from the personal signal is ϵ or (ii) the belief from the personal signal is $1-\epsilon$ and from the sample signal is at most $2x - 1 + \epsilon$. Thus, $F_a(x) = \epsilon + (1-\epsilon)F_a(2x - 1 + \epsilon)$. It follows that the value of F_a at $d_{n,k} \geq \frac{1}{2}$ is determined uniquely by $F_a(2d_{n,k} - 1 + \epsilon)$ where, by definition, $2d_{n,k} - 1 + \epsilon = d_{n-2^{k-1},k-1}$. Thus, the values of F_a on D_k are determined uniquely by the values on D_{k-1}. It follows that F_a is uniquely determined on the dense set D and the claim is proved. ■

To illustrate the proof, consider $\epsilon = 1/3$:

points	values of F
D_0	$F(1/3) = 0, \ F(2/3) = 1$
$D_1 - D_0$	$F(1/2) = 1/3$
$D_2 - D_1$	$F(\frac{5}{12}) = \frac{1}{9}$, $F(\frac{7}{12}) = \frac{5}{9}$
$D_3 - D_2$	$F(\frac{9}{24}) = \frac{1}{27}$, $F(\frac{11}{24}) = \frac{5}{27}$, $F(\frac{13}{24}) = \frac{11}{27}$, $F(\frac{15}{24}) = \frac{19}{27}$

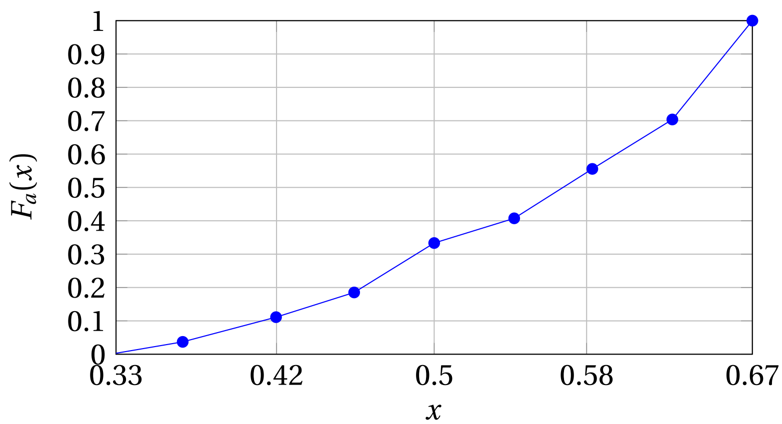

Figure 1: Plot of $F_a(x)$.

Note that the function F_a is neither convex nor concave.

In the Max procedure, it is of course essential to specify whether the operation is performed on the probability of state a or of state b. Here, we choose state a. In particular, we assume that when the new agent's belief from the private signal for state a is α and the sampled belief for state a is β, his inference is that the probability of state a is $max\{\alpha, \beta\}$. This specification does affect which distributions are stable.

Claim B *In each state, a distribution is stable if and only if it assigns probability equal to zero to any belief strictly below* $1 - \epsilon$.

Proof. Obviously, such a distribution is stable as the outcome of the sample is always the larger of the sample signal and the private signal and thus the beliefs of a new agent are distributed according to the distribution in the population.

Let G_s be a stable distribution at state s and let m_s be the probability that beliefs are strictly below $1-\epsilon$ according to G_s. At state s, a new agent will have a belief below $1-\epsilon$ only if (i) according to his personal signal the probability of state a is ϵ, which happens with probability δ_s (which is either ϵ or $1-\epsilon$ and thus strictly below 1), and (ii) his sample signal is below $(1-\epsilon)$ (which has probability m_s in state s). However, $m_s \delta_s \neq m_s$ unless $m_s = 0$. Therefore, no stable pair of distributions assigns positive probability to beliefs below $1-\epsilon$. ■

We have assumed that the Max procedure operates on the same state, state a, regardless of the realisation of the personal signal. Alternatively, the procedure could depend on the state "declared" by the personal signal rather than on a predetermined one. For instance, it could be applied to the belief for state a when the personal signal declares "a" and to the belief for state b when the personal signal declares "b". Using the above argument, one can show that a pair (F_a, F_b) is stable if and only if F_a assigns probability equal to zero to any belief strictly below $1-\epsilon$ and F_b assigns probability equal to zero to any belief strictly above ϵ. Note that in this case a pair of distributions that reveals the true state is stable, as in the case of Bayesian updating.

3. Voting: an Example

In models with incomplete information, agents aggregate multiple signals that are either observed privately and directly or inferred from equilibrium. In this section, we present a simple example of strategic voting that illustrates the aggregation of private and equilibrium signals using some of the procedures discussed in the introduction.

Two jurors must determine the guilt of a defendant. Initially, they believe that the probability of the accused being guilty is $s = 0.1$. Each juror receives a conditionally independent binary signal about the guilt or innocence of the defendant which is incorrect with probability $\alpha = 0.1$. The jurors vote simultaneously on the defendant's guilt by choosing either Y or N. The defendant is found guilty only if both jurors vote Y.

The jurors' payoffs are depicted in the two matrices below: a juror receives 1 if the decision is correct when the defendant is guilty, 8 if the decision is correct when the defendant is innocent, and 0 in any other case. Thus, each juror prefers a conviction only if his belief that the defendant is innocent does not exceed $1/9$.

Not Guilty	N	Y
N	8	8
Y	8	0

Guilty	N	Y
N	0	0
Y	0	1

Obviously, there is an equilibrium in which both jurors always vote N. If the jurors are Bayesian, then there is also an equilibrium in which they vote according to their signals. To see this, note that a juror cares only about the event in which he is pivotal, that is, when the other juror chooses Y. If he receives a signal declaring innocence and the other juror a signal declaring guilt (and thus votes Y), he believes that the defendant is guilty with probability $\frac{s(1-\alpha)\alpha}{(1-s)(1-\alpha)\alpha+s\alpha(1-\alpha)} = s = 0.1$ and it is indeed optimal for him to vote N. Conversely, with a signal declaring guilt, conditional on the other juror getting the same signal, the probability of guilt is $\frac{s(1-\alpha)^2}{s(1-\alpha)^2+(1-s)\alpha^2} = \frac{0.1(0.9)^2}{0.1(0.9)^2+0.9(0.1)^2} = 0.9$ and it is optimal for him to vote Y. Thus, voting Y if and only if the signal declares guilt is an equilibrium. A guilty defendant is convicted with probability 0.81 and an innocent one with probability 0.01.

Pivotal reasoning suggests a natural way of applying some of the non-Bayesian procedures we found in our survey. Assume that each juror forms his belief about the defendant's guilt by aggregating the beliefs from two signals: his own private signal and the "pivotal" event that the other agent votes Y. The latter signal is an "equilibrium" signal and its information content is determined by the equilibrium profile. Suppose that there is an equilibrium in which the jurors vote according to their private signals. Naturally, with Bayesian updating we would reproduce the above conclusion. Each of the signals on its own increases the probability that the defendant is guilty to $\frac{s(1-\alpha)}{s(1-\alpha)+(1-s)\alpha} = 0.5$. Then, both the Avg and Max procedures generate the belief that the defendant is guilty with probability 0.5, while the M^c procedure generates the belief that the defendant is guilty with probability 0.75. Thus, on receiving a signal declaring guilt, a juror using any of those procedures will choose N. We conclude that the unique equilibrium is non-informative.

In the next section we expand this example to a version of the familiar strategic voting game of Duggan and Martinelli (2001).

4. The voting game in Duggan and Martinelli (2001)

A panel of n jurors is to determine whether a defendant is guilty or innocent. The prior probability of guilt is denoted as s. The jurors vote simultaneously. Each juror votes either Y (guilty) or N (not guilty) and the defendant is found guilty only when the jurors unanimously vote Y. Prior to voting, each juror receives a private signal in the form of a number in $[0,1]$ about the guilt of the defendant and then votes whether or not to convict. The signals are identically distributed and conditionally independent across the jurors. The cdf of a signal conditional on the defendant being guilty (innocent) is F (G). The cdfs have continuous density functions f and g, respectively. The following restrictions are imposed throughout:[4]

(i) $f(0)=0$, $g(1)=0$, $f(t)>0$ for all $t>0$ and $g(t)>0$ for all $t<1$.

(ii) $\frac{f(t)}{g(t)}$ is strictly increasing.

A juror prefers that the defendant be convicted if he believes that the probability of guilt is at least z, a number in $(0,1)$. This is equivalent to each juror maximizing a vNM utility function that is equal to 1 if the correct decision is made when the defendant is guilty; equal to $\lambda=\frac{z}{1-z}$ if the correct decision is made when the defendant is innocent; and equal to 0 if the incorrect decision is made. Therefore, a natural welfare function when each juror votes Y if and only if his observed signal is at least α, is:

$$W^n(\alpha)=s(1-F(\alpha))^n+(1-s)\lambda(1-(1-G(\alpha))^n).$$

We assume that $z\geq 1/2\geq s$. That is, the ex-ante belief that the defendant is guilty is not stronger than the belief that he is innocent and the standards for conviction are higher than the standards for acquittal.

In a strategic environment with incomplete information such as this one, the strategies of the "other players" generate events that can be interpreted by a player as "signals" and are aggregated with other exogenous signals to formulate posterior beliefs.[5] In the standard approach to binary voting models, a voter's best reply depends on his own information and the event in which his vote is pivotal. Accordingly, we assume that each

[4]Condition (ii) implies that $\frac{F(t)}{G(t)}<\frac{f(t)}{g(t)}<\frac{1-F(t)}{1-G(t)}$ for all $t\in(0,1)$, $\frac{F(t)}{G(t)}$ and $\frac{1-F(t)}{1-G(t)}$ are strictly increasing, and together with condition (i), $\lim_{t\to 1}\frac{1-F(t)}{1-G(t)}=\infty$.

[5]For an alternative perspective on belief formation in Bayesian games, see Eyster and Rabin (2005).

juror aggregates two signals: his own private signal and the event that he is pivotal. The probability of guilt given the private signal is exogenous. The probability of guilt given that a juror is pivotal is determined in equilibrium. We assume that each juror applies a signal-aggregation procedure in order to form a belief about the defendant's guilt and votes to convict if the belief is above the threshold z.

We refer to juror i's belief that the defendant is guilty – given his private signal t and conditional on him being pivotal given that all players vote Y if and only if their signal is above σ – as his C-belief and denote it by $\mu_i(t,\sigma)$. The C-beliefs depend on the procedures – Bayesian or non-Bayesian – used by the jurors to update their beliefs.

To summarize, the model is the tuple $\langle n, s, z, F, G, (\mu_i)\rangle$ where n is the number of jurors, s is the common prior probability that the defendant is guilty, z is the common minimal belief as to the guilt of the defendant for which a conviction is optimal, F and G are the distributions of each private signal given that the defendant is respectively guilty or not guilty, and μ_i is juror i's C-belief.

It remains to define the equilibrium concept. A *(symmetric) equilibrium* is a cutoff $\sigma^* < 1$ such that for every juror i we have that $\mu_i(t,\sigma^*) \le z$ for all $t < \sigma^*$ and $\mu_i(t,\sigma^*) \ge z$ for all $t > \sigma^*$. That is, in equilibrium whenever a juror votes N he believes that the probability of the defendant being guilty is sufficiently low (weakly below z) and whenever he votes Y he believes that it is sufficiently high (weakly above z). We require that $\sigma^* < 1$ in order to exclude the discussion of the non-informative equilibrium (which always exists) in which all players vote N regardless of their signals and no player is ever pivotal.

This model was not chosen because it is particularly realistic[6] but because it is standard, simple and well known. Other applications could have served the purpose. In models of mechanism design, auctions, bargaining, and pricing with rational expectations, agents also aggregate multiple "signals" that are either observed directly or inferred from equilibrium, as in the voting model.

We first review the voting model under the standard approach where for each i the C-beliefs are defined by Bayesian updating:

$$\mu_i(t,\sigma) = \frac{sf(t)(1-F(\sigma))^{n-1}}{sf(t)(1-F(\sigma))^{n-1} + (1-s)g(t)(1-G(\sigma))^{n-1}}.$$

[6]We agree with the criticism that the application of pivotal reasoning in voting is rather unrealistic (see Esponda and Vespa (2014) for experimental evidence).

Since μ_i is strictly increasing in t for a fixed $\sigma < 1$, σ^* is an equilibrium if and only if it satisfies:

$$\mu_i(\sigma^*, \sigma^*) = z.$$

Clearly, the equilibrium is unique.[7] For later use, we rearrange the equilibrium condition. Let $r^k(\theta)$ be the probability that the defendant is guilty conditional on k jurors with a common cutoff θ voting Y:

$$r^k(\theta) = \frac{s(1-F(\theta))^k}{s(1-F(\theta))^k + (1-s)(1-G(\theta))^k},$$

and set $r^k(1) = 1$, which is the limit of $r^k(\theta)$ as $\theta \to 1$. The function r^k is strictly increasing and $r^k(\theta) < r^{k+1}(\theta)$ for any $k \geq 1$ and $0 < \theta < 1$. Then, σ^* is an equilibrium if and only if it satisfies:

$$r^{n-1}(\sigma^*) = \frac{zg(\sigma^*)}{zg(\sigma^*) + (1-z)f(\sigma^*)}.$$

It follows from Duggan and Martinelli (2001) that the equilibrium condition also characterizes welfare maximization with a common cutoff. As the number of jurors increases, the optimal cutoff decreases and converges to 0. Furthermore, the probability of an incorrect decision given the optimal cutoff converges to 0. That is, the equilibrium level of welfare under the standard Bayesian approach converges to the "first-best", namely $s + (1-s)\lambda$.

We now consider the case in which all or some of the players are non-Bayesian and use one of the procedures for signal aggregation described in the introduction. These procedures use a formula that is a function of the two probabilities of the relevant state conditional on each signal. In our case:
(i) the probability $p(t)$ that the defendant is guilty given the juror's own signal t:

$$p(t) = \frac{sf(t)}{sf(t) + (1-s)g(t)}$$

[7] Our assumptions imply that $\mu_i(0,0) = 0$, and the function $\mu_i(t,t)$ is increasing, continuous and converges to 1 as $t \to 1$. Therefore, there is a unique α satisfying the equation $\mu_i(\alpha, \alpha) = z$.

(ii) the probability $r^{n-1}(\theta)$ that the defendant is guilty given that the juror is pivotal and the rest of the jurors use the cutoff θ.

Note that p is strictly increasing and satisfies $p(t) < r^k(t)$ for any $t \in (0,1)$ and $k \geq 1$.

We will see that replacing the Bayesian approach with other methods of signal aggregation (Avg or Max) discussed in the introduction yields significantly different results.

4.1 The Avg game

Assume that all jurors use the Avg procedure, that is, juror i's C-belief is:

$$\mu_i(t,\sigma) = \frac{1}{2}(p(t) + r^{n-1}(\sigma)).$$

Since C-beliefs are monotonic in t, equilibria are characterized by the solutions of the equation $\frac{1}{2}(p(\beta) + r^{n-1}(\beta)) = z$. The following claim states that the equilibrium cutoff is above the Bayesian one and as the number of jurors increases the probability that any defendant is convicted converges to zero.

Claim C *In the Avg game:*
(i) There is a unique equilibrium denoted by β^*.
(ii) $\beta^* \geq \sigma^*$ with strict inequality unless $z = s = \frac{1}{2}$.
(iii) Denote by $\beta^*(n)$ the equilibrium in the game with n jurors. If $z > \frac{1}{2}$, then as $n \to \infty$:
(a) $\beta^*(n) \to \bar{\beta}$ where $p(\bar{\beta}) + 1 = 2z$; and
(b) the level of welfare converges to $\lambda(1-s)$ (the defendant is almost never convicted).

Proof. (i) The function $\frac{1}{2}(p(\beta)+r^{n-1}(\beta))$ is increasing in β and ranges from $\frac{s}{2}$ to 1. Hence there is a unique β^* at which it equals z. Clearly, this is the unique equilibrium.
(ii) Since $z \geq s$:

$$p(\sigma^*) + r^{n-1}(\sigma^*) = \frac{sf(\sigma^*)}{sf(\sigma^*) + (1-s)g(\sigma^*)} + \frac{zg(\sigma^*)}{zg(\sigma^*) + (1-z)f(\sigma^*)} \leq$$

$$\frac{zf(\sigma^*)}{zf(\sigma^*) + (1-z)g(\sigma^*)} + \frac{zg(\sigma^*)}{zg(\sigma^*) + (1-z)f(\sigma^*)} \leq 2z.$$

The first inequality holds strictly unless $s = z = \frac{1}{2}$ given that $z \geq \frac{1}{2} \geq s$. The second inequality follows from $\frac{x}{zx+(1-z)y} + \frac{y}{zy+(1-z)x} \leq 2$ which holds with equality only for $z = \frac{1}{2}$ or $x = y$ (or $z = 1$). Therefore, one of the inequalities will hold strictly unless $s = z = \frac{1}{2}$.

By the monotonicity of p and r^{n-1}, it follows that $\beta^* \geq \sigma^*$ with strict inequality unless $z = s = \frac{1}{2}$.

(iii) Since the function $\frac{1}{2}(p(\beta) + r^{n-1}(\beta))$ is increasing in β and n, the sequence $\beta^*(n)$ is decreasing. Since $p(\beta^*(n)) \geq 2z - 1 > 0$, $\beta^*(n)$ is bounded away from 0 and the equilibrium probability that the defendant will be found guilty converges to 0 as $n \to \infty$. Therefore, the sequence $\beta^*(n)$ converges to the unique solution of $p(\beta) + 1 = 2z$. ■

4.2 The Max game

Suppose that all jurors use the Max procedure, that is, for all i:

$$\mu_i(t, \theta) = max\{p(t), r^{n-1}(\theta)\}.$$

The following claim states that the equilibrium cutoff is always below the Avg cutoff and also below the Bayesian one when n is large. As the number of jurors increases, the equilibrium level of welfare approaches that in the Avg game, namely the level of welfare in the case of a panel of jurors that never convicts, despite the fact that convictions occur with probability larger than s.

Claim D *In the Max game:*
(i) There exists a unique equilibrium γ^* satisfying $r^{n-1}(\gamma^*) = z$.
(ii) $\gamma^* \leq \beta^*$.
(iii) Let $\gamma^*(n)$ be the equilibrium in the game with n jurors. There exists a sequence $\eta(n)$ that converges to ∞ such that for any n, $\sigma^*(n) > \gamma^*(n)$ if and only if $\lambda < \eta(n)$.
(iv) As n increases, the probability of convicting a guilty defendant converges to 1 and that of convicting an innocent one converges to $\frac{s}{(1-s)\lambda}$. The level of welfare converges to $\lambda(1-s)$.

Proof. (i) The unique solution γ^* of $r^{n-1}(\gamma) = z$ is an equilibrium: if a juror receives a signal below γ^*, then his C-belief is z and he is indifferent between voting Y and voting N; if he receives a signal above γ^*, then his C-belief is at least z and voting Y is optimal.

There is no other equilibrium:

(a) A common cutoff $\underline{\gamma} < \gamma^*$ is not an equilibrium since a juror with a signal $t \in (\underline{\gamma}, \gamma^*)$ has a C-belief equal to $max\{p(t), r^{n-1}(\underline{\gamma})\}$ which is less than z since $p(t) < p(\gamma^*) < r^{n-1}(\gamma^*) = z$ and $r^{n-1}(\underline{\gamma}) < r^{n-1}(\gamma^*) = z$. Such a juror prefers to vote N.

(b) A common cutoff $\bar{\gamma} > \gamma^*$ is not an equilibrium since in that case any juror's C-belief is at least $r^{n-1}(\bar{\gamma}) > z$ and thus voting N is not optimal given any signal.

(ii) The assertion follows from the Avg equilibrium condition $\frac{1}{2}(p(\beta^*) + r^{n-1}(\beta^*)) = z$ and the fact that $r^{n-1}(t) > p(t)$ for all $t \in (0,1)$.

(iii) $\gamma^*(n)$ is the solution to $\frac{r^{n-1}(t)}{1-r^{n-1}(t)} = \lambda$ and $\sigma^*(n)$ is the solution to $\frac{r^{n-1}(t)}{1-r^{n-1}(t)}\frac{f(t)}{g(t)} = \lambda$. The two LHS functions are increasing and have the same value only at the point $\bar{t}$ such that $f(\bar{t}) = g(\bar{t})$. Define $\eta(n) = \frac{r^{n-1}(\bar{t})}{1-r^{n-1}(\bar{t})}$. The sequence converges to infinity since $F(\bar{t}) < G(\bar{t})$. Then, $\lambda < \eta(n)$ iff $\gamma(n) < \bar{t}$ iff $\frac{f(\gamma(n))}{g(\gamma(n))} < 1$ iff $\sigma^*(n) > \gamma^*(n)$.

(iv) Since $\gamma^*(n) < \sigma^*(n)$ for large n, the probability that a guilty defendant is found guilty goes to 1. The ratio of convictions of guilty defendants to those of innocent defendants is $\frac{s(1-F(\gamma^*(n)))^n}{(1-s)(1-G(\gamma^*(n)))^n} = \frac{r^{n-1}(\gamma^*(n))}{1-r^{n-1}(\gamma^*(n))}\frac{1-F(\gamma^*(n))}{1-G(\gamma^*(n))}$ which converges to λ by the equilibrium condition. Therefore, the probability of an innocent defendant being found guilty converges to $\frac{s(1-z)}{(1-s)z}$ and the level of welfare converges to $s + (1-s)\lambda(1 - \frac{s}{1-s}\frac{1-z}{z}) = \lambda(1-s)$. ■

Comment (The Min Game): By an argument similar to that in Claim D, the only equilibrium of the game in which all jurors follow the Min procedure is δ^* satisfying $p(\delta^*) = z$. Since the equilibrium is independent of n and $\delta^* > 0$, the probability of conviction goes to zero as the number of jurors increases, and the level of welfare converges to $(1-s)\lambda$, as in the Avg and Max games.

Comment (The Mixed Bayesian and *Avg* Game): Suppose that $n\kappa$ jurors are Bayesian and the rest (at least two) use Avg. An extension of the equilibrium definition specifies α^* and β^* (both less than 1) where α^* is the common cutoff of the Bayesian players and β^* is the common cutoff of the Avg players. Assuming that $z > \frac{1}{2} > s$ one can show that an equilibrium exists. Moreover, as $n \to \infty$, (i) any sequence of equilibria $(\alpha^*(n), \beta^*(n))$ converges to $(0, \bar{\beta})$ where $p(\bar{\beta}) = 2z - 1$ and (ii) the equilibrium probability of conviction converges to zero.

Comment (Aggregation of n Signals): An alternative approach to modelling belief formation with non-Bayesian procedures is that each juror aggregates n signals rather than two: his own signal and additional distinct signals, one for each Y vote cast by the other jurors. That is, a juror i who receives the signal t and knows that all of the other jurors follow a common cutoff θ forms his C-belief by aggregating:
(i) his own signal t (according to which the defendant is guilty with probability $p(t)$); and
(ii) $n-1$ other signals: for each juror $j \neq i$ the signal that he is voting Y, that is, j's private signal, is at least θ (for any one of these signals the defendant is guilty with probability $r^1(\theta)$).
Thus, under the Avg approach, $\mu_i(t,\theta) = \frac{1}{n}(p(t)+(n-1)r^1(\theta))$ and under the Max approach, $\mu_i(t,\theta) = max\{p(t), r^1(\theta)\}$.

We verified that the conclusions reached in this section remain essentially valid. In particular, in the equilibria of the Avg and Max games, the probability of conviction goes to zero as the number of jurors becomes large.

5. Survey evidence

The intuition that people fail to recognize that two moderate and conditionally independent signals of a rare event aggregate to significant evidence of the event was confirmed in several discussions with colleagues familiar with probability theory, many of whom were surprised by this feature of Bayesian updating. We also conducted a survey which showed that only a small minority of people reach the correct conclusion and that the majority use one of the formulae mentioned in Section 1.

The main survey was conducted at the Center for Experimental Social Science at New York University (NYU) and the Centre for Behavioural and Experimental Social Science at the University of East Anglia (UEA).[8] Students registered with the labs were invited to participate[9] and each participant was assigned randomly to answer one of five

[8] We are indebted to the Centre for Behavioural and Experimental Social Science at The University of East Anglia, and especially to Theodore Turocy, and to the Center for Experimental Social Science at New York University, and especially to Anwar Ruff and Shayne Trotman, who were so generous with their time.

[9] At UEA, one out of ten students received 20 pounds for participating and at NYU each student received \$10. Students were not incentivized in any other way. We do not report the results for each center separately since the differences were negligible.

questions.[10] The questions were similar to the one presented in the introduction except for two main differences. First, we provided a non-technical description of conditional independence. Second, in four of the five questions, the ex-ante frequency of the trait of "1%" was replaced with "a very small proportion". Here is Q1 as it appeared on the survey:

Q1: A very small proportion of the newborns in a certain country have a specific genetic trait.
Two screening tests, A and B, have been introduced for all newborns in order to identify this trait. However, the tests are not precise. A study has found that:
70% of the newborns who are found to be positive according to test A have the genetic trait (and conversely 30% do not).
20% of the newborns who are found to be positive according to test B have the genetic trait (and conversely 80% do not).
The study has also found that when a newborn has the genetic trait, a positive result in one test does not affect the likelihood of a positive result in the other. Likewise, when a newborn does not have the genetic trait, a positive result in one test does not affect the likelihood of a positive result in the other.

Suppose that a newborn is found to be positive according to both tests. What is your estimate of the likelihood (in %) that this newborn has the genetic trait?

In Q2, we changed the probabilities to **80%** and **60%** instead of 70% and 20%. In Q3, we changed the underlying story by replacing newborns with **undergraduates** who are either continuing on to graduate school or not. A positive result on a test is replaced by a student having taken a certain course. The proportion of students who continue on to graduate school is 70% for those who took course A and 20% for those who took course B. The results for Q1, Q2 and Q3 are presented in Table 1.

Note that under the assumption that "a very small proportion" is interpreted as being

[10] Prior to this survey we conducted a pilot on the platform arielrubinstein.org/gt, a website for carrying out pedagogical experiments in choice theory and game theory, in which almost all of the participants are current or past students in game theory courses. No monetary incentives were provided other than a few participants being randomly chosen to receive $40 regardless of their answers. No significant difference were detected between the pilot's results and those reported below.

at most 5%, a correct answer in Q1 and Q3 is at least 91% and in Q2 it is at least 98%. The proportions of those answers appear in the first column.

	"correct"	n *(MRT)*	M	Avg	Max	$> Max$ (M^c)	*Other*
$Q1_{\text{genetic, 70-20}}$	3%	93 (118s)	14%	20%	14%	20% (4%)	32%
$Q2_{\text{genetic, 80-60}}$	1%	91 (107s)	21%	27%	15%	18% (11%)	19%
$Q3_{\text{students, 70-20}}$	1%	97 (98s)	7%	27%	20%	21% (6%)	25%

Table 1: Results for Q1, Q2 and Q3.

As shown in the table, a majority (about 60%) of the participants used one of the four formulae: M, Avg, Max or M^c. Around 20% chose an answer strictly above the larger of ϕ_1 and ϕ_2, which is qualitatively correct. There do not appear to be any significant differences in the results between the three questions. The median response time (MRT) was also quite similar in all three, ranging from 98 to 118 seconds.

Q4 was used to check the consistency of the approaches used by the participants. We asked them to simultaneously estimate the likelihood of the trait given two different pairs of test results: Alice tests positive on two tests with accuracy of 70% and 20%, respectively and Bob tests positive on two tests with accuracy of 50% and 40%, respectively. We observe a considerable degree of consistency. Of the 92 participants, about 60% were consistent in their use of one of the four formulae: M (15%), Avg (30%), Max (13%) and M^c (2%). Furthermore, 8% of the participants consistently gave answers above Max which differed from the answer according to M^c. There were no correct answers.

Q5 is identical to Q1 except that the base rate is 20% (rather than "a very small proportion"). As noted earlier, when the base rate is equal to ϕ_1, the correct answer is ϕ_2. Nonetheless, all four formulae were used in this question. About 17% gave the right answer but it is unclear whether it was for the right reason or was based on the Max formula. About 13% gave the base-rate probability (20%) as their answer. One-sixth of the participants chose an answer strictly above both ϕ_1 and ϕ_2, although in this question any answer other than 70% is qualitatively incorrect.

	"correct"	n (MRT)	M	20%	Avg	Max	$> Max$ (M^c)	*Other*
Q5	17%	94 (121s)	10%	13%	9%	17%	16% (3%)	36%

Table 2: Results for Q5.

6. Summary

Individuals often use non-Bayesian procedures to process multiple signals based on simple intuitive formulae. We showed that replacing Bayesian agents with agents who follow one of these alternatives can have significant effects on the equilibrium analysis in standard economic models.

Nonetheless, one of the main conclusions which was mentioned already in the introduction, is a practical one: when processing conditional independent signals using the Bayesian paradigm, two relatively weak signals that a rare state has been realized can be combined to produce very strong evidence.

The fact that people fail to recognize that two moderate and conditionally independent signals of a rare event aggregate to significant evidence for the event, was confirmed not only in our survey but also in several discussions with colleagues familiar with probability theory, many of whom – like us – were surprised by this feature of Bayesian updating. This accidental evidence was the main motivation for the paper.

References

Bar-Hillel, Maya (1980). “The base-rate fallacy in probability judgments”. *Acta Psychologica, 44*(3), 211-233.

Benjamin, Daniel J. (2019). “Errors in Probabilistic Reasoning and Judgment Biases.” Chapter for the Handbook of Behavioral Economics (edited by Doug Bernheim, Stefano DellaVigna, and David Laibson). Elsevier Press.

Budescu, David V. and Hsiu-Ting Yu (2006). “To Bayes or not to Bayes? A Comparison of Two Classes of Models of Information Aggregation”. *Decision Analysis, 3(3)*, 145-162.

Duggan, John and César Martinelli (2001). “A Bayesian model of voting in juries”. *Games and Economic Behavior, 37*(2), 259-294.

Esponda, Ignacio and Emanuel Vespa (2014). “Hypothetical Thinking and Information Extraction in the Laboratory”. *American Economic Journal: Microeconomics, 6,* 180-202.

Eyster, Erik and Matthew Rabin (2005). “Cursed Equilibrium”. *Econometrica*, 73(5), 1623-1672.

Gigerenzer, Gerd and Wolfgang Gaissmaier (2011). “Heuristic Decision Making”. *The Annual Review of Psychology, 62*, 451-482.

Kahneman, Daniel and Amos Tversky (1973). “On the psychology of prediction”. *Psychological Review, 80*(4), 237-251.

Milgrom, Paul R. and Robert J. Weber (1982). “A Theory of Auctions and Competitive Bidding”. *Econometrica 50(5)*, 1089-1122.

Porat, Ariel and Eric Posner (2012). “Aggregation and law”. *Yale Law Journal*, 122(1), 2-69.