An introduction to Prediction Markets[1]

Luca Iandoli
University of Naples Federico II, Italy

Abstract. Markets can be seen as a good example of wisdom of the crowd. In the following notes I provide some ground for this perspective. Furthermore I present Prediction Markets as the current technology that can be used to leverage the "wisdom of the crowds" and their applications for decision making.

1. Can groups be more intelligent than single individuals?

Under certain assumptions, groups can be smarter than their average member. Since the definition of what "smarter" means depends on the task to be carried out or the problem to be solved, in this paper we will focus our attention on one specific task: the prediction of future events. The accuracy of collective predictions is surprisingly high. Sometimes large random groups of people can be even more accurate than small teams of experts. Usually the following assumptions must be satisfied for collective predictions to be accurate (Surowiecki, 2004):

1) people must have diverse predictive models;
2) they are independent (they are not allowed to influence each other or do so limitedly);
3) prediction process is decentralized.

Example of wise crowds - class experiment: guess my weight and height.

Why groups are accurate in collective prediction? Three possible models for information aggregation:

a) the competent minority ("who wants to be a millionaire" example) - canceling out noise: while incompetent guesses produce random errors that compensate each other, true information is biased and so it emerges;
b) aggregating pieces: the solution is the assembling of partially correct answers.

To work properly all the above models assume that: i) at least some people have (pieces of) correct information; ii) if they are not informed they can seek for information, under adequate incentives; iii) people are willing to disclose truthful information. Implicitly all the models assume some degree of diversity among people in the group. Implicitly all models assume some degree of accuracy for a crowd to be wise.
An apparent paradox is why diversity leads to convergence to the best solution.

Social learning Example: designing pedestrian paths in Harvard courtyard, ants' systems, …

The explanation is that diversity helps production of many more good ideas. For diversity to work, however, there must be some selection criteria against which the degree of fitness of each solution

[1] These notes are a summary of different sources and have been created for class use only. Do not use or quote for other purposes without the author's permission.

can be tested. Natural evolution is a good example: sexual reproduction and casual mutations generate variations, then the environment selects the fittest. Competitive markets work very much the same way: many different versions of a same product are proposed by companies, customers select the ones who make a better fit with their needs. Selection of the best alternative is also helped in this phase by imitation, provided that imitation follows competent choice and does not happen prematurely. In other words imitation is a very efficient mechanism for social learning because it saves us the effort of resolving a problem each time from scratch by pigging back on the efforts made by *others we trust*. On the other hand too much or too early imitation does not help because it suppresses diversity.

Scientists and mathematicians have proposed some theoretical explanations of why diversity works. Probably one of the oldest and most notable results is Condorcet's theorem (1785). Put simply the theorem states that the probability for a crowd to make an accurate prediction through majority voting approaches 1 as the number of individuals n in the crowd increases, *provided that each individual in the crowd is more likely to be right than wrong (he/she votes for the correct answer with probability higher than .5)*. The theorem also says that when this individual probability is less than .5 the probability of the crowd to be *wrong* approaches 1 as n increases. In other words, large and enough informed crowds can be very smart, but large and incompetent crowds are really a disaster.
There is a very important assumption behind the theorem: individual guesses must be produced in an independent way, i.e. individuals are not supposed influence each other (e.g., by exchanging information or collaborating in any way).

The diversity prediction theorem elaborated by Scott Page shows that collective wisdom is always a combination of individual accuracy and diversity. In other words we need diverse predictive models but this diversity is leveraged only if people in the crowd are smart enough. The theorem states that:

$$\textit{Collective Error} = \textit{Average Individual Error} - \textit{Prediction Diversity}$$

If AIE is not small we need to counterbalance it with some diversity. If people are all perfectly accurate, they do not need to be any diverse (diversity = 0, but Individual error = 0). This result is useful because it says that we can compensate ignorance to some extent by adding diversity. Little, good diversity always helps in this case because, being PD ≥ 0 and AIE does not increase, thus

$$\textit{Collective Error} < \textit{Average Individual Error}$$

A metaphor can provide another interpretation of the above results and show why diversity *and* ability need to work together: the search of food by ants (which in general is a metaphor for the search of solution in a complex decision spaces). Diversity allows for the increase in amount of space around the nest being explored because diverse ants will likely move along diverse directions (in the solution space formulation these correspond to the number of alternative dimensions of the problem and to the different complementary solutions strategy exhibited by different problem solvers). Ability allows for: i) ants that know where to go (proper identification of relevant dimensions in the solution space, e.g. north/south and east-west coordinates); ii) ants that know how to search (i.e., they are able to collect & filter relevant information and can recognize effective solutions – food in the metaphor - when they see it).

2. When crowds are not smart?

When people are allowed to communicate freely we have a dilemma: collaboration can lead to the improvement of models through open criticism as well as to undesirable results like amplification of conflict or herding behavior. Common group pitfalls are:

a) *Balkanization and polarization*: crowds tend to disarticulate in groups holding opposite beliefs (balkanize); the members of a faction of like-minded people tend to develop even more extreme opinions than the one they held individually prior to joining the group; For instance, people holding anti-abortion opinion tend to become even more pro-life after they engage in a discussion with other people having the same belief.
b) *Information cascades*: information propagates quickly in social network by word of mouth and imitation without being evaluated. Some research models equate this propagation to virus contagion. If the information being diffused is wrong the crowd will dramatically increase the speed through which incorrect facts and beliefs propagate. Again, this is typically the result of low diversity and lack of independence.
c) *Hidden profiles*: many studies have confirmed that individuals may not share all the information they have when they are in a group, but instead they will focus on the information that group members already have in common.

Interestingly, polarization, imitation, and hidden profiles all imply a reduction of diversity and are the result of the natural human tendency to conform to increase group cohesion and reduce risk of social marginalization.
Diversity, however, may not be beneficial at all if it is about ends and not means (Page). Diversity about ends indicates that people differ with respect to what is valuable, fair, and desirable for them or for the group. People with different values may have incompatible visions of given problems. Usually this kind of diversity brings to conflicts and polarization. Public discourse on sensitive topics such legal abortion, death penalty, gun control, and the like are examples of such value-based debates.
Diversity about means implies that different people agree about the ends but use different problem solving strategies to achieve the same objectives. Diversity about means is the kind of diversity that has to be nurtured. Diversity about means also subsumes that people end up using a shared language and perhaps have some similar previous experience about a problem. We could find desirable to know what marketing people think of a product and have them cooperate with production people, while it does not make much sense in most of the cases to have a biologist and an expert of ancient Greece literature to solve a problem in the area of biochemistry (however, it makes sense apparently that someone who is passionate in calligraphy can become the founder of one of the most successful computer companies! The reference is to Steve Jobs and how his passion for calligraphy led him to design fonts and ultimately the WYSIWYG interface that made Apple computers stand out from the competition).

Definition of Prediction markets

Prediction markets are tools that leverage the wisdom of the crowds. They come through virtual trading software platforms and are used in distributed networks of decision makers, sometimes even in the form of public markets on the Internet (see tradesport, holliwood stock exchange, ideafutures, inkling, University of Iowa Prediction markets, …).
PMs are a special type of market, often known as “information market” or “event futures.” Analytically, these are markets where participants trade in contracts whose payoff depends on unknown future events.

How they work

Substantially, as real markets do. In a prediction market, payoffs are tied to the outcomes of future events. Participants trade contracts associated to the occurrence of a given event (e.g. Will a Democrat President be elected in 2020? Who will win the UEFA Champions league? Will Italian GDP increase over 1.5% in 2018?). The market exchange of contracts determines their price: in general, the higher the price of a contract, the higher the confidence of the market in the future occurrence of the associated event. Participants trade with real or virtual money; in any case they bet on the outcome they think is more likely (this does not exclude speculation attempts if some traders have hidden information or are just experienced traders whose objective is to make money through speculation).
In general 3 types of contract are available (see following table): winner-take-all, index and spread.

Table 1: Contract Types—Estimating Uncertain Quantities or Probabilities

Contract	Example	Details	Reveals market expectation of…
Winner-takes-all	Event y: Al Gore wins the popular vote	Contract costs \$$p$ Pays \$1 if and only if event y occurs Bid according to value of \$$p$	Probability that event y occurs, $p(y)$
Index	Contract pays \$1 for every percentage point of the popular vote won by Al Gore	Contract pays \$$y$.	Mean value of outcome y: $E[y]$
Spread	Contract pays even money if Gore wins more than y^*% of the popular vote.	Contract costs \$1 Pays \$2 if $y>y^*$ Pays \$0 otherwise. Bid according to the value of y^*.	Median value of y.

In most prediction markets, the mechanism that matches buyers to sellers is a continuous double auction, with buyers submitting bids and sellers submitting prices, and with the mechanism executing a trade whenever the two sides of the market reach a mutually agreeable price.

Why they work

For several reasons:

Efficient market hypothesis. Much of the enthusiasm for prediction markets derives from the **efficient markets hypothesis**. In a truly efficient market, the market price will be the best predictor of the event, and no combination of available polls or other information can be used to improve on the market-generated forecasts. This statement does not require that all individuals in a market be rational, as long as the marginal trade in the market is motivated by rational trading. Of course, it is unlikely that prediction markets are perfectly efficient, but a number of successes in these markets, both with regard to public events like presidential elections and within firms, have generated substantial interest.

Garbage-in/Garbage-out: the role of information. Another theoretical justification comes from studies on group decision-making and social choice. The well known Condorcet's theorem states that the probability a group of *n* individuals arrives at a correct decision increases with *n* if single voters are correct with a probability $p > ½$ (if instead $p < ½$ the probability of a wrong decision approaches 1 when *n* increase). If $p > ½$, the higher p the lower the number of people we need for a correct decisions.
Condorcet theorem and efficient market hypothesis have a lot in common: both are based on rationality and independence of participants. Both require large numbers to work properly (the market needs liquidity through a high number of traders since only a minority of participant usually drive the market). Both require availability of information ($p > ½$ means that voters are more likely to be informed on what is the right outcome than not). However the market adds on the top of Condorcet theory a monetary incentive: when people risk their money and can gain profits through betting, they will have an incentive to collect good information. The incentive will cause p to grow. In other words, the power of prediction markets derives from the fact that they provide incentives for truthful revelation, research and information discovery, while the market provides an "algorithm" to aggregate opinions. As such, these markets are unlikely to perform well when there is little useful intelligence to aggregate or when public information is selective, inaccurate, biased, or misleading (e.g. through deliberate manipulation).

Uncertainty and diversity. Trade also requires some disagreement about likely outcomes. Disagreement is unlikely among fully rational traders with common priors. It is more likely when traders are overconfident in the quality of their private information or their ability to process public information or when they have priors that are sufficiently different to allow them to agree to disagree. Unlike polls, markets do not give the same weight to each person, but it's like they assign more weights to predictions made by people who are willing to bet more on an outcome. Those are usually the people that have more information and competence, while badly informed forecasters will be ruled out by the fear to loose their money. In this sense the market improves accuracy since it favors a sort of natural selection of the best predictive models.
But market incentives may also encourage diversity when the solution is not so obvious. The nature of market reward is such that traders will gain a lot when they bet on unexpected outcomes: if the correct answer is obvious and everybody knows it, all the bets will go in the same direction and the payoff will be virtually zero. If instead we have uncertainty and some buyers are more confident than others in a surprise and eventually they are right, they will get a very high payoff. Unfortunately this is also the kind of reasoning that may trigger speculative behaviors (buy low to sell high), but speculation is less likely to occur in prediction markets and in any case it tends to have a short duration.
So the market provides both incentives for accuracy and for diversity.
Let's quickly compare prediction markets with some alternative decision support systems: polls and experts. Polls assign the same weight to every respondents and, by doing so, do not filter out bad information. Expert-based methods do the opposite: they over rely on the knowledge of a few selected people by concentrating weight only on these subjects (even just one in some case). One expert is more likely to be accurate than the average person but, by definition, it is not diverse from herself. Following Page's theorem we miss all the benefits coming form diversity. So if experts are rights, we don't need diversity, but if they are wrong we have a recipe for disaster.
So, it's fair to say that PMs are somewhere in the middle in the continuum between polls and the expert.

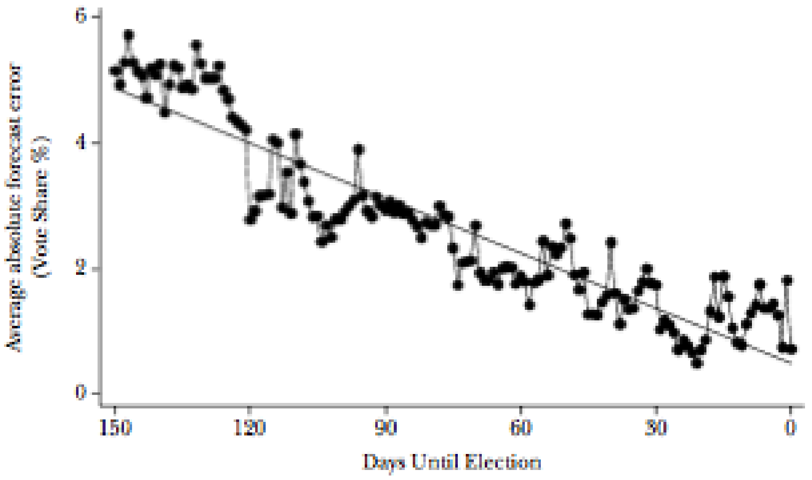
Applications

Examples:

- Public markets: Forecasting of political events (e.g. elections, decisions), economic statistics, movies successes, sports competitions
- Private companies markets: Increasing use of PM tools in companies for distributed forecasting related to internal or external events that are of interest for the organization: will project P end within the deadline? Will sales of product X be higher than a certain value? Which product should we launch (opinions market)?

From a number of studies there is evidence that the market has both yielded very accurate predictions and also outperformed large-scale polling organizations (see in the political domain, Berg, Forsythe, Nelson and Reitz (2001)). Accuracy in many cases has proven to be higher than those produced by polls. See fig. 1 for an example: the graph also shows how the accuracy of the market prediction improves as information is revealed and absorbed as the election draws closer, when more information accumulates and is made available for decision making. In contrast polls can vary, but they tend to oscillate randomly following the crowd's mood.

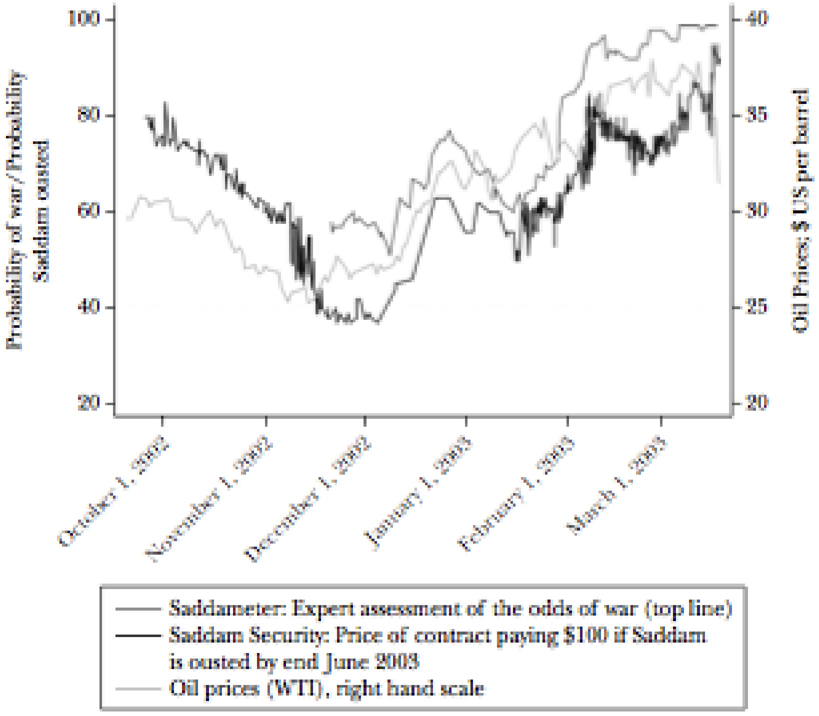
Figure 1
Information Revelation Through Time



(Source Wolfers and Zitzewitz, 2004)

Prediction markets perform well also when compared with expert estimates (see the example in fig. 2 in which experts and market predictions are compared toward the likelihood of a war in Iraq to remove Saddam from power). If we compare prediction performance with costs of polling experts VS setting up a prediction market, in many cases PMs result much more cost effective and quicker. Figure 2 is a good example of how to link the time series of expectations elicited in prediction markets with time series of other variables (inferences). For instance, in Leigh, Wolfers and Zitzewitz (2003), movements in the Saddam Security are interpreted as an index for the risk of war and interpreted the co-movement with the oil price shown in Figure 2 as a causal relationship, concluding that war led to a $10 per barrel increase in oil prices.

Figure 2
The Saddam Security



Source: Trade-by-trade Saddam Security data provided by Tradesports.com; Saddameter from Will Saletan's daily column in Slate.com.

Evidence of performances and low cost of such tools have prompted some companies to adopt PM for business forecasting. A large number of company's employees can be involved in collective forecasts. Usually participants trade with virtual money that is eventually converted in prizes for best forecasters. Forecasts based on PMs in companies however may suffer from less accuracy (like being over-optimistic, see Wolfers and Zitzewitz study at Google's). One explanation is that in a closed group there is a higher chance to violate the assumption about the independence of traders (in other words traders may influence each other or just be less diverse than in an open market).

Futures market in finance are based on the same idea: a future contract is a standardized contract to buy or sell a specific commodity of standardized quality at a certain date in the future and at a market-determined price. Futures are not direct securities like companies stocks, but a type of derivative contract (i.e. title whose value depends on the value of another asset, also called the underlying assets – e.g. oil). Futures can also been applied to non-commodities like currencies or other financial goods. A futures contract gives the holder the obligation to exercise the contract at the established date (e.g. sell the contract). Another kind of contract called "option" instead gives the right to exercise but not the obligation). Futures contracts may be sold or bought before the expiration date on the market.

Do PMs always work well?

Prediction markets do seem to display some of the deviations from perfect rationality that appear in other financial markets. There is substantial evidence from psychology and economics suggesting that people tend to overvalue small probabilities and undervalue near certainties.

Another behavioral bias reflects the tendency of market participants to trade according to their desires, rather than objective probability assessments. Strumpf (2004) provides evidence that certain New York gamblers are more likely to bet on the Yankees.
A further possible limitation of prediction market pricing arises if speculative bubbles drive prices away from likely outcomes.

In addition to the possibility of irrational behavior the main limitation of a market is that it does not give any insight on what is the knowledge traders use behind their market decisions. In other words, a prediction market only aggregates information to tell us which is the probability of occurrence of an event in the future, but it will not say why the event is going to happen. Markets distill knowledge into price, but they do not capture the reasoning behind market behaviors. Sometime this can be obvious, sometime not so trivial. The market may process information effectively but the way this happens is a black box.

Finally markets aggregate the knowledge that is already somehow available to their members; they are not able to support market participants in developing new insights and supporting collective creativity. For that we need interaction, as in a conversation, but we know that there is no guarantee that interaction will lead to positive outcomes, actually, quite the opposite how we discussed above when presenting the typical group decision making pitfalls.

Suggested readings

Scott Page (2007), The difference: how the power of diversity creates better groups, firms, schools and societies. Princeton University Press.
James Surowiecki (2004). The Wisdom of Crowds. New York: Doubleday Press.
Justin Wolfers and Eric Zitzewitz (2004). Prediction Markets. Journal of Economic Perspectives—Volume 18, Number 2, Pages 107–126.
Guido Romeo, Alla Ricerca di Senso, Nova – Il Sole 24 ore