Information Aggregation

diego aranega

CHAPTER 6

Information Aggregation

1. A common theme in economic theory is that markets aggregate information well: the price reflects all the information available to the agents participating in the market. The result is surprising. Given the number of dimensions on which uncertainty could bear, how can a price mechanism, a one-dimensional instrument, achieve this *tour de force*? The truth is that in general it cannot.

2. Information aggregation is a challenge because in modeling a market, one assumes that there is no direct information sharing between participants. Participants make bids, based on possibly poor estimates of the value of the item for sale, and the transaction price is determined by market rules as a function of bids. How, then, can information transmission come about? We explain below the logic behind information aggregation. Following this, we shall explain the logic behind information aggregation in markets, and in nonmarket mechanisms such as voting.

3. *An urn example.* Think, first, of an urn filled with coins of various sizes and colors. Your objective is to determine the total value of the coins present in the urn, and come up with a point estimate of that value. Your problem is that you cannot be sure of the number of coins, nor of the nominal value of each coin.

An object for sale is like this urn. You don't know with certainty all its characteristics. Once you acquire the object, you learn its true value but, before that, you have only an estimate of that value. That estimate may be overly optimistic or overly pessimistic.

Now imagine a group of people, all staring at the same urn and willing to guess the total value of the coins in the urn. Also imagine that you are told that, among all participants, you have the most optimistic estimate. This information is likely to undermine your faith in your estimate, and it is plausible that you revise it to a more conservative figure. And presumably, one can expect that the larger the group, the larger the change in your estimate.

Information aggregation is precisely about this adjustment. Others have estimates or perceptions that are related to some degree to the actual value of the urn. Being told something about others' estimates could, in principle, improve mine.

4. A critical question is the extent of the adjustment. If I adjust too much, I may end up being too pessimistic. If I don't adjust enough, I may remain overly optimistic. The literature on information aggregation generally portrays agents who misperceive the value of an asset, but who manage to perfectly undo their initial misperceptions when confronted with information about others' estimates. This information about others' estimates may differ depending on context (the number of assets for sale, the number of buyers, the mechanism by which assets are sold). But across these variants, the objective is the same – to explain or show that markets lead to the asset being sold at a price that is close to the true value of the asset, despite the fact that each agent starts with a misperception of that value. We explain below the logic of these papers.

5. *A stylized illustration.* We denote by v the true value of the urn. We consider n agents staring at the urn. We model ignorance about v by assuming that each one gets a perception z_i of that value, and by defining a particular joint distribution ω over the value v and the perceptions z_i. Specifically, let us assume that

$$z_i = \eta_i v$$

where v is drawn from an exponential distribution, and where each η_i is an estimation factor drawn independently from the same distribution having positive support on $[0.9, 1.1]$. Roughly, this means that agents make an estimation error, but this error, when positive, cannot exceed 10 percent of the value of the urn.

Imagine now that n is very large, that you have the highest estimate, and that you are informed of this. You are asked your best prediction of v, based on your estimate z. Call b that best prediction.[1] Since there are many agents, the estimate of the most optimistic agent exceeds v by a factor that is, most likely, very close to 1.1: that is, $z \simeq 1.1v$. If told that you have the highest estimate, then, knowing the structure of the model, you can infer that you are very likely to be optimistic by a factor close to 1.1, so your best prediction is:

$$b \simeq \frac{z}{1.1} \simeq \frac{1.1v}{1.1} = v$$

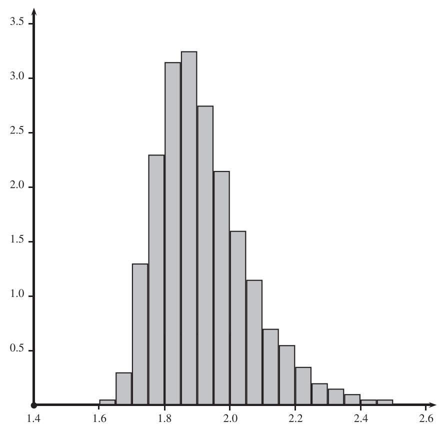
Thus, being told that you are most optimistic among a large group of people permits you to almost perfectly infer the error in your perception. You end up with an almost perfect prediction of the value.

[1] Call I the event where you are the most optimistic agent. For an agent who fully exploits the structure of the model, the best prediction is by definition the number that minimizes $E((b-v)^2 \mid z, I)$, or equivalently, $b = E(v \mid z, I)$.

The example is a caricature, but it explains the process of information aggregation in models. That one should take his initial estimate with caution sounds right. The degree to which one ought to shade one's estimate would seem to be a difficult matter in practice. Knowing the structure of the model, the task becomes trivial: you may infer almost exactly your initial estimation error.

6. *Unbounded errors.* Consider an example that is less of a caricature, in which errors are possibly unbounded. For example, assume that η_i is lognormally distributed. Then, being told that you are the most optimistic bidder does not permit you to perfectly infer the size of the estimation error.

Consider for example the case where $\log \eta_i$ is distributed normally with expectation 0 and standard deviation $\sigma = 0.2$. With a thousand participants, the maximum estimate may vary substantially. The following graph plots the empirical distribution of the largest estimation factor $\eta^0 = \max_i \eta$, obtained through $10,000$ independent trials, each involving 1,000 participants.

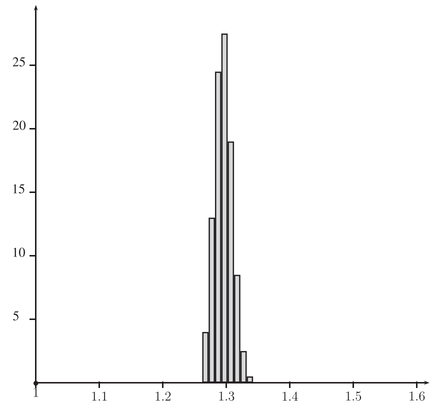


On average across these trials, the largest factor is 1.92, and the standard deviation is 0.14. This means that when you are told that your estimate is highest, you should adjust your initial estimate downward, dividing it by a factor close to 2. The resulting estimate, however, remains a noisy approximation of value, because the distribution over the highest factor is not concentrated on a particular number.

Assume, now, that you are told that among all participants, your estimate is not the highest, but that it coincides with the 10^{th} percentile: only 10 percent of bidders have an estimate higher than yours.

Then, this information, jointly with the knowledge of the distribution over errors, permits you to almost perfectly infer the size of your error when the number of participants is large. Call $\eta^{0.1}$ the estimation factor of the participant for whom 10 percent of participants have a larger estimation factor than his. For a population of $1,000$ participants, this means that his estimate is ranked

101^{st}. The following figure shows the empirical distribution of $\eta^{0.1}$ (obtained through 10,000 independent trials, each involving 1,000 participants).



The graph shows that $\eta^{0.1}$ is much more concentrated than $\overline{\eta}$. The average value of $\eta^{0.1}$ is 1.29, and the standard deviation is 0.014.

Of course, the conclusions above draw heavily on the presumption that the agent knows the distribution from which estimation errors are drawn. Different values for the variance σ^2 would generate quite different predictions for the factor $\eta^{0.1}$. With $\sigma = 0.3$, the average value of $\eta^{0.1}$ increases to 1.47.

Information aggregation works insofar as the distribution over errors is known with precision. With uncertainty about the distribution over errors, being told about the ranking of one's estimate does not pin down with precision the size of one's error.

7. *Markets.* In a market, nobody tells you how your estimate compares with others', if it is the highest, or if it coincides with the 10^{th} percentile. However, with bidders that differ only in their estimate z_i but are otherwise identical, one expects that higher estimates will translate into higher bids. So if there is a single good for sale, we expect the most optimistic bidder to win. And if there are 130 identical objects being sold to 1,300 bidders, each willing to acquire only one object, we expect the bidders having the highest 130 estimates to win.[2] In particular, the least optimistic bidder among those who win has an estimate that coincides with the 10^{th} percentile.

One contribution of theoretical models has been to prove that for an agent bidding in a uniform auction,[3] it is optimal for him to behave as if he was pivotal, that is, as if he was winning by a zero margin. The reason is that in

[2] This is a simple market setup inspired by Pesendorfer and Swinkels (2000). In practice, and for most markets, agents would be allowed to bid for more than one unit.

[3] In a uniform auction, all winning bidders pay the same price, equal to the highest bid among losers.

other events, small changes to his bid are of no consequence to his chance of winning or losing, nor to the price he pays in the event he wins.

This means that he should behave as if he was approximately the least optimistic among the winners (or most optimistic among the losers). With many participants, this precisely pins down how he ought to correct his initial estimate: when he knows the distribution over perception errors (his own and that of others), being the least optimistic among the winners tells with precision the size of the mistake that he makes.

To illustrate simply why transaction prices must be close to v in a uniform auction, assume that each potential bidder i chooses to bid a fraction λ_i of his estimate:

$$b_i = \lambda_i z_i.$$

Look for a symmetric equilibrium in which all bidders use the same shading factor λ^*.[4] With 130 goods sold to 1,300 participants, the price must settle on the 131^{st} largest bid, that is, on a bid equal to $\lambda^* \eta^{0.1} v$. If $\lambda^* \eta^{0.1}$ happens to be above 1 for most realizations, winners pay too much, providing each bidder an incentive to forgo participating in the auction or to decrease λ_i to win less often. If the quantity $\lambda^* \eta^{0.1}$ happens to be below 1 for most realizations, winners pay too little, and some bidders will find that a larger λ_i profitably increases the chance of winning. With a large number of bidders, $\eta^{0.1}$ is concentrated on the expectation $E\eta^{0,1}$ and, in a symmetric equilibrium, it must be that:

$$\lambda^* \simeq \frac{1}{E\eta^{0.1}},$$

and the price, equal to $\lambda^* \eta^{0.1} v$, must thus be concentrated on v.

8. *Partial versus complete information about estimates.* The information revealed in a price mechanism is partial: bidders don't have access to the estimate that *each particular participant* receives. Why isn't this an issue?

This is not an issue because by construction, we built a model in which there is a stable relationship (i.e., "almost one-to-one") between the value of the asset v and the μ_0-largest estimate, for $\mu_0 = 0.1$. The inference that one makes from observing the μ_0-largest estimate is thus perfect, and observing the whole empirical distribution over estimates (which amounts to observing the μ-largest estimate for all values of μ) cannot further improve this inference. Many factors can alter this conclusion. We mention two.

Variations in dispersion of errors. Depending on the characteristics of the asset, the variance of the error term might differ substantially. Agents could,

[4] We only provide below a necessary condition on λ^*. Proving that such a symmetric equilibrium exists is more involved. We shall come back to this issue in Chapter 16.

of course, be aware of that possibility. But having a precise idea of these variations across assets seems unrealistic.

Heterogeneity. Gauging the quality of one's estimate and how it compares to others' is a second key issue in markets. The model circumvents these difficulties by assuming symmetric agents. Departing from that benchmark likely creates further difficulties, as agents' perceptions of the characteristics of the asset may be correlated, partially driven by their own reading of these characteristics, with variations in readings across assets generating further sources of dispersion and bias of the estimates.

9. *Aggregating many versus few estimates.* Aggregating two expert opinions is difficult. If I ask a young French demographer and a renowned Congolese historian about life expectancy in Congo, I might get two quite different responses. Deciding how to weigh these expert opinions is a challenge, because the quality or bias of any given expert opinion is difficult to assess with precision. Why would this issue mysteriously disappear once we move to a large number of opinions, expert or not. Do we believe that getting a million opinions would improve our assessment of life expectancy in Congo?

The same issue arises in markets. For a given asset, some experts are presumably better qualified than others at estimating its value. I care about whether those participating actively in the market are those actually getting better estimates. And I also care about the fraction of optimistic bidders participating in the market, as their participation may affect prices. Do we believe that as the number of participants increases, the fraction of optimistic bidders will remain stable across different kinds of sales?

10. *The wisdom of crowds.* Notwithstanding previous comments, the belief that large crowds make it easy to get to the truth is widespread, and it is often referred to as the *wisdom of the crowd*. The usual story behind the idea stems from a single field experiment by Francis Galton at a cattle market. Galton asked participants to estimate the weight of a given cow. Most quotes were off track by a substantial amount. However, the median quote got it (almost) right.

One could take the lack of replication of the Galton experiment as suggesting that the wisdom of crowds is perhaps not so prevalent. Of course, it may simply be that few have attempted to replicate it. As always, inference is difficult.

11. *Voting under majority rule.* An implication of the wisdom-of-crowds logic is that majority voting aggregates information nicely. Imagine that a group considers the possible adoption of a new alternative A, comparing it to the status quo S. All group members are assumed to have the same preferences: the problem is not one of aggregating preferences, but of aggregating information.[5]

[5] This is a simple voting setup inspired by Feddersen and Pesendorfer (2000).

Define u as a measure of the difference in value between A and S and let $v = \exp u$. The group would like to select A if and only if u is positive or, equivalently, if and only if v is above 1. However, members are ignorant of the exact value of v. We model this by assuming that v is a random variable, and that each member gets a noisy estimate of v:[6]

$$z_i = \eta_i v.$$

This means that when η_i is above 1, the member is optimistic, and when η_i is below 1, he is pessimistic.

This simple model illustrates what is behind the wisdom of crowds. If each η_i is independently drawn from the same distribution and if $\Pr(\eta_i > 1) = 1/2$, a large group will always get it right. In a large group, optimists (i.e., $\eta_i > 1$) and pessimists (i.e., $\eta_i < 1$) are in almost equal numbers. Whenever A is better ($v > 1$), not only the optimists vote for A, but also the mildly pessimistic as well, and the alternative A gets majority support.

The wisdom-of-crowds hypothesis may seem suspicious, as one can think of many reasons why crowds' judgements would be subject to systematic biases. Nevertheless, the typical voting model in economics goes one step beyond. It portrays agents who behave as though they knew with precision the distribution over the errors that each member makes. We examine the consequence of this assumption below.

12. *Voting under qualified majority rules.* We now have in mind that the new alternative A passes if and only if it gets sufficient support. "Sufficient" can mean different things. We consider qualified majority rules: A passes if a fraction q of the population votes for it.

At first glance, a more stringent majority rule should make alternative A more difficult to pass, because it then takes fewer people to reject A. The contribution of economics models has been to point out that individuals adjust (or should adjust) their decisions to the qualified majority rule considered: somebody willing to support S under the majority rule might be less inclined to do so if a very large support is needed to get A through. A negative vote carries more weight, and one may be reluctant to cast a negative vote unless the evidence that one receives against A is strong enough.

The problem is similar to the urn question. Think of the case where unanimity for A is required, so that any individual has the option to veto A through a single vote for S. I might have received evidence against A, but if all others received evidence more favorable to A than I did, I may want to reconsider my position, and avoid vetoing A.

As for the urn, the difficulty is the degree to which one should adjust the decision. When will I think that my evidence against A is strong enough that it justifies voting for S? *What exactly does strong enough mean*?

[6] An alternative formulation could be $z_i = u + \varepsilon_i$. We adopt the multiplicative formulation to make the comparison with earlier formalizations easier.

Economic models provide a clear-cut answer. A sophisticated voter should adjust his behavior to the voting mechanism, tailoring the adjustment to the distributions over errors. The optimal adjustment induces perfect information aggregation: A passes if and only if it is worthwhile, irrespective of the particular voting rule considered.

13. *A formal illustration.* Consider the qualified majority rule q in which adopting A requires support by at least a fraction q of voters (with $q < 1/2$). Assume agent i supports A if and only if his signal is strong enough, that is:

$$z_i > \lambda_i.$$

The threshold λ_i characterizes agent i's strategy, and one can compute the expected gains associated with the strategies $(\lambda_1, \ldots, \lambda_n)$ and look for a symmetric equilibrium λ^*. The insight that the voting rule affects voting behavior will be reflected in the fact that the equilibrium threshold λ^* decreases when the majority requirement q increases.

Assume all voters other than i use the same strategy λ^*. To voter i, the only event he cares about is when he is pivotal, that is, when his choice determines the outcome. This happens when among all other voters, exactly $k-1$ vote for A, where k satisfies $(k-1)/n < q < k/n$, which roughly means that the q-highest estimate is close to λ^*, hence that $v\eta^q \simeq \lambda^*$.[7]

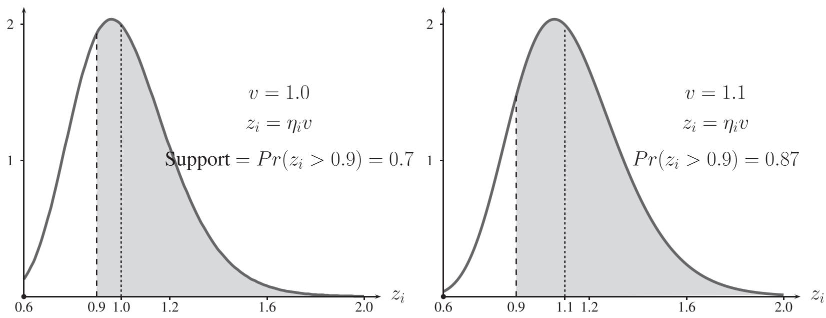
With a large number of voters, η^q is concentrated on the scalar $\overline{\eta}^q$ for which $\Pr(\eta_i > \overline{\eta}^q) \equiv q$. So in the event that i is pivotal, one may infer the value of v almost perfectly: $v \simeq \lambda^*/\overline{\eta}^q$. If voter i infers that v almost certainly exceeds 1, he should vote for A (unless z_i is extremely low). If he infers $v < 1$, he should vote against A (unless z_i is extremely high). The only case when his optimal threshold is not extreme is when the inference is $v \simeq 1$, hence equilibrium requires $\lambda^* \simeq \overline{\eta}^q$, which decreases when q increases.

To illustrate this numerically, assume $\log \eta_i$ is normally distributed with standard deviation equal to $\sigma = 0.2$. When the majority rule is 1/2, the equilibrium threshold λ^* is 1. When the necessary majority q increases to 0.7, the equilibrium threshold λ^* decreases to 0.9.

The graphs on the next page illustrate the latter case (i.e., $q = 0.7$ and $\lambda^* = 0.9$), showing the extent of actual support when $v = 1$ and when $v = 1.1$. In these graphs, the curve defines the density over estimates when $v = 1$ (on the left) and when $v = 1.1$ (on the right).[8] All group members getting an estimate above $\lambda^* = 0.9$ vote in favor of A. The shaded area indicates the probability that a given individual gets a draw above 0.9. With large numbers, the support for A approximately coincides with that probability. The threshold λ^* is precisely set so that when $v = 1$, support exactly coincides with 0.7. When $v = 1.1$, support is 0.87, above the required majority 0.7.

[7] Recall that η^q designates the estimation factor of the participant for whom q percent of participants have a larger estimation factor than his.

[8] In both cases, v is the median estimate – half of the voters get an estimate above v.



14. Perfect information aggregation across various voting systems can be seen as an embellishment of the wisdom of crowds. As for the urn and markets, it builds on the assumption that not only the fraction of optimists remains the same across instances, but also that, for any given x, the degree of optimism of the x-percent-most-optimistic agents remains stable as well.

Rephrasing this in terms of Galton's cow experiment, one could tell a farmer the following:

"You know what, among all the many estimates that I have received, there are exactly 13 percent of estimates that are more optimistic than yours…";

With the farmer soon replying:

"Ah, okay, my mistake. I know what the correct figure is then. Thanks."

15. While much emphasis is generally put on the performance of institutions in aggregating information, the literature recognizes that when uncertainty bears on more than one dimension, a price or voting mechanism will often not aggregate information well, opening the path for understanding how performance differs across institutions.

One interpretation of our critique is that information aggregation results hinge on the presumption that uncertainty bears on a single dimension. In general however, there is no reason to think that uncertainty would only bear on one dimension. For mathematical convenience, we describe precisely the distribution from which a given variable is drawn. Once we recognize that there is no reason that this distribution should be known, we need to introduce the possibility that uncertainty bears on more than one dimension, and as a result, institutions will generally fail to aggregate information well.

Further Comments

Galton. *We referred to Galton's cow experiment*[9] *because it sometimes lends support to the idea that the median opinion would be the wisest, closest to the*

[9] Galton (1907).

truth. Note, however, that Galton's main interest was not to elicit some "truth," but to improve how a jury would assess damages, with the objective of finding a method that would select the fairest compromise between necessarily diverse opinions.[10] *His idea was that a fair compromise would be better achieved by eliciting the median estimate among jury members, rather than by asking each jury member his estimate and computing the average.*

Part of his objection was that reporting a precise figure for damages is too difficult,[11] *and interestingly, Galton proposed a method to elicit the median estimate which does not rely on agents reporting such a precise figure. Based on the presumption that estimates are normally distributed around the median, he advocates choosing two numbers x and $y > x$, and asking jury members to report whether they think damages should be below x, between x and y or above y. Fitting these answers to a normal distribution, one can elicit the median estimate (simultaneously with the dispersion of opinions).*

Condorcet*. Formal analysis comparing voting institutions, with the aim of producing better decisions (say judicial ones) goes back at least to Condorcet. In relation to the ideas developed in this chapter, and the issue as to whether larger crowds would help, Condorcet comments that larger consent that a proposition is true may only prove this fact (that the consent is larger), without bringing new certitude to the truth of the proposition itself.*[12]

Market efficiency hypothesis*. In the 1970's, with the development of information economics, a number of authors addressed the question of whether markets could be efficient, with prices perfectly revealing the value of the underlying assets. The issue often takes the form of analyzing a stylized market composed of a set of identically informed and a set of identically uninformed agents, and determining whether prices reflect the information held by informed agents.*[13] *In these stylized models, price is a simple function*

[10] With respect to the question of finding "the most suitable amount of money to be granted for any particular purpose" [such as compensation of damages], Galton (1899, page 638) asks: "How is the medium amount to be ascertained, which is the fairest compromise between many different opinions?" So it is not the "true" level of compensation but the fairest, that one is looking for.

[11] His other objection to using the average opinion is that this leaves undue influence or voting power to those coming up with absurdly large or small ones.

[12] Condorcet, Essai sur l'application de l'analyse à la probabilité des décisions (page 148 (cxlviii): "le jugement de tous les hommes, en faveur de cette proposition, ne peut produire une probabilité plus grande [...] Si ceux qui y croient se bornent aux preuves données avant eux, et n'en cherchent pas de nouvelles, leur consentement, en supposant qu'il put produire une certitude, prouverait seulement qu'il est certain que ce fait, que cette proposition sont probables."
[The [unanimous] judgment of all men in favor of this proposal cannot produce a larger probability [...] If all that believe in it restrict attention to the existing evidence, without looking for new ones, their [unanimous] approval, if it ever were to produce a certitude, could only prove one, that it is certain that the proposal is probable].

[13] See, for example, Kihlstrom and Mirman (1975) and Green (1973).

of the information held by the informed traders, and of some other parameter reflecting "market conditions." To the extent that this extraneous market parameter remains unknown to agents, prices cannot reveal perfectly the information held by informed traders. Grossman and Stiglitz (1980) pointed out that the existence of some residual uncertainty is critical to the functioning of markets, because otherwise no trader would have an incentive to pay the cost of acquiring information in the first place.

In this chapter, we also ask whether prices can reflect the underlying value of the asset in the presence of a population of diversely informed agents. Our emphasis, however, is not on the theoretical impossibility of markets that perfectly reveal prices, but on the practical difficulty/impossibility of determining what to make of each and everyone's opinion even if it were freely available. Said differently, "market conditions" include the particular way that information and opinions are dispersed across agents, and one suspects that these market conditions (hence prices) are subject to shocks unrelated to underlying values.

References

Condorcet, N. de (1785). *Essai sur l'application de l'analyse à la probabilité des décisions*, Paris.

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

Galton, F. (1899). The median estimate. *Report of the British Association for the Advancement of Science*, 638–640.

Galton, F. (1907). Vox Populi, *Nature*, 1949(75), 450–451.

Green, J. (1973). Information, efficiency, and equilibrium. Discussion Paper No. 284, Harvard: Harvard Institute of Economic Research.

Grossman, S. J. and Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, 70(3), 393–408.

Kihlstrom, R. E. and Mirman, L. J. (1975). Information and market equilibrium. *The Bell Journal of Economics*, 6(1), 357–376.

Pesendorfer, W. and Swinkels, J. M. (1997). The loser's curse and information aggregation in common value auctions. *Econometrica*, 65(6), 1247–1281.