# 6.254: Game Theory with Engineering Applications
# Guest Lecture: Social Choice and Voting Theory

Daron Acemoglu
MIT

May 6, 2010

# Outline

- Social choice and group decision-making
- Arrow's Impossibility Theorem
- Gibbard-Satterthwaite Impossibility Theorem
- Single peaked preferences and aggregation
- Group decisions under incomplete information

- Reading:
- Microeconomic Theory, MasColell, Whinston and Green, Chapters 21 and 23.

# Social Choice Functions

- Recall that mechanism design, in an environment with $H$ players each with possible type $\theta_i \in \Theta_1$ and a set of feasible alllocations $Y$, started with a social choice function

$$f : \Theta_1 \times ... \times \Theta_H \to Y.$$

- But where does this social choice function come from?
- Presumably, it reflects some "social objective" such as fairness or efficiency. But how do we arrive to such an objective?
- **More general question:** How do groups make collective/group/political decisions?
- Two sets of issues:
  - Aggregating up to collective preferences from individual preferences.
  - Using dispersed information of the group efficiently.
  - Allocation of "power" in a group (not discussed in this lecture).

# Setup

- Let us investigate these questions in a slightly more general setup, where we take the types, the $\theta$'s, to be rankings over all possible allocations.
- For this purpose, let us restrict attention to a society with a finite set of individuals $\mathcal{H}$, with the number of individuals denoted by $H$, and a finite set of allocations denoted by $\mathcal{P}$.
- Individual $i \in \mathcal{H}$ has an *indirect utility function* defined over choices available to the group or "policies" $p \in \mathcal{P}$

$$U\left(p; \alpha_i\right),$$

  where $\alpha_i$ indexes the utility function (i.e., $U\left(p; \alpha_i\right) = U_i\left(p\right)$).
- The **bliss point** of individual $i$ is defined as:

$$p\left(\alpha_i\right) = \arg\max_{p \in \mathcal{P}} \; U(p; \alpha_i).$$

# Preferences More Generally

- Individual $i$ weakly prefers $p$ to $p'$,

$$p \succeq_i p'$$

and if he has a strict preference,

$$p \succ_i p'.$$

- Assume: *completeness*, *reflexivity* and *transitivity* (so that $z \succeq_i z'$ and $z' \succeq_i z''$ implies $z \succeq_i z''$).

# Collective Preferences?

- **Key question:** Does there exist welfare function $U^S(p)$ that ranks policies for this group (or society)?
- Let us first start with a simple way of "aggregating" the preferences of individuals in the group: majoritarian voting.
- This will lead to the Condorcet paradox.

# The Condorcet Paradox

- Imagine a group consisting of three individuals, 1, 2, and 3, three choices and preferences

$$\begin{array}{ll} 1 & a \succ c \succ b \\ 2 & b \succ a \succ c \\ 3 & c \succ b \succ a \end{array}$$

- Assume "open agenda direct democracy" system for making decisions within this group.

A1. *Direct democracy.* The citizens themselves make the policy choices via majoritarian voting.
A2. *Sincere voting.* Individuals vote "truthfully" rather than strategically.
A3. *Open agenda.* Citizens vote over pairs of policy alternatives, such that the winning policy in one round is posed against a new alternative in the next round and the set of alternatives includes all feasible policies.

- What will happen?

# The Condorcet Paradox

- It can be verified that $b$ will obtain a majority against $a$.
- $c$ will obtain a majority against $b$.
- But $a$ will obtain a majority against $c$.
- Thus there will be a **cycle**.

# Towards Collective Preferences

- How general is the Condorcet cycle?
- Arrow's Impossibility Theorem: very general.
- Suppose that the set of feasible policies is some finite set $\mathcal{P}$.
- Let $\Re$ be the set of all transitive orderings on $\mathcal{P}$, that is, $\Re$ contains information of the form $p_1 \succeq_i p_2 \succeq_i p_3$ or $p_1 \succeq_i p_2 \succ_i p_3$, or or $p_1 \succeq_i p_2 \sim_i p_3$ and so on, and imposes the requirement of transitivity on these individual preferences.
- An individual ordering $R_i$ is an element of $\Re$, that is, $R_i \in \Re$.
- Since our society consists of $H$ individuals, $\rho = (R_1, ..., R_H) \in \Re^H$ is a *preference profile.*
- Also $\rho_{|\mathcal{P}'} = \left(R_{1|\mathcal{P}'}, ..., R_{H|\mathcal{P}'}\right)$ is the society's preference profile when alternatives are restricted to some subset $\mathcal{P}'$ of $\mathcal{P}$.

# Restrictions on Collective Preferences I

- Let $\Im$ be the set of all reflexive and complete binary relations on $\mathcal{P}$ (but notice *not necessarily* transitive).
- A social ordering $R^S \in \Im$ is therefore a reflexive and complete binary relation over all the policy choices in $\mathcal{P}$:

$$\phi : \Re^H \rightarrow \Im.$$

- We have already imposed "**unrestricted domain**," since no restriction on preference profiles.
- A social ordering is weakly Paretian if

$$\left[p \succ_i p' \text{ for all } i \in \mathcal{H}\right] \Longrightarrow p \succ^S p'.$$

# Restrictions on Collective Preferences II

- Given $\rho$, a subset $\mathcal{D}$ of $\mathcal{H}$ is **decisive between** $p, p' \in \mathcal{P}$, if

$$\left[p \succeq_i p' \text{ for all } i \in \mathcal{D} \text{ and } p \succ_{i'} p' \text{ for some } i' \in \mathcal{D}\right] \Longrightarrow p \succ^S p'$$

- If $\mathcal{D}' \subset \mathcal{H}$ is decisive between $p, p' \in \mathcal{P}$ for *all* preference profiles $\rho \in \Re^H$, then it is *dictatorial between* $p, p' \in \mathcal{P}$.
- $\mathcal{D} \subset \mathcal{H}$ is **decisive** if it is decisive between any $p, p' \in \mathcal{P}$
- $\mathcal{D}' \subset \mathcal{H}$ is dictatorial if it is dictatorial between any $p, p' \in \mathcal{P}$.
- If $\mathcal{D}' \subset \mathcal{H}$ is dictatorial and a singleton, then its unique element is a **dictator.**

# Restrictions on Collective Preferences III

- A social ordering satisfies independence from irrelevant alternatives, if for any $\rho$ and $\rho' \in \Re^H$ and any $p$, $p' \in \mathcal{P}$,

$$\rho_{|\{p,p'\}} = \rho'_{|\{p,p'\}} \Longrightarrow \phi(\rho)_{|\{p,p'\}} = \phi(\rho')_{|\{p,p'\}}.$$

- This axiom states that if two preference profiles have the same choice over two policy alternatives, the social orderings that derive from these two preference profiles must also have identical choices over these two policy alternatives, regardless of how these two preference profiles differ for "irrelevant" alternatives.
    - While this condition (axiom) at first appears plausible, it is in fact a reasonably strong one. In particular, it rules out any kind of interpersonal "cardinal" comparisons—that is, it excludes information on how strongly an individual prefers one outcome versus another.

# Arrow's Impossibility Theorem

## Theorem

*(***Arrow's (Im)Possibility Theorem***) Suppose there are at least three alternatives. Then if a social ordering, $\phi$, is transitive, weakly Paretian and satisfies independence from irrelevant alternatives, it must be dictatorial.*

- An immediate implication of this theorem is that any set of minimal decisive individuals $\mathcal{D}$ within the society $\mathcal{H}$ must either be a singleton, that is, $\mathcal{D} = \{i\}$, so that we have a dictatorial social ordering, or we have to live with intransitivities.
- Also implicitly, political power must matter. If we wish transitivity, political power must be allocated to one individual or a set of individuals with the same preferences.
- How do we proceed?$\longrightarrow$ Restrict preferences or restrict institutions.

# Proof of Arrow's Impossibility Theorem I

- Suppose to obtain a contradiction that there exists a non-dictatorial and weakly Paretian social ordering, $\phi$, satisfying independence from irrelevant alternatives. Contradiction in two steps.
- **Step 1:** Let a set $\mathcal{J} \subset \mathcal{H}$ be **strongly decisive** between $p_1, p_2 \in \mathcal{P}$ if for any preference profile $\rho \in \Re^H$ with $p_1 \succ_i p_2$ for all $i \in \mathcal{J}$ *and* $p_2 \succ_j p_1$ for all $j \in \mathcal{H} \backslash \mathcal{J}$, $p_1 \succ^S p_2$ ($\mathcal{H}$ itself is strongly decisive since $\phi$ is weakly Paretian).
- We first prove that if $\mathcal{J}$ is strongly decisive between $p_1, p_2 \in \mathcal{P}$, then $\mathcal{J}$ is dictatorial (and hence decisive for all $p, p' \in \mathcal{P}$ and for all preference profiles $\rho \in \Re^H$).
- To prove this, consider the restriction of an arbitrary preference profile $\rho \in \Re^H$ to $\rho_{|\{p_1, p_2, p_3\}}$ and suppose that we also have $p_1 \succ_i p_3$ for all $i \in \mathcal{J}$.

# Proof of Arrow's Impossibility Theorem II

- Next consider an alternative profile $\rho'_{|\{p_1,p_2,p_3\}}$, such that $p_1 \succ'_i p_2 \succ'_i p_3$ for all $i \in \mathcal{J}$ and $p_2 \succ'_i p_1$ and $p_2 \succ'_i p_3$ for all $i \in \mathcal{H} \backslash \mathcal{J}$.
- Since $\mathcal{J}$ is strongly decisive between $p_1$ and $p_2$, $p_1 \succ'^S p_2$. Moreover, since $\phi$ is weakly Paretian, we also have $p_2 \succ'^S p_3$, and thus $p_1 \succ'^S p_2 \succ'^S p_3$.
- Notice that $\rho'_{|\{p_1,p_2,p_3\}}$ did not specify the preferences of individuals $i \in \mathcal{H} \backslash \mathcal{J}$ between $p_1$ and $p_3$, but we have established $p_1 \succ'^S p_3$ for $\rho'_{|\{p_1,p_2,p_3\}}$.
- We can then invoke independence from irrelevant alternatives and conclude that the same holds for $\rho_{|\{p_1,p_2,p_3\}}$, i.e., $p_1 \succ^S p_3$.
- But then, since the preference profiles and $p_3$ are arbitrary, it must be the case that $\mathcal{J}$ is dictatorial between $p_1$ and $p_3$.

# Proof of Arrow's Impossibility Theorem III

- Next repeat the same argument for $\rho_{|\{p_1,p_2,p_4\}}$ and $\rho'_{|\{p_1,p_2,p_4\}}$, except that now $p_4 \succ_i p_2$ and $p_4 \succ'_i p_1 \succ'_i p_2$ for $i \in \mathcal{J}$, while $p_2 \succ'_j p_1$ and $p_4 \succ'_j p_1$ for all $j \in \mathcal{H} \backslash \mathcal{J}$.
- Then, the same chain of reasoning, using the facts that $\mathcal{J}$ is strongly decisive, $p_1 \succ'^S p_2$, $\phi$ is weakly Paretian and satisfies independence from irrelevant alternatives, implies that $\mathcal{J}$ is dictatorial between $p_4$ and $p_2$ (that is, $p_4 \succ^S p_2$ for any preference profile $\rho \in \Re^H$).
- Now once again using independence from irrelevant alternatives and also transitivity, for any preference profile $\rho \in \Re^H$, $p_4 \succ_i p_3$ for all $i \in \mathcal{J}$.
- Since $p_3, p_4 \in \mathcal{P}$ were arbitrary, this completes the proof that $\mathcal{J}$ is dictatorial (i.e., dictatorial for all $p, p' \in \mathcal{P}$).

# Proof of Arrow's Impossibility Theorem IV

- **Step 2:** Given the result in Step 1, if we prove that some individual $h \in \mathcal{H}$ is strongly decisive for some $p_1, p_2 \in \mathcal{P}$, we will have established that it is a dictator and thus $\phi$ is dictatorial. Let $\mathcal{D}_{ab}$ be the strongly decisive set between $p_a$ and $p_b$.
- Such a set always exists for any $p_a$, $p_b \in \mathcal{P}$, since $\mathcal{H}$ is itself a strongly decisive set. Let $\mathcal{D}$ be the minimal strongly decisive set (meaning the strongly decisive set with the fewest members).
- This is also well-defined, since there is only a finite number of individuals in $\mathcal{H}$.
- Moreover, without loss of generality, suppose that $\mathcal{D} = \mathcal{D}_{12}$ (i.e., let the strongly decisive set between $p_1$ and $p_2$ be the minimal strongly decisive set).
- If $\mathcal{D}$ a singleton, then Step 1 applies and implies that $\phi$ is dictatorial, completing the proof.

# Proof of Arrow's Impossibility Theorem V

- Thus suppose that $\mathcal{D} \neq \{i\}$. Then, by unrestricted domain, the following preference profile (restricted to $\{p_1, p_2, p_3\}$) is feasible

$$\begin{array}{cc} \text{for } i \in \mathcal{D} & p_1 \succ_i p_2 \succ_i p_3 \\ \text{for } j \in \mathcal{D} \setminus \{i\} & p_3 \succeq_j p_1 \succ_j p_2 \\ \text{for } k \notin \mathcal{D} & p_2 \succ_k p_3 \succ_k p_1. \end{array}$$

- By hypothesis, $\mathcal{D}$ is strongly decisive between $p_1$ and $p_2$ and therefore $p_1 \succ^S p_2$.
- Next if $p_3 \succ^S p_2$, then given the preference profile here, $\mathcal{D} \setminus \{i\}$ would be strongly decisive between $p_2$ and $p_3$, and this would contradict that $\mathcal{D}$ is the minimal strongly decisive set.

# Proof of Arrow's Impossibility Theorem VI

- Thus $p_2 \succsim^S p_3$. Combined with $p_1 \succ^S p_2$, this implies $p_1 \succ^S p_3$. But given the preference profile here, this implies that $\{i\}$ is strongly decisive, yielding another contradiction.
- Therefore, the minimal strongly decisive set must be a singleton $\{h\}$ for some $h \in \mathcal{H}$. Then, from Step 1, $\{h\}$ is a dictator and $\phi$ is dictatorial, completing the proof.

# Gibbard-Satterthwaite Theorem

- Another issue: so far, we have assumed that people will express their preferences **truthfully**. But in the same way that we have to ensure truthfulness in implementing mechanisms, we have to make sure that our social choice rules provide incentives to report preferences truthfully.
- We say that a social ordering $\phi : \Re^H \to \Im$ is strategy proof if when $\phi$ is being implemented, all individuals have a **dominant strategy** of representing their preferences truthfully.
- More explicitly, we now have a game, in which each individual reports preference profile $\hat{R}_i \in \Re$ but $\hat{R}_i$ need not be the same as the true preference of this individual, $R_i$.
- **Question:** What types of restrictions does strategy proofness impose?

# Gibbard-Satterthwaite Theorem

- Suppose again unrestricted domain for the preferences and denote a social ordering by $\phi$.

**Theorem**

**(Gibbard-Satterwhite Theorem)** *Suppose there are at least three alternatives in $\mathcal{P}$, and suppose that $\phi(\Re) = \mathcal{P}$ (full range). Then if $\phi$ is strategy proof, it must be dictatorial.*

- This theorem therefore implies that even without the requirement of Arrow Theorem (in particular, weak Paretian and independence of irrelevant alternatives), unless we restrict the set of allowable preferences, only dictatorial social choice functions are allowed.

# Proof of Gibbard-Satterthwaite Theorem

- The proof of the Gibbard-Satterthwaite Theorem follows from the following two lemmas.

**Lemma**

*If $\phi(\Re) = \mathcal{P}$ and $\phi$ is strategy proof, then it is weakly Paretian.*

**Lemma**

*If $\phi(\Re) = \mathcal{P}$ and $\phi$ is strategy proof, then it satisfies independence from irrelevant alternatives.*

- Given these two lemmas, the theorem follows from Arrow's Theorem.
- Throughout the proofs, $\rho = (R_1, ..., R_H)$ refers to the preference profile represented by $\succ_1, ..., \succ_H$ and the alternative profile $\rho = (R_1', ..., R_H')$ is represented by $\succ_1', ..., \succ_H'$, and so on. $\hat{\rho} = \left(\hat{R}_1, ..., \hat{R}_H\right)$ refers to the reported profile

# Proof of First Lemma

- Suppose to obtain a contradiction that that $\phi$ is strategy proof but not weakly Paretian. Then there exists $p \succ_i p'$ for all $i$, but $\phi(\rho) = p'$. Since $\phi(\Re) = \mathcal{P}$, there exists $\rho'$ such that $\phi(\rho') = p$.
- Let $p_n = \phi(R_1', ... R_n', R_{n+1}, ..., R_H)$. Clearly $p_0 = p'$ and $p_H = p$. Then

$$n^* = \min\{H \geq n > 0 : p_n \succ_n p_{n-1}\}$$

  exists (suppose not, then $p' = p_0 \succeq_1 p_1 \succeq_2 p_2 ... \succeq_{n-1} p_{n-1} \succeq_n p_n = p$, which contradicts $p \succ_i p'$ for all $i$ together with transitivity).
- By the unrestricted domain, the profile $\rho''$ such that $R_i'' = R_i'$ for all $i < n^*$ and $R_i'' = R_i$ for all $i \geq n^*$ is possible. Given this profile and truthtelling by all $i$, individual $n^*$ strictly prefers to report $\hat{R}_{n^*} = R_n'$ instead of $\hat{R}_{n^*} = R_n$, contradicting strategy proofness.

# Proof of Second Lemma

- We will now prove this lemma for the case where preference profiles belong to $\bar{\Re}^H \subset \Re^H$, where $\bar{\Re}$ includes only strict preferences.
- Suppose to obtain a contradiction that that $\phi$ is strategy proof but does not satisfy independence of irrelevant alternatives. Then there exist $\rho$, $\rho' \neq \rho$, $p$ and $p' \neq p$ (given the restriction to $\bar{\Re}$, this implies that $p \succ_i p'$ or $p' \succ_i p$ for each $i$) such that:
  - $\rho$ and $\rho'$ have the same ranking over $p$ and $p'$ for all $i$ (i.e., $p \succ_i p' \Longrightarrow p \succ_i p'$ and $p' \succ_i p \Longrightarrow p' \succ_i p$).
  - $\phi(\rho) = p$ and $\phi(\rho') = p'$.
- But then again defining $p_n = \phi(R_1', \ldots R_n', R_{n+1}, \ldots, R_H)$, there exists

$$n^* = \min\{H \geq n > 0 : p_n \neq p_{n-1}\}.$$

- Since $\rho \in \bar{\Re}$, either $p_{n^*} \succ_{n^*} p_{n^*-1}$ or $p_{n^*-1} \succ_{n^*-1} p_{n^*}$, and in either case $n^*$ can strictly gain by misreporting (either when true preferences are $R_{n^*}$ or $R'_{n^*}$), contradicting strategy proofness.

# Proof of Gibbard-Satterthwaite Theorem

- The two lemmas together imply that when preference profiles belong to $\bar{\mathfrak{R}}^H$, strategy proofness implies a dictatorial social choice function, say with the dictator given by individual $i^*$.
- The proof of theorem is completed by showing that on the domain $\mathfrak{R}^H$, if $i^*$ is not a dictator, then either $i^*$ or another individual would have a strict incentive to misreport.

# The Condorcet Winner

- We can avoid the Condorcet paradox when there is a Condorcet winner.

### Definition

*A* **Condorcet winner** *is a policy $p^*$ that beats any other feasible policy in a pairwise vote.*

# Single-Peaked Preferences

- Suppose $\mathcal{P} \subset \mathbb{R}$.

**Definition**

*Consider a finite set of $\mathcal{P} \subset \mathbb{R}$ and let $p(\alpha_i) \in \mathcal{P}$ be individual i's unique bliss point over $\mathcal{P}$. Then, the policy preferences of citizen i are* **single peaked** *iff:*

$$\text{For all } p'', p' \in \mathcal{P}, \text{ such that } p'' < p' \leq p(\alpha_i) \text{ or } p'' > p' \geq p(\alpha_i),$$
$$\text{we have } U(p''; \alpha_i) < U(p'; \alpha_i).$$

- Essentially strict quasi-concavity of $U$
- Median voter: rank all individuals according to their bliss points, the $p(\alpha_i)$'s. Suppose that $H$ odd. Then, the median voter is the individual who has exactly $(H-1)/2$ bliss points to his left and $(H-1)/2$ bliss points to his right.
- Denote this individual by $\alpha_m$, and his bliss point (ideal policy) by $p_m$.

# Median Voter Theorem

## Theorem

**(The Median Voter Theorem)** *Suppose that $H$ is an odd number, that A1 and A2 hold and that all voters have single-peaked policy preferences over a given ordering of policy alternatives, $\mathcal{P}$. Then, a Condorcet winner always exists and coincides with the median-ranked bliss point, $p_m$. Moreover, $p_m$ is the unique equilibrium policy (stable point) under the open agenda majoritarian rule, that is, under A1-A3.*

- This also immediately implies:

## Corollary

*With single peaked preferences, there exists a social ordering $\phi$ that satisfies independence from irrelevant alternatives and that is transitive, weakly Paretian and non-dictatorial.*

# Proof of the Median Voter Theorem

- The proof is by a "separation argument".
- Order the individuals according to their bliss points $p(\alpha_i)$, and label the median-ranked bliss point by $p_m$.
- By the assumption that $H$ is an odd number, $p_m$ is uniquely defined (though $\alpha_m$ may not be uniquely defined).
- Suppose that there is a vote between $p_m$ and some other policy $p'' < p_m$.
- By definition of single-peaked preferences, for every individual with $p_m < p(\alpha_i)$, we have $U(p_m; \alpha_i) > U(p''; \alpha_i)$.
- By A2, these individuals will vote sincerely and thus, in favor of $p_m$.
- The coalition voting for supporting $p_m$ thus constitutes a majority.
- The argument for the case where $p'' > p_m$ is identical.

# Median Voter Theorem: Discussion

- Odd number of individuals to shorten the statement of the theorem and the proof.
- It is straightforward to generalize the theorem and its proof to the case in which $H$ is an even number.
- More important: does it also help us against the Gibbard-Satterthwaite Theorem?
- The answer is Yes.
- In particular, with single peaked preferences, **sincere voting** (truthful revelation of preferences) is optimal, which implies **strategy proofness**.

# Strategic Voting

A2$'$. *Strategic voting.* Define a **vote function** of individual $i$ in a pairwise contest between $p'$ and $p''$ by $v_i(p', p'') \in \{p', p''\}$. Let a voting (counting) rule in a society with $H$ citizens be $V{:}\{p', p''\}^H \rightarrow \{p', p''\}$ for any $p', p'' \in \mathcal{P}$.
Let $V(v_i(p', p''), v_{-i}(p', p''))$ be the policy outcome from voting rule $V$ applied to the pairwise contest $\{p', p''\}$, when the remaining individuals cast their votes according to the vector $v_{-i}(p', p'')$, and when individual $i$ votes $v_i(p', p'')$.
Strategic voting means that

$$v_i\left(p', p''\right) \in \arg \max_{\tilde{v}_i(p', p'')} U\left(V\left(\tilde{v}_i\left(p', p''\right), v_{-i}\left(p', p''\right)\right); \alpha_i\right).$$

- Recall that a *weakly-dominant* strategy for individual $i$ is a strategy that gives weakly higher payoff to individual $i$ than any of his other strategies regardless of the strategy profile of other players.

# Median Voter Theorem with Strategic Voting

### Theorem

**(The Median Voter Theorem With Strategic Voting)** *Suppose that H is an odd number, that A1 and A2′ hold and that all voters have single-peaked policy preferences over a given ordering of policy alternatives, $\mathcal{P}$. Then, sincere voting is a weakly-dominant strategy for each player and there exists a unique weakly-dominant equilibrium, which features the median-ranked bliss point, $p_m$, as the Condorcet winner.*

- Notice no more “open agenda”. Why not?
- Why emphasis on weakly-dominant strategies?

# Proof of the Median Voter Theorem with Strategic Voting

- The vote counting rule (the political system) in this case is majoritarian, denoted by $V^M$.
- Consider two policies $p', p'' \in \mathcal{P}$ and fix an individual $i \in \mathcal{H}$.
- Assume without loss of any generality that $U(p'; \alpha_i) \geq U(p''; \alpha_i)$.
- Suppose first that for any $v_i \in \{p', p''\}$, $V^M(v_i, v_{-i}(p', p'')) = p'$ or $V^M(v_i, v_{-i}(p', p'')) = p''$, that is, individual $i$ is *not* pivotal.
- This implies that $v_i(p', p'') = p'$ is a best response for individual $i$.
- Suppose next that individual $i$ is pivotal, that is, $V^M(v_i(p', p''), v_{-i}(p', p'')) = p'$ if $v_i(p', p'') = p'$ and $V^M(v_i(p', p''), v_{-i}(p', p'')) = p''$ otherwise. In this case, the action $v_i(p', p'') = p'$ is clearly a best response for $i$.
- Since this argument applies for each $i \in \mathcal{H}$, it establishes that voting sincerely is a weakly-dominant strategy and the conclusion of the theorem follows.

# Strategic Voting in Sequential Elections

- But sincere voting no longer optimal in dynamic situations.

$$\begin{array}{ll} 1 & a \succ b \succ c \\ 2 & b \succ c \succ a \\ 3 & c \succ b \succ a \end{array}$$

- These preferences are clearly single peaked (e.g., alphabetical order).
- Dynamic voting set up: first, a vote between $a$ and $b$. Then, the winner goes against $c$, and the winner of this is the social choice.
- Sincere voting will imply that in the first round players 2 and 3 will vote for $b$, and in the second round, players 1 and 2 will vote for $b$, which will become the collective choice.
- However, when players 1 and 2 are playing sincerely, in the first round player 3 can deviate and vote for $a$ (even though she prefers $b$), then $a$ will advance to the second round and would lose to $c$.
- Consequently, the social choice will coincide with the bliss point of player 3. What happens if all players are voting strategically?

# Strategy Proofness with Single Peaked Preferences

- The above example notwithstanding, single peaked preferences are sufficient to ensure strategy proofness. In particular:

### Theorem

*Suppose preferences are single peaked. Then there exists a social choice rule $\phi$ that is strategy proof.*

- Moreover, it can be shown that in this case strategy proofness implies that the social choice rule must be an "augmented median voter" rule, which essentially selects the median from a list of the bliss points of all individuals augmented by additional choices (so a dictatorial social choice rule as well as certain other rules are augmented median voter rules).

# Group Decisions under Incomplete Information

- Consider the following **common value** decision problem by a group, similar to a jury problem.
- Each of $n$ individuals have a prior $\pi$ that a defendant is guilty, denoted by $\theta = G$. (Or this could be some other underlying state relevant for the decision).
- The alternative is $\theta = I$ (for innocent).
- In addition, each individual 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

$$\begin{aligned} \Pr(s = g \mid \theta = G) &= p, \text{ and} \\ \Pr(s = i \mid \theta = I) &= q \end{aligned}$$

# Decisions and Payoffs

- Suppose that the group requires **unanimity** to take a decision $x = G$. This is a natural assumption for juries, but also applies in many situations in which there is a **status quo**.
- Suppose also that each member $j$ of the group has the following payoff:

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

- This payoff allows for a wrong conviction to have a different cost than a wrong acquittal.
- It also implies that the "optimal" decision is

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

# Bayesian Nash Equilibrium

- The Bayesian Nash equilibrium here has to take into account that others will vote according to their signal, which is also informative.
- Throughout let $j \neq 1$. Then

$$u_1(x, \theta) = \begin{cases} 0 & \text{if } x_j = G \text{ for all } j \neq 1 \text{ and } x_1 = G \text{ and } \theta = G \\ 0 & \text{if } x_j = I \text{ for some } j \neq 1 \text{ or } x_1 = I \text{ and } \theta = I \\ -z & \text{if } x_j = G \text{ for all } j \neq 1 \text{ and } x_1 = G \text{ and } \theta = I \\ -(1-z) & \text{if } x_j = I \text{ for some } j \neq 1 \text{ or } x_1 = I \text{ and } \theta = G \end{cases}$$

- **The paradox of Nash equilibrium**: suppose all others vote according to their signal (i.e., $x_j = s_j$ for all $j \neq 1$). Then the utility of individual 1 can be written as

$$u_1 = \begin{cases} 0 & \text{if } s_j = G \text{ for all } j \neq 1 \text{ and } x_1 = G \text{ and } \theta = G \\ 0 & \text{if } s_j = I \text{ for some } j \neq 1 \text{ or } x_1 = I \text{ and } \theta = I \\ -z & \text{if } s_j = G \text{ for all } j \neq 1 \text{ and } x_1 = G \text{ and } \theta = I \\ -(1-z) & \text{if } s_j = I \text{ for some } j \neq 1 \text{ or } x_1 = I \text{ and } \theta = G \end{cases}$$

# Convicting the Innocent

- In light of this, the relevant probability for an individual to vote according to his signal is

$$\Pr(\theta = G \mid s_j = g \text{ for all } j \neq 1 \text{ and } s_1 = i)$$

- Why? Because if, when all individuals are voting following their signals, $s_j = i$ for some $j \neq 1$, individual 1 is not **pivotal**. His decision does not matter.

# Convicting the Innocent (continued)

- Now, we have from Bayes rule

$$
\begin{aligned}
& \Pr(\theta = G \mid s_j = g \text{ for all } j \neq 1 \text{ and } s_1 = i) \\
= \; & \frac{\Pr(s_j = g \text{ for all } j \neq 1 \text{ and } s_1 = i \mid \theta = G)\Pr(\theta = G)}{\begin{array}{c}\Pr(s_j = g \text{ for all } j \neq 1 \text{ and } s_1 = i \mid \theta = G)\Pr(\theta = G) \\ +\Pr(s_j = g \text{ for all } j \neq 1 \text{ and } s_1 = i \mid \theta = I)\Pr(\theta = I)\end{array}}
\end{aligned}
$$

# Convicting the Innocent (continued)

- Or in other words

$$
\begin{aligned}
& \Pr\left(\theta = G \mid s_j = g \text{ for all } j \neq 1 \text{ and } s_1 = i\right) \\
= & \frac{(1-p)\,p^{n-1}\pi}{(1-p)\,p^{n-1}\pi + q\,(1-q)^{n-1}(1-\pi)} \\
= & \frac{1}{1 + \frac{q}{1-p}\left(\frac{1-q}{p}\right)^{n-1}\frac{1-\pi}{\pi}}.
\end{aligned}
$$

- Since $p > 1/2 > 1-q$, for $n$ large, this number is close to 1. Therefore, for any $z < 1$, **it would be optimal to vote to convict even if you have a signal that the defendant is innocent.**

# Convicting the Innocent (continued)

- Of course,the above argument suggests that all individuals vote in according to their signals cannot be an equilibrium.
- In general, there exists a mixed strategy equilibrium, in which all individuals vote to convict when $s_j = g$, and mix with probability $\alpha \in (0, 1]$ to convict when $s_j = i$. This mixed equilibrium is found by setting

$$\Pr(\theta = G \mid x_j = g \text{ for all } j \neq 1 \text{ and } s_1 = i) = z$$

given the mixed strategy profile of others.

# Convicting the Innocent (continued)

- Namely, from the same argument, we have

$$
\begin{aligned}
& \Pr\left(\theta = G \mid x_j = g \text{ for all } j \neq 1 \text{ and } s_1 = i\right) \\
= \; & \frac{(1-p)\left(p+(1-p)\,\alpha\right)^{n-1}\pi}{(1-p)\left(p+(1-p)\,\alpha\right)^{n-1}\pi + q\left(1-q\left(1-\alpha\right)\right)^{n-1}\left(1-\pi\right)} \\
= \; & \frac{1}{1+\frac{q}{1-p}\left(\frac{1-q(1-\alpha)}{p+(1-p)\alpha}\right)^{n-1}\frac{1-\pi}{\pi}} \\
= \; & z.
\end{aligned}
$$

- Naturally, the probability that an innocent defendant will be convicted can be quite high and is increasing in $n$.

# Convicting the Innocent (continued)

- In particular,

$$\alpha = \frac{pK(n) - (1-q)}{q - (1-p)K(n)},$$

where

$$K(n) \equiv \left( \frac{\pi}{1-\pi} \frac{(1-p)(1-z)}{qz} \right)^{1/(1-n)}.$$

- Clearly, as $n \to \infty$, $K(n) \to 1$, so that $\alpha \to 1$, and the innocent is convicted with a very high probability.
- Interestingly, it can also be shown that the larger is the jury the more likely is the innocent to be convicted.
- This model therefore illustrates potential problems that group decisions can face.
- Of course, in this case, directly communicating signals will solve the problem. However, in general such communication would also need to be strategic (another topic for another course...).

6.254 Game Theory with Engineering Applications
Spring 2010