14.770: Introduction to Political Economy
Lecture 3: Voting and Information Aggregation

Daron Acemoglu

MIT

September 13, 2018

Voting for Being Pivotal

- Suppose that voters are strategic — they vote because they think they may be pivotal and are "hyper rational" so that they can understand the likelihood of being so.

- If we have a model of pure redistributive politics with two options, then each voter will vote for the option that maximizes his or her utility (with the usual arguments after ruling out weakly dominated strategies).

- But what if there is also a "common interest" element?

- In this case, each voter would like to maximize his or her utility, but this involves taking into account when he or she will be pivotal conditional on the state. Similar to common value auctions.

- In this lecture, we will develop these ideas formally and then provide some evidence as to whether these kinds of considerations will matter in actual voting decisions.

The Condorcet Jury Theorem

- The first person to think about such issues was again Condorcet.
- Condorcet reasoned about the jury problem, where all jurors have the same interests, and would like to convict a defendant if he is guilty.
- But each has incomplete information (say a signal about the underlying state of nature).
- Condorcet reasoned that if they all pool their information — say by voting *sincerely* — then with a sufficiently large jury, the law of large numbers will kick in and the dispersed information of the jurors will be well aggregated.
- So voting acts as a good way of information aggregation.
- This point was picked up about a century later by Francis Galton, who developed the idea of the "wisdom of the crowd" and provided fascinating evidence consistent with it.

A Modern Jury Problem

- But let's dig a little bit deeper into this (following Fedderson and Pesendorfer, 1998).
- There are n jury members who have to decide whether to convict a defendant.
- There are no conflicts of interest — all jury members would like to convict the defendant if he is guilty, denoted by the underlying state $\theta = G$, but not if he is innocent, $\theta = I$.
- Each jury starts with a common prior that the defendant is guilty with probability $\pi \in (0, 1)$.
- Then receives a signal $s = \{g, i\}$ (for example, from their reading of the evidence presented at the trial). Suppose that the signals are conditionally independent and identically distributed and satisfy

$$\Pr\left(s = g | \theta = G\right) = p \text{ and } \Pr\left(s = i | \theta = I\right) = q; \; q, p > 0.5 \ .$$

Unanimity

- The key assumption is that the jury requires unanimity to reach the verdict of $x = G$.

- Let the vote of juror j be denoted by $v_j \in \{g, i\}$. Then $x = G$ if $v_j = g$ for all j.

- Suppose also that each member j of the group has the following payoff:

$$u_j(x, \theta) = \begin{cases} 0 \text{ if } x = \theta \\ -z \text{ if } x = G \text{ and } \theta = I \\ -(1-z) \text{ if } x = I \text{ and } \theta = G \end{cases}$$

  where $z \in [0; 1]$.

- This in particular implies that convicting an innocent defendant has a higher negative payoff when z is greater (leading to more conservative decisions).

Best Responses

- When will a juror vote to convict?
- Suppose first that the juror expects not to be *pivotal* — meaning that her vote doesn't matter. This will in particular happen when other jurors have already voted to acquit (since the jury requires unanimity to convict). In such cases her vote doesn't matter, so voting $v = G$ has no payoff implications.
- Instead, her vote matters (if and only) if she is pivotal, meaning that all $n - 1$ other jurors have voted to convict.
- In this case, she would like to induce a collective decision (a jury verdict) such that

$$x = I \text{ if } \Pr(\theta = G | \text{information set}) \leq z.$$

- This simply says that given the costs of convicting an innocent, she would only like to convict the defendant if the probability that he is guilty is greater than z.

Optimal Conviction

- To simplify the discussion, let's assume that

$$\Pr(\theta = G | s_j = g \text{ for all } j) = \frac{1}{1 + \left(\frac{1-q}{p}\right)^n \frac{1-\pi}{\pi}} > z \qquad (1)$$

so that when all information is against the defendant and if jurors had access to this information, they would be confident enough to convict him.

Sincere Voting

- Let us now focus on the case where all jurors both "sincerely" and consider the problem of juror 1 who has received signal $s_1 = i$.
- The key objects we need to compute is
  $P_1 = \Pr(\theta = G | s_j = g$ for all $j \neq 1$ and $s_1 = i)$. Why?
- Under sincere voting, this probability is

$$P_1 = \frac{1}{1 + \frac{q}{1-p} \left(\frac{1-q}{p}\right)^{n-1} \frac{1-\pi}{\pi}}.$$

- Does sincere voting make sense?
- First suppose that $P_1 < z$, then together with our above assumption, this condition ensures that sincere voting is an equilibrium (and in some sense the jury system works well). Why?
- But this condition is unlikely to hold together with (1). Now suppose that $P_1 > z$. What happens?

Bayesian Nash Equilibrium

- Let us now understand how the Bayesian-Nash equilibrium works when $P_1 > z$. (We note that this will always be the case when n is large. Is 12 large?).
- Then sincere voting is not an equilibrium.
- But clearly, voting to convict always cannot be in equilibrium either.
- The Bayesian-Nash equilibrium will then be in mixed strategies. In particular, suppose that $v_j = g$ if $s_j = g$, but also

$$v_j = g \text{ with probability } \gamma \text{ if } s_j = i.$$

- For such an equilibrium, we need each juror to be indifferent between voting guilty and innocent when they receive $s_j = i$. In other words,

$$\tilde{P}_1 = \Pr(\theta = G | v_j = g \text{ for all } j \neq 1 \text{ and } s_1 = i) = z.$$

Bayesian Nash Equilibrium in the Limit

- It can be shown (as it is intuitive) that as n increases, the probability of convicting the defendant converges increases, and as $n \to \infty$, it converges to a positive number.

- Thus large juries will over-convict — they also convict the guilty with probability 1.

- Why?

- Essentially, each juror finds it optimal to rely on the implicit information that if her vote is pivotal, it must be that others have voted to convict, and that's pretty good evidence that the defendant is guilty.

- Put differently, no one wants to be a contrarian and acquit when others are voting to convict.

Lessons

- Voting in common-interest in complete information situations will be very different than what we have seen so far.
- If voters are "hyper-rational" to be able to make such inferences, they will have a tendency to distort their information (thus not engage in "sincere voting").
- But this may also involve major inefficiencies, very different from Condorcet's Jury Theorem.
- Does this mean voting is always a very bad way of aggregating information? Well, yes and no.

A Model of a Large Election

- Feddersen and Pesendorfer (1996) consider the following environment.
- There our two states of nature, $\theta = \{0, 1\}$, and two policy choices of candidates, $x \in \{0, 1\}$.
- There are three types of voters, denoted by elements of the type space $T = \{0, 1, i\}$.
- The first two are committed voters and will always choose $x = 0$ or $x = 1$ either because of distributional or ideological reasons.
- The last one designates "independent" voters, which we normally think as the "swing voters". These independents have preferences given by

$$U_i(x, \theta) = -\mathbb{I}(x \neq \theta),$$

where $\mathbb{I}(x \neq \theta)$ is the indicator function for the position of the candidate from being different than the state of nature.
- This implies that the voters received negative utility if the "wrong" candidate is elected.

A Common Value Model (continued)

- A candidate (policy) that obtains an absolute majority is chosen. If both options obtain the same number of votes, then one of them is chosen at random.

- Let us suppose, without loss of any generality, that the prior probability that the true state is $\theta = 0$ is $\alpha \leq 1/2$, so that state $\theta = 1$ is more likely ex ante.

- To make the model work, there needs to be some uncertainty about the preferences of other voters. One way to introduce this is to suppose that how many other voters there are (meaning how many other voters could potentially turn out to vote) and what fractions of those will be committed types are stochastically generated. (This is the assumption first developed in Myerson and Weber, 1993).

Uncertainty

- Suppose, in particular, that the total number of voters is determined by Nature taking $N + 1$ independent draws from a potentially large pool of voters.

- At each draw, an actual voter is selected with probability $1 - p_\phi$. This implies that the number of voters is a stochastic variable with the binomial distribution with parameters $(N + 1, 1 - p_\phi)$.

- Conditional on being selected, an agent is independent with probability $p_i / (1 - p_\phi)$, is committed to $x = 0$ with probability $p_0 / (1 - p_\phi)$, and is committed to $x = 1$ with probability $p_1 / (1 - p_\phi)$.

- Therefore, the numbers of voters of different types also follow binomial distributions.

Uncertainty (continued)

- The probability vector (p_ϕ, p_i, p_0, p_1), like preferences and the prior probability α, is common knowledge.
- Finally, each agent knows her type and also receives a signal $s \in S = \{0, 1, \phi\}$, where the first two entries designate the actual state, i.e., $\theta = 0$ or $\theta = 1$, so that conditional on receiving the signal values the agent will know the underlying state for sure.
- The last entry means that the agent receives no relevant information and this event has probability q.
- This formulation implies that some voters will potentially be fully informed, but because all events are stochastic, whether there is indeed such an agent in the population or how many of them there are relative to committed types is not known by any of the voters.
- Voting truthfully is not necessarily optimal for independents. In fact they may prefer to abstain rather than vote according to their information (priors or some other source of signals that are not certain).

Strategies

- A pure strategy here is simply

$$\sigma : T \times S \rightarrow [\phi, 0, 1] \,,$$

where ϕ denotes abstention.
- Clearly, $\sigma(0, \cdot) = 0$ and $\sigma(1, \cdot) = 1$ (for committed voters).
- Moreover, it is also clear that $\sigma(i, z) = z$ for $z \in \{0, 1\}$, meaning that independent informed voters will vote according to their (certain) posterior.
- This implies that we can simply focus on the decisions by uninformed independent voters, denoted by

$$\tau = \left(\tau_0, \tau_1, \tau_\phi \right),$$

which correspond to the probabilities that they will vote for $x = 0$, $x = 1$ and abstain, respectively. Recall that though "uninformed," these voters have posteriors that are not equal to $1/2$, thus have relevant information.

Swing Voter's Curse

- The key observation in the analysis of this model is that, as in the jury problem, an individual should only care about his or her vote conditional on being pivotal.

- Since they do not obtain direct utility from their votes and only care about the outcome, their votes when there is a clear majority for one or the other outcome are irrelevant.

- But this implies that one has to **condition** on a situation in which one is pivotal in a large election.

- This happens (in the unlikely event) where either an equal number of agents have voted for each choice, or one of the two choices is winning with only one vote.

Swing Voter's Curse (continued)

- This intuition is sufficient to establish the following proposition, which captures the idea of the "**swing voter's curse**".
- Let $U(x, \tau)$ be the expected utility of an uninformed independent agent to choose $x \in \{0, 1, \phi\}$, when all other independents are using (symmetric) mixed strategies given by τ.

Proposition

Suppose that $p_\phi > 0$, $q > 0$ and that N is greater than 2 and even. Then if $U(1, \tau) = U(0, \tau)$, then all uninformed independent voters abstain.

Intuition

- If $U(1, \tau) = U(0, \tau)$, meaning that an uninformed voter is indifferent between voting for either candidate (policy), then he or she must prefer to abstain.

- By continuity, we could also show that if $|U(1, \tau) - U(0, \tau)| < \varepsilon$ for ε sufficiently small, then the same conclusion will apply. This is despite the fact that uninformed voters actually have relevant information, because the prior α can be arbitrarily small.

- Intuitively, when a voter expects the same utility from the two options available to him or her, then abstaining and leaving the decision to another voter who is more likely to be informed is better.

- This is despite the fact that the voter may be leaving the decision to a committed type.

- Different from the implications of models in which swing voters are "powerful".

Implications

- The implication is that useful information will be lost in the elections, and this is the essence of the "swing voter's curse".
- Nevertheless, Feddersen and Pesendorfer also show that in large elections information still aggregates in the sense that the correct choice is made with arbitrarily high probability. In particular:

Proposition

Suppose that $p_\phi > 0$, $q > 0$ and $p_i \neq |p_1 - p_0|$, then for every $\varepsilon > 0$, there exists $\bar{N}$ such that for $N > \bar{N}$, the probability that the correct candidate gets elected is greater than $1 - \varepsilon$.

- The idea of this result is that as the size of the electorate becomes large, uninformed independents mix between the "disadvantaged" candidate and abstaining, in such a way that informed independents become pivotal with very high probability.

Discussion

- Results depend on "hyper rational voters". Is this realistic?

- On the other hand, the resulting voting rule may be "simple": abstain if you do not have strong information. But this conclusion is still follows from a complicated reasoning and sometimes mixed strategies are necessary.

- How to interpret the result that the correct action will be taken in large elections?

Any Voter's Curse

- But if voters are strategic in this fashion and vote just to be pivotal, turnout will be extremely low with even trivial costs of voting.
- Turnout has to be low in particular in order to make each voter be pivotal with a sufficiently high probability.
- No way of explaining turnout rates of 20 or 30% in large elections (let alone 60 or 70%).

Evidence?

- We will discuss evidence in the next lecture.
- But it's worth mentioning the work by Battaglini, Morton and Palfrey (2008, 2010), which looks at voting behavior in reasonable-sized lab experiments with common values (as with the model here).
- They find support for two of the key features here:
  - Swing voter's curse: abstention by low information independent voters.
  - Swing voter's cunning strategy: they mix in a way to encourage more informed independence to be pivotal (and this cunning strategy is stronger when there is greater imbalance between committed voters as theory would suggest).