14.773 Political Economy of Institutions and Development.
Lectures 10 and 11. Information, Beliefs and Politics

Daron Acemoglu

MIT

March 11 and 13, 2014.

Introduction

- Why do people vote?
  - Because they think they will be pivotal.
  - Because they obtain utility from voting (but if so, how do they vote?)
  - Because they wish to express their opinions (again, if so, is this along the lines of their narrow interests?)

- What do people infer about candidates from their policies and past performance? How do they form beliefs about future policies?

- Central questions for understanding functioning of voting systems (and little empirical evidence).

- In the lecture, we take the voting motive as given and study the interaction between information and political outcomes.

Voting for Being Pivotal

- Suppose that voters 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, 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.

A Common Value Model

- 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

- 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.

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 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, which happens when 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?

Information-Behavior Feedbacks

- In Feddersen and Pesendorfer, the relevant information is about some abstract state of the world.
- In practice, one important dimension is how distortionary redistributive policies will be.
- If voters think that these are very distortionary, then they may choose low redstribution. But then the society may not learn about true consequence of redistributive policies.
- This idea is investigated in an "overlapping generations" model by Piketty.
- To avoid the "swing voter's curse," Piketty assumes that each individual votes according to what they think would maximize "social welfare" and does not try to infer information of others from their votes (formally, heterogeneous priors and some "myopia").

Redistribution and Mobility: Model

- An individual i of generation t has utility

$$U_{it} = y_{it} - \frac{1}{2\alpha} e_{it}^2,$$

where $y_{it} \in \{0, 1\}$ can be thought of as success or failure, and e_{it} is the effort level. The set of individuals is denoted by $\mathcal{I}$ and is taken to be odd for simplicity.

- Suppose that success depends on effort and also on

$$\mathbb{P}\left(y_{it} = 1 \mid e_{it} = e \text{ and } y_{it-1} = 0\right) = \pi_0 + \theta e,$$

and

$$\mathbb{P}\left(y_{it} = 1 \mid e_{it} = e \text{ and } y_{it-1} = 1\right) = \pi_1 + \theta e,$$

where $\pi_1 \geq \pi_0$.

- The gap between these two parameters is the importance of "inheritance" in success, whereas θ is the importance of "hard work".

- The vector of parameters (θ, π_0, π_1) is unknown.

Model (continued)

- At any given point in time, individuals will have a posterior over this policy vector μ_{it}, shaped by their dynasty's prior experiences as well as other characteristics in the society that they may have observed.

- The only policy tool is a tax rate on output, which is then redistributed lump sum.

- Let total output under tax rate τ be $Y(\tau)$.

- This implies that given an expectation of a tax rate τ, an individual with a successful or unsuccessful parent denoted by $z = 1$ or $z = 0$ will choose

$$
\begin{aligned}
e_z(\tau, \mu) \quad \in \quad & \arg\max_e \mathbb{E}_\mu[(1 - \pi_z - \theta e)\,\tau\, Y(\tau) \\
& + (\pi_z + \theta e)\,((1 - \tau) + \tau\, Y(\tau))] - \frac{1}{2\alpha} e^2,
\end{aligned}
$$

where the expectation is over the parameters.

Effort and Voting

- It can be easily verified that

$$e_z\left(\tau, \mu\right) = e\left(\tau, \mathbb{E}_\mu \theta\right) = \alpha\left(1 - \tau\right) \mathbb{E}_\mu \theta.$$

- Therefore, all that matters for effort is the expectation about the parameter θ.

- Now given this expectation, individuals will also choose the tax rate by voting.

- Individuals vote for the tax rates that maximizes "expected social welfare" $\mathbb{E}_{\mu_{it}} V_t$ (why is this conditional on μ_{it}?).

- Given the quadratic utility function, it can be verified that individuals have single peaked preferences, with bliss point given by

$$\tau\left(\mu_{it}\right) \in \arg\max \mathbb{E}_{\mu_{it}} V_t.$$

- An application of the median voter theorem then gives the equilibrium tax rate is the median of these bliss points.

Evolution of Beliefs

- How will an individual update their beliefs? Straightforward application of Bayes rule gives the evolution of beliefs.
- For example, for an individual $i \in \mathcal{I}$ with a successful or unsuccessful parent denoted by $z = 1$ or $z = 0$, starting with beliefs μ_{it}, with support $S[\mu_{it}]$, we have that for any $(\theta, \pi_0, \pi_1) \in S[\mu_{it}]$, we have

$$\mu_{it+1}(\theta, \pi_0, \pi_1) = \mu_{it}(\theta, \pi_0, \pi_1) \frac{\pi_z + \theta e\left(\tau_t, \mathbb{E}_{\mu_{it}}\theta\right)}{\int \left[\pi'_z + \theta' e\left(\tau_t, \mathbb{E}_{\mu_{it}}\theta\right)\right] d\mu_{it}}.$$

- Note that individuals here are not learning from the realized tax rate, simply from their own experience. This is because individuals are supposed to have "heterogeneous priors". They thus recognize that others have beliefs driven by their initial priors, which are different from theirs and there is no learning from initial priors.
- Is this just to consequence of heterogeneous priors?

Evolution of Beliefs (continued)

- Standard results about Bayesian updating, in particular from the *martingale convergence theorem*, imply the following:

Proposition

The beliefs of individual $i \in \mathcal{I}$, μ_{it}, starting with any initial beliefs μ_{i0} almost surely converges to a stationary belief $\mu_{i\infty}$.

- But if beliefs converge for each dynasty, then the median also converges, and thus equilibrium tax rates also converge.

Proposition

Starting with any distribution of beliefs in the society, the equilibrium tax rate τ_t almost surely converges to a stationary tax rate τ_∞.

Limits of Learning

- The issue, however, is that this limiting tax rate need not be unique, because the limiting stationary beliefs are not necessarily equal to the distribution that puts probability 1 on truth.

- The intuition for this is the same as "self confirming" equilibria, and can be best seen by considering an extreme set of beliefs in the society that lead to $\tau = 1$ (because effort doesn't matter at all).

- If $\tau = 1$, then nobody exerts any effort and there is no possibility that anybody can learn that effort actually matters.

Limits of Learning (continued)

- The characterization of the set of possible limiting beliefs is straightforward.
- Define $M^*(\tau)$ be the set of beliefs that are "self consistent" at the tax rate τ in the following sense:
- For any $\tau \in [0, 1]$, we have

$$
\begin{aligned}
M^*(\tau) \quad = \quad & \{\mu : \text{for all } (\theta, \pi_0, \pi_1) \in S[\mu], \\
& \pi_z + \theta e\left(\tau, \mathbb{E}_\mu \theta\right) = \pi_z^* + \theta^* e\left(\tau, \mathbb{E}_\mu \theta\right) \\
& \text{for } z = 0, 1 \text{ and } (\theta^*, \pi_0^*, \pi_1^*) \in S[\mu]\}.
\end{aligned}
$$

- Intuitively, these are the set of beliefs that generate the correct empirical frequencies in terms of upward and downward mobility (success and failure) given the effort level that they imply.
- Clearly, if the tax rate is in fact τ and $M^*(\tau)$ is not a singleton, a Bayesian cannot distinguish between the elements of $M^*(\tau)$: they all have the same observable implications.

Limits of Learning (continued)

- Now the following result is immediate.

Proposition

Starting with any initial distribution of beliefs in society $\{\mu_{i0}\}_{i \in \mathcal{I}}$, we have that

1. *For all $i \in \mathcal{I}$, $\mu_{i\infty}$ exists and is in $M^*(\tau_\infty)$, and*
2. *τ_∞ is the median of $\{\tau(\mu_{i\infty})\}_{i \in \mathcal{I}}$.*

- This proposition of course does not rule out the possibility that there will be convergence to beliefs corresponding to the true parameter values regardless of initial conditions. But it is straightforward from the above observations establish the next result:

Limits of Learning (continued)

Proposition

Suppose $\mathcal{I}$ is arbitrarily large. Then for any $\{\mu_{i\infty}\}_{i \in \mathcal{I}} \in M^*(\tau_\infty)$ such that τ_∞ is the median of $\tau(\mu_{i\infty})$, there exists a set of initial conditions such that there will be convergence to beliefs $\{\mu_{i\infty}\}_{i \in \mathcal{I}}$ and tax rate τ_∞ with probability one.

- This proposition implies that a society may converge and remain in equilibria with very different sets of beliefs and these beliefs will support different amounts of redistribution.
- Different amounts of redistribution will then lead to different tax rates, which "self confirm" these beliefs because behavior endogenously adjusts to tax rates.

Interpretation

- Therefore, according to this model, one could have the United States society converge to a distribution of beliefs in which most people believe that θ is high and thus vote for low taxes, and this in turn generates high social mobility, confirming the beliefs that θ is high.

- Many more Europeans believe that θ is low (and correspondingly $\pi_1 - \pi_0$ is high) and this generates more redistribution and lower social mobility.

- Neither Americans nor Europeans are being "irrational".

Discussion

- How to interpret these results?
- Perhaps a good approximation to the formation of policemen individuals are not "hyper rational".
- But why don't different societies learn from each other?
- How likely is this process to lead to multiple stable points?

Voting and Experimentation

- Information is in general acquired dynamically, as a result of past political choices.
- Example: Economic or social reforms
  - Reforms make winners and losers, whose identities are unknown ex ante.
  - Fernandez and Rodrik (1991): resistance to trade liberalization because of losers' fear that they will not be compensated.
- But in a dynamic context, there are new effects that make political actors even more averse to the information and experimentation.
- Strulovici (2010): two novel reasons for this:
  - Loser trap (can't return to status quo).
  - Winner frustration (can't exploit new alternative).

Illustration

- Ann, Bob and Chris go to the restaurant every week-end.
- They always choose their restaurant by majority rule.
- A new restaurant has opened.
- If any one of them could choose *alone* future restaurants, he or she would try the new one now.
- However, it is possible that all three will vote against trying this restaurant.

Illustration (continued)

- Experimentation with new alternatives is less attractive when one has to share power.
- Sharing control induces two opposite control loss effects, which have different implications.
    - *Loser trap*. If Ann and Bob like the new restaurant, they will impose it to Chris in the future, even if he does not like it.
    - *Winner frustration*. If only Ann likes the new restaurant, she will be blocked by Bob and Chris. So the "risk" of trying a new restaurant need not be rewarded even for those who do turn out to like it.
- Majority-based experimentation is also shorter than the socially efficient outcome.
- New winners induce more experimentation from remaining voters.

Model: Single Agent Problem

- Safe (S) and risky (R) actions.
- R can be good or bad. Agent type initially unknown.
- Continuous time with fixed discount rate, infinite horizon.
- At each instant, one action (S or R) is chosen.

Model: Single Agent Problem (continued)

- Payoffs:

$S \rightarrow s > 0$

$R \nearrow$ $\quad bad : 0$
$\quad \searrow$ $\quad good :$ lump sums > 0 at Poisson arrival times

- bad (loser) $<$ safe $<$ good (winner).
- Bayesian updating of beliefs:

$$\frac{dp_t}{dt} = -\lambda p_t (1 - p_t)$$

where λ arrival rate of good outcome from the risky action and p_t belief at time t that risk action is good (or the agent is of good type).

Model: Single Agent Problem (continued)

- Equilibrium: Experiment up to some level of belief $p^{SD} < p^{\text{myopic}}$
- This is because of the option value of experimentation.

Model: Collective Decision-Making

- N (odd) agents.
- Publicly observed payoffs.
- Types are iid. Initially, $Prob[good] = p_0$ for all.
- Arrival times also independent across agents.
- State variables (k, p) where k is number of *sure winners*, and $p = Prop[good]$ for *unsure voters*.
- Equilibrium concept: *Markov Voting Equilibrium*
- At any time, chose the action preferred by majority (given that the same rule holds in the future).
- Equilibrium can be solved by backward induction on number of sure winners.

Markov Voting Equilibrium

- A *Majority Voting Equilibrium (MVE)* is a mapping
  $C : (k, p) \rightarrow \{S, R\}$ such that $C = R$ if $k > k_N = (N-1)/2$
  and $C = R$ if $k \leq k_N$ and

  $$pg + \lambda p[w(k+1, p) - u(k, p)] + \qquad (1)$$
  $$\lambda p(n-1)[u(k+1, p) - u(k, p)] - \lambda p(1-p)\frac{\partial u}{\partial p} > s,$$

  where u and w are the value of functions of unsure voters and sure
  winners when voting rule C determines future votes.

Collective Decision-Making: Structure of Equilibrium

- Now threshold belief $p^G(k)$ for stopping when there has been k people revealed to be of good type until now.
- Monotonicity: $p^G(k)$ is decreasing in k.
- Intuition: Good news for any one prompts remaining unsure voters to experiment more.
  - Why? Suppose to the contrary that experimentation stops when a new winner is observed.
  - Then, risky action pays lower expected payoffs and has no option value.
  - Therefore, experimentation was not optimal when the news arrived: contradiction.

Collective Decision-Making: Comparison

- We have that $p^G(k)$ is always greater than what social planner maximizing utilitarian welfare would choose.
- This is because of loser trap and winner frustration.

Comparative Statics

- - Experimentation decreases if N increases (enough): $p(k, N)$ almost increases in N.
  - Agents behave myopically as $N \rightarrow \infty$
- For N above some threshold, agents prefer safe action even if trying risky action would immediately reveals types.

Alternative Rules

- Suppose R requires unanimous approval.

- This gets rid of the loser trap.

- However, this increases winner frustration, since R is less likely to be played in the long run.

- Which rule performs better depends on the relative strengths of the two effects.

Populism

- Capturing a variety of related concepts
- Albertazzi and McDonnell (2008):

    "an ideology which pits a virtuous and homogeneous people against a set of elites and dangerous 'others' who are together depicted as depriving (or attempting to deprive) the sovereign people of their rights, values, prosperity, identity and voice"

- Hawkins (2003) about the rise of Chavez:

    "If we define populism in strictly political terms—as the presence of what some scholars call a charismatic mode of linkage between voters and politicians, and a democratic discourse that relies on the idea of a popular will and a struggle between 'the people' and 'the elite'—then Chavismo is clearly a populist phenomenon."

What is Populism?

- Populist policies (not just rhetoric):
  - Budget deficits, mandatory wage increases, price controls, overvalued exchange rates, expropriation of foreign investors / large businesses.
- Costly to businesses, but also costly to the population at large.
- Dornbush and Edwards (1991):

  "Populist regimes have historically tried to deal with income inequality problems through the use of overly expansive macroeconomic policies. These policies, which have relied on deficit financing, generalized controls, and a disregard for basic economic equilibria, have almost unavoidably resulted in major macroeconomic crises that have ended up hurting the poorer segments of society."

Populism vs. Median Voter

- Are these policies what the "median voter" wants?
- Perhaps, but Dornbusch and Edwards's definition and the fact that middle classes and lower middle classes suffer on their populist policies suggests may be not.
- The fact that populist policies are often to the left of the "median voter" cannot be explained solely by personal biases of the populist politician.
  - such biased politician would fail to be reelected.

Populism and Popularity

- Most populist regimes are "popular," at least for quite a while.
- Popularity of populist regimes even allows leaders to violate constitutional norms:
    - most of Latin American postwar leaders post term-limited (often by one term), but many violated the rules.
    - this should not be the case if they are known to involve highly inefficient policies
- Also interestingly, many of the populist politicians or parties, at least in Latin America, often end up choosing policies consistent with the interests of traditional elites
    - E.g.: PRI in Mexico, the policies of traditional parties in Venezuela and Ecuador, Fujimori's reign in Peru, Menem in Argentina.

Possible Definition

- Populism = policy to the left of median voter's ideal policy
                but still popular
- Why would this be the case? *An informational theory.*
- One-dimensional policy space
- Two points of attraction for politician
  - median voter's preferences
  - elite's preferences, excercised through bribes
  - (personal preferences if partisan)
- Normally, policy should lie between median voter's and elite's ideal points.
- But there are informational reasons for policy to be to the left of the median voter— i.e., populist.

A Political Theory

- Major concern of the median voter under weak institutions: a politician is secretly biased to the right or being disproportionately influenced by the elite (e.g., through bribery, corruption or lobbying).
- Relevant for the Latin American context.
- Politicians will move to the left to signal that they are not closet right-wingers or in the pockets of the traditional elites.
- Then: moderate politicians will necessarily adopt populist policies and even right-wingers (or corrupt politicians) may adopt such policies.
- Intuition: it is the threat of excessive elite influence under weaker institutions that leads to populist policies.

Policy Space and Voters

- One-dimensional policy space
- Two periods, 1 and 2
- Two groups of voters
  - majority (poor), with bliss point $\gamma^p = 0$
  - minority (elite), with bliss point $\gamma^r = r > 0$
  - results identical if there is a distribution of preferences with median at $\gamma = 0$
- Voters care about policy only
  - Person with bliss point γ gets utility

  $$u(x_1, x_2) = -\sum_{t=1}^{2} (x_t - \gamma)^2$$

  from policies x_1 and x_2 in periods 1 and 2
- Elections are decided by median voter who is poor

Politicians

- Politicians' utility in each period depends on:
    - policy
    $$v = -\alpha \left(x - \gamma\right)^2 \ \ldots$$
    - office
    $$\ldots \ + W\mathbf{I}_{\{\text{in office}\}} \ \ldots$$
    - bribes
    $$\ldots \ + B$$
- Two types of politicians
    - share μ has $\gamma = 0$ ("**moderate**")
    - share $1 - \mu$ has $\gamma = r$ ("**right-winger**")
- We start with $B = 0$

Timing

1. Politician chooses first-period policy $x_1 \in \mathbb{R}$.
2. Population gets a noisy signal $s = x_1 + z$.
3. Median voter decides whether to replace the current politician with a random one drawn from the pool.
4. In the second period, the politician (the incumbent or the new one) chooses policy $x_2 \in \mathbb{R}$.
5. Everyone learns the realizations of both policies and gets payoffs.

Noisy Signal

- Noise z has a distribution with support on $(-\infty, +\infty)$ with c.d.f. $F(z)$ and p.d.f. $f(z)$.
- Density $f(z)$ is assumed to be an even (i.e., symmetric around 0) function, which is everywhere differentiable and satisfies $f'(z) < 0$ for $z > 0$.
  - the density function f is single-peaked
- Noise z is sufficiently high and well-behaved:

$$\left| f'(z) \right| < \frac{1}{\frac{r^2}{2} + \frac{W}{2\alpha}} \text{ for all } z.$$

  - implies $\Pr\left(|z| > \frac{r}{4}\right) > \frac{1}{4}$
  - implies $f(0) < \frac{2}{r}$
  - holds for $\mathcal{N}(0, \sigma^2)$ if σ^2 is sufficiently high, i.e., $\sigma^2 > \frac{\frac{r^2}{2} + \frac{W}{2\alpha}}{\sqrt{2\pi e}}$.

Equilibrium Concept
Period 2

- Perfect Bayesian equilibrium in pure strategies

- In period 2:
  - moderate politician chooses $x_2 = 0$
  - right-wing politician chooses $x_2 = r$

- Median voter prefers to have moderate politician in period 2
  - incumbent reelected if and only if his posterior that he is moderate is at least μ

Period 1: Elections

- Suppose that in equilibrium:
  - moderate politicians choose $x_1 = a$
  - right-wing politicians choose $x_1 = b > a$ (proved in the paper that this is always the case).
- For median voter who gets signal s, posterior probability that politician is moderate equals

$$\hat{\mu} = \frac{\mu f (s - a)}{\mu f (s - a) + (1 - \mu) f (s - b)}$$

  - It exceeds μ if and only if

$$s < \frac{a + b}{2}$$

- The probability of reelection if policy is x equals

$$\pi (x) = F \left(\frac{a + b}{2} - x \right)$$

Period 1: Policy Choices

- Moderate politician maximizes

$$\max_{x} -\alpha x^2 + W\pi(x) - (1-\mu)\alpha r^2(1-\pi(x))$$

  - he loses αr^2 in period 2 only if right-wing politician comes to power
  - FOC must hold at $x = a$:

$$-2\alpha a - \left(W + (1-\mu)\alpha r^2\right) f\left(\frac{b-a}{2}\right) = 0$$

- Right-wing politician maximizes

$$\max_{x} -\alpha(x-r)^2 + W\pi(x) - \mu\alpha r^2(1-\pi(x))$$

  - FOC at $x = b$:

$$-2\alpha(b-r) - \left(W + \mu\alpha r^2\right) f\left(\frac{b-a}{2}\right) = 0$$

Equilibrium
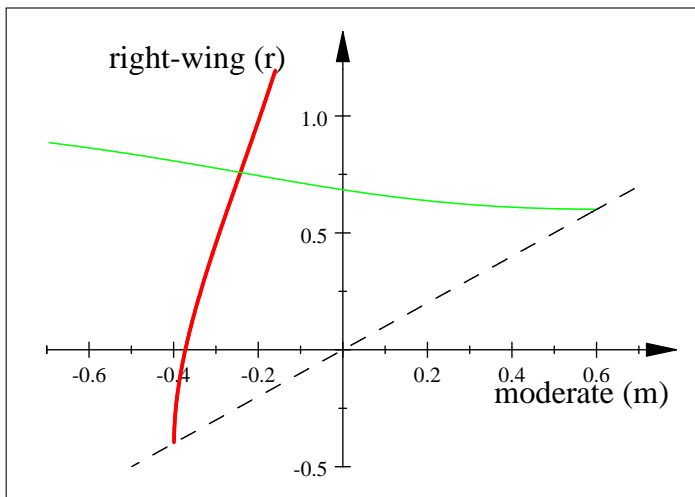
Figure: Reaction curves of moderate (red/thick) and right-wing (green/thin) politicians.

Solution

- In equilibrium, $a < 0$
    - moving from $x_1 = 0$ to $x_1 < 0$ causes second-order loss
    - but first-order gain due to higher chance of reelection
- $b < r$ for the same reason
- This moves a left even further

- For moderate politicians: a right-wing alternative necessitates populist bias!
- This would be true even if $W = 0$
   - reelection is valuable as it allows to influence second-period policy

Comparative Statics
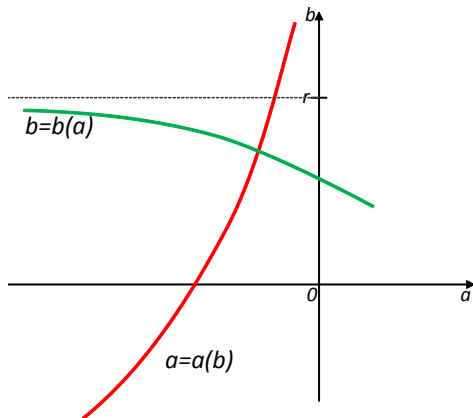
Comparative Statics (continued)

- Populist bias is stronger if
  - W is higher (i.e., politicians value being in office more)
  - α is lower (i.e., changing political positions is relatively costless for politicians)
  - μ is lower (i.e., moderate politicians are rarer)
- This holds even if W increases or α decreases for only one type of politician
  - e.g., higher W for pro-elite politicians makes them move left
  - and then pro-poor politicians move left as well

Comparative Statics (continued)
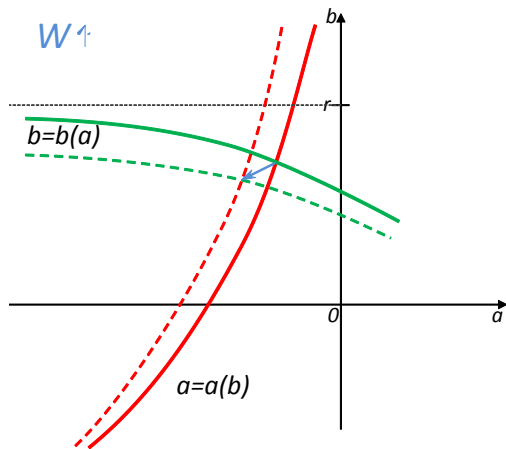
Comparative Statics (continued)

- Also, under additional conditions on distribution F, populist bias is stronger if:
  - r is greater (i.e., greater polarization).
  - two competing effects:
    1. benefits from reelection to both types of politicians is greater, which leads to more signaling;
    2. cost of signaling is also higher to right-wingers.
       Additional conditions ensure that the first effect dominates.

- Populist bias would be weaker if elitist politicians could commit to $b = r$

Populism of Right-Wing Politicians

- If $W = 0$, then $0 < b < r$
  - $x_1 < 0$, $x_2 = r$ is dominated even by $x_1 = r$, $x_2 = 0$
  - hence switching to $x_1 = r$ is better even if it guaranteed losing elections
- If $W > 0$, then $b < 0$ is possible
  - if office is very valuable per se, *all* politicians will be populists!

Cycles of Conflict

- Conflict (between ethnic groups, religious groups, countries, ideologies, social classes, rival individuals) is endemic.
- Why? Part of it may be related to incorrect information ("misperceptions") and relatedly to **fear** of actions, intentions and behavior of the other party as Thucydides emphasized long ago.
- Often continuing cycles of conflict between different groups. Partly related to information:

Group A's actions look aggressive
$\implies$ Group B thinks Group A is aggressive
$\implies$ Group B acts aggressively
$\implies$ Group A thinks Group B is aggressive
$\implies$ Group A acts aggressively . . .

Examples

- Spirals in the World Horowitz (2000) on ethnic conflict:
  "The fear of ethnic domination and suppression is a motivating force for the acquisition of power as an end . . . The imminence of independence in Uganda aroused 'fears of future ill-treatment' along ethnic lines. In Kenya, it was 'Kikuyu domination' that was feared; in Zambia, 'Bemba domination'; and in Mauritius, ['Hindu domination'] . . . "

- Serbo-Croation War (DellaVigna et al, 2011).

- Protestant-Catholic Conflict in Northern Ireland.

- Trade (Guiso, Sapienza, and Zingales, 2009, Bottazzi, Da Rin, and Hellmann, 2011).

- Political polarization (Sunstein, 2006).

Ebbs and Flows of Conflict

- But not ever-lasting continuous conflict.
- Ethnic conflict in Africa way down in last 20 years.
- France and Germany not on brink of war, and trade a lot.
- Conflict and distrust in Balkans greatly diminished.
- Political polarization in U.S. was probably as bad or worse in first third of 20th century.

Idea for Cycles

- Once Groups A and B are both acting aggressively, aggression becomes uninformative of their true types.
- Once this happens, one group will experiment with cooperation, which causes trust to restart.
- Conflict spirals cannot last forever, because if they did the informational content of conflict would eventually dissipate.

Model

- Timing and Actions 2 groups, A and B. Time $t = 0, 1, 2, \ldots$.
- Overlapping generations.
- At time t, one active player: player t.
- Player t takes pair of actions $(x_t, y_t) \in \{0, 1\} \times \{0, 1\}$.
- t even $\implies$ player t from Group A.
- t odd $\implies$ player t from Group B.

Model: Information

- Before player t takes actions, observes noisy signal $\tilde{y}_{t-1} \in \{0, 1\}$.

$$Pr(\tilde{y}_{t-1} = 1 | y_{t-1} = 1) = 1 - \pi$$
$$Pr(\tilde{y}_{t-1} = 1 | y_{t-1} = 0) = 0.$$

- Each group is either normal or bad.
- If normal, all representatives are normal types.
- If bad, all representatives are bad types.
  $Pr(\text{bad}) = \mu_0 \in (0, \mu^*)$.

Model

- For bad player t, playing $(x_t = 0, y_t = 0)$ is dominant strategy.

- For normal player t, utility function is

$$u(x_t, \tilde{y}_{t-1}) + u(\tilde{y}_t, x_{t+1}).$$

- Assume "subgame" between neighboring players is coordination game, and $(1, 1)$ is Pareto-dominant equilibrium: $u(1, 1) > u(0, 1)$, $u(0, 0) > u(1, 0)$, $u(1, 1) > u(0, 0)$.

Equilibrium

- What happens in (sequential) equilibrium?
- Normal player t plays $x_t = 1$ if and only if $\tilde{y}_{t-1} = 1$.
- So normal player 0 plays $y_0 = 1$ if and only if μ_0 is below some threshold μ^*:

$$\mu^* \equiv \left(\frac{u(1,1) - u(0,0)}{u(1,1) - u(1,0)} \right).$$

- If normal player 1 sees $\tilde{y}_0 = 1$, learns other group is good, and plays $y_1 = 1$.
- If normal player 1 sees $\tilde{y}_0 = 0$, posterior belief that other group is bad rises to

$$\mu_1 = \frac{\mu_0}{\mu_0 + (1 - \mu_0)\pi} > \mu_0.$$

- Plays $y_1 = 0$ if and only if $\mu_1 > \mu^*$. Holds if π small.

Equilibrium (continued)

- Equilibrium Suppose up to time t normal players play $y_t = 0$ when $\tilde{y}_{t-1} = 0$.
- Then normal player t's posterior when $\tilde{y}_{t-1} = 0$ is

$$\mu_t = \frac{\mu_0}{\mu_0 + (1 - \mu_0)(1 - (1 - \pi)^t)}.$$

- Observe that μ_t is decreasing in t, $\mu_t \to \mu_0$ as $t \to \infty$, and $\mu_0 < \mu^*$.
- But this implies that there is first time t at which $\mu_t \leq \mu^*$. Call it T.

- Normal player T plays $y_T = 1$ **even if he sees a bad signal**.

- But now normal player $T + 1$ faces same problem as player 1.
- This implies a **cycle of conflict.**

Equilibrium (continued)

Proposition

Assume $\mu_0 < \mu^*$ and $\mu_t \neq \mu^*$ for all t.
Then the baseline model has a unique sequential equilibrium.
It has the following properties:

- *At every time $t \neq 0 : \text{mod} : T$, normal player t plays good actions ($x_t = 1, y_t = 1$) if she gets the good signal and plays bad actions ($x_t = 0, y_t = 0$) if she gets the bad signal.*

- *At every time $t = 0 : \text{mod} : T$, normal player t plays the good action $x_t = 1$ toward player $t - 1$ if and only if she gets the good signal, but plays the good action $y_t = 1$ toward player $t + 1$ regardless of her signal.*

- *Bad players always play bad actions ($x_t = 0, y_t = 0$).*
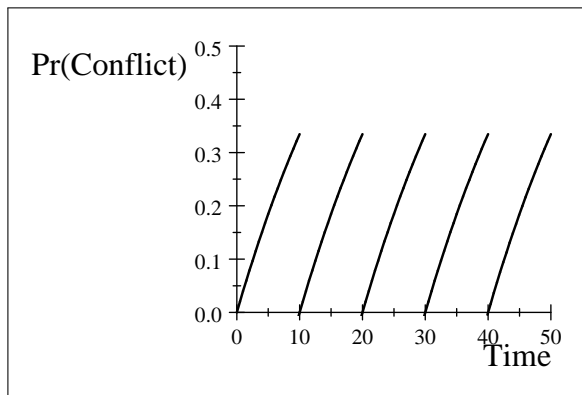
Equilibrium (continued)

Figure: A Cycle of Conflict

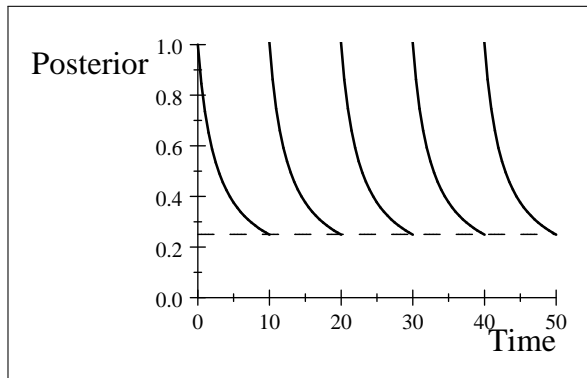# Equilibrium (continued)

Figure: The Corresponding Cycle of Beliefs

Comparative Statics

Proposition

The cycle period T has the following properties:

- *It is increasing in $u(0,0)$, decreasing in $u(1,0)$, and decreasing in $u(1,1)$.*
- *It is increasing in the prior probability of the bad type μ_0.*
- *It is decreasing in the error probability π.*

Comparative Statics (continued)

Proposition

Welfare If player t's payoff is u_t, define social welfare to be

$$lim_{N \to \infty} \frac{1}{N} \sum_{t=0}^{N} u_t.$$

Suppose both groups are normal. Then:

- *The limit of social welfare as $\pi \to 0$ is less than the efficient level $2u(1, 1)$.*
- *For any sequence $(\pi_n, \mu_{0,n})$ converging to $(0, 0)$ as $n \to \infty$, the limit of social welfare as $n \to \infty$ equals the efficient level $2u(1, 1)$.*

- The limit of no misperception is not the same as the perfect information game because any conflict lasts so much longer in that limit.

Conclusion

- Important feedbacks between beliefs and political/public actions.
- Important high-level questions are:
  - Does the presence of political economy lead to biased or less accurate learning/belief formation?
  - Does imperfect information exacerbate political economy conflicts? Does it lead to new types of inefficiencies?
  - Are there feedback cycles leading from bad politics to bad information to bad politics?
  - How can these issues be empirically operationalized?