I.1 Voter Behavior

Voters have to decide whether or not to turnout to vote, who to vote for if they do turnout, and how much information to acquire about the options they face.

Economic theorists have addressed all these issues.

I.1.i Voter Turnout

Turnout refers to the fraction of eligible voters who show up to vote.

In the U.S., there is considerable variation in turnout both across and within types of elections.

Turnout is obviously key to understanding elections because who shows up determines who wins.

In addition, from a strategic viewpoint parties are anxious to “bring out their base” and this may impact the type of candidates they run and/or the policy stances they take.

There is a huge academic literature documenting and trying to understand voter turnout - see Feddersen (2004) for an excellent review.

Some Facts about Turnout

Voter turnout in the U.S. in the 2008 presidential election was around 62%.

Voter turnout in the 2010 midterm elections was around 41%.

There is variation across states - for example, in the 2010 midterms turnout was 35% in New York and 55% in Maine.

Turnout in local elections, such as school board elections, can be extremely low (e.g., 10%) when they are held separately from other elections.

Turnout tends to be higher for close races, where closeness is measured by pre-election polls.

The likelihood of voting is positively correlated with education, age, income, religiosity, and being married.

Some people are very regular voters, others sporadic, and still others never vote.

Turnout can be influenced by campaigns.

Gerber and Green have done a series of field experiments whereby they arrange with campaigns to provide different “treatments” to different groups of voters (see for example their paper in the 2000 *American Journal of Political Science*).

Gerber and Green find, for example, that door-to-door campaigning is very effective at getting people to turn out.

Sending campaign material through the mail or calling people on the phone is less effective.

The Calculus of Voting Model

The standard economic approach to the voting decision is the *Calculus of Voting Model*.

This model treats the decision to vote as a cost-benefit calculation.

Consider a voter deciding whether to vote in an election between two candidates A and B.

Suppose that the voter will obtain a utility level V_A if A is elected and a utility level V_B if B is elected.

Differences in these utilities will reflect differences in the policies the candidates are expected to pursue and also such things as how attractive or annoying the voter finds the candidate.

Assume that $V_A > V_B$ and let $\Delta V = V_A - V_B$.

Suppose that the utility cost of going to the polling place is c.

This cost will depend on things like the weather, how busy the voter is that day, how the voter is feeling, etc.

According to the calculus of voting model, the voter will decide to vote if

$$p\Delta V + d \geq c.$$

The variable p is the probability that the voter assigns to the event that his vote will be instrumental in bringing about A's victory.

This is referred to as the probability that the voter is *pivotal*.

The variable d measures the *non-instrumental benefit* the voter gets from voting; i.e., any benefit he gets which is independent of the potential impact of his vote on the outcome of the election.

To get a theory of voter turnout it is necessary to understand how these variables vary across elections.

Most of the literature focuses on variations in the $p\Delta V$ term.

More recent literature has turned to the d term.

We will discuss both.

The Pivotal-Voter Model

The *Pivotal-voter model* endogenizes p via game theoretic techniques.

The basic idea is that p depends on the strategic decisions of voters as to whether to turnout.

It cannot be an equilibrium for everyone to vote for then p will be too small, but it cannot be an equilibrium for no-one to vote because then p will be too high.

To illustrate how such models work, consider a community that is holding a referendum.

There are n citizens, indexed by $i \in \{1, ..., n\}$.

These citizens are divided into supporters and opposers of the referendum.

Supporters are each willing to pay b for the proposed change, while opposers are each willing to pay x to avoid it.

Each citizen knows whether he is a supporter or an opposer, but does not know the number of citizens in each category.

All citizens know that the probability that a randomly selected individual is a supporter is μ.

Citizens must decide whether to not to vote in the referendum.

If they do vote, supporters vote in favor and opposers vote against.

If the number of votes in favor of the referendum is at least as big as the number against, the proposed change is approved.

Each citizen's cost of voting on the day of the referendum is ex ante uncertain.

This reflects the fact that the voting cost will depend on idiosyncratic stuff such as how the citizen as feeling, whether his car is working, etc.

We will model this by assuming that citizen i's cost of voting c_i is the realization of a random variable distributed on $[0, c_{\text{max}}]$ with CDF $F(c_i)$.

Each citizen observes his own voting cost.

However, he only knows that the costs of his fellows are the independent realizations of $n-1$ random variables distributed on $[0, c_{\text{max}}]$ according to $F(c_i)$.

The pivotal-voter model assumes that the only benefit of voting is the instrumental benefit of changing the outcome.

Since the probability of being pivotal depends upon who else is voting, voting is a strategic decision.

Accordingly, the situtation is modelled as a *game of incomplete information*.

The incomplete information concerns the preferences and voting costs of the other voters.

A *strategy* for a citizen i is a function which for each possible realization of his voting cost specifies whether he will vote or abstain.

The equilibrium concept is *Bayesian Nash equilibrium*.

Roughly speaking, every citizen must be happy with his strategy given he knows (i) what strategies others are playing and (ii) the statistical information μ and $F(c_i)$.

Equilibrium

We look for a *symmetric* equilibrium in which all supporters use the same strategy and all opposers use the same strategy.

With no loss of generality, we can assume that supporters and opposers use "cut-off" strategies that specify that they vote if and only if their cost of voting is below some critical level.

Accordingly, a symmetric equilibrium is characterized by a pair of numbers γ_s^* and γ_o^* representing the cut-off cost levels of the two groups.

To characterize the equilibrium cut-off levels, consider the situation of some individual i.

Suppose that the remaining $n-1$ individuals are playing according to the equilibrium strategies; i.e., if they are supporters (opposers) they vote if their voting cost is less than γ_s^* (γ_o^*).

Let $\rho(v_S, v_O; \gamma_S^*, \gamma_O^*)$ denote the probability that v_S of the $n-1$ individuals vote in support and v_O vote in opposition when they play according to the equilibrium strategies.

We discuss how to compute this below.

Recall that the referendum passes if and only if at least as many people vote for it as vote against it.

Thus, if citizen i is a supporter, he will be pivotal whenever v of the $n-1$ other individuals vote in opposition and $v-1$ vote in support.

In all other circumstances, his vote does not impact the outcome.

Accordingly (assuming n is even), the expected benefit of i voting is

$$\sum_{v=1}^{n/2} \rho(v-1, v; \gamma_S^*, \gamma_O^*) b.$$

Individual i will wish to vote if this expected benefit exceeds his cost of voting.

Accordingly, in equilibrium

$$\sum_{v=1}^{n/2} \rho(v-1, v; \gamma_s^*, \gamma_o^*) b = \gamma_s^*. \qquad (1)$$

If citizen i is an opposer, he will be pivotal whenever v of the $n-1$ other individuals vote in opposition and v vote in support.

In all other circumstances, his vote does not impact the outcome.

Accordingly, the expected benefit of i voting is

$$\sum_{v=0}^{n/2-1} \rho(v, v; \gamma_s^*, \gamma_o^*) x.$$

In equilibrium, we have that:

$$\sum_{v=0}^{n/2-1} \rho(v, v; \gamma_s^*, \gamma_o^*)x = \gamma_o^*. \qquad (2)$$

Equations (1) and (2) give us two equations in the two unknown equilibrium variables γ_s^* and γ_o^*.

Existence of an equilibrium pair γ_s^* and γ_o^* for any given values of the exogenous parameters is not a problem (see Palfrey and Rosenthal (1985)), but there might in principle be multiple solutions.

Any equilibrium (γ_s^*, γ_o^*) is going to generate a probability distribution over turnout levels.

It is also going to generate a probability distribution over whether the referendum passes or fails.

Technical detail

To compute equilibria we need to know the function $\rho(v_s, v_o; \gamma_s^*, \gamma_o^*)$.

Let $P(s)$ denote the probability that s of the $n-1$ other citizens are supporters: i.e.,

$$P(s) = \binom{n-1}{s} \mu^s (1-\mu)^{n-1-s}.$$

If there are s supporters, the probability that $v_s \in \{1, ..., s\}$ vote in support is

$$\binom{s}{v_s} F(\gamma_s^*)^{v_s} (1 - F(\gamma_s^*))^{s-v_s}.$$

Similarly, the probability that $v_o \in \{1, ..., n-1-s\}$ vote in opposition is

$$\binom{n-1-s}{v_o} F(\gamma_o^*)^{v_o} (1 - F(\gamma_o^*))^{n-1-s-v_o}.$$

Thus,

$$\rho(\cdot) \;=\; \sum_{s=v_s}^{n-1-v_o} \binom{s}{v_s} F(\gamma_s^*)^{v_s}(1-F(\gamma_s^*))^{s-v_s} \cdot \binom{n-1-s}{v_o} F(\gamma_o^*)^{v_o}(1-F(\gamma_o^*))^{n-1-s-v_o} P(s).$$

Critiques of the pivotal-voter model

This theory of turnout has been heavily criticized over the years.

Critics argue that given the actual number of citizens who choose to vote in large scale elections (such as U.S. presidential elections), the actual probability that a citizen will be pivotal is miniscule.

For example, suppose that a voter knows that 10,000 other voters will definitely vote in an election between two candidates A and B.

Then even if all these voters are equally likely to vote for A as for B (i.e., a really close election), then it can be shown that p is less than 0.006.

With 100,000 other voters definitely voting p is less than 0.002 and with 1,000,000 other voters it is less than 0.0006.

Thus, in order for observed outcomes to be consistent with the fact that very large numbers of citizens vote, the costs of voting for a large group of citizens must clearly be minuscule.

The critics argue that not only does this seem unlikely, but if it were the case, then it is not clear how to explain the variation in turnout observed in the data.

This criticism is formalized by Palfrey and Rosenthal (1985).

They consider a pivotal-voter model in which the range of voting costs is $[c_{\text{min}}, c_{\text{max}}]$ where $c_{\text{min}} < 0 < c_{\text{max}}$.

Thus, some fraction of citizens get a positive benefit from voting because of non-instrumental benefits.

Roughly speaking, they show that the equilibrium cut-offs γ_s^* and γ_o^* converge to 0 as the number of citizens n becomes large.

Thus, in large electorates, the only people who vote are those who get a non-instrumental benefit.

To explain variation in turnout in large scale elections, we need to explain why this fraction changes.

How about small-scale elections?

Coate, Conlin and Moro (2008) analyze how well the pivotal-voter model explains turnout in Texas liquor referenda.

These liquor referenda are very small scale with the number of potential voters often less than 1000.

They find that the model is capable of predicting turnout in the data fairly well, but tends, on average to predict closer elections than are observed in the data.

The pivotal-voter logic implies that elections must be expected to be close even if there is a significant difference between the sizes of the groups supporting the candidates or the intensity of their preferences.

With a very small number of eligible voters, elections that are expected to be close ex ante may end up not close ex post because of sampling error.

For example, an unexpectedly large number of eligible voters may favor one side of the issue or, alternatively, a disproportionate number of eligible voters on one side of the issue may receive low voting cost realizations.

As the number of eligible voters increases, this sampling error very quickly disappears and elections that are expected to be close ex ante will be close ex post.

However, in the data, winning margins can be significant even in the larger elections.

This suggests that the model does not work for either large or small-scale elections.

Defense of the pivotal-voter model

Despite these criticisms, the pivotal-voter theory remains a core model in formal political science.

It is in many respects the simplest and most natural way of thinking about turnout.

Moreover, the basic idea that turnout should be higher in elections that are expected to be close is borne out in the data (see, for example, Shachar and Nalebuff (1999)).

In addition, Levine and Palfrey (2007) find in their experimental study of voter turnout that three key comparative static predictions of the theory are bourne out in the data.

These are: (i) the *size effect* whereby turnout goes down in large elections,

(ii) the *competition effect* whereby turnout is higher in elections that are expected to be close, and

(iii) the *underdog effect* whereby turnout rates are higher among voters supporting the less popular alternative.

Duffy and Tavits (2008) run a voting experiment that allows them to elicit peoples' beliefs about the likelihood that their votes will be pivotal.

They find support for the idea that people who believe that there is higher chance their vote will be pivotal are more likely to vote.

They also find that people substantially overestimate the chance that their votes will be pivotal.

This seems consistent with the fact that popular rhetoric encouraging citizens to participate politically typically stresses the idea that every individual's vote counts.

Still, surveys find that while people do overestimate their chances of being pivotal, they are not way off on this.

Expressive Voting

Dissatisfaction with the pivotal-voter model has led researchers to turn to theorize about the determinants of the d-term (i.e., the non-instrumental benefit of voting).

One way to think about d is that it is the benefit the voter gets from *expressing* his preferences.

According to the expressive view, voting is like cheering at a football game: you do not cheer because you think it is going to impact the outcome, you do it because the game is exciting and cheering is fun.

Under this view, it is natural to assume that d will depend on how strongly the voter feels about the candidates and also how close the election is going to be.

After all, people cheer harder when they care a lot about their team winning and also when the game is close.

Thus, assuming that in a closer election p will be higher, we might write $d = f(\Delta V, p)$ where $f(\cdot)$ is a function with the property that $\partial f / \partial \Delta V > 0$ and $\partial f / \partial p > 0$.

Such a model would then predict that turnout would be higher in elections which are close and in which people feel strongly about the candidates.

It would also be natural to assume that being contacted by a candidate or having more media attention devoted to a race would raise peoples' expressive benefits.

Thus this view is capable of generating predictions consistent with the facts discussed above.

Voting as Civic Duty

Another way to think about the d term is that it represents the payoff of doing your civic "duty".

The idea is that we are taught in civics classes that we should vote and thus we get a "warm-glow" feeling when we do vote (or perhaps avoid feeling guilty!).

Survey evidence reveals that a significant fraction of voters (50%) feel that voting is a moral obligation and would feel guilty if they did not vote.

However, the fact that people turn out in much lower rates in local elections suggests that either the payoff from doing one's duty depends on the type of election under consideration, or perhaps that it is not one's duty to vote in school board elections or elections for town sheriff.

All this suggests that to operationalize this civic duty view we need (i) a theory of what one's duty as a citizen is, and (ii) a theory which tells us what determines the warm-glow from doing one's civic duty.

A Rule Utilitarian View of Duty

One interesting perspective on (i) (i.e., what doing one's duty is) comes from the idea of behaving as a *rule utilitarian*.

A rule utilitarian follows the rule of behavior that if everyone else also followed would maximize aggregate societal utility.

For example, consider the decision as to whether to throw your MacDonald's burger wrapper out of your car.

A standard economic agent would throw it out of his car, reasoning that his one wrapper would do little damage to the environment, but would save him the trouble of putting it in the trash when he got home.

A rule utilitarian would think through to the environmental damage that would happen if everyone threw their wrappers out of their cars and would thus refrain from so doing.

The interesting thing is that the rule utilitarian would not necessarily always vote.

A simple example will illustrate the point.

Consider a society of 2 citizens, Mr 1 and Mr 2.

Suppose that the society is holding a referendum to approve some policy reform.

The reform will be approved if at least one citizen votes in favor.

Each citizen gets a benefit of $B > 0$ if the reform is approved, so that both citizens favor the reform.

Each citizen's cost of voting on the day of the referendum is ex ante uncertain.

This reflects the fact that the voting cost will depend on idiosyncratic stuff such as how the citizen as feeling, whether his car is working, etc.

We will model this by assuming that Mr i's cost of voting c_i is the realization of a random variable uniformly distributed on the interval $[0, c_{\text{max}}]$.

This implies the probability that c_i is less than any given $c \in [0, c_{\text{max}}]$ is just c/c_{max}.

A *voting rule* in this context is a cut-off voting cost c_r below which a citizen will choose to vote.

Lets figure out the rule that a rule utilitarian would use.

This will be the rule that would maximize aggregate societal utility if both voters followed it.

If the cut-off is c_r, the reform will pass if either c_1 or c_2 is less than c_r.

The probability of this is

$$\left(\frac{c_r}{c_{\max}}\right)^2 + 2\left(\frac{c_r}{c_{\max}}\right)(1 - \frac{c_r}{c_{\max}})$$

The first term is the probability that both citizens have voting costs less than c_r and the second term is the probability that one citizen has a voting cost less than c_r.

Cancelling terms, this probability reduces to

$$2\left(\frac{c_r}{c_{\max}}\right) - \left(\frac{c_r}{c_{\max}}\right)^2$$

If a citizen follows this voting rule, his expected voting cost is

$$\int_0^{c_r} c\left(\frac{1}{c_{\max}}\right) dc + \int_{c_r}^{c_{\max}} 0\left(\frac{1}{c_{\max}}\right) dc$$
$$= \frac{c_r^2}{2c_{\max}}$$

Aggregate societal utility under the voting rule c_r is equal to the expected benefits from the reform less the expected voting costs.

Thus, aggregate societal utility is

$$\left(2\left(\frac{c_r}{c_{\max}}\right) - \left(\frac{c_r}{c_{\max}}\right)^2\right) 2B - \frac{c_r^2}{c_{\max}}.$$

The first order condition for the optimal c_r is

$$\left(2\left(\frac{1}{c_{\max}}\right) - 2\left(\frac{c_r}{c_{\max}^2}\right)\right) 2B = 2\frac{c_r}{c_{\max}}$$

Solving this for c_r we find that

$$c_r = \left[\frac{2B}{c_{\mathsf{max}} + 2B}\right] c_{\mathsf{max}}$$

Observe that $c_r < c_{\mathsf{max}}$, which proves that the optimal rule is such that sometimes citizens do not vote.

The optimal rule balances the social benefits from voting with the costs.

The probability that each citizen votes is increasing in B and decreasing in c_{max}.

It follows that the expected level of turnout in the referendum is higher the greater is B and the lower is c_{max}.

If we assume that both voters behave as rule utilitarians, we thus get a nice theory of turnout.

This example is special in the sense that all citizens benefit from the policy reform.

In most elections, there is disagreement and the role of the election is to resolve this disagreement.

The rule utilitarian perspective can be generalized to deal with this case.

Feddersen and Sandroni (2006) develop a theory of turnout in two candidate elections based on the idea that rule utilitarians differ in the candidate who they believe maximizes aggregate utility.

They assume that there are two groups of rule utilitarians with differing views.

One group believes that one candidate is best for the country, the other group believes the other candidate is best.

Individuals in each group choose their voting rule taking as given the behavior of the other group.

A related model is investigated by Coate and Conlin (2004) who assume that individuals are motivated to vote by the ethical desire to "do their part" to help their side win.

Thus, individuals follow the voting rule that, if followed by everyone else on their side of the issue, would maximize their side's aggregate utility.

Individuals therefore act as *group rule-utilitarians*, with their groups being those who share their position.

Coate and Conlin (2004) analyze how well this group rule-utilitarian model explains turnout in Texas liquor referenda.

They find that the comparative static predictions of the model are consistent with the data.

They also structurally estimate the model and show that it out-performs a very simple expressive voting model in which a voter's benefit from expressing a preference just depends on how strongly he feels about the issue.

Thus, thinking about how a rule utilitarian would vote seems to provide a promising theory of what one's duty as a citizen is when it comes to voting.

This still leaves us with the second part of the problem (i.e., (ii)) which is to find a theory which tells us what determines the warm-glow from doing one's civic duty.

Gerber, Green and Larimer (2008) have a very interesting empirical finding that mailings promising to publicize individuals' turnout to their neighbors led to substantially higher turnout.

This perhaps suggests that the warm-glow from doing one's duty partly comes from knowing that others know we did our duty.

I.1.ii Voting in Multicandidate Elections

In elections with three or more candidates, there is an important distinction between *sincere* and *strategic* voting.

To illustrate the distinction, consider an election in which there are three candidates A, B, and C.

Suppose there are n voters labeled voter 1, voter 2, etc.

Let voter i's utility if candidate J is elected be denoted V_{iJ}.

Assume that all voters vote and let $\alpha_i \in \{A, B, C\}$ denote voter i's decision as to who to vote for; i.e., if $\alpha_i = A$, voter i votes for candidate A.

Let $(\alpha_1, ..., \alpha_n)$ denote the voting decisions of all the voters.

Assume the outcome of the election is decided by *plurality rule* which means that the candidate with the most votes is elected.

Further assume that if there are two or more candidates with the most votes (i.e., a tied election), ties are broken by the toss of a fair coin.

Given this, for any given $(\alpha_1, ..., \alpha_n)$ we can figure out the probability that each candidate will be elected.

Thus, if candidate A has the most votes he will be elected with probability 1 and the other two candidates will be elected with probability 0.

If candidates A and B tie for the most votes, then candidates A and B will be elected with probability $1/2$ and candidate C will be elected with probability 0.

Let these election probabilities be denoted by $P_A(\alpha_1, ..., \alpha_n)$, $P_B(\alpha_1, ..., \alpha_n)$, and $P_C(\alpha_1, ..., \alpha_n)$.

Sincere Voting

Voter i is said to be *voting sincerely* if he casts his vote for the candidate he most prefers.

Formally, this means that

$$\alpha_i \in \arg\max\{V_{i\alpha} : \alpha \in \{A, B, C\}\}.$$

[If you are not familiar with the argmax notation it is just the set of arguments that maximize the function in question. Thus, for a function $f(x)$ defined on some set X, the set $\arg\max\{f(x) : x \in X\}$ is the set of values of x that maximize the function f on the set X.]

Thus, if $V_{iA} > V_{iB} > V_{iC}$, and i is voting sincerely then $\alpha_i = A$.

If $V_{iA} = V_{iB} > V_{iC}$, and i is voting sincerely then $\alpha_i \in \{A, B\}$.

Strategic Voting

Voter i is said to be *voting strategically* if he casts his vote so as to maximize his expected payoff.

Formally, given the voting decisions $(\alpha_1, ..., \alpha_n)$, voter i is voting strategically if

$$\alpha_i \in \text{arg max} \left\{ \begin{array}{c} \sum_J P_J(\alpha_1, ., \alpha_{i-1}, \alpha, \alpha_{i+1}, ., \alpha_n) V_{iJ} \\ : \alpha \in \{A, B, C\} \end{array} \right\}.$$

Thus, with strategic voting, the voter takes into account that his vote might impact the outcome of the election.

The key point to note is that voting sincerely may be inconsistent with voting strategically.

For example, let $n = 5$, let $V_{1A} > V_{1B} > V_{1C}$ and let $(\alpha_2, \alpha_3, \alpha_4, \alpha_5) = (B, B, C, C)$.

Then if $\alpha_1 = A$ voter 1 is voting sincerely but not strategically.

The reason is that voter 1's expected payoff from voting for A is

$$\left(\frac{1}{2}\right) V_{1B} + \left(\frac{1}{2}\right) V_{1C}$$

This is less than his expected payoff if he votes for B which is V_{1B}

Note also that with strategic voting, a voter's vote will depend on what other voters are doing.

For example, if $(\alpha_2, \alpha_3, \alpha_4, \alpha_5) = (B, B, A, A)$, voter 1 voting strategically would imply that $\alpha_1 = A$.

If all voters are voting strategically then the voting decisions $(\alpha_1, ..., \alpha_n)$ form a *Nash Equilibrium* of the n player voting game in which voter i chooses a strategy $\alpha_i \in \{A, B, C\}$ and has payoff function

$$\sum_J P_J(\alpha_1, ..., \alpha_n) V_{iJ}.$$

Strategic voting is therefore a *game theoretic* notion.

Sincere or Strategic Voting?

From the perspective of modelling elections, should we assume sincere or strategic voting?

There are many real world examples in which people appear to be behaving strategically.

People commonly justify their decisions for not voting for candidates who are long shots as not wanting to "waste their vote".

For example, in the 2000 presidential election between Gore, Bush, and Nader, many Nader supporters voted for Gore.

This reflects a strategic logic.

However, assuming strategic voting is much more complicated since it makes voting a game theoretic rather than a decision theoretic problem.

Moreover, strategic voting typically fails to deliver a unique prediction.

For example, in the above example let voters' payoffs be:

$$\begin{array}{c} V_{1A} > V_{1B} > V_{1C} \\ V_{2A} > V_{2B} > V_{2C} \\ V_{3B} > V_{3A} > V_{3C} \\ V_{4C} > V_{4B} > V_{4A} \\ V_{5C} > V_{5A} > V_{5B} \end{array}$$

Then, the following voting decisions are consistent with strategic voting: $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) = (A, A, A, C, C)$ and $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) = (B, B, B, C, C)$.

Moreover, unless the election is "close", the best response requirement has no bite.

Thus, $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) = (C, C, C, C, C)$ is consistent with strategic voting!

In game theoretic language, there are "multiple" Nash equilibria.

To try to resolve this problem in voting games it is customary to impose additional requirements on equilibrium

Typically, we require that individuals do not play weakly dominated strategies - this rules out $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) = (C, C, C, C, C)$

We might also use iterated weak dominance to help us rule out possibilities - this can eliminate candidates from contention. For example, suppose that

$$\begin{array}{l} V_{1A} > V_{1B} > V_{1C} \\ V_{2A} > V_{2B} > V_{2C} \\ V_{3A} > V_{3B} > V_{3C} \\ V_{4B} > V_{4A} > V_{4C} \\ V_{5B} > V_{5A} > V_{5C} \end{array}$$

Then the first round of eliminating weakly dominated strategies eliminates C from contention.

The second round implies that $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) = (A, A, A, B, B)$

Another approach is to introduce population uncertainty to give the best response requirement more bite.

This helps matters somewhat but does not resolve the basic indeterminancy exhibited in the above example where voters can coalesce around one or another candidate (i.e., $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) = (A, A, A, C, C)$ and $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) = (B, B, B, C, C)$).

Such indeterminacy seems somewhat realistic.

For example, in the 2008 Democratic primary, would anti-Hilary Clinton Democrats vote for Obama or Edwards?

This was by no means clear early in the campaign.

In reality, such issues seems to be determined dynamically with the help of polls and some coordination among like-minded citizens.

Also candidates may withdraw to shift their supporters to a like-minded candidate.

Myatt (2007) uses techniques from the theory of global games to resolve the indeterminacy.

Sophisticated vs Simple Sincere Voting

Even in a two-candidate election, there some ambiguity in defining what is meant by sincere voting.

This is the case when the elected candidate will determine policy collectively with others - for example, when electing a legislator to a legislature.

To illustrate, consider electing a U.S. senator.

Consider a centrist voter choosing between a moderate Republican and a left wing Democrat.

Suppose that policy-making in the Senate is determined by the party that holds the majority of seats and reflects a compromise between the Senators in the majority party.

Suppose that the Republicans currently hold the majority in the Senate and that most of the Republican senators are right wing Republicans.

Further suppose that most of the Democrat senators are moderate Democrats.

Simple sincere voting would involve the voter voting for the candidate who is closest to his own position; i.e., the Republican.

Sophisticated sincere voting would involve the voter anticipating how the candidate would impact policy-making in the Senate and voting for the candidate whose election would yield the preferred policy outcome.

This could involve voting for the Democrat.

The logic is as follows: electing the Democrat could help switch the Senate from Republican to Democrat control and, since most Democrat senators are moderates, this would result in moderate Democrat policies as opposed to right wing Republican policies.

This logic appeared to be behind the loss of moderate Republican Senator Lincoln Chaffee (Rhode Island) in the 2006 mid-term elections.

A related phenomenon is "*ticket splitting*", which arises when a voter votes for, say, the Democratic presidential candidate and the Republican congressional candidate.

This is quite a common phenomenon in fact and the literature discusses at length what might be going on.

It could be just simple sincere voting and reflect the fact that candidates in the same party have different ideologies.

It could alternatively be sophisticated sincere voting in which voters are trying to achieve the right balance between the ideology of Congress and the president.

For example, a moderate voter may prefer a divided government (e.g., Democrat congress and Republican president) to a unified government (i.e., all Republican or all Democrat).

Finally, it could actually be real strategic voting where voters are taking into account the probability of impacting the outcome.

See Morton's book (pp523-529), Burden and Kimball (1998), Degan and Merlo (2009), and Lacy and Paolini (1998) for more on this topic.

I.1.iii Voter Information

In addition to deciding whether or not to vote and to deciding for whom to vote, voters also have to decide how much information to acquire.

How much do voters actually know about the candidates and/or issues they are voting on?

In general, surveys reveal that voters do not know much at all about politics.

Around 50% of Americans do not know that each state has two senators and 40% cannot name either of their senators.

Over 50% cannot name their congressman and only about 50% know which party controls the House.

The fact that voters do not know much is perfectly consistent with an economic approach.

If acquiring information about politics is costly for voters, then we should expect them to remain *rationally ignorant.*

The benefit of acquiring political information is small because there is almost no chance of a voter's decision changing the outcome.

Thus, we would only expect voters to have such information as can be costlessly acquired (say through tv commercials) or which they enjoy acquiring (like information about scandals)

Significance of Rational Ignorance

Despite rational ignorance, many political scientists argue that voters are able to choose which candidate or policy option is best for them.

There are three distinct arguments.

First, there are plenty of easily acquired signals that voters can use to make the correct choices.

Such signals include party affiliations, newspaper endorsements, endorsements from interest groups, advice from informed friends and relatives, etc.

Second, those who are uninformed are less likely to participate.

There is a great deal of evidence that the less information people have about the options on the ballot the less likely they are to show up to vote.

Indeed, it is common for people not to vote when they lack information even when voting is costless.

This is evidenced by the phenomenon of “roll off” - in which voters voting in “bundled elections” (i.e., those held on the same day) do not vote on the “down ticket” (i.e., less high profile) races.

Thus, people will often vote for president, governor, U.S. senator, etc, but not for state assemblyman or town sheriff.

The interpretation is that people know nothing about these down ticket races and feel it inappropriate to express an opinion even though it would be costless to do so.

Third, even when people think they know what is best for them but are wrong, one can appeal to the so-called *miracle of aggregation.*

Suppose that voters are choosing between candidate A and candidate B, and suppose that candidate A is the better candidate.

If 100% of voters know that candidate A is better, then candidate A will obviously win the election.

But if only 1% of voters know that candidate A is better, then candidate A will still win the election provided that the remaining voters have *unbiased beliefs*.

By unbiased beliefs, I mean they are just as likely to believe A is better as to believe B is better.

With unbiased beliefs, an almost completely ignorant electorate makes the same choice as a fully informed electorate.

This is the miracle of aggregation.

Unbiased Beliefs and Rational Updating

Economic models of voting tend to assume that, while voters may not be fully informed, they have unbiased beliefs, and update these beliefs rationally given the available evidence.

Rational updating means that as information is revealed, voters change their beliefs in the direction of the truth.

In reality, voters may not have unbiased beliefs and update rationally if they enjoy holding certain beliefs.

Suppose that voters like to believe that the budget deficit can be eliminated by eliminating waste, fraud, and abuse.

Then they may repeatedly choose candidates who tell them that the deficit can be eliminated by eliminating waste, fraud, and abuse, over candidates who tell them that taxes need to be raised.

If deficits continue to grow, voters will not necessarily change their beliefs because the personal gain from having the correct beliefs is tiny and the utility loss from having to give up their beliefs may be large.

The gain from having the correct belief is costless since their vote is very unlikely to be pivotal.

The key point is that the individual does not bear a personal cost from having incorrect political beliefs: the cost is a social cost.

This differs from having incorrect beliefs about products: if you believe that Fiats are reliable cars, you buy one and end up spending a lot of time in the shop.

For more on this argument see Bryan Caplan's book: *The Myth of the Rational Voter*.

Voting with Incomplete Information

How would we expect incomplete information to impact voter behavior?

Feddersen and Pesendorfer (1996) present a famous game theoretic analysis of this.

They construct a very simple model which captures the idea that some voters may have better information than others.

They characterize what happens in the model and find some very interesting results.

Their model assumes that people are motivated to vote solely by the instrumental desire to influence the outcome.

There are two "states of nature" indexed by $z \in \{0, 1\}$ and two candidates indexed by $x \in \{0, 1\}$

There are three types of citizens indexed by $t \in \{0, 1, i\}$

Types 0 and 1 are *partisans* who prefer candidates 0 and 1 independent of the state of nature.

Type i are *independents* whose preferences over candidates are given by

$$U(x; z) = \{ \begin{array}{cc} -1 & \text{if } x \neq z \\ 0 & \text{if } x = z \end{array}$$

Thus, independents want to match the candidate with the state.

There are three key assumptions:

A.1) The state of nature is uncertain: the probability that $z = 0$ is α which is less than $1/2$.

Thus, candidate 1 is more likely to be the best candidate for the independents.

A.2) The number and type of citizens is uncertain.

The way this is modelled is by assuming that the number and type of citizens is determined before the election according to the following process.

There are $N + 1$ independent draws.

In each draw there is probability $1 - p_{\emptyset}$ of getting a citizen where $p_{\emptyset} > 0$.

This creates uncertainty in the size of the population which means that there is always a positive probability that any citizen's vote is pivotal.

Conditional on getting a citizen it is an independent with probability $\frac{p_i}{1-p_\emptyset}$, a type 1 partisan with probability $\frac{p_1}{1-p_\emptyset}$, and a type 0 partisan with probability $\frac{p_0}{1-p_\emptyset}$..

Thus, p_i, p_1, and p_0 measure the expected sizes of the groups in the population.

A.3) Some citizens have more information than others.

The way this is modelled is that each citizen receives a *signal* $s \in \{0, 1-\alpha, 1\}$.

If a citizen receives signal s he knows that Pr$\{z = 1\} = s$.

Thus, if a citizen receives the signal $s = 1$, he knows that the state is 1.

If he receives the signal $s = 0$, he knows that the state is 0.

If he receives the signal $s = 1 - \alpha$, he just knows that the probability that the state is 0 is α.

The probability that a citizen is *informed* (i.e., $s \in \{0, 1\}$) is q.

Equilibrium

Each citizen chooses one of three actions: abstain, vote for candidate 0, or vote for candidate 1.

The candidate with the most votes wins and in the event of a tie, each candidate wins with probability 1/2.

F & P analyze symmetric Nash equilibria in which all citizens with type (t, s) choose the same strategy

All citizens except the uninformed independents (i.e., types $(i, 1-\alpha)$) have a strictly dominant strategy.

Partisans vote for their preferred candidate and informed independents vote for the candidate that matches the state.

Results

F & P characterize the equilibrium of the game for N large.

Proposition 1. *Suppose* $q > 0$ *and* $p_i(1-q) < |p_0 - p_1|$. *Then equilibrium involves uninformed independents voting for the candidate with the smallest partisan base; i.e., candidate* 0 *if* $p_0 < p_1$ *and candidate* 1 *if* $p_1 < p_0$

The condition that $p_i(1-q) < |p_0 - p_1|$ tells us that the fraction of uninformed independents is "small".

The novel idea here is that if you do not have full information you should condition your voting decision on what must be true about the state of the world when your vote makes a difference.

If a candidate has less partisan support but nonetheless gets almost as many votes as his opponent he must be being supported by the informed independents.

In this case, he is the best candidate from the viewpoint of the uninformed!

Observe that when uninformed independents are voting for candidate 0 they are voting for the candidate their own information suggest is likely to be worse.

Thus, they are ignoring their own signals!

Proposition 2. *Suppose $p_0 = p_1$. Then equilibrium involves uninformed independents abstaining.*

The condition $p_0 = p_1$ tells us that the two candidates have (in expectation) the same sized partisan bases.

In this case, the uninformed independents should just abstain - thereby delegating the decision to the informed independents.

This is a formal model of "roll off" which we discussed earlier.

Proposition 3. *Suppose $q > 0$ and $p_i(1-q) \geq |p_0 - p_1|$. Then equilibrium involves uninformed independents mixing between voting for the candidate with the smallest partisan base and abstaining.*

Discussion

The basic insight in the F & P paper is familiar from optimal bidding behavior in common value auctions.

If you are bidding in an auction and do not have full information about the object, you should condition your bid on what must be true about others' valuations if your bid is the highest.

If you just bid your signal, you will end up over-paying for the object if you win it - a result known as the *winner's curse*.

In the voting context, the idea has many interesting implications.

For example, it suggests that requiring juries to vote unanimously to convict someone of a crime may lead people to ignore their private information.

F & P also show that for N large the election perfectly *aggregates information* in the sense that the outcome of the election is the same as it would be if all citizens were perfectly informed.

This is another take on the miracle of aggregation.

The F & P analysis has spawned an experimental literature which has studies whether the theoretical predictions are bourne out in the lab.

The "roll-off" prediction of Proposition 2 seems intuitive.

The "off-setting voting" prediction described in Propositions 1 and 2 seems much less intuitive and there is only mixed support for it.

See Battaglini, Morton, and Palfrey (2010) and Esponda and Vespa (2011) for conflicting evidence.

I.2 Candidate Behavior

Candidates are citizens who have chosen to run for elected office.

Candidates have to decide how to campaign.

Once in office, they have to decide how to vote on legislation and which policy issues to pursue.

Economic theory has largely focused on candidate competition for office.

The objective is to understand the positions that candidates will take.

Understanding candidates' positions is seen as an important step in understanding the policies that eventually emerge from the political process.

We will discuss two candidate competition with Downsian and policy-motivated candidates.

We will also discuss the citizen-candidate model and candidate policy choice with re-election concerns.

I.2.i Two Candidate Competition

We begin with the classic Downsian model of two candidate electoral competition.

This model is named after Anthony Downs who wrote a well-known book *An Economic Theory of Democracy* in the 1950s.

In this model, two candidates whose objective is to win office compete by staking out ideological positions in a one dimensional space.

A) The Downsian Model

There are two candidates, A and B

Each candidate $J \in \{A, B\}$ must choose an ideology $i_J \in [0, 1]$ on which to run

The interpretation is that 0 is the most extreme left-wing ideology and 1 the most extreme right-wing ideology

This one dimensional notion of ideology is quite natural, since it is commonplace to talk of candidates as being centrist, extreme right, moderate right, etc.

The notion of choosing ideology is also natural, since it is commonplace to talk of candidates moving to the center, appealing to the base, etc.

If elected, candidates are assumed to pursue policies consistent with the ideology they have run on in the election.

Citizens have ideologies distributed over the interval $[0, 1]$.

Let $F(i)$ be the fraction of voters with ideology $\leq i$. Assume $F(0) = 0$, and $dF/di > 0$.

Let i_m denote the ideology of the *median voter*; that is, the voter exactly in the middle of the ideology distribution.

Formally, i_m is defined by $F(i_m) = 1/2$.

A voter with ideology i obtains utility $u(i_J; i)$ if candidate J is elected.

The utility function $u(i_J; i)$ is assumed to satisfy the *single-crossing property*.

To explain this, let i and i' be two voters such that i' is more right wing than i (i.e., $i < i'$) and suppose that candidate A is to the left of candidate B (i.e., $i_A < i_B$).

Then the utility function u satisfies the single-crossing property if whenever voter i likes candidate B at least as much as candidate A, voter i' strictly prefers candidate B; i.e.,

$$u(i_B; i) \geq u(i_A; i) \Rightarrow u(i_B; i') > u(i_A; i')$$

and whenever voter i' likes candidate A at least as much as candidate B, voter i strictly prefers candidate A; i.e.,

$$u(i_A; i') \geq u(i_B; i') \Rightarrow u(i_A; i) > u(i_B; i).$$

This property is very intuitive and should not be considered much of a restriction.

One example of such a utility function is $u(i_J; i) = -|i_J - i|$.

When voters have these utility functions they are said to have *distance preferences.*

Another example is $u(i_J; i) = -(i_J - i)^2$.

When voters have these utility functions they are said to have *quadratic preferences.*

Candidates get a payoff $r > 0$ if they win and 0 otherwise.

Thus, candidates have no ideological preferences and are purely *office motivated*; i.e., they just want to win.

Each candidate J simultaneously announces his ideology $i_J \in [0, 1]$

Each voter votes for the candidate whose ideology he prefers.

If a voter is indifferent he votes for each candidate with equal probability.

The candidate with the most votes is elected and pursues policies consistent with his announced ideology.

If each candidate has the same number of votes, then the election is decided by the toss of a fair coin.

Equilibrium

Relabelling the candidates if necessary, assume $i_A \leq i_B$

Assuming $i_A < i_B$, let $i^*(i_A, i_B)$ be the ideology of the voter who is just indifferent between candidates A and B.

Formally, $i^*(i_A, i_B)$ is defined by the equation $u(i_A; i^*) = u(i_B; i^*)$

If voters have distance or quadratic preferences, then

$$i^*(i_A, i_B) = \frac{i_A + i_B}{2}.$$

Then given the single-crossing assumption, all those voters for whom $i < i^*(i_A, i_B)$ vote for A, while those voters for whom $i > i^*(i_A, i_B)$ vote for B.

If $\pi(i_A, i_B)$ is the probability that candidate A wins, then

$$\pi(i_A, i_B) = \begin{cases} \frac{1}{2} & \text{if } i_A = i_B \text{ or if } i^*(i_A, i_B) = i_m \\ 1 & \text{if } i^*(i_A, i_B) > i_m \\ 0 & \text{if } i^*(i_A, i_B) < i_m \end{cases}.$$

A pair of ideologies is an *equilibrium* if each candidate is happy with his choice given what the other candidate is choosing.

Formally, (i_A^*, i_B^*) is an equilibrium if

$$i_A^* \in \arg\max \pi(i_A, i_B^*)r$$

and

$$i_B^* \in \arg\max[1 - \pi(i_A^*, i_B)]r.$$

This corresponds to the usual notion of a Nash equilibrium.

Results

Median Voter Theorem (i_A^*, i_B^*) *is an equilibrium if and only if* $i_A^* = i_B^* = i_m$*; that is, each candidate adopts the ideology of the median voter.*

Proof There are two parts to the proof of this theorem.

The first part is to show that if (i_A^*, i_B^*) is an equilibrium, then it must be the case that $i_A^* = i_B^* = i_m$.

To this end, let (i_A^*, i_B^*) be an equilibrium.

It is clear that $\pi(i_A^*, i_B^*) = 1/2$ because each candidate can guarantee that he wins with probability $1/2$ by just adopting the same ideology as his opponent.

Suppose that $i_A^* < i_B^*$.

Then, it follows from the fact that $\pi(i_A^*, i_B^*) = 1/2$ that the median voter must be indifferent between the two candidates; i.e., that $i^*(i_A^*, i_B^*) = i_m$.

But now imagine that candidate A deviates from the equilibrium by proposing the median voter's ideology.

Then, the median voter strictly favors A and hence $i^*(i_m, i_B^*) > i_m$.

It follows that $\pi(i_m, i_B^*) = 1$ implying that A would win with probability 1 - which contradicts the fact that (i_A^*, i_B^*) is an equilibrium.

Thus, we conclude it must be the case that $i_A^* = i_B^*$.

If $i_A^* = i_B^* \neq i_m$ then either $i_A^* = i_B^* < i_m$ or $i_A^* = i_B^* > i_m$.

Consider only the former case, since the latter is similar.

If $i_A^* = i_B^* < i_m$ candidate B could deviate by moving his position to i_m.

Then, $i^*(i_A^*, i_m) < i_m$ implying that $\pi(i_A^*, i_m) = 0$.

This means that candidate B would win with probability 1 - which contradicts the fact that (i_A^*, i_B^*) is an equilibrium.

This completes the proof of the first part.

The second part of the proof is to show that if $i_A^* = i_B^* = i_m$ then (i_A^*, i_B^*) is an equilibrium.

This follows from the fact that any candidate who deviated from i_m would lose and hence reduce his expected payoff. QED

Substantively, this model predicts that competition leads candidates to move to the political center.

In terms of policy, the result suggests that policies in a representative democracy will be those that would be preferred by middle of the road voters.

Empirical Evidence

There have been many attempts to test the Median Voter Theorem.

One of the most convincing is the effort by Gerber and Lewis in the 2004 *Journal of Political Economy*.

They use voting data from Los Angeles County to estimate the distribution of voter ideologies district by district.

In particular, they have votes on both ballot propositions and candidate elections which allows them to do a convincing job of estimating voter ideologies.

They then estimate the ideology of the winning candidates from legislator voting records.

They look at U.S. House and California Assembly races.

They find little support for the idea that the ideology of winning candidates should match the ideology of the median voter in their constituency.

In particular, the ideology of winning candidates can diverge significantly from the median voter's ideology in heterogeneous districts (i.e., districts with a lot of variance in citizen ideologies).

Winning Republicans are to the right of the median voter in their district, while winning Democrats are to the left.

This is consistent with casual empiricism and the findings of most who have looked carefully at the issue.

The one exception is a recent paper by Ferreira and Gyourko in the 2010 *Quarterly Journal of Economics*.

They compare policies in cities with Republican and Democrat mayors.

They use a *regression discontinuity design* which compares policies in cities which elected a Democrat mayor by a very small margin with those who elected a Republican mayor by a very small margin.

The idea behind this research design is that these two groups of cities should be basically quite similar, except for the partisan affiliation of the mayor.

If the Median Voter Theorem is right, both Democrat and Republican mayors should implement basically the same policies.

This means that there should be no difference between the policies in these two groups of cities, which is what they find.

B) Policy-motivated Candidates

The Downsian assumption that candidates are purely office-motivated is strange.

Indeed, it is logically inconsistent because candidates must be citizens and citizens are presumed to have policy preferences.

What would happen if we modified the Downsian model by assuming that the two candidates A and B had, respectively, utility functions $u(i_J; t_A)$ and $u(i_J; t_B)$ with "true" ideologies t_A and t_B where $t_A < i_m < t_B$?

In fact, it can be shown that the Median Voter Theorem is robust to this extension.

However, if candidates face uncertainty in the location of the median voter, their policy preferences do matter.

Candidates move to the center, but not all the way.

Consider the Downsian model and assume for simplicity that voters have quadratic preferences; i.e., $u(i_J; i) = -(i_J - i)^2$

Change the Downsian model by assuming that the candidates have quadratic preferences with true ideal points $t_A = t$ and $t_B = 1 - t$ where $t < 1/2$.

Candidates still get a non-policy related reward to holding office $r > 0$ if they win.

To capture the idea that candidates do know perfectly the distribution of voter preferences, assume that candidates believe that the ideology of the median voter i_m is uniformly distributed on $[1/2 - \varepsilon, 1/2 + \varepsilon]$ where $\varepsilon \in (0, 1/2 - t)$.

This means that the probability that ideology of the median voter is less than any $x \in [1/2 - \varepsilon, 1/2 + \varepsilon]$ is

$$\frac{x - (1/2 - \varepsilon)}{2\varepsilon}$$

If candidates choose ideologies i_A and i_B where $i_A < i_B$, then as before candidate A wins if $i_m < i^*(i_A, i_B)$.

As noted above, with quadratic preferences

$$i^*(i_A, i_B) = \frac{i_A + i_B}{2}.$$

Thus, the probability that candidate A wins is the probability that i_m is less than $\frac{i_A+i_B}{2}$.

Assuming that $\frac{i_A+i_B}{2} \in (1/2-\varepsilon, 1/2+\varepsilon)$, this is given by

$$\begin{aligned}\pi(i_A, i_B) &= \frac{\frac{i_A+i_B}{2} - (1/2 - \varepsilon)}{2\varepsilon} \\ &= \frac{1}{2} + \frac{i_A + i_B - 1}{4\varepsilon}.\end{aligned}$$

A pair of ideologies (i_A^*, i_B^*) is an *equilibrium* if

$$\begin{aligned}i_A^* \in\ & \arg\max \pi(i_A, i_B^*)[r - (i_A - t)^2] \\ & -(1 - \pi(i_A, i_B^*))(i_B - t)^2\end{aligned}$$

and

$$\begin{aligned} i_B^* \in & \arg\max -\pi(i_A^*, i_B)(1-t-i_A)^2 \\ & +(1-\pi(i_A^*, i_B))[r-(1-t-i_B)^2]. \end{aligned}$$

This is a Nash equilibrium as before, but it is more complicated because candidates care about policy.

We say that the equilibrium (i_A^*, i_B^*) is *symmetric* if the two candidates positions are mirror images of each other; that is, if $i_A^* = 1 - i_B^*$.

Proposition *Assume that* $r < 2\varepsilon(1-2t)$*. Then, if* (i_A^*, i_B^*) *is a symmetric equilibrium*

$$i_A^* = 1 - i_B^* = \frac{r/2 + 1/2 - t(1-2\varepsilon)}{2\varepsilon + 1 - 2t}.$$

Proof Candidate A's first order condition is

$$\frac{\partial \pi(i_A^*, i_B^*)}{\partial i_A}[r-(i_A^*-t)^2+(i_B^*-t)^2] = \pi(i_A^*, i_B^*)2(i_A^*-t).$$

Using the expression for $\pi(i_A^*, i_B^*)$ and the fact that the equilibrium is symmetric, this implies that

$$\frac{1}{4\varepsilon}[r - (i_A^* - t)^2 + (i_B^* - t)^2] = (i_A^* - t).$$

Using the fact that $i_A^* = 1 - i_B^*$, we can rewrite this as:

$$\frac{1}{4\varepsilon}[r - (i_A^* - t)^2 + (1 - i_A^* - t)^2] = (i_A^* - t).$$

Expanding this, yields

$$\frac{1}{4\varepsilon}[r + 4i_A^* t + 1 - 2i_A^* - 2t] = (i_A^* - t).$$

We can solve this equation for i_A^* yielding

$$i_A^* = \frac{\frac{r}{2} + \frac{1}{2} - t(1 - 2\varepsilon)}{2\varepsilon + 1 - 2t}.$$

QED

Thus, we see that the candidates move toward the expected ideology of the median voter (since $\frac{r/2+1/2-t(1-2\varepsilon)}{2\varepsilon+1-2t} > t$) but not all the way as long as $r < 2\varepsilon(1-2t)$.

The degree to which candidates move to the center is positively related to the non-policy related reward to holding office (r) and negatively related to the degree of uncertainty in the location of the median voter (ε).

It is also the case that the closer the candidates' true ideologies are to the center (i.e., the higher is t), the closer are the equilibrium positions to the center.

C) Further Developments

Research on models of two candidate competition continues.

An influential recent contribution is by Groseclose (2001) who assumes that candidates have different *valence* characteristics.

A valence characteristic is an exogenous characteristic like honesty, good looks, or intelligence which *all* voters value.

Formally, valence characteristics are introduced by assuming that a voter with ideology i obtains utility $u(i_J; i) + v_J$ if candidate J is elected where v_J measures candidate J's valence.

It turns out that if one candidate has a valence advantage, this can change the equilibrium quite significantly.

Krasa and Polborn (2011) assume that candidates have different abilities with respect to policy-making.

For example, some candidates are good at cutting taxes and others are good at managing large government programs.

This also changes the equilibrium in interesting ways.

A further interesting recent contribution is Kartik and McAfee (2007)

They take a standard Downsian model and introduce unobservable candidate "character".

They motivate this extension from the perceived importance of character in U.S. presidential politics.

The idea is that candidates with character do not pander to public opinion and this makes the position-taking game a signalling game - you do not want to appear to be a panderer even if you have no character.

There are two candidates, indexed by $J \in \{A, B\}$

Each candidate J must choose an ideology $i_J \in [0, 1]$

The fraction of voters with ideology $\leq i$ is $F(i)$ where $F(0) = 0$ and $dF/di > 0$.

A voter with ideology i obtains a payoff $u(i_J; i)$ from ideology i_J, where the function $u(\cdot; i)$ satisfies the single-crossing property.

The median voter's ideology is i_m.

The new feature is that each candidate can have "character" or not.

Let $C_J = 1$ if candidate J has character and $C_J = 0$ if he does not.

Candidates know whether they have character but voters do not.

All they know is that the probability that $C_J = 1$ is $b \in (0, 1)$.

If a candidate has character his ideology choice i_J is non-strategic - it is the realization of a random variable with CDF $H(i)$ and density $h(i)$.

This formalizes the idea that candidates with character do not pander.

If a candidate does not have character then he chooses i_J strategically in the usual way; i.e., he seeks to maximize the probability of winning.

Voters obtain an additional benefit λ from electing a candidate with character (why this is, is not modelled).

Because of this, candidates without character would like to signal they have character.

If they just choose the median voter's ideology then voters will conclude that they do not have character.

Thus, their position choice becomes a signal and a signalling game results.

Equilibrium

If candidate J does not have character his strategy is described by $G_J(i)$; i.e., $G_J(i)$ is the probability he chooses an ideology less than i.

If candidate J has character his strategy is just $H(i)$.

At the time of voting, voters are not sure whether each candidate has character.

Let $\Phi_J(i)$ represent voters' beliefs concerning the probability that candidate J has character given he selects position i.

A *symmetric equilibrium* consists of a candidate strategy $G(i)$ and voter beliefs $\Phi(i)$ such that: i) the strategy is optimal for the candidates given the voting behavior implied by the voters' beliefs, and ii) the voters' beliefs are rational given the candidate strategy.

Proposition. *In equilibrium, strategic candidates mix over positions. In particular, the probability that they choose i_m is zero. Moreover,* i) *candidates without character win with a higher probability than candidates with character; ii) voters' posterior belief on character is single-troughed around the median; that is,* $\Phi(i)$ *is decreasing for $i < i_m$ and increasing for $i > i_m$; iii) as λ goes to ∞, $G(i)$ converges to $H(i)$; and iv) as b goes to* 0, *$G(i)$ converges to the distribution that puts point mass on i_m.*

The paper is interesting in that it identifies an important reason why even the most cynical candidates will not pander totally to the median voter.

However, it does not explain why voters value candidates with "character" or why candidates with character would wish to ignore voters' preferences.

If you are interested in this aspect of the argument see the paper by Callander (2008).

I.2.ii The Citizen-Candidate Approach

Once we recognize that candidates have policy preferences, we must address the question of what determines these preferences.

Thus, in the model just considered what determines t_A and t_B?

To understand this, we have to consider the decision of citizens of whether to run for office.

This also raises the question of the number of candidates who decide to run.

The *citizen-candidate model* seeks to explain the number of candidates running and their policy preferences.

In so doing, it offers a very different vision of candidate behavior: the key assumption of the approach is that candidates must run on their true ideologies.

There are two distinct justifications for this assumption.

First, candidates may prefer to be honest and therefore find it costly to misrepresent their true beliefs.

Second, once elected, candidates' behavior may plausibly be driven more by their true ideologies than the ideologies they announce in the campaign.

If this is the case, voters will recognize that what matters for predicting policy choices will be what the candidate truly believes rather than what he annouces.

The assumption that candidates must run on their true ideologies is fundamentally different than that underlying the prior models which assume candidates can reposition themselves at will.

Casual empiricism seems to suggest that there is some truth to both positions.

Candidates certainly sometimes seem to change their positions (i.e., to flip-flop) and some voters seem to believe them (e.g., Mitt Romney's conversion to a social conservative).

However, the controversy which arises following a flip-flop certainly suggests that voters feel that candidates ought to present their true policy preferences.

There is some empirical work that is relevant to this discussion (see, for example, Lee, Morretti, and Butler (2004)).

In support of the citizen-candidate model, it does not seem to be the case that legislators change their positions (as measured by voting records) when their constituency changes (say, via redistricting).

Moreover, survey evidence suggests that repositioning leads voters to distrust candidates' announced positions and their integrity.

The Model

The citizen-candidate approach models elections as a three stage game.

In Stage 1, citizens decide whether or not to enter the race as a candidate.

Running entails a sunk cost which may be thought of as the time devoted to running a campaign.

In Stage 2, citizens vote over the set of self-declared candidates.

Voting is assumed not to be costly, so everybody votes.

Voting is strategic in some treatments of the model and sincere in others.

In Stage 3, the candidate with the most votes is elected (plurality rule) and follows his/her true ideology when in office.

If there is a tie, the winning candidate is determined by the toss of a fair coin.

When voting in Stage 2, citizens rationally anticipate that the winning candidate will follow his/her true ideology and this determines their payoffs from the different candidates.

Importantly, citizens are assumed to know each candidate's true ideology.

In Stage 1, candidates are assumed to perfectly anticipate how citizens will vote for any given set of candidates.

Roughly speaking, an *equilibrium* is a set of candidates such that, given perfect anticipation of voting behavior, every citizen who is a candidate is better off being in the race given who else is in the race and every citizen who is not a candidate is better off not in the race.

Results

The advantage of the citizen-candidate model is that it endogenizes the number of candidates and their policy positions.

The disadvantage of the model is that it does not yield a unique prediction.

There are many different possible equilibria, some involving spoiler candidates who run just to prevent other candidates from winning.

We will just look at the predictions of the model concerning equilibria with two candidates.

This will enable us to compare predictions with the previous models.

To facilitate comparisons, lets assume the usual set-up in which citizens have ideologies distributed over the interval $[0, 1]$ and utility functions satisfying single crossing.

Assume that running for office entails a cost c and that the winning candidate gets a non-policy related benefit of holding office $r < c$.

First assume that citizens vote sincerely (as in Osborne and Slivinski (1996)).

Two candidates with ideologies i_A and i_B running against each other is an equilibrium if (i) candidates A and B want to run against each other and (ii) there is no third candidate C who would gain from entering the race.

The first prediction of the model is that i_A and i_B will be on opposite sides of the median voter's ideology and the median voter will be indifferent between the two candidates; i.e., $i_A < i_m < i_B$ and $u(i_A; i_m) = u(i_B; i_m)$

The reason for this is that both candidates must stand a chance of winning.

This is because if a candidate knew he was going to lose, he would be better off dropping out of the race and saving the cost of running.

The model also predicts that the ideologies of the two candidates can neither be too far apart nor too close together.

The ideologies cannot be too far apart, otherwise a centrist candidate will be able to enter and win the race.

For example, if citizens have quadratic preferences, a candidate with ideology i_m who entered would obtain a vote share

$$F(\frac{i_m + i_B}{2}) - F(\frac{i_m + i_A}{2}).$$

Assuming $c - r$ is small, this vote share must be less than

$$\max\{F(\frac{i_m + i_A}{2}), 1 - F(\frac{i_m + i_B}{2})\},$$

otherwise such a median candidate would enter.

This limits how far apart i_A and i_B can be.

The ideologies cannot be too close together because if they were, one candidate would be better off dropping out and letting the other candidate be elected (since $r < c$).

For example, with quadratic preferences, it must be the case that

$$\frac{1}{2}(i_A - i_B)^2 > c - \frac{r}{2}$$

The basic picture in terms of the positions of the candidates looks very similar to the predictions of the candidate competition model with policy preferences and voter uncertainty.

If we assume strategic voting (as in Besley and Coate (1997)), it is no longer the case that the candidates cannot be too far apart.

With strategic voting even if the two candidates are at the extremes, a centrist candidate would not necessarily be able to enter and win.

The reason is that in a three way race, centrist voters might continue to view the race as a contest between the two extremist candidates and be reluctant to switch their votes to the centrist candidate for fear of "wasting their votes".

This result is notable because it shows that extremism can arise even with a very competitive looking political system.

I.2.iii Candidate Policy Choice with Re-election Concerns

In the citizen-candidate approach, candidates can make no policy commitments in the campaign.

They cannot credibly commit to do anything other than maximize their true policy preferences when elected.

However, this analysis is purely static and it seems reasonable that even if campaign promises are non-credible, the *threat* of re-election might effectively constrain politicians' choices when in office.

This brings us to *political agency models* which study the choices of politicians facing the threat of re-election.

In these models, the politician is the *agent* and the voters are the *principals*.

The maintained hypothesis is that the politician's preferences (may) differ from that of the voters.

These models differ from standard principal-agent models in that the principals have only a very crude incentive structure - they can either re-elect the politician or remove him from office.

There are two strands of work in the political agency tradition.

Early papers assumed that politicians were all identical and that voters used *retrospective voting* to discipline them.

Thus, voters view all politicians as equally good or bad, but, in a conscious effort to provide discipline, remove underperforming candidates from office.

Later work assumes that politicians are different and that voters use forward-looking voting rules when deciding to re-elect them; i.e., they vote for the incumbent if and only if they believe he is a better candidate than the challenger

In the first class of models, incentives for politicians are *explicit*.

Politicians know that they will not be re-elected unless they perform above a certain standard.

In the second class, incentives are *implicit*.

Politicians know that they will not be re-elected if they are viewed as less attractive than their challengers.

But they can manipulate voters' beliefs about their characteristics (preferences and/or capabilities) via their performance.

The second class of models are basically signalling models and are a little harder to work with than the first class.

Nonetheless, the assumption that all politicians are the same is unappealing.

Moreover, casual empiricism suggests that politicians care greatly about their reputations with voters and this influences their choices (votes on bills, decisions as to which policy issues to pursue, etc).

These factors have made the second class of political agency models more popular and they are now used quite widely in political science.

We will consider the paper by Banks and Sundaram (1998) which nicely illustrates the logic of this second class of models.

See Ferejohn (1986) for a nice example of the first class of models.

Banks and Sundaram's Model

There are an infinite number of periods indexed by $t = 0, ..., \infty$

There are an infinite number of politicians and a single representative voter

Each politician can hold office for a maximum of two periods (i.e., there is a two term limit).

When in office, a politician puts in effort on behalf of the voter and this influences the voter's income.

At the end of each period, there is an election.

If the incumbent is in his second term, the election is contested by two randomly drawn challengers.

If the incumbent is in his first term, he runs against a randomly drawn challenger.

If a first term incumbent fails to be re-elected, he cannot run again.

Politicians come in different "types".

The possible politician types are $\{\omega_1, ..., \omega_n\}$ where $\omega_1 < ... < \omega_n$.

Higher types like to work harder on behalf of the voter, so ω_1 is the worst type and ω_n the best.

The fraction of type ω_i politicians in the population is $\pi_i > 0$.

In any period t, the incumbent chooses some effort level $a_t \in [a_{\text{min}}, a_{\text{max}}]$.

This effort determines stochastically the voter's income y_t.

Specifically, y_t is the realization of a random variable with CDF $F(y\,|a_t)$ and density $f(y\,|a_t)$

The voter observes y_t but not a_t.

The voter is also unable to observe politicians' types.

Assumption. (i) (*Non-moving Support*) *The set* $\{y : f(y\,|a) > 0\}$ *is independent of* a *for all* $a \in [a_{\mathsf{min}}, a_{\mathsf{max}}]$*.* *(ii) (Monotone Likelihood Ratio Property) For all* $a, a' \in [a_{\mathsf{min}}, a_{\mathsf{max}}]$ *with* $a < a'$ $f(y\,|a')/f(y\,|a)$ *is increasing in* y.

These are standard assumptions in principal-agent models.

Part (ii) implies that higher levels of effort make higher incomes for the voter more likely.

During period t, the incumbent obtains a payoff of

$$r + a_t - (\kappa - \omega_i)a_t^2$$

if he is of type ω_i where $r > 0$ and $\kappa > \omega_n$

Note that the optimal level of effort for a type ω_i is $1/2(\kappa - \omega_i)$, so that higher types prefer higher efforts.

Assume that $1/2(\kappa - \omega_1) > a_{\text{min}}$ and $1/2(\kappa - \omega_n) < a_{\text{max}}$.

A politician not in office gets a per-period payoff of 0

Both the voter and politicians have discount rate δ.

Thus higher types like higher actions and get more utility from holding office for any given action.

Results

In an open seat election, the voter has no basis for choosing between the two contestants.

He therefore makes a random choice.

In an election between a first term incumbent and a challenger, the voter has the information obtained from observing the incumbent's first term performance as measured by his income y.

Proposition. *There exists an equilibrium in which the voter employs a cut-off voting rule whereby he re-elects a first period incumbent if* $y \geq y_C$. *In this equilibrium, second term politicians of type* ω_i *choose effort level* $a_{i2} = 1/2(\kappa - \omega_i)$. *First term politicians of type* ω_i *choose effort level* $a_{i1} > 1/2(\kappa - \omega_i)$. *In either term, higher types extert more effort; i.e., for* $t = 1$ *or* 2, $a_{it} < .... < a_{nt}$.

Thus, all politician types perform better in their first than their second term and higher types perform better in both terms than lower types.

The voter employs a cut-off rule because he infers from having higher income that the politician is more likely to have put in higher effort and therefore is more likely to be a higher type.

Why does the voter ever want to re-elect an incumbent to a second term given that a first term incumbent will perform better?

The reason is due to sorting: second term incumbents are likely to be higher types.

Why do first term politicians put in more effort than in their second term?

The reason is that they want to get re-elected and they know they have to generate at least y_C for the voter.

This provides an incentive which is absent in their second term.

In essence, first term politicians are trying to create a good reputation for themselves.

Empirical Evidence

This model predicts a *term limit effect*: politicians behave differently when they can and cannot run for for re-election.

This raises the empirical question of whether politicians do indeed behave differently in their last terms.

One interesting study is that by Besley and Case in the 1995 *Quarterly Journal of Economics*.

They exploit the fact that in almost half the U.S. states, governors face a binding term limit.

They find that when governors are in their last term, state taxes and spending are higher.

This suggests that in their early terms, governors are trying to signal that they are more fiscally conservative then they actually are.

I.3 Legislatures

In almost all practical applications, policy is made by a *legislature* of elected representatives.

There is a large literature on legislative decision making.

This literature focuses on trying to predict the policies that would be selected by a majority rule legislature composed of legislators who disagree on what the optimal policy should be.

Legislators' are assumed to have different policy preferences because they are elected by different districts.

Legislative Decision-Making Model

Consider a legislature consisting of n legislators indexed by $i = 1, ..., n$.

Assume n is odd and ≥ 3 and suppose that the legislature operates by majority rule.

Suppose the legislators have to choose some policy p from some set of alternative policies P.

Legislator i's utility if policy p is selected is $V_i(p)$.

This utility function will reflect both the legislator's ideology and also how the policy will impact him and his constituents.

What policy would we expect the legislature to choose?

Condorcet Winners

A policy $p^* \in P$ is said to be a *Condorcet Winner* if it would defeat or tie any other policy in a pairwise majority vote.

Assuming that legislators abstain if they are indifferent between alternatives, this requires that for any policy $p \in P$ the number of legislators who prefer p^* to p is at least as large as the number who prefer p to p^*.

Formally, the requirement is that for all $p \in P$

$$\#\{i \,|V_i(p^*) > V_i(p)\} \geq \#\{i \,|V_i(p) > V_i(p^*)\}.$$

If a unique Condorcet winner exists, it is reasonable to think that the legislature would select it.

After all, for any other policy, there will exist another policy that could be proposed that would defeat it.

Thus, if the legislature were about to implement that policy, a legislator could offer an amendment proposing a majority-preferred policy and we might expect that amendment to be supported.

So will a Condorcet winner exist?

This depends on the nature of the set of alternative policies P and the legislators' utility functions.

Example 1

Suppose that the legislature is making a discrete decision, such as whether to declare war or not.

The set of policy alternatives then just consists of two elements a and b where, say, a means declare war and b means not declare war.

In this case, there will always exist a Condorcet winner - it is just whichever option is favored by the majority of legislators.

Example 2

Suppose that the legislators are choosing the aggregate level of spending for a particular department (e.g., defense).

Then the set of alternative policies will be $P = \{p : p \geq 0\}$

In this case, there are an infinite number of policy alternatives.

Let legislator i's ideal spending level be p_i^*; that is,

$$p_i^* = \arg\max\{V_i(p) : p \geq 0\}.$$

Label the legislators such that

$$p_1^* < \ldots\ldots < p_n^*.$$

Presumably, if the policy is defense spending, left wing Democrats will prefer a smaller level and conservative Republicans a higher level.

Thus, for a policy issue like this, legislators' preferred positions will reflect their underlying ideologies.

Let i_m denote the *median legislator*; that is, the legislator such that as many legislators prefer a lower as prefer a higher spending level.

Suppose that the legislators' utility functions satisfy the single-crossing property.

Thus, if i and i' are two legislators such that i' prefers more spending than i (i.e., $i < i'$) and spending level p is smaller than p', then

$$V_i(p') \geq V_i(p) \Rightarrow V_{i'}(p') > V_{i'}(p)$$

and

$$V_{i'}(p) \geq V_{i'}(p') \Rightarrow V_i(p) > V_i(p').$$

Proposition *In Example 2, the policy preferred by the median legislator is a Condorcet winner.*

Example 3

Suppose that the legislators are dividing up a fixed budget between projects located in the legislators' districts.

Then a policy is a vector $(p_1, ..., p_n)$ where p_i is the spending in legislator i's district.

The set of alternative policies is

$$P = \{(p_1, ..., p_n) \in \Re_+^n : \sum_{i=1}^{n} p_i = B\},$$

where B denotes the budget.

It may be reasonable to assume that

$$V_i(p_1, ..., p_n) = p_i.$$

Thus, each legislator would prefer that all the projects are located in his own district.

This is an example of a so-called *distributive policy* - the policy issue involves distributing a benefit between different constituencies.

For such issues, ideology is less important than how the policy impacts a legislator's district.

In this example, there are not only an infinite number of policy alternatives as in Example 2 but set of policy alternatives is multi- as opposed to uni-dimensional.

Proposition *In Example 3, there exists no Condorcet winner.*

Proof Let $(p_1, ..., p_n) \in P$.

Assume by relabeling as necessary that $p_1 > 0$.

Then $(p_1', ..., p_n') = (0, p_2 + \frac{p_1}{n-1}, ..., p_n + \frac{p_1}{n-1})$ is preferred by legislators 2 through n to $(p_1, ..., p_n)$.

QED

Discussion

The non-existence of a Condorcet winner arising in Example 3 is a general phenomenon.

Whenever we are looking at distributive policies (or any other multi-dimensional policies), we are very likely to run into this type of problem.

There is a large literature that establishes this (for discussion and analysis see, for example, McKelvey (1976)).

This means that while looking for a Condorcet winner is a natural thing to do, it is often not going to deliver a prediction.

This is an interesting result in its own right since it suggests that in many situations the notion of the "will of the majority" is not really well-defined.

One approach to making predictions for distributive policy problems like those in Example 3, is to build more structured models of how the legislators interact.

Thus, we need to specify which legislators get to make policy proposals, when voting takes place, etc.

The structure will “induce” an equilibrium leading to the idea of *structure-induced equilibrium* (Shepsle).

Example 3 Revisited

To illustrate, lets go back to Example 3 and make some assumptions about how decision-making takes place.

Assume at the beginning of the legislative session, one legislator is selected to make a policy proposal.

Such a proposal would be a proposed division of the budget $(p_1, ..., p_n)$

All legislators then vote for or against the proposal.

If a majority of legislators vote for the proposal, it is implemented, and the legislators move on to deal with some other policy issue.

If a majority of legislators vote against the proposal, another legislator is selected to make a policy proposal.

All legislators then vote for or against this proposal and if a majority vote for it is implemented.

If a majority of legislators vote against the proposal, another legislator is selected to make a policy proposal, etc, etc.

This process continues until a proposal is passed.

It is natural to assume that the legislators would rather pass a proposal sooner rather than later.

To capture this, assume that if the policy $(p_1, ..., p_n)$ is agreed to after t proposal rounds, legislator i obtains utility $\delta^{t-1}p_i$, where $\delta < 1$.

What would happen under these rules?

Much will depend on the rule for selecting proposers.

The simplest case to analyze is when each legislator is equally likely to be chosen.

Thus, in each proposal round, each legislator is selected to be proposer with probability $1/n$.

Under this proposer selection rule, we can figure out what might happen.

Equilibrium in Example 3

The structure we have described gives rise to a game between the legislators.

The game is similar to the alternating offer bargaining model used in economics.

The differences are that there are more than two players and majority agreement is required instead of unanimous agreement.

The game is referred to as a *legislative bargaining game* (Baron and Ferejohn (1989)).

A strategy for a legislator describes the proposal that he will make if selected and how will he vote on any proposal.

In principle, strategies may depend in complicated ways on the entire history of play.

We restrict attention to *stationary* strategies which do not depend on the proposal round or the history of play.

This rules out strategies whereby a proposing legislator discriminates against legislators who have previously proposed and offered his districts nothing.

Formally, legislator i's pure strategy is described by the proposal he will make if selected $(p_1^i, ..., p_n^i)$, and a cut-off x_i such that i will vote to accept any proposal that gives his district more than x_i.

It will also be necessary to allow legislators to mix over the proposals they will make.

We look for a symmetric equilibrium in which each legislator treats other legislators in a symmetric way and has the same cut-off.

To figure out what a symmetric stationary equilibrium looks like, consider the problem of the legislator selected to be proposer at the beginning of some proposal round t.

He obviously wants to get his proposal passed but to do so in a way that gives his district the largest share of the budget.

He therefore needs to form a *minimum winning coalition* consisting of himself and $(n-1)/2$ other legislators.

A minimum winning coalition (mwc for short) is a coalition of the smallest size necessary to pass legislation.

In a symmetric equilibrium, we know that all legislators will have the same cut-off x.

Thus, the proposer will propose that the district of each legislator in his mwc gets x and that his district gets the remaining budget which is $B - x(n-1)/2$.

The districts of the legislators who are not in his mwc get nothing.

We just need to solve for the equilibrium cut-off x.

Consider the voting decision of a legislator in the mwc.

By definition, x must be just sufficient to get him to vote for the proposal.

If the proposal passes, the legislator gets a payoff of $\delta^{t-1}x$.

If the proposal is rejected, we go to proposal round $t+1$.

In equilibrium, the legislator knows that the round $t+1$ proposer will propose giving x to the districts of $(n-1)/2$ other legislators and $B - x(n-1)/2$ to his own district.

The legislator knows that there is a probability $1/n$ he will be the round $t+1$ proposer.

Moreover, in a symmetric equilibrium, there must be a probability $(n-1)/2n$ that he will be one of the legislators in the round $t+1$ proposer's mwc.

This requires that the round $t+1$ proposer will select members of his mwc randomly.

We conclude that the legislator's equilibrium expected payoff if we get to proposal round $t+1$ is

$$\delta^t \left[\frac{B - x(n-1)/2}{n} + \frac{(n-1)}{2n} x \right].$$

Since x must be just sufficient to get the legislator to vote for the round t proposal, it must be that

$$\delta^{t-1} x = \delta^t \left[\frac{B - x(n-1)/2}{n} + \frac{(n-1)}{2n} x \right]$$

Solving for x, we obtain

$$x = \delta \frac{B}{n}.$$

Thus, we conclude that in any proposal round t the equilibrium proposal will involve the proposer giving $\delta \frac{B}{n}$ to the districts of $(n-1)/2$ randomly selected legislators and $B[1 - \delta \frac{(n-1)}{2n}]$ to his own district.

The proposal will be supported by the proposer and the $(n-1)/2$ other legislators in the mwc.

In equilibrium, the first proposal made will pass, so the legislature will not waste time on the issue.

Notice that it is important for the logic of the symmetric equilibrium that proposers choose their mwcs randomly and thus the composition of the mwc will be uncertain.

This is why we have to allow for mixed strategies.

Further Discussion

The above analysis shows that, even when a Condorcet winner does not exist, we can make predictions if we are willing to make assumptions about how legislative decision-making takes place.

The difficulty is knowing what assumptions to make.

It is very hard to actually model the rules governing legislative decision-making, since many of these may be informal, and in any case, they will likely to be too complicated to write down in a model.

Moreover, these rules will vary across different legislative bodies.

Nonetheless, there are certain standard predictions that tend to emerge from models of legislative decision-making in situations of distributive policy-making.

First, the proposing legislator (or legislators if a committee is in charge of proposals) gets extra benefits for his district.

Second, benefits are enjoyed by a mwc of legislators.

Third, which legislators are in the mwc is uncertain.

See Eraslan (2001) for a clean theoretical treatment of the Baron-Ferejohn legislative bargaining model.

See Volden and Wiseman (2007) for a study of legislative bargaining in which there are public goods as well as transfers.

See Morelli (1999) for a slightly different approach to modelling legislative bargaining.

See Frechette, Kagel, and Morelli (2005) for experimental work on legislative bargaining.

See Knight (2005) for empirical evidence of "proposal power" in the allocation of Federal highway dollars.

Legislative Norms

An alternative (now less popular) approach to making predictions in distributive policy settings is to assume that legislators develop norms of behavior to overcome the potential instability associated with the lack of a Condorcet winner.

Weingast came up with the idea of a "norm of universalism" (see Weingast, Shepsle, and Johnson (1981) for discussion).

Suppose that the legislators are choosing the amount to be spent on (say) transportation projects in each legislator's district.

Then a policy will be a vector $(p_1, ..., p_n)$ where p_i is the spending in legislator i's district.

There is no predetermined budget, so the set of alternative policies is

$$P = \{(p_1, ..., p_n) : p_i \geq 0 \text{ for all } i\}.$$

Assume that transportation spending is funded by a uniform tax on all districts $T = \sum_{i=1}^{n} p_i/n$.

Assume that

$$V_i(p_1, ..., p_n) = B(p_i) - T,$$

where $B(0) = 0$ and $B(\cdot)$ is increasing and strictly concave.

There is no Condorcet winner in this Example (Prove yourself)

The *norm of universalism* works as follows: each legislator chooses the amount of public spending that he would like for his district.

These desired spending levels are then passed unanimously by the legislature.

The norm is “I will vote for your pet projects, with the understanding that you will vote for mine” or “you scratch my back and I will scratch yours”

Let $(p_1^*, ..., p_n^*)$ be the equilibrium spending levels.

Then,

$$p_i^* = \arg\max V_i(p_1^*, ..p_i.., p_n^*).$$

This implies that $p_i^* = p^*$ where

$$B'(p^*) = 1/n.$$

This leads to the *Law of 1/n* which says that, ceteris paribus, the greater the number of legislative districts, the greater the amount of public spending.

Interestingly, there is a considerable amount of evidence for the Law of 1/n - see Baqir (2002) for evidence from U.S. cities.

Equilibrium spending is too high when compared to the surplus maximizing per district spending level which is $p_i^o = p^o$ where

$$B'(p^o) = 1.$$

The theory suggests that all districts will get spending, which distinguishes it from the predictions of legislative bargaining models.

Despite seeming eminently sensible, the postulated behavior of the legislators is hard to make sense of.

At the time of voting for the omnibus bill, all the legislators would be better off agreeing to reduce all their spending levels.

This observation has led the model to lose favor in the theoretical literature.

I.4 Interest Groups

Interest groups are groups of citizens who share common policy objectives and seek to influence policy in a coordinated way.

Important examples in American politics are the NRA, the AARP, the AMA, the Sierra Club, the NAACP, the ACLU, Common Cause, and the Humane Society.

We will discuss how and why interest groups are formed and then how they influence policy determination.

We will consider two different avenues of influence: lobbying and campaign contributions.

I.4.i Interest Group Formation

There are many different groups of citizens who share common policy objectives.

For example, college students who share an interest in lower tuition at state universities, university professors who share an interest in more government funding for research, pet lovers who share an interest in government subsidization of no-kill animal shelters and spaying and neutering programs, smokers who share an interest in lower cigarette taxes, etc, etc.

What determines which groups will form an interest group and seek to influence policy?

This is a difficult problem for the economic approach to politics because interest group formation involves the voluntary provision of what is essentially a public good for potential group members.

The economic approach suggests that public goods will not be voluntarily provided or, if they are, will be provided at extremely low levels.

This is because of the well-known *free rider problem*.

To illustrate the problem, suppose that an interest group entrepreneur decides to try and form an environmental interest group.

Suppose that there are n environmentally concerned citizens and the entrepreneur asks each to contribute to the interest group.

Let citizen i's contribution be denoted c_i and the total contributions to the interest group be denoted $C = \sum_i c_i$.

Let the stringency of environmental regulation chosen by the government be denoted by r.

Assume that the environmental interest group will use its contributions to try to get more stringent regulations.

Thus, assume that $r = r_0 + r_1 C$ where r_0 and r_1 satisfy $r_1 > r_0/n \geq 0$

Let citizen i's payoff be

$$\ln r - c_i.$$

Thus, citizen i likes more stringent regulation, but does not like contributing.

The per-citizen contribution that maximizes the aggregate utility of the n environmentally concerned citizens is

$$c^o = \arg\max n \left[\ln\left(r_0 + r_1 n c\right) - c\right].$$

This satisfies the first order condition

$$\frac{r_1 n}{r_0 + r_1 n c^o} = 1.$$

Solving this, we see that the optimal per-capita contribution is

$$c^o = \frac{r_1 n - r_0}{r_1 n}.$$

This implies that the optimal aggregate contribution to the interest group is

$$C^o = nc^o = \frac{r_1 n - r_0}{r_1}.$$

However, this is not the aggregate contribution that will be made in equilibrium.

Define c^* to be an *equilibrium* per-citizen contribution if each citizen would choose to make contribution c^* if he knew that all the other citizens were also choosing c^*.

This is just the standard notion of Nash equilibrium.

Formally, this definition of equilibrium implies that

$$c^* = \arg\max_{c \geq 0} \ln\left(r_0 + r_1\left[(n-1)c^* + c\right]\right) - c.$$

If $c^* > 0$, it satisfies the first order condition

$$\frac{r_1}{r_0 + r_1 n c^*} = 1.$$

This implies that

$$c^* = \frac{r_1 - r_0}{r_1 n}.$$

If $r_1 < r_0$, then c^* must equal 0.

Thus, the equilibrium per-capita contribution is

$$c^* = \begin{cases} 0 & \text{if } r_1 \leq r_0 \\ \frac{r_1 - r_0}{r_1 n} & \text{if } r_1 > r_0 \end{cases}$$

This implies that the equilibrium aggregate contribution is

$$C^* = \begin{cases} 0 & \text{if } r_1 \leq r_0 \\ \frac{r_1 - r_0}{r_1} & \text{if } r_1 > r_0 \end{cases}$$

We conclude from this that either the interest group will not form or, if it does, the contributions it will get will be extremely small relatively to the optimal level C^o.

The problem is created by citizens free-riding on the contributions of others.

Olson's Ideas

So how can we explain the fact that interest groups do form?

Mancur Olson suggested two explanations in a famous book *The Logic of Collective Action*.

First, he argued that small groups who have large potential benefits from influence activities will be better able to overcome the free rider problem.

Thus, for example, industries which consist of relatively few firms should be able to coordinate their activities.

This contrasts with consumer interest groups, which should be rather rare.

The theoretical case for this argument is not completely clear (see Esteban and Ray (2001)).

In our model, for example, while equilibrium per-capita contributions are decreasing in n, the aggregate contribution is actually independent of n.

More relevant for Olson's argument may be the theory of repeated games which shows that the free-rider problem can be overcome in dynamic contexts.

The optimal outcome may be sustained by the threat that if one player free-rides, the others will revert to free-riding behavior in the future.

Coordinating on this good outcome may be easier in small groups because it is easier for group members to monitor the behavior of other group members.

Thus, in industries which consist of relatively few firms, if one firm is not paying its fair contribution to the industry's influence activities, the other firms should be able to observe this and react accordingly.

Second, Olson argued that interest groups who can supply their members private goods will be more likely to be successful.

Possible private goods consist of buttons, decals, bumper stickers, mugs, magazines, newsletters, discounts on insurance, discounts on hotels, stuffed toys, etc.

To illustrate this idea, suppose that our interest group entrepreneur in exchange for a contribution c offers an amount of interest group-related private goods γc.

Suppose that the cost to the entrepreneur of providing a unit of one of these private goods is β.

This means that if total contributions are C, the revenues the interest group has left over for influencing policy are $C(1 - \beta\gamma)$

Suppose that citizen i's payoff is

$$\ln r + \alpha \ln x_i - c_i,$$

where x_i denotes interest group-related private goods and $\alpha > 0$.

Then the equilibrium per-citizen contribution c^* solves the problem

$$c^* = \arg\max_{c \geq 0} \begin{array}{c} \ln \{r_0 + r_1 [((n-1)c^* + c)(1-\beta\gamma)]\} \\ +\alpha \ln \gamma c - c \end{array}.$$

The first order condition is

$$\frac{r_1(1-\beta\gamma)}{r_0 + r_1 n c^*(1-\beta\gamma)} + \frac{\alpha}{c^*} = 1$$

In general this is a quadratic equation in c^*, but to make it a linear equation lets assume that $r_0 = 0$.

Then we have a very simple solution

$$c^* = \frac{1}{n} + \alpha.$$

This implies that aggregate contributions are

$$C^* = 1 + \alpha n$$

and profits are

$$(1+\alpha n)\,(1-\beta\gamma).$$

Observe that as long as the entrepreneur keeps γ small, he can get significantly more contributions for the interest group in this way.

Discussion

There seems to be some truth to Olson's ideas.

Industry-based interest groups do seem common (tobacco companies, automobile manufacturers, boat manufacturers, beverage producers, etc, etc).

Citizen-based groups do indeed offer members private goods in exchange for contributions.

The AARP offers a magazine and numerous discounts, the NRA provides a magazine and accidental death insurance, the World Wildlife Fund offers stuffed tigers and opportunities to adopt animals.

However, Olson's private good theory does not explain why these private goods cannot be provided by entrepreneurs who do not have the additional cost burden associated with influence activities.

Nor does his private good theory give us a particularly sharp answer to our question of which interest groups will form.

Moreover, in general, Olson's concern with free-riding is a bit too pessimistic regarding human nature.

Free-riding has been subject to intense scrutiny in laboratory experiments and people contribute much more than theory predicts.

On the basis of this evidence, we would expect a significant fraction of people to contribute to interest groups who champion issues they care about.

Why people do this is not really clear - they may be choosing to behave as rule utilitarians as we discussed earlier in the context of voting.

The bottom line is that the literature has not really produced a compelling answer to the question of interest group formation.

From now on, we will assume that interest groups exist and analyze how they influence policy outcomes.

I.4.ii Interest Group Influence by Lobbying

Interest groups in the U.S. spend a considerable amount lobbying Congress and Federal Agencies.

They do this either by hiring the services of a lobbying firm or by employing their own "in house" lobbyist.

OpenSecrets.org reports that in 2008 the total amount spent on lobbying the U.S. federal government was $3.3bn and there were 14,800 registered lobbyists.

Presumably, interest groups would not spend all this money on lobbying unless it was effective.

A large literature explores why and how lobbying works.

Informational Lobbying

One strand of literature views lobbying as providing information to legislators concerning the impact of proposed policy changes on their constituents.

In many circumstances interest groups will have much more information on how a proposed policy will impact their members than do politicians.

Providing such information can change legislators' decisions.

For example, suppose that a legislator is thinking of voting for eliminating a tariff on a particular manufactured good and one consideration is how this will impact the domestic industry in his district.

Suppose that he would favor elimination if the harm to the domestic industry is low and be against if it was high.

Suppose that he does not know for sure whether the harm is low or high, but given his prior beliefs, is inclined to favor tariff elimination.

Then, the industry's interest group can change the legislator's preferences by convincing him that the harm is actually likely to be high.

Of course, the legislator should rationally be a bit sceptical of the interest group's claims.

An important concept in the theoretical literature is therefore the *provability* of information.

The literature distinguishes full provability, partial provability, and no provability.

With full provability, an interest group can prove the validity of its information - for example, present a statistical study of the effect of removing the tariff.

With no provability, it is unable to do this.

With partial provability it can present partial proof.

With full provability the issue is whether an interest group will choose to reveal its information; that is, which facts will it keep to itself and which facts will it convey to policy-makers.

With no provability, the issue is how an interest group *credibly* convey its information.

The majority of the theoretical literature focuses on the case of no provability, so called "cheap talk".

The literature shows that even in this case, and interest group can influence policy as long as its preferences are not too far apart from the policy-maker's.

The Cheap Talk Model

There is a single policy maker and a single interest group.

The policy maker has to choose the level of some policy $p \in \Re$

The optimal level of the policy depends upon the the *state of nature* $\theta \in [\theta_{\mathsf{min}}, \theta_{\mathsf{max}}]$

The state of nature is the realization of a random variable with CDF $F(\theta)$ and density $f(\theta)$.

Assume that $f(\theta) > 0$ for all $\theta \in [\theta_{\mathsf{min}}, \theta_{\mathsf{max}}]$

The state of nature is observed by the interest group but not the policy maker.

The policy maker's preferences are

$$U(p, \theta) = -(p - \theta)^2.$$

The interest group's preferences are

$$G(p, \theta) = -(p - \theta - \delta)^2,$$

where $\delta > 0$.

Thus, the interest group prefers a larger level of the policy than the policy-maker and δ measures the degree of non-alignment.

Game and Equilibrium

The interest group observes the state of nature θ

The interest group then sends some message $m \in M$ to the policy-maker - M is the *message space*

The policy-maker observes the message m and chooses some policy p

A *strategy* for the interest group is a function $\sigma : [\theta_{\mathsf{min}}, \theta_{\mathsf{max}}] \to M$.

A *strategy* for the policy-maker is a function $\rho : M \rightarrow \Re$

The policy-maker also has *beliefs* which are represented by the conditional CDF $H(\theta \,|m)$ with density $h(\theta \,|m)$.

Thus, $H(\theta' \,|m)$ is the probability that the policy-maker assigns to the state being less than θ' given that he hears the message m

An *equilibrium* is $\{\sigma; \rho; H\}$ such that σ is optimal for the interest group given ρ; ρ is optimal for the policy-maker given his beliefs H; and H is consistent with σ where possible.

Results

Note that the policy-maker's strategy is given by

$$\rho(m) = \arg\max \int_{\theta_{\mathsf{min}}}^{\theta_{\mathsf{max}}} U(p,\theta) h(\theta\,|m) d\theta$$

The interest group's strategy satisfies

$$\sigma(\theta) = \arg\max_{m \in M} G(\rho(m), \theta)$$

When $\sigma(\theta') = m$ for some θ', the policy-maker's beliefs are

$$h(\theta\,|m) = \frac{f(\theta)\,\mathsf{Pr}\{\sigma(\theta) = m\}}{\int_{\theta_{\mathsf{min}}}^{\theta_{\mathsf{max}}} f(\theta')\,\mathsf{Pr}\{\sigma(\theta') = m\}}$$

Proposition 1. *Let* $M = [\theta_{\text{min}}, \theta_{\text{max}}]$*. Then there exists no equilibrium in which* $\sigma(\theta) = \theta$*; i.e., there exists no fully revealing equilibrium.*

Proof. To see this suppose that $\sigma(\theta) = \theta$.

Then, the policy-maker would believe that the true state was θ with probability one whenever he saw the message θ and would choose the policy

$$\rho(\theta) = \arg\max U(p, \theta)$$

But we know that the interest group prefers a higher level of the policy for any given θ and hence has an incentive to over-exaggerate.

Thus, $\theta \neq \arg\max G(\rho(\theta), \theta)$. QED

Proposition 2. *For any message space* M *there always exists a babbling equilibrium in which for all* m*,* $h(\theta\,|m) = f(\theta)$ *and*

$$\rho(m) = \text{arg}\max_p \int_{\theta_{\text{min}}}^{\theta_{\text{max}}} U(p,\theta) f(\theta) d\theta$$

Proof. Roughly speaking, if the policy-maker ignores the interest group then the interest group can send any message it likes.

More formally, if $\rho(m)$ is independent of m then it is a best response for the interest group to send the same message for any θ; i.e., for some m', $\sigma(\theta) = m'$ for all θ

But then $h(\theta\,|m') = f(\theta)$ and we can choose the out-of-equilibrium beliefs such that $h(\theta\,|m) = f(\theta)$ for all $m \neq m'$. QED

Crawford and Sobel (1982) showed that any equilibrium can be represented as a *partition equilibrium*.

In such an equilibrium, the interest group chooses one of a finite number n messages, say $m \in \{m_1, ..., m_n\}$

The policy-maker interprets the message m_i to mean that $\theta \in [\theta_{i-1}, \theta_i]$ where $\theta_0 = \theta_{\text{min}}$ and $\theta_n = \theta_{\text{max}}$

Thus, for all $\theta \in [\theta_{i-1}, \theta_i]$

$$h(\theta \,|m_i) = \frac{f(\theta)}{F(\theta_i) - F(\theta_{i-1})},$$

while for all $\theta \notin [\theta_{i-1}, \theta_i]$, $h(\theta \,|m_i) = 0$

The boundaries of the intervals are determined by the requirement that

$$G(\rho(m_i), \theta_i) = G(\rho(m_{i+1}), \theta_i)$$

Crawford and Sobel establish that if an equilibrium with $n \geq 2$ messages exists there exists one with $n-1$. (Note the babbling equilibrium has $n = 1$)

They also show that there exists some n_C such that $n \leq n_C$.

The larger is n_C the greater is the revelation of information.

n_C depends on the preference alignment between the interest group and the policy-maker.

Example

To illustrate, suppose that θ is uniformly distributed so that $F(\theta) = (\theta - \theta_{\text{min}})/(\theta_{\text{max}} - \theta_{\text{min}})$.

Lets look for a partitition equilibrium of size n.

Thus, the message space is $\{m_1, ..., m_n\}$ and the policy-maker interprets the message m_i to mean that $\theta \in [\theta_{i-1}, \theta_i]$ where $\theta_0 = \theta_{\text{min}}$ and $\theta_n = \theta_{\text{max}}$.

Thus, for all $\theta \in [\theta_{i-1}, \theta_i]$

$$h(\theta\,|m_i) = \frac{1}{\theta_i - \theta_{i-1}},$$

while for all $\theta \notin [\theta_{i-1}, \theta_i]$, $h(\theta\,|m_i) = 0$.

This means that

$$\rho(m_i) = \arg\max - \int_{\theta_{i-1}}^{\theta_i} (p - \theta)^2 \frac{d\theta}{\theta_i - \theta_{i-1}},$$

which implies that

$$\rho(m_i) = \frac{\theta_i + \theta_{i-1}}{2}.$$

The boundaries of the intervals are determined by the requirement that

$$G(\rho(m_i), \theta_i) = G(\rho(m_{i+1}), \theta_i),$$

or

$$-([\frac{\theta_i - \theta_{i-1}}{2} + \delta])^2 = -(\frac{\theta_{i+1} - \theta_i}{2} - \delta)^2,$$

or

$$([\frac{\theta_i - \theta_{i-1}}{2} + \delta]) = (\frac{\theta_{i+1} - \theta_i}{2} - \delta).$$

This implies that for all $i = 1, ..., n-1$, we have that

$$\theta_{i+1} = 2\theta_i + 4\delta - \theta_{i-1}$$

with the end-point conditions $\theta_0 = \theta_{\text{min}}$ and $\theta_n = \theta_{\text{max}}$.

This second-order difference equation implies that for all $i = 1, ..., n$

$$\theta_i = \frac{i}{n}\theta_{\mathsf{max}} + \frac{n-i}{n}\theta_{\mathsf{min}} - 2i(n-i)\delta.$$

For this to make sense, we must have that $\theta_1 > \theta_{\mathsf{min}}$ which requires that

$$2n(n-1)\delta < \theta_{\mathsf{max}} - \theta_{\mathsf{min}}.$$

Thus n_c is the largest integer such that the above inequality holds.

Observe that $lim_{\delta \to 0} n_c = \infty$ as common sense suggests.

This type of model can be extended to include two interest groups - see Krishna and Morgan (2001)

It can also be extended to have a multidimensional policy space - see Battaglini (2003)

One can also extend the model to have costly information acquisition and so on.

Lobbying as Legislative Subsidy

A recent paper by Hall and Deardoff (2006) offers a slightly different take on what lobbying does.

They develop a theory based on the following five assumptions.

Assumption 1: For a legislator to have much influence on policy, he must work at it.

Assumption 2: Legislators' resources are scarce.

Assumption 3: At any given point in time, individual legislators care about influencing more than one policy.

Assumption 4: Legislators care about some issues more than others.

Assumption 5: Relative to legislators, lobbyists are specialists.

Given these assumptions, the main elements of Hall and Deardoff's theory can be represented as a simple consumer choice problem.

The legislator cares about "making progress" on some target issue a and on an aggregate of other issues o.

By making progress, is meant moving policy closer to their preferred level by introducing a bill, making an amendment to an existing bill, influencing a government agency, fighting against a bad bill, etc.

Letting P_a denote progress on issue a and P_o progress on other issues, we can write the legislator's utility function as $U(P_a, P_o)$.

The legislator must chooses how much effort to devote to issue a and the other issues.

Let e_a denote effort devoted to issue a and e_o denote effort devoted to other issues.

The constraint that the legislator's resources are scarce can be modeled by assuming that

$$e_a + e_o \leq e_{\mathsf{max}}.$$

Assuming that $P_a = \gamma e_a$ and $P_o = \gamma e_o$ for some $\gamma > 0$, the legislator's problem in the absence of lobbying can be posed as

$$\begin{gathered} \max U(P_a, P_o) \\ P_a + P_o \leq \frac{e_{\mathsf{max}}}{\gamma} \end{gathered}$$

This is a standard consumer choice problem which can be represented diagrammatically with the usual indifference curve and budget line diagram.

Without lobbying there will be some optimal choice (P_a^*, P_o^*).

Lobbyists for issue a are modeled as making the legislator's effort for issue a more productive.

Thus, when there are lobbyists for issue a, $P_a = (1 + s)\gamma e_a$ where $s > 0$.

Lobbyists make the legislator's efforts more productive because they can provide assistance in terms of in-depth policy analysis, expertise, help with speech writing, focus groups, information concerning other legislators, etc.

The legislator's problem with lobbying then becomes

$$\max U(P_a, P_o)$$
$$\frac{P_a}{1+s} + P_o \leq \frac{e_{\max}}{\gamma}$$

This amounts to a rotation of the budget line which induces the legislator to choose a point $(P_a^*(s), P_o^*(s))$ which involves $P_a^*(s) > P_a^*$.

In this way, lobbying leads to greater progress being made on issue a.

The problem with this analysis is that it does not recognize that collective action by legislators is necessary to make progress on an issue.

Thus just focusing on a single legislator is not fully satisfying, since no such legislator can make progress alone.

It may be possible to extend the theory to capture this.

Evidence

The above two stories explain theoretically why lobbying legislators may produce results.

There have been few successful empirical attempts to measure the returns to lobbying.

One interesting effort is by de Figueiredo and Silverman (2006) who look at university lobbying.

They measure the outcome variable as the amount of educational spending earmarked to go to fund projects at a university.

The vast majority of federal money financing academic research is allocated by federal agencies (such as the NSF and NIH) using competitive peer reviews.

However, a significant chunk (around 10% at the time de Figueiredo and Silverman wrote their article) is allocated via earmark.

de Figueiredo and Silverman ran regressions with earmarked spending to University x as the dependent variable and the amount spent on lobbying by University x as the independent variable.

They find no returns to lobbying for universities whose Congressman or Senator is not on either the Senate or House Appropriations Committtees.

However, the average university whose senator is on the Senate Appropriations Committee receives an average return on $1 of lobbying spending in the range of $11-$17.

The average university whose congressman is on the House Appropriations Committee receives an average return on $1 of lobbying spending in the range of $20-$36.

These are very significant returns!

I.5.iii Interest Group Influence by Campaign Contributions

A campaign contribution is an amount of money given to a politician to help him finance his campaign to get elected.

Such contributions are used by politicians to finance campaign advertisements, town hall meetings, voter turnout efforts, etc.

According to OpenSecrets.org, the total amount spent in the 2008 election cycle in House, Senate and Presidential races was $3.1bn.

Contributions are made by individual citizens and *Political Action Committees* (PACs).

When an interest group, union, or corporation wants to contribute to federal candidates or parties, it must do so through a PAC.

Examples of PACs are the Sierra Club Political Committee, the League of Conservation Voters Action Fund, the National Rifle Association Political Victory Fund, the National Association of Realtors PAC, the Planned Parenthood Action Fund, the American Beverage Association PAC, the American Beer Companies PAC, and the American Automobile Manufacturers PAC.

Corporations and unions may not contribute directly to federal PACs, but they may pay for the administrative costs of an affiliated PAC.

Corporate-affiliated PACs may only solicit contributions from executives, shareholders, and their families, while union-affiliated PACs may only solicit contributions from members.

Overall, PACs account for less than thirty percent of total contributions in U.S. Congressional races, and considerably less in presidential races.

The amount that individuals and PACs can give to any one candidate is limited.

The individual limit per candidate is \$2400 per election and the PAC limit per candidate is \$5000 per election.

For more information go to the website of the Federal Election Commission (FEC).

All the individual U.S. states have their own regulations concerning state elections.

A large academic literature addresses how campaign contributions by interest groups impact policy outcomes.

The topic is a complex one since a good theory should (i) explain how and why campaign spending impacts voter behavior, and (ii) why contributors give to candidates.

These are issues on which common sense and the empirical literature offer no clear guidance.

Campaign Spending and Voter Behavior

In the theoretical literature, there are two main ways to think about how and why campaign spending impacts votes.

First, spending buys the votes of "noise voters".

Noise voters are analogous to noise traders in financial markets.

These are voters who are voting for random non-policy related reasons.

Campaign advertising is simply assumed to attract these voters.

Second, spending allows candidates to provide information about their accomplishments and policy positions to uninformed voters.

In this view, voters are rational and update their beliefs based on the information they see in campaign advertising.

Gerber, Gimbel, Green, and Shaw (2011) run an interesting field experiment where they try to assess the impact of campiagn advertising on voters' preferences.

They worked with the Rick Perry campaign in the 2006 Texas Gubernatorial Race.

Perry, the incumbent, was running for re-election and the experiment took place during the early months of the campaign (Jan 2006), before the full field of candidates had become clear.

The experiment randomized the deployment of an advertisment which listed some of the accomplishments of Texas during Perry's first term.

Using surveys of treated voters, they found that the advertising campaign caused a large initial boost in support for Perry, but this effect diminished rapidly over time.

They argue that this up and then down pattern is not consistent with opinion change due to informational updating.

Independent of why campaign spending impacts voter behavior, there is the question of the magnitude of the effects.

In fact, there is a good deal of debate empirically on the impact of campaign spending on votes (see Gerber (2004) for a review).

Estimating a causal effect is a difficult econometric problem because of the endogeneity of campaign spending.

For example, a good candidate will presumably attract more contributions leading to a positive correlation between votes and money.

On the other hand, an incumbent who is facing a high quality challenger will presumably seek more contributions and spend more, than an incumbent facing a low quality challenger, leading to a negative correlation between votes and money.

The general conclusions are that campaign spending by incumbents in the U.S. House does not have a significant effect on their vote share, whereas spending by challengers does.

By contrast, spending by both incumbents and challengers in U.S. Senate races has a significant effect, although spending by challenegers is more effective.

It is not clear how sensible the result concerning House incumbents is, given how much time incumbents devote to raising contributions.

Why do Interest Groups Give?

On the question of why contributors give, the central issue is whether contributions are used to *influence* politicians, or whether contributors just give to help get their preferred politicians *elected*.

Influence can take the form of *buying votes* or *buying access*.

Access means the opportunity to present persuasive information.

The idea is that if a politician has received a contribution from someone, a politician is more likely to agree to listen to a contributor's argument concerning a proposed piece of legislation.

Economic approaches to politics tend to focus on the influence motive for giving as opposed to the electoral motive.

This is because giving to help get a preferred candidate elected is like making a voluntary contribution to a public good for the supporters of this candidate.

As we have already seen, the economic approach suggests that public goods will not be voluntarily provided or, if they are, will be provided at extremely low levels.

However, as we have seen, people do contribute more to public goods than the economic approach would suggest.

Ansolabehere, de Figueiredo, and Snyder (2003) make a strong case against the practical importance of the influence motive in U.S. national politics.

They argue that if politicians could be so influenced, then given the size of federal government spending, there would be way more money in U.S. politics than there is.

On the other hand, Jayachandran (2006) documents significant negative effects on the stock market values of firms contributing soft money to the Republican party from Senator Jeffords 2001 switch from the Republican party which tipped the balance of the Senate.

"Soft money" refers to unregulated contributions to national parties that were since banned by the McCain-Feingold Act of 2002.

One interpretation of her finding is that the stock market anticipated that these firms would now have less influence over policy because the party they supported no longer controlled the agenda.

Also Mian, Sufi, and Trebbi (2009) document an association between higher campaign contributions from the financial services industry and voting in favor of a bill which transfered money from taxpayers to the industry (The Emergency Economic Stabilization Act of 2008).

The suggestion is that these contributions bought votes.

While suggestive, neither of these papers conclusively establish an influence motive.

It could be that the contributions were simply given because of the policy preferences of the party or legislators.

There is quite a bit of anecdotal evidence from Congressmen and Senators in favor of the idea that contributions do buy access.

That is, these politicians report that they do feel obligated to meet with those citizens who have supported them with significant contributions.

Presumably, if an interest group receives better access this should translate into obtaining better policy outcomes, but this part of the link is not well-established empirically.

The Common Agency Model

The simplest economic model of interest group influence by campaign contributions is the *common agency model*.

The model assumes that there are n interest groups competing to influence a politician who must choose some policy $p \in P$.

The idea is that the politician is already in office and having to make policy decisions, but needs campaign contributions to run for re-election in the future.

If the politician receives an aggregate contribution C from the interest groups and chooses policy p he receives a payoff $V_0(p) + C$.

The idea is that the politician cares about the policy decision he makes but also values contributions for his re-election campaign.

The latter is modeled in a reduced form way by just assuming that the politician's payoff depends on his aggregate contributions.

If interest group i makes a contribution C_i and the politician chooses policy p, its payoff is $V_i(p) - C_i$.

Thus, interest groups do not like making contributions.

Game and Equilibrium

The interaction between the interest groups and the politician is modelled as a two-stage game.

In *Stage 1* each interest group i independently and simultaneously offers the politician a contribution schedule $C_i(p)$ which describes the contribution it will make conditional on the policy choice p that he makes.

In *Stage 2* the politician chooses policy taking into account the implications of his choice for the contributions he will receive.

Given the n interest groups' contribution schedules $C_1(\cdot), ..., C_n(\cdot)$, the politician will choose a policy from the set

$$M(C_1(\cdot), ..., C_n(\cdot)) = \arg\max_{p \in P} [V_0(p) + \sum_{i=1}^{n} C_i(p)].$$

Thus, the politician will trade off his own policy preferences with the impact of his policy choice on the contributions that he will receive.

The contribution schedules $C_1^*(\cdot), ..., C_n^*(\cdot)$ and the policy choice p^* represent an *equilibrium* if (i)

$$p^* \in M(C_1^*(\cdot), ..., C_n^*(\cdot))$$

and (ii) there is no interest group i, contribution schedule $C_i(\cdot)$, and policy choice $p \in M(C_1^*(\cdot), .., C_i(\cdot), .., C_n^*(\cdot))$ such that

$$V_i(p) - C_i(p) > V_i(p^*) - C_i^*(p^*).$$

Thus, in equilibrium we require (i) that the politician be choosing an optimal policy given the contribution schedules he faces, and (ii) that no interest group can alter its contribution schedule to achieve a preferred outcome.

Results

Consider first the case of a single interest group.

If there were no contributions, the politician would simply choose his preferred policy

$$p_0 = \arg\max_{p \in P} V_0(p).$$

The interest group can get the politician to choose any policy p as long as it provides him with a contribution C_1 such that

$$V_0(p) + C_1 \geq V_0(p_0).$$

The interest group's problem can therefore be posed as:

$$\begin{array}{rl} \max_{(p,C_1)} & V_1(p) - C_1 \\ s.t. & V_0(p) + C_1 \geq V_0(p_0) \end{array}$$

Clearly, the interest group will set

$$C_1 = V_0(p_0) - V_0(p).$$

Substituting this into the interest group's objective function, we see that the optimal policy p^* satisfies

$$p^* = \arg\max_{p \in P} V_0(p) + V_1(p).$$

Note that the policy maximizes the sum of the interest group and politician's policy payoffs.

The contribution paid to the politician will then be

$$C_1^* = V_0(p_0) - V_0(p^*).$$

The model delivers a unique prediction in terms of the equilibrium policy and the amount of the contribution actually paid.

There are many possible contribution schedules that the interest group could use to induce the politician to choose p^*.

One example is

$$C_1(p) = \begin{cases} C_1^* & \text{if } p = p^* \\ 0 & \text{if } p \neq p^* \end{cases} .$$

Matters are more complicated when there are more than one interest group.

In general, there will be many different equilibria.

Example 1

$P = \{0, 1\}$; $n = 3$

$$\begin{array}{ll} V_0(0) = 0 & V_0(1) = 0 \\ V_1(0) = 0 & V_1(1) = 3 \\ V_2(0) = 0 & V_2(1) = 2 \\ V_3(0) = 4 & V_3(1) = 0 \end{array}$$

One equilibrium is $p^* = 1$ and

$$\begin{array}{ll} C_1(0) = 0 & C_1(1) = 3 \\ C_2(0) = 0 & C_2(1) = 1 \\ C_3(0) = 4 & C_3(1) = 0 \end{array}$$

Another equilibrium is $p^* = 1$ and

$$\begin{array}{ll} C_1(0) = 0 & C_1(1) = 2 \\ C_2(0) = 0 & C_2(1) = 2 \\ C_3(0) = 4 & C_3(1) = 0 \end{array}$$

The source of the multiplicity in this instance is a certain arbitrariness in how the costs of providing policy $p^* = 1$ are divided between interest groups 1 and 2.

The multiplicity in this example just concerns the allocation of the contributions.

Is it possible to get multiple policy choices?

Example 2

$P = \{0, 1, 2\}$; $n = 3$

$$\begin{array}{lll} V_0(0) = 0 & V_0(1) = 0.5 & V_0(2) = -8 \\ V_1(0) = 0 & V_1(1) = 10 & V_1(2) = -60 \\ V_2(0) = 0 & V_2(1) = 15 & V_2(2) = -40 \\ V_3(0) = 0 & V_3(1) = 20 & V_3(2) = 60 \end{array}$$

One equilibrium is $p^* = 0$ and

$$\begin{array}{lll} C_1(0) = 33 & C_1(1) = 0 & C_1(2) = 0 \\ C_2(0) = 19 & C_2(1) = 0 & C_2(2) = 0 \\ C_3(0) = 0 & C_3(1) = 0 & C_3(2) = 60 \end{array}$$

One equilibrium is $p^* = 1$ and

$$\begin{array}{lll} C_1(0) = 0 & C_1(1) = 12 & C_1(2) = 0 \\ C_2(0) = 0 & C_2(1) = 19.5 & C_2(2) = 0 \\ C_3(0) = 0 & C_3(1) = 0 & C_3(2) = 40 \end{array}$$

Notice that $p = 0$ is a Pareto inefficient policy choice!

The problem here is one of coordination failure - interest groups 1 and 2 should coordinate to shift promises over to option 1.

To deal with this problem of multiple equilibria, Bernheim and Whinston (1986) argued for focusing on a special class of equilibria - *truthful equilibria*

The contribution schedule $C_i(\cdot)$ is *truthful relative to the policy* p' if for all $p \in \mathcal{P}$

$$C_i(p) = \max\{0, V_i(p) - V_i(p') + C_i(p')\}$$

Note that if the contribution schedule $C_i(\cdot)$ is truthful relative to the policy p', it reflects (on the margin) the interest group's true willingness to pay for the policy: i.e., if $C_i(p) > 0$

$$C_i'(p) = V_i'(p)$$

The equilibrium $C_1^*(\cdot), ..., C_n^*(\cdot)$ and p^* is *truthful* if for each interest group i $C_i^*(\cdot)$ is truthful relative to p^*.

Bernheim and Whinston show that for each interest group i and any contribution schedules for the other interest groups, the set of best responses for i contains a truthful strategy.

They also show that truthful equilibria exist and are coalition proof.

Bernheim and Whinston also provided a complete characterization of truthful equilibria.

In particular, they show that if the equilibrium $C_1^*(\cdot)$, ..., $C_n^*(\cdot)$ and p^* is truthful, then

$$p^* = \arg\max_{p \in P}[V_0(p) + \sum_{i=1}^{n} V_i(p)]$$

Thus, the equilibrium policy choice maximizes the sum of the politician's and interest groups' policy utility.

This result is taken as the distinctive prediction of the common agency model.

It provides a concrete way of thinking about how campaign contributions bias policy choices in favor of interest groups.

Unfortunately, there is still multiplicity with respect to contributions and in most applications the equilibrium contributions made by interest groups will not typically be uniquely defined.

Bernheim and Whinston's paper provides a complete characterization of the contributions that are possible in a truthful equilibrium.

Alternative Perspectives

The common agency model takes a very cynical view of the influence of campaign contributions.

Contributions are essentially bribes paid by interest groups in exchange for policy choices.

There is not much evidence to justify such a cynical view.

The model also makes very strong assumptions about the ability of interest groups to commit to deliver campaign contribution promises.

That is, once the politician has chosen p, what is to stop the interest groups reneging on their promise to contribute $C_i(p)$?

Fox and Rothenberg (2009) present an interesting political agency model showing how an interest group can have significant influence on policy outcomes without bribery or even spending any money at all.

This influence arises from the credible threat to contribute to an incumbent's challenger if the incumbent votes against their position.

For example, suppose that you are a congressman who has to decide whether to vote for a gun-control bill.

You know that if you vote in favor, the NRA will realize that you are anti-gun and support your opponent in the next primary or general election.

As a consequence, you abstain or vote against and the gun-control bill fails without the NRA spending a dime.

I.5 Political Parties

Political parties are groups of like-minded citizens who collectively organize to elect candidates.

Party members provide support to their candidates in the form of money and time.

Through their nominations, parties also provide candidates with a brand name which provides useful information to voters about candidates' ideologies (see Snyder and Ting (2002)).

In the legislature, political parties can pressure their candidates to further the ideological goals of the party.

In this section, we briefly consider how parties impact candidate competition and legislative decision-making.

We take as given party membership, which seems a reasonable short run assumption.

For models with endogenous party membership, see Jackson and Moselle (2002), Levy (2004) and Roemer.

I.5.i Parties and Candidate Competition

In the U.S., political parties select candidates through *primary elections*.

This is distinct from a system in which candidates are selected by party elites.

The type of primary elections vary by state.

In *closed primaries* only party members are allowed to vote, in *open primaries* voters can choose in which primary to vote on the day of the election.

States are approximately evenly divided between open and closed primaries (Gerber and Morton (1998)).

Primaries perform the function of winnowing the field of candidates down to two, one from each party.

Independent or third party candidates are not common in the U.S. and are rarely successful when they do run.

Primary elections can be straightforwardly incorporated into our earlier models of campaign competition.

We just need to categorize voters according to whether they are registered Democrats, registered Republicans, or unaffiliated.

Thus, we assume as before that citizens have ideologies distributed over the interval $[0, 1]$ and we let $F(i)$ be the fraction of voters with ideology $\leq i$.

But now we classify each voter as either a registered Democrat, a registered Republican, or unaffiliated.

Let γ_D be the fraction of voters who are registered Democrats and γ_R be the fraction of voters who are registered Republicans.

Let $F_D(i)$ be the fraction of Democrats with ideology $\leq i$ and $F_R(i)$ be the fraction of Republicans with ideology $\leq i$.

It is natural to assume that Democrats are to the left of the median voter and Republicans to the right.

We can identify a Democrat median voter i_{mD} and a Republican median voter i_{mR} in the obvious way (i.e., i_{mD} and i_{mR} satisfy $F_D(i_{mD}) = 1/2$ and $F_R(i_{mR}) = 1/2$).

It will be the case that $i_{mD} < i_m < i_{mR}$, where i_m is the aggregate median voter.

We also need to make an assumption as to how voters in primary elections vote.

In particular, in primary elections the distinction between simple and sophisticated sincere voting is important.

With simple sincere voting, voters in the primary just vote for the candidate who they like the best.

With sophisticated sincere voting, voters take into account how well candidates are going to perform in the general election.

This may involve voting for a less preferred candidate because he has a better chance in the general election.

For example, even if you really liked Dennis Kucinich, you might have avoided voting for him in the 2008 Democratic primaries on the grounds that there was no chance he could win the general election.

Downsian Candidates and Primaries

With simple sincere voting, the Downsian model predicts that candidates in closed primary elections would adopt the ideology of the median party member and then the ideology of the median voter in the general election.

This is indeed quite consistent with the conventional wisdom concerning candidate behavior in primaries and general elections.

Candidates are often said to be moving to the center after primary victories.

Nonetheless, the idea that candidates can completely change their positions once they win the primary is not very plausible.

With sophisticated sincere voting, whatever candidates said about their ideologies during the primary elections would be irrelevant as primary voters would anticipate that the candidates would move to the center once elected.

Thus, the Downsian perspective does not lend much insight in this case.

Policy-motivated Candidates and Primaries

The most natural way to incorporate primaries into the model of policy-motivated candidates is to assume that primaries determine the true ideologies of the two candidates competing in the general election.

Recall that in the model presented in I.2.ii, these true ideologies were denoted t_A and t_B.

Thus, we can think of primary voters choosing between candidates on the basis of their true ideologies.

The winning candidates will then moderate their true ideologies in the general election as in the model of I.2.ii.

In this case, whether we assume primary voters vote sincerely in a simple or sophisticated way will not matter.

With either form of voting, primary voters will want to vote for candidates who share their true ideologies.

In a two-candidate closed primary race, therefore, the candidate whose true ideology is closest to that of the party median voter will win.

This provides some motivation for assuming that the true ideologies t_A and t_B in the model of I.2.ii reflect those of the party median voters; i.e., $t_A = i_{mD}$ and $t_B = i_{mR}$.

Under this assumption, for any given specification of voters' party affiliations, we obtain a complete model of electoral competition which incorporates both primary and general elections.

Citizen-Candidates and Primaries

In a citizen-candidate model with closed primaries, party members would select a candidate from those citizens who have chosen to run in their party's primary election.

In the general election, all voters would vote between the candidates chosen in the two primaries.

Primaries would actually simplify the model because they would ensure that the general election would only involve two candidates (assuming the costs of running as an independent candidate are prohibitive).

Whether primary voters vote sincerely in a simple or sophisticated way will matter in the citizen-candidate model.

With simple sincere voting, primary voters will vote for the candidate whose true ideology is closest to their own.

With sophisticated sincere voting, primary voters will take into account that the true ideology of the winning candidate will determine his success in the general election.

They will therefore prefer candidates more moderate than themselves.

Jackson, Mathevet, and Mattes (2007) provide a formal treatment of this model.

Open versus Closed Primaries

The logic of both the policy-motivated candidate and the citizen-candidate models suggests that open primaries will be more likely to produce candidates closer to the center than closed primaries.

This is because moderate candidates in, say, the Democrat party, can pick up votes from Independents and moderate Republicans.

There is some empirical evidence in support of this prediction from both Congressional and state-level elections (see, for example, Gerber and Morton (1998)).

On the other hand, in open primaries there is the fear that, say, Republicans might vote for the weakest Democrat candidate thereby paving the way for an easy victory for the Republican nominee.

There is some anecdotal evidence of this type of "crossover voting", but not systematic evidence that it is significant.

I.5.ii Parties and Legislative Decision-Making

Legislators from the same political party can further the ideological objectives of party members by voting as a block.

To illustrate, consider a legislature consisting of 5 legislators indexed by $j = 1, ..., 5$.

Suppose that legislators have to decide on some ideological issue p.

Legislator j has ideal point p_j^* and quadratic preferences $-(p - p_j^*)^2$

The ideal points are ordered such that

$$p_1^* < p_2^* < p_3^* < p_4^* < p_5^*$$

Suppose that legislators 1, 2, and 3 belong to the D Party and legislators 4 and 5 belong to the R Party.

Thus, the D Party has a majority of seats in the legislature.

The Condorcet Winner is p_3^*, so assume that this is the outcome that would arise with no parties.

Suppose that legislator 2's ideology is close to the median member of the D Party, so that the majority of party members prefer p_2^* to p_3^*

The D party can *pressure* legislator 3 to support p_2^* against policy proposals to the right of p_2^*.

It can do this by threatening to withhold campaign money and support from legislator 3 or by threatening to run an opponent against him in the next primary.

If successful, this pressure will lead to the party's preferred policy outcome p_2^*.

When parties can control the voting decisions of their legislators in this way, the political system is said to have *strong parties*.

The U.S. political system is well known for having *weak parties* - it is unusual for party members to vote cohesively within their ranks.

For example, in the 105th Congress (1997-1999) only 11% of all roll call votes in the House saw at least 90% of the Democrats voting one way and at least 90% of the Republicans voting the other.

Gate-keeping Power

Even with weak parties, the majority party in a legislature can influence policy if it effectively controls what issues are brought up.

This is referred to as *gate-keeping power*.

In any legislative session, there are way more issues to deal with than time to deal with them.

Thus, in any legislative session, there are many important issues that are never brought up: for example, immigration reform has not been considered recently by the U.S. Congress despite its central importance.

To illustrate the significance of gate-keeping power, consider our earlier 5 legislator example.

Suppose that the rules of the legislature are that the majority party controls whether the legislature considers the ideological issue p.

Further suppose that the decision to bring up the issue is made by the leader of the majority party who is elected by the legislators in his party.

It makes sense to assume that the D party legislators will elect legislator 2 since he is the Condorcet winning legislator in the set $\{1, 2, 3\}$.

Now consider legislator 2's decision to bring up issue p.

Suppose that the *status quo* level of the policy is p_0 - this is the level that will prevail if no action is taken.

Suppose that if the issue is brought up, then because of weak parties, the median legislator will prevail and the policy will be changed to p_3^*.

Then legislator 2 will bring the issue to the floor only if he prefers p_3^* to p_0.

Gate-keeping power thus ensures that policies will gradually gravitate in the direction favored by legislator 2.

Assuming that legislator 2's preferences are close to the median D party member's policy, this means that policies will move towards those favored by the majority of D party members.

Of course, this analysis begs the question of why the legislature has a rule under which the majority party controls the agenda.

For more on this see Diermeier and Vlaicu (2011).

I.6 The Media

The political role of the media is to provide information to citizens concerning policy options, candidate positions, and the performance of politicians.

TV broadcasts and newspaper stories are probably the most important source of information concerning these issues.

The information the media provides impacts voters' beliefs and hence preferences over candidates and policies.

These voter preferences determine election outcomes and thus policy outcomes.

There is plenty of evidence that the media matters for voters' information.

Survey evidence reveals that the more voters are exposed to the media, the more likely they are to be informed about candidates.

Snyder and Stromberg (2010) in a paper not on the reading list ("Press Coverage and Political Accountability" published in the *Journal of Political Economy*) show that voters living in a district with less press coverage of their congressman, are less informed about their congressman.

Media coverage matters not only for elections, but also day to day policy-making.

Eisensee and Stromberg (2007) explore how the extent of news coverage of foreign disasters determines the extent of U.S. federal disaster relief.

They show that the extent of U.S. relief depends on whether the disaster occurs at the same time as other unrelated news-worthy events (e.g., the Olympic games).

This suggests that relief decisions are driven by news coverage of disasters and that other newsworthy events crowds out such coverage.

The threat of negative media exposure and/or the promise of positive media exposure can also incentivize elected politicians.

Gentzkow, Glaeser and Goldin (2004) document the historical development of informative newspapers in the U.S. and show that it mattered for the reduction of corruption.

Snyder and Stromberg (2010) show that in districts with less press coverage, congressmen work less for their constituents.

Moreover, such districts obtain less federal spending.

The central issue in the literature is how well does the media do its job of providing accurate information to citizens?

The key issue here is *media bias*.

Intuitively, this refers to the idea that media outlets slant their reporting of the news in one or the other ideological direction.

Media bias has received a huge amount.

We review three issues: i) measuring media bias; ii) the impact of media bias on political behavior; and iii) the determinants of media bias.

I.6.i Measuring Media Bias

Media bias is something that people love to talk about and debate, but how can it be quantified?

There have been many attempts to measure media bias.

We discuss two recent efforts.

Groseclose and Milyo

Groseclose and Milyo (2005) estimate ideological scorces for a number of different media outlets.

To compute these scores, the authors count the times that a particular media outlet cites various think tanks and policy groups.

They then compare this with the times that members of Congress cite the same groups in speeches in the House and Senate.

Think tanks include institutions like AEI, Brookings Institution, Heritage Foundation, Economic Policy Institute, and policy groups include interest groups like NAACP, NRA, and Sierra Club.

To understand Groseclose and Milyo's procedure, suppose that there are just two think tanks cited in congressional speeches and the media: AEI and Brookings.

Suppose that a particular news show mentions AEI twice as much as Brookings.

Then the authors assign this show the same ideological score that is assigned to a member of Congress who mentions AEI twice as much as Brookings in his/her speeches.

The relevant ideological scores are those assigned to members of Congress by *Americans for Democratic Action* - they are known as *ADA scores*.

Groseclose and Milyo's results show a strong liberal bias.

All of the new outlets they examine have ideological scores to the left of the average member of Congress except Fox News' Special Report and the Washington Times.

CBS Evening News and the New York Times received scores far to the left of center.

The most centrist news outlets were ABC's Good Morning America, CNN's NewsNight, and PBS NewsHour.

The only surprising result was that the Wall Street Journal was far to the left.

Note, however, that Groseclose and Milyo's procedure only uses news stories and excludes editorials and opinion pieces.

Gentzkow and Shapiro

Gentzkow and Shapiro (2010) build on Groseclose and Milyo's work in a study of U.S. daily newspapers.

They use an index of media slant that measures whether a newspaper's language is more similar to a congressional Republican or Democrat.

To do this they measure the set of all phrases used by members in the 2005 Congressional record and identify those that are used much more frequently by one party than the other.

For example, strongly Republican phrases were "War on Terror", "Death Tax", and "Tax Relief".

Strongly Democrat phrases were "War in Iraq", "Estate Tax", and "Tax Break".

They then index newspapers by the extent to which the use of politically charged phrases in their news coverage resembles the use of the same phrases in the speech of congressional Republicans or Democrats.

The motivation for this measure is that the language chosen by congressional Republicans and Democrats is presumably intended to persuade listeners to support their agendas.

If a news show uses similar language to, say, congressional Republicans, it is natural to assume that it wishes to persuade listeners to support the congressional Republican agenda.

Getnzkow and Shapiro's measure is broadly in agreement with that of Groseclose and Milyo.

They find that the New York Times, Washington Post, and Los Angeles Times are fairly similar to each other and similar to a fairly liberal congressman.

They find that USA Today is closer to the center and that the Washington Times is significantly to the right.

They also find that the Wall Street Journal is fairly right leaning, which seems a more sensible conclusion than that of Groseclose and Milyo.

I.6.ii Does Media Bias Matter?

Given the evidence that different media sources have different political slants, it is important to ask how this impacts citizens' political behavior.

Intuitively, it is not clear that bias will have any impact at all.

After all, rational readers should simply adjust for a newspaper or TV show's bias.

Thus, if the New York Times is reporting that the Democrats Health Care Reform bill is expected to save billions of dollars, a rational reader should be more susupicious of such a claim than if the Washington Times were making it.

There are three recent studies that speak to the issue of the impact of bias.

Gerber, Karlan, and Bergan

Gerber, Karlan and Bergan (2009) develop a field experiment to investigate how newspapers matter for voter behavior.

One month prior to the 2005 Virginia gubernatorial election, they administered a survey to find out people's political views and whether they subscribed to the Washington Post or Washington Times.

From the group who received neither, they randomly assigned people to three treatments: receive a free subscription of the Washington Times; receive a free subscription of the Washington Post; and receive nothing.

People received their free subscriptions for a 5 week period, beginning 3 weeks before the election.

As we noted above, the Washington Times is very conservative and the Washington Post is liberal.

The authors then conducted a survey of these households 1 week after the gubernatorial race.

This race was won by Democrat Tim Kaine with 51.7% of the vote.

Those assigned the Washington Post were 8% more likely to vote for Tim Kaine when compared with the control group who did not receive free subscriptions.

There was not a symmetric impact for those assigned the Washington Times; i.e., those assigned the Times were not more likely to vote for the Republican candidate Jerry Kilgore.

The authors argue that the asymmetric results could reflect the fact that this was just a very bad period for Republicans.

Indeed, during this period, Bush's approval rating fell by 4%.

Thus, during this period, greater exposure to the news might naturally lead voters to shift to the Democrats.

So the rightward bias effect of the Times could have been offset by this general shift.

The authors also found leftward shifts in public opinion on specific issues as a result of receiving the Post.

DellaVigna and Kaplan

DellaVigna and Kaplan (2007) show that households exposed to the Fox News Channel were more likely to vote Republican.

How did they establish this?

Between 1996 and 2000, the Fox News Channel was introduced in the cable programming of 20% of US towns.

The Fox News Channel is significantly to the right of all the other mainstream TV networks (CNN, ABC, NBC, CBS)

The towns in which it was introduced seemed basically random.

Given this randomness, the introduction of the Fox News Channel constitutes a *natural experiment*.

The authors investigate if those towns in which the Fox News Channel was present in 2000 gained Republican vote share in the 2000 elections as compared with the 1996 elections.

They computed the change in Republican vote share in the 1996 and 2000 Presidential elections and conclude that the Republicans gained between 0.4 and 0.7 of a percentage point.

The way this was established was by comparing the change in vote share with that in neighboring towns in which the Fox News Channel was not introduced.

Remember that the 2000 Presidential election was between Bush and Gore and was extremely close.

Gore got more of the popular vote, but marginally lost the electoral college vote.

The state of Florida was decisive and the vote totals were very close.

Thus, the Fox News effect was large enough to have been decisive in Florida.

The authors did the same exercise for U.S. Senate elections and got an even bigger effect (0.8 of a percentage point).

Since there is not much coverage of Senate elections on FNC, they took this as evidence of FNC creating a general boost for Republicans.

The effect seemed to have come from turnout rather than existing voters switching party lines.

Thus, FNC mobilized Republican voters.

Chaing and Knight

Chaing and Knight (2011) study the influence of newspaper endorsements.

It is commonplace for U.S. newspapers to publish candidate endorsements on their editorial pages.

These endorsements are made close to the election day and often contain some discussion justifying the endorsement.

Chaing and Knight show that these endorsements are influential in the sense that voters are more likely to support the recommended candidate.

The degree of this influence, however, depends on the bias of the newspaper.

Voters are more influenced by left-leaning newspaper endorsements of Republican candidates and by right-leaning newspaper endorsements of Democratic candidates.

This is consistent with voters being uninformed but rationally adjusting for a newpaper's bias.

Chaing and Knight also find that centrist voters are more likely to be influenced by endorsements which also makes sense intuitively.

I.6.iii Determinants of Bias

Why do media outlets have different bias?

There are two possible avenues: supply side and demand side.

One supply side argument is that bias stems from journalists and reporters.

Only 8% of national journalists decribe themselves as conservative, while 32% describe themselves as liberal.

This compares with 36% of the U.S. population who describe themselves as conservative and 19% who describe themselves as liberal.

An alternative supply side argument is that the bias may stem from owners of newspapers or TV stations who are pushing an ideological agenda.

Concern over this possibility has led to regulations that try to limit the concentration of ownership of newspapers and TV stations.

Demand side arguments suggest that newspapers and TV shows are biased because their readers and viewers like it that way.

These arguments assume that readers enjoy holding certain beliefs and that they like seeing these confirmed by the media.

We already discussed the possibility that voters might enjoy holding certain beliefs, when we discussed voter information.

To rationalize this demand side story we just need to add the assumption that newspapers and TV shows can slant stories towards these beliefs (which they certainly can).

See Mullainathan and Shleifer (2005) for a model along these lines.

Gentzkow and Shapiro's study comes down strongly in favor of the importance of demand side considerations in explaining newspaper bias.

They suggest that citizens get politically biased newspaper coverage because thats exactly what they want.

II. Voting Rules and Electoral Systems

There are many different ways in which communities can elect leaders and make collective choices.

In this section, we discuss some of these ways.

We first identify a number of different voting rules, which are ways of choosing a single alternative from a set of alternatives.

We then discuss some general theoretical results concerning voting rules.

Finally, we discuss alternative electoral systems, which are ways of electing legislatures.

II.1 Voting Rules

Consider a community with n citizens indexed by $i \in \{1, ..., n\}$.

Suppose that this community must choose some alternative a from a finite set of alternatives $A = \{a_1, ..., a_m\}$.

This set of alternatives may be policy options or candidates.

Citizen i's utility if alternative a is selected is $V_i(a)$.

A *voting rule* is a method for choosing an alternative.

The simplest voting rule is *plurality rule* - each citizen has one vote to cast and the alternative with the most votes wins.

A common alternative to plurality rule is *majority rule with run-offs* - a plurality rule election is held and if no alternative has an absolute majority, a run-off election between the two top vote getters is held.

A more sophisticated alternative to plurality rule is *approval voting* in which citizens can vote for as many alternatives as they like and the one with the most votes wins.

Even more sophisticated is the *Borda count* in which citizens rank the alternatives.

Assuming that there are m alternatives, citizens assign 1 vote to their first choice; $\frac{m-2}{m-1}$ votes to their second choice; $\frac{m-3}{m-1}$ votes to their third choice; etc.

A still more sophisticated system, is the *single transferable vote* (also known as *Hare voting*).

Under this system, voters first rank all the alternatives.

If any alternative is ranked first by a majority, it wins.

If no winner exists, the alternative with the fewest first place votes is eliminated.

Ballots are then re-tabulated as if that alternative never existed.

The procedure is continued until one alternative is ranked first among the un-eliminated alternatives by a majority.

The key point to note about all these different rules, is that they produce different outcomes.

To see this we consider a very clever example, which is taken from Shepsle's textbook.

The example was apparently invented by Joseph Malkevitch.

A Clever Example

n is 55 and $A = \{a_1, a_2, a_3, a_4, a_5\}$.

Citizens rankings of the alternatives take one of six possible forms.

These six rankings are labelled I, II, III, IV, V, and VI and are illustrated in the following table.

I	II	III	IV	V	VI
18	12	10	9	4	2
a_1	a_2	a_3	a_4	a_5	a_5
a_4	a_5	a_2	a_3	a_2	a_3
a_5	a_4	a_5	a_5	a_4	a_4
a_3	a_3	a_4	a_2	a_3	a_2
a_2	a_1	a_1	a_1	a_1	a_1

The columns provide the rankings.

Thus, for a citizen i with ranking I,

$$V_i(a_1) > V_i(a_4) > V_i(a_5) > V_i(a_3) > V_i(a_2).$$

The number of citizens with each ranking is recorded in the second row.

Thus, there are 18 citizens who have ranking I, 12 citizens who have ranking II, etc, etc...

Lets figure out the outcome under the different voting rules.

We assume sincere voting throughout.

Plurality Rule

Under plurality rule, the votes for each option are

$$\begin{array}{ccccc} a_1 & a_2 & a_3 & a_4 & a_5 \\ 18 & 12 & 10 & 9 & 6 \end{array}$$

The winner is therefore a_1.

Majority Rule with Run-Offs

Under majority rule with run offs, the first round votes are

a_1	a_2	a_3	a_4	a_5
18	12	10	9	6

Thus, alternatives $\{a_3, a_4, a_5\}$ are eliminated.

The run off votes are

a_1	a_2
18	37

The winner is therefore a_2.

Single Transferable Vote

In the first round, alternative a_5 is eliminated since it is ranked first by only 6 voters.

In the second round, with a_5 eliminated, the rankings are

$$\begin{array}{cccccc} 18 & 12 & 10 & 9 & 4 & 2 \\ a_1 & a_2 & a_3 & a_4 & a_2 & a_3 \\ a_4 & a_4 & a_2 & a_3 & a_4 & a_4 \\ a_3 & a_3 & a_4 & a_2 & a_3 & a_2 \\ a_2 & a_1 & a_1 & a_1 & a_1 & a_1 \end{array}$$

Alternative a_4 is eliminated since it is ranked first by only 9 voters.

In the third round, with a_4 eliminated, the rankings are

$$\begin{array}{cccccc} 18 & 12 & 10 & 9 & 4 & 2 \\ a_1 & a_2 & a_3 & a_3 & a_2 & a_3 \\ a_3 & a_3 & a_2 & a_2 & a_3 & a_2 \\ a_2 & a_1 & a_1 & a_1 & a_1 & a_1 \end{array}$$

Alternative a_2 is eliminated since it is ranked first by only 16 voters.

In the fourth round, with a_2 eliminated, the rankings are

18	12	10	9	4	2
a_1	a_3	a_3	a_3	a_3	a_3
a_3	a_1	a_1	a_1	a_1	a_1

Alternative a_3 is the winner.

Borda Count

Under Borda Count, the votes are

$$\begin{array}{lc}
a_1 & 18 \\
a_2 & 12 + \left(\frac{3}{4}\right) 14 + \left(\frac{1}{4}\right) 11 = 25.25 \\
a_3 & 10 + \left(\frac{3}{4}\right) 11 + \left(\frac{1}{4}\right) 34 = 26.75 \\
a_4 & 9 + \left(\frac{3}{4}\right) 18 + \left(\frac{2}{4}\right) 18 + \left(\frac{1}{4}\right) 10 = 34 \\
a_5 & 6 + \left(\frac{3}{4}\right) 12 + \left(\frac{2}{4}\right) 37 = 33.5
\end{array}$$

The winner is therefore a_4.

Approval Voting

The outcome under approval voting is ambiguous because it is unclear how far down their rankings citizens will go before they decide not to approve.

All that is clear is that they will not approve their least preferred alternative but will approve their most preferred alternative.

Suppose they all approve their top three.

Then the votes are

a_1	18
a_2	26
a_3	21
a_4	45
a_5	55

The winner is therefore a_5.

Discussion

In this example, each voting rule produces a different outcome!

The example beautifully illustrates the point that the voting rule matters.

This raise the question of which is the “best” voting rule?

II.2 General Results

We now discuss two general results relating to voting rules which are relevant for thinking about the question of a best voting rule.

II.2.i Arrow's Impossibility Theorem

Arrow was concerned with the problem of constructing a social ranking of alternatives from a set of individual rankings.

The different voting rules that we have discussed represent different ways of doing this.

They all provide ways of going from individual rankings to a social ranking.

For example, in the Clever Example, the individual rankings are those given in the initial table.

The social ranking under plurality rule is

$$a_1 \succ a_2 \succ a_3 \succ a_4 \succ a_5,$$

where the notation $\succ$ means is ranked higher than.

The social ranking under the Borda count is

$$a_4 \succ a_5 \succ a_3 \succ a_2 \succ a_1.$$

Arrow wanted to find "good" ways of constructing these social rankings.

Clearly, this is an interesting problem, since if we can agree on what are good and not so good ways to construct social rankings we may then be able to judge between different voting rules.

The way Arrow approached the problem was to specify a set of conditions that he thought any good method of social ranking should satisfy.

He then proved that there was no method which satisfied these conditions.

Arrow's framework

Arrow's underlying framework was similar to the one we used to discuss voting rules, except that he assumed that citizens simply have rankings over the alternatives rather than utility functions.

Recall from micro, that rankings (or equivalently preferences) are the more fundamental concept.

Utility functions are just a way of representing preferences.

Thus, we have a community with n citizens indexed by $i \in \{1, ..., n\}$.

This community must choose some alternative a from a finite set of feasible alternatives $A = \{a_1, ..., a_m\}$.

The number of alternatives is m and we assume $m \geq 3$.

Each citizen i has a ranking or preference over these alternatives denoted $\succeq_i$.

The notation $a_j \succeq_i a_k$ means that citizen i ranks alternative a_j at least as high as alternative a_k (or equivalently *weakly prefers* alternative a_j to alternative a_k).

The notation $a_j \succ_i a_k$ means that $a_j \succeq_i a_k$ but that it is not the case that $a_k \succeq_i a_j$; i.e., citizen i ranks alternative a_j higher than alternative a_k (or equivalently *strictly prefers* alternative a_j to alternative a_k).

We assume that each citizen i's ranking is *complete* and *transitive*.

Complete means that citizen i can rank any two alternatives.

Formally, for any two alternatives a_j and a_k, either $a_j \succeq_i a_k$ or $a_k \succeq_i a_j$.

Transitive means that if $a_j \succeq_i a_k$ and $a_k \succeq_i a_l$ then $a_j \succeq_i a_l$.

Transitivity is a basic requirement of rationality.

Arrow's problem was to find a method of constructing from individual rankings $(\succeq_1, \succeq_2, ..., \succeq_n)$ a social ranking $\succeq_S$.

Arrow's conditions

Condition U (Universal Domain): the method should produce a complete and transitive social ranking for any possible vector of complete and transitive individual rankings $(\succeq_1, \succeq_2, ..., \succeq_n)$.

Condition P (Pareto Optimality): if each citizen ranks alternative a_j higher than alternative a_k, then the social ranking must rank a_j higher than alternative a_k (i.e., if for each citizen i, $a_j \succ_i a_k$ then $a_j \succ_S a_k$).

Condition I (Independence of Irrelevant Alternatives): if alternatives a_j and a_k stand in a particular relationship to one another in each individual's ranking and this relationship does not change, then neither may the social ranking.

This is the case even if individual rankings over other (irrelevant) alternatives change.

The way Condition I is formally stated is as follows.

Let $(\succeq_1, \succeq_2, ..., \succeq_n)$ and $(\succeq'_1, \succeq'_2, ..., \succeq'_n)$ be two vectors of individual rankings and let $\succeq_S$ and $\succeq'_S$ be the associated social rankings.

Suppose that for each citizen i $a_j \succeq_i a_k$ if and only if $a_j \succeq'_i a_k$ then it must be the case that $a_j \succeq_S a_k$ if and only if $a_j \succeq'_S a_k$.

Condition D (Nondicatorship): there is no individual i whose own ranking dicatates the social ranking independent of the other individuals in the community.

Formally, there is no individual i such that the social ranking associated with any vector of individual rankings $(\succeq_1, \succeq_2, ..., \succeq_n)$ always equals $\succeq_i$.

Arrow's theorem

Arrow's Impossibility Theorem. *There exists no method of constructing a social ranking that simultaneously satisfies Conditions U, P, I, and D.*

The proof is involved, so we will not try to go through it.

There is a proof in Mas-Colell, Green and Whinston (1995).

We will just illustrate the theorem in action.

Arrow's theorem in action

The way to appreciate the theorem is to see how different methods of constructing social rankings violate the various conditions.

Majority rule

Suppose we construct a social ranking in the following way: $a_j \succeq_S a_k$ if for a majority of citizens $a_j \succeq_i a_k$.

This method violates Condition U; i.e., it does not always produce a complete and transitive social ranking.

Suppose there are three alternatives and three citizens with rankings as follows:

1	2	3
a_1	a_2	a_3
a_3	a_1	a_2
a_2	a_3	a_1

Then $a_1 \succ_S a_3$, $a_3 \succ_S a_2$, but $a_2 \succ_S a_1$.

The difficulty which arises here is that voters' preferences over alternatives are not "single-peaked".

Random choice

We randomly assign the alternatives labels - alternative #1, alternative #2, etc.

Then the social ranking is

$$a_1 \succeq_S a_2 \succeq_S \ldots\ldots \succeq_S a_m$$

This violates Condition P, since the ranking is completely insensitive to individuals' rankings.

Borda count

This method violates Condition I.

Let $n = 7$ and $A = \{a_1, a_2, a_3, a_4\}$

Consider the following ranking.

$$\begin{array}{ccc} 3 & 2 & 2 \\ a_3 & a_2 & a_1 \\ a_2 & a_1 & a_4 \\ a_1 & a_4 & a_3 \\ a_4 & a_3 & a_2 \end{array}$$

The top row refers to the number of citizens with the ranking in question and the column refers to the ranking.

Under the Borda count, the votes are

$$\begin{array}{ll} a_1 & 3\left(\frac{1}{3}\right)+2\left(\frac{2}{3}\right)+2=4.333 \\ a_2 & 3\left(\frac{2}{3}\right)+2=4 \\ a_3 & 3+2\left(\frac{1}{3}\right)=3.666 \\ a_4 & 2\left(\frac{1}{3}\right)+2\left(\frac{2}{3}\right)=2 \end{array}$$

In particular note that $a_1 \succ_S a_2$ and $a_2 \succ_S a_3$.

Now consider the following ranking, where all that has changed is the relative position of a_4

$$\begin{array}{ccc} 3 & 2 & 2 \\ a_3 & a_2 & a_1 \\ a_2 & a_1 & a_3 \\ a_1 & a_3 & a_2 \\ a_4 & a_4 & a_4 \end{array}$$

Under the Borda count, the votes are

$$
\begin{array}{ll}
a_1 & 3\left(\frac{1}{3}\right)+2\left(\frac{2}{3}\right)+2=4.333 \\
a_2 & 3\left(\frac{2}{3}\right)+2+2(\frac{1}{3})=4.666 \\
a_3 & 3+2\left(\frac{1}{3}\right)+2(\frac{2}{3})=5 \\
a_4 & 0
\end{array}
$$

Thus, $a_2 \succ_S a_1$ and $a_3 \succ_S a_2$.

Thus, the social rankings of these options has been flipped.

Nonetheless, comparing across the two sets of individual rankings we see that in both rankings a_1 and a_2 are ranked in the exact same way (relative to each other).

Similarly, in both rankings a_2 and a_3 are ranked in the exact same way (relative to each other).

Monarchy

Suppose we have a King (say, citizen 1) and we construct a social ranking in the following way: $a_j \succeq_S a_k$ if $a_j \succeq_1 a_k$.

This method violates Condition D.

Discussion

Arrow's Theorem is considered an important result.

However, there is some question about the relevance of Condition I.

In particular, it is not clear why Condition I should be considered a vital condition for a method of constructing a social ranking to satisfy.

For example, there seems nothing particularly egregious about the fact that a_3 becomes the most preferred outcome in the Borda count example.

II.2.ii Gibbard-Satterthwaite Theorem

In considering the problem of constructing a social ranking of alternatives from a set of individual rankings, Arrow implicitly assumed that individuals would report their true rankings accurately.

However, it is fairly obvious that in some circumstances individuals will have an incentive to misreport their rankings.

Lets consider two motivating examples.

Motivating examples

Plurality rule example

We have already seen this type of issue arise when we discussed the distinction between sincere and strategic voting in plurality rule elections with three or more candidates.

Let us translate the example that we used in Part I to illustrate the distinction between sincere and strategic voting into the Arrow framework.

Let $n = 5$ and let $A = \{a_1, a_2, a_3\}$.

The 5 citizens' rankings are illustrated in the following table.

$$\begin{array}{ccccc} a_1 & a_2 & a_2 & a_3 & a_3 \\ a_2 & a_1 & a_1 & a_1 & a_1 \\ a_3 & a_3 & a_3 & a_2 & a_2 \end{array}$$

If citizens report their rankings truthfully, the outcome under plurality rule is a tie between a_2 and a_3 which would result in each of these alternatives being chosen with equal probability.

However, this gives citizen 1 an incentive to report that he ranks alternative a_2 first.

Borda count example

Consider the Borda count example that we just used to illustrate the violation of Condition I.

Let $n = 7$ and $A = \{a_1, a_2, a_3, a_4\}$

The rankings are

$$\begin{array}{ccc} 3 & 2 & 2 \\ a_3 & a_2 & a_1 \\ a_2 & a_1 & a_4 \\ a_1 & a_4 & a_3 \\ a_4 & a_3 & a_2 \end{array}$$

The top row refers to the number of citizens with the ranking in question and the column refers to the ranking.

The votes are

$$
\begin{array}{ll}
a_1 & 3\left(\frac{1}{3}\right)+2\left(\frac{2}{3}\right)+2=4.333 \\
a_2 & 3\left(\frac{2}{3}\right)+2=4 \\
a_3 & 3+2\left(\frac{1}{3}\right)=3.666 \\
a_4 & 2\left(\frac{1}{3}\right)+2\left(\frac{2}{3}\right)=2
\end{array}
$$

This means alternative a_1 wins.

Suppose that one of the citizens with the first type of ranking reports instead a ranking $a_2 \succ a_3 \succ a_4 \succ a_1$.

Then we end up with the following table

2	1	2	2
a_3	a_2	a_2	a_1
a_2	a_3	a_1	a_4
a_1	a_4	a_4	a_3
a_4	a_1	a_3	a_2

The votes are

$$
\begin{array}{cc}
a_1 & 2\left(\frac{1}{3}\right)+2\left(\frac{2}{3}\right)+2=4 \\
a_2 & 2\left(\frac{2}{3}\right)+1+2=4.333 \\
a_3 & 2+\left(\frac{2}{3}\right)+2\left(\frac{1}{3}\right)=3.333 \\
a_4 & \left(\frac{1}{3}\right)+2\left(\frac{1}{3}\right)+2\left(\frac{2}{3}\right)=2.333
\end{array}
$$

The winning alternative is now a_2.

By misreporting, the citizen flips the outcome from a_1 to a_2.

Since the citizen prefers a_2 to a_1, he has the incentive to misreport.

The theorem

These examples raise the following interesting question: is there any voting rule which does not in some circumstances provide citizens with incentives to misreport their rankings?

The Gibbard-Satterthwaite Theorem tells us that the answer is no.

Gibbard and Satterthwaite work within the same general framework as Arrow.

However, to formalize their question, they distinguish between a citizen's *true* ranking over alternatives, denoted as before by $\succeq_i$, and his *reported* ranking, denoted $\succeq_i^r$.

They define a voting rule as a function which maps citizens reported rankings into alternatives.

Thus $f(\succeq_1^r, \succeq_2^r, ..., \succeq_n^r)$ is the alternative chosen when the citizens report rankings $(\succeq_1^r, \succeq_2^r, ..., \succeq_n^r)$.

They restrict attention to voting rules which have the property that for each possible alternative $a_j \in \{a_1, ..., a_m\}$ there exists some vector of individual rankings for which the rule would pick a_j.

This seems unobjectionable - for example, suppose that all individuals rank a_j the highest.

They define a voting rule to be *manipulable* if there exists some citizen i, some reported ranking $\succeq_i^r$, and some vector of true rankings $(\succeq_1, .., \succeq_i, ..., \succeq_n)$ such that

$$f(\succeq_1, .., \succeq_i^r, ..., \succeq_n) \succ_i f(\succeq_1, .., \succeq_i, ..., \succeq_n)$$

Intuitively, when the true rankings are $(\succeq_1, .., \succeq_i, ..., \succeq_n)$ citizen i gains from reporting the ranking $\succeq_i^r$ rather than reporting his true ranking $\succeq_i$ when everyone else reports their true ranking.

In game-theoretic language, if a voting rule is manipulable, reporting truthfully is not a dominant strategy equilibrium for some vector of true rankings.

Finally, they define a voting rule to be *dictatorial* if there exists some citizen i such that whatever the citizens report, the voting rule always chooses the outcome that citizen i ranks highest.

Gibbard-Satterthwaite Theorem. *Every non-dictatorial voting rule is manipulable.*

Discussion

The Gibbard-Satterthwaite Theorem is considered important in the literature.

The point seems unobjectionable - any voting rule is going to sometimes give people an incentive to misreport their preferences.

The point is arguably not so surprising, but the way the problem is formalized is nice and paved the way for the mechanism design literature.

II.2.iii Further work on voting rules

The two general results we have described are both negative results.

Arrow's theorem tells us that any voting rule will for some configurations of citizens' preferences produce what are arguably pathological results.

Gibbard and Satterthwaite's theorem tells us that any voting rule will for some configurations of citizens' preferences provide incentives to vote in a way that does not reflect their sincere preferences.

The results are often interpreted as suggesting that there is no best voting rule and that it is sort of pointless to look for one.

Nonetheless, communities must choose the rules that they will use to decide and hence the question still remains as to which one should they choose.

More recent literature has tried to be more practically orientated and compare the relative merits of different voting rules.

Cox (1987) considers Downsian competition between an exogenous number (3 or more) of candidates and provides a comparison of various voting rules.

He shows that there are no symmetric equilibria under plurality rule, but with the Borda count there exists a unique symmetric equilibrium in which all candidates adopt the position of the median voter.

Myerson (1999) argues that plurality rule provides candidates with greater incentive to create programs that transfer resources to small minorities of voters than do rules like Borda count.

These analyses assume: (i) sincere voting; (ii) Downsian candidates; and (iii) a fixed number of candidates.

Myerson (1999) also compares plurality rule and approval voting with strategic voting and three Downsian candidates.

He shows that approval voting yields median outcomes, while plurality rule need not.

The problem is coordination failure among voters stemming from the wasted vote idea.

Osborne and Slivinski (1996) compare outcomes under plurality rule and majority rule with run-offs in a citizen-candidate model with sincere voting.

They find that two candidate elections are "more likely" under plurality rule and that candidates' positions are "less differentiated" under majority rule with run-offs.

Dellis and Oak (2003) compare outcomes under plurality rule and approval voting in a citizen-candidate model with strategic voting.

Interestingly, they find that approval voting is capable of generating similar anomalies to plurality rule.

They do, however, present conditions under which approval voting leads to more moderate outcomes.

Comparison of voting rules is an important area of ongoing research.

There is plenty of comparisons to do, but results are hard to find.

II.3 Electoral systems

To this point, we have focused on methods communities can use to choose a single alternative.

We now broaden the discussion to consider methods to choose a legislature.

If we have single-seat districts (as in U.S. congressional districts), we just need to elect a single candidate in each district and we can use any of the voting rules that we have already discussed.

Matters get more complicated if we have multi-seat districts.

In some countries, for example, the whole country elects the whole legislature (Isreal and the Netherlands) so there is only one district!

Suppose that we have to elect $k \geq 2$ candidates from a district.

Single non-transferable vote

The simplest system would be the *single non-transferable vote*.

Under this system, citizens have one vote to cast and the k candidates with the most votes are elected.

This system is used in Japan for the lower house, where each district has three to five seats.

While this system generalizes plurality rule, it creates new strategic issues - for example, citizens may not cast their vote for the candidate they like most if they think he is already going to be in the top k.

In addition, political parties face a difficult choice in knowing how many candidates to put up - a large party may end up with few seats if it puts up too many candidates and voters split their votes among them.

Limited vote

A variant of this system is the *limited vote*, under which each citizen has m votes (where m is between 2 and k) and the k candidates with the most votes are elected.

For example, in a district with four seats, voters may be allowed to vote for two candidates.

This system is used in local elections in the U.S. and also to elect the senate in Spain.

Single transferable vote

The *single transferable vote* system discussed above can be generalized to handle multi-seat districts.

It is used by Ireland and Malta to elect representatives to their lower houses - they have districts with three to five seats.

One needs to define a *quota* q of first place votes above which any candidate gets a seat and also a *transfer rule* by which a winning candidate's surplus is redistributed to the other candidates.

We go through the same procedure as described in Part II.1 until we find the first candidate who has more than q first place votes in the set of uneliminated candidates.

This candidate then takes one of the seats and is eliminated.

If his first place votes were s then his surplus votes $s-q$ must be redistributed to the other candidates.

Suppose that candidate i is in second place on l_i of these s ballots.

Then, in the next round he is awarded $l_i(\frac{s-q}{s})$ first place votes.

The weighting ensures that the sum of all candidates' scores remains equal to the total number of votes cast.

To illustrate, suppose we have 3 candidates, A, B, & C, 100 voters, and 2 seats.

We set $q = 34$.

Suppose that the voters rankings are

50	25	25
A	B	C
B	A	B
C	C	A

Then, A is elected in the first round and his surplus is $50 - 34 = 16$.

Thus, the second round votes are $25 + 50(\frac{16}{50})$ for B and 25 for C.

B is then elected in the second round.

The single transferable vote system is very complicated and not practical in circumstances in which voters are electing many representatives.

In such circumstances, the most common electoral system is some form of *proportional representation* (PR).

Proportional representation

When candidates are put forward by competing political parties, *List-PR* systems are common.

Under these systems, the alternatives on the ballot are lists of candidates, each list associated with a different party.

List PR systems differ in terms of whether citizens can vote for candidates or for parties.

In a *closed-list PR system*, citizens can vote only for parties and then seats are allocated to parties in proportion to their vote share.

If, say, Party A gets 10 seats then the first 10 candidates on Party A's list are allocated these seats.

Each party chooses the order in which to list candidates.

This gives party leadership a large amount of power and gives rise to *strong parties* (as defined in Part I.5).

The way the seats are allocated among parties is straightforward.

For example, suppose that we have a five seat district, five parties compete and the percent vote distribution among them is 48.5-29-14-7.5-1.

The quota required for a seat is 20%.

The largest party is allocated two seats and this leaves a surplus of 8.5.

The second largest party is allocated one seat and this leaves a surplus of 9.

The remaining two seats are allocated to the third party and the second party - so the overall seat distribution is 2-2-1-0-0

In an *open-list PR system*, citizens can not only vote for parties but also which candidates from a party's list they prefer.

For example, in Belgium, the candidates are listed on the ballot and the voter places an optional "x" by his/her preferred candidates.

The rules then specify what percentage of votes cast a candidate must have in individual preferences to be moved up the list.

In this way, open-list PR allows voters to change the order in which party leaders have ranked candidates.

II.4 Comparing electoral systems

It is obviously interesting to think about the outcomes that different systems produce.

There is a huge political science literature devoted to this task.

A major focus is the comparison of proportional representation with plurality rule.

Under the former parties receive seat shares proportional to their votes and under the latter the party with the most votes in each district gets the seat.

There are two general claims commonly made about these systems.

Claim 1: number of parties

The first claim is that plurality rule leads to two party competition, while proportional representation leads to multiple parties.

This is known as *Duverger's Law*.

The logic underlying this claim is based on the idea of voters not wanting to waste their votes by voting for small parties under plurality rule.

This logic showed up in our discussion of why it could be an equilibrium for two extremist candidates to run against each other in the citizen-candidate model.

Under proportional representation, seats are proportional to votes and hence voting for a small party does not lead to a wasted vote.

Indeed, in proportional representation systems, small parties can have large influence if they form part of the governing coalition.

There is some evidence for this claim - there do tend to be less political parties in plurality rule systems and, in some instances (such as the U.S.) there are just two main parties.

Claim 2: types of policies

The second claim is that plurality rule provides parties with more incentive to use tax dollars to implement narrowly targeted transfer programs rather than providing universally valued public programs.

The logic behind this prediction is that plurality rule gives rise to a discontinuous relationship between seats and votes, whereby a party can get a whole lot more seats if it can increase its vote share to just slightly more than its closest rival.

This makes it highly valuable to attract the votes of small constituencies and hence leads to an abundance of special programs designed to appeal to small constituencies.

By contrast, in proportional representation systems, the relationship between seats and votes is continuous.

There is also some evidence for this claim: spending on broadly targeted social programs tends to be higher in proportional representation systems.

Related formal models

Lizzeri and Persico (2001) is an effort to come up with a formalization of Claim 2.

They study two-party competition for legislative seats under plurality rule (the party with the most votes gets all the seats) and proportional representation (seats are awarded in proportion to votes).

They show that plurality rule provides parties with more incentive to use tax dollars to implement narrowly targeted transfer programs rather than providing a universally valued public good.

One limitation of Lizzeri and Persico is their model of proportional representation assumes two parties.

One might imagine that if proportional representation systems allows more parties, then some of the small parties might represent small constituencies.

In general, progress on modelling policy formation in PR systems with many parties has been quite limited.

Baron and Diermier (2001) present an interesting model of elections and government formation in PR systems.

Their model has three policy-motivated parties.

Parties cannot commit and voters must anticipate the possible formation of a coalition government if one party does not have a majority of seats.

Government formation is modelled as a bargaining process.

Morelli (2004) undertakes the even more ambitious task of trying to compare both policy choice and party formation under plurality rule and proportional representation.

One of his goals is to try and formalize Claim 2.

His model incorporates strategic voters, strategic candidates, and strategic parties.

In fact, he does not find consistent results in terms of which system has the most parties, but proportional representation produces more centrist policies under some conditions.

Again, using formal models to explore the implications of different systems and testing their predictions empirically is an active area of research.

For more work in this style see Coate and Knight (2011), Persson, Roland, and Tabellini (2000), and Persson and Tabellini's book chapter in W & W.

III. Political Distortions

An important goal of public choice theory is to understand how political decision-making distorts policy choices away from those choices that would be made by the benevolent governments from economics textbooks.

Such governments intervene to correct market failures and to redistribute income to improve some appropriate notion of equity.

To the extent that policies emerging from the political process are similar to those that would be chosen by benevolent governments, it is not necessary to study politics: the *prescriptions* of normative theories of government can simply be interpreted as positive *predictions*.

There is a tradition in economics along these lines and so economists interested in politics have long been interested in understanding political distortions.

In these lectures we first briefly review some key concepts from the normative study of economic policy-making.

We then identify ways in which policies emerging from the political process may violate standard notions of equity.

Next we discuss reasons why politically-determined policies may be inefficient.

Finally, we discuss the implications of political distortions for the role of government.

III.1 Socially Optimal Policies

Consider a community with n citizens indexed by $i \in \{1, ..., n\}$.

Suppose that this community must choose some policy p from a set of feasible policies P.

Citizen i's utility if policy p is selected is $V_i(p)$.

What would be a good policy for this community?

III.1.i Pareto efficiency

The simplest requirement is that the policy be Pareto efficient.

Formally, a policy p^* is *Pareto efficient* if (i) it is feasible and (ii) there exists no alternative feasible policy p such that $V_i(p) \geq V_i(p^*)$ for each citizen i with the inequality holding strictly for at least one citizen.

As a requirement for a good policy, Pareto efficiency is certainly something that we would want.

The problem with Pareto efficiency is that a lot of policies can be Pareto efficient, so the requirement does not really narrow the field much.

Diagrammatic representation

Pareto efficient policies can be represented graphically in a two person community.

The *utility possibility set* is the set

$$\{(V_1(p), V_2(p)) : p \in P\}\,.$$

The *utility possibility frontier* is the outer envelope of this set.

Any point on the utility possibility frontier is associated with a Pareto efficient policy.

III.1.ii Social welfare functions

To go beyond the idea of Pareto efficiency, we have to be willing to make interpersonal comparisons of citizens' utilities.

That is, we have to be able to compare gains and losses of utility across citizens.

This is done with a social welfare function.

A *social welfare function* is a function which tells us for any given vector of citizen utility levels what the level of "social welfare" will be.

Mathematically, a social welfare function is represented as a real-valued function defined over citizens' utilities $W(V_1, ..., V_n)$.

Types of social welfare functions

The best known social welfare function is the *Utilitarian* social welfare function which is

$$W(V_1, ..., V_n) = \sum_i V_i.$$

This assumes social welfare is just the sum of individual citizens' utilities.

Another famous social welfare function is the *Rawlsian* social welfare function which is

$$W(V_1, ..., V_n) = \min_i V_i.$$

This assumes social welfare is the utility of the worst off individual in society.

A useful social welfare function that encompasses these two as special cases is the *Iso-elastic* social welfare function which is

$$W(V_1, ..., V_n) = \frac{1}{1-\alpha} \sum_i V_i^{1-\alpha}$$

where α is a positive constant.

When $\alpha = 0$, this just equals the Utilitarian social welfare function.

As α gets bigger and bigger, the Iso-elastic social welfare function approaches the Rawlsian social welfare function.

Diagrammatic representation

Social welfare functions can be represented graphically with *Social Indifference Curves*.

In a two-person economy, a social indifference curve is the set of utility pairs (V_1, V_2) satisfying the equation $W(V_1, V_2) = U$ for some constant U.

By varying the constant U we trace out a family of social indifference curves.

The Utilitarian social welfare function has indifference curves which are downward sloping straight lines, sloping left to right at a 45^o angle.

The Rawlsian social welfare function has L-shaped indifference curves which kink at the 45^o line.

The Iso-elastic social welfare function has convex indifference curves which lie between those of the Utilitarian and Rawlsian social welfare functions.

III.1.iii Socially optimal policies

Socially optimal policies are those that maximize social welfare as measured by the social welfare function.

Once we have agreed upon a social welfare function, we obtain the socially optimal policy (or policies) by solving the problem

$$\begin{gathered}\max W(V_1(p), ..., V_n(p)) \\ s.t.\ p \in P.\end{gathered}$$

Diagrammatically, the socially optimal policy is that associated with the point at which the social indifference curve is tangent to the utility possibility frontier.

Socially optimal policies are not only Pareto efficient but also allocate utility optimally across citizens.

What allocating utility optimally means will depend upon the social welfare function.

Thus, the socially optimal policy will depend on the specific social welfare function.

III.1.iv Political distortions

Political decision-making is going to lead the community to select some equilibrium policy and this policy will generate some utility allocation for the citizens.

Political distortions arise when the equilibrium policy diverges from the socially optimal policy.

There are two distinct types of distortions.

The first are *political inefficiencies* or *political failures* in which the equilibrium policy is not even Pareto efficient.

The second are *political inequities* in which the equilibrium policy is Pareto efficient but not socially optimal.

III.2 Politics and Equity

We begin by discussing political inequities.

It is easy to see that politically determined policies need not be socially optimal according to commonly used social welfare functions.

Recall our discussion of legislatures, which is where most policy issues are determined.

We saw that for ideological policy issues, such as the aggregate level of spending on a policy, or the scale of some regulation, the policy preferred by the median legislator would be likely to be chosen.

To the extent that the median legislator's preferences reflects the median citizen's preferences (for example, as in the Downsian model), this suggests that on ideological policy issues, politically determined policies will reflect the preferences of the median citizen.

But maximizing the preferences of the median citizen is typically inconsistent with maximizing social welfare.

To illustrate, suppose that the policy in question is spending on some public good (like parks, nature trails, or firework displays).

Suppose that citizens have different preferences for the public good but the costs are shared equally through taxation.

Specifically, assume that

$$V_i(p) = \theta_i \ln p - \frac{p}{n}.$$

The feasible set of policies is just the non-negative numbers.

Citizen i's preferred level of the public good is

$$p_i^* = \arg\max \left\{ \theta_i \ln p - \frac{p}{n} \right\}.$$

The preferred level p_i^* satisfies the first order condition

$$\frac{\theta_i}{p} - \frac{1}{n} = 0$$

implying that

$$p_i^* = n\theta_i.$$

Label the citizens so that $\theta_1 < ... < \theta_n$.

If m is the median citizen, the preferred policy of the median voter is $p_m^* = n\theta_m$.

Compare this with the Utilitarian solution which is $p_U^* = \sum_i \theta_i$.

These two will coincide if and only if the *median* equals the *mean* preference; that is, $\theta_m = \sum_i \theta_i / n$.

This is possible, but unlikely.

Stark divergencies between the two solutions arise when there are differences in *intensity* of preferences.

The median solution is completely independent of how strong the preferences are of those citizens below and above the median.

Practically speaking, these problems arise most clearly when society is dealing with issues in which those with the minority view feel most strongly about the issue.

Good examples might be gay marriage, affirmative action, and immigration reform.

It may be that interest group activity may partially resolve these issues along the lines suggested by the common agency model.

Nonetheless, as we pointed out, there is no particularly good reason to suppose that interest groups will form when there is demand for them.

When it comes to distributive policy issues, there is even less reason to believe that outcomes will be equitable.

As we saw in the legislative bargaining model, proposers get extra benefits and benefits are shared by minimum winning coalitions.

Socially optimal policies, by contrast, will typically involve equal division.

III.2.i Politically determined income redistribution

Just pointing out that politically determined policies may not be socially optimal only takes us so far.

A deeper question is what determines how income gets redistributed in political systems.

There is a large literature on this which is surveyed in the chapter by Londregan in W & W.

We will discuss a couple of the classic papers in this literature: Meltzer and Richards (1981) and Lindbeck and Weibull (1987).

Meltzer and Richard's Model

There are a continuum of individuals and two goods - consumption and leisure.

Individuals get utility from consumption x and work l according to the utility function

$$x - \frac{l^{(1+\frac{1}{\varepsilon})}}{1+\varepsilon},$$

where $\varepsilon > 0$ is the elasticity of labor supply.

Individuals differ in their income generating ability.

An individual with ability a earns income $y = al$ if he works an amount l.

The range of ability levels is $[a_{\text{min}}, a_{\text{max}}]$.

Let $F(a)$ denote the fraction of individuals with ability less than or equal to a.

Let a_{med} denote the median ability and let a_{mean} denote the mean ability.

The government taxes income at rate t and redistributes the revenue via a uniform transfer T.

In reality, this uniform grant is supposed to correspond to public programs with universal benefits like the social security system.

Given the tax rate t, an individual with ability a works an amount

$$
\begin{aligned}
l^*(t;a) &= \text{arg max}(1-t)al - \frac{l^{(1+\frac{1}{\varepsilon})}}{1+\varepsilon} \\
&= (\varepsilon(1-t)a)^{\varepsilon}.
\end{aligned}
$$

Note that work effort is decreasing in the tax rate.

Let $v(t, a)$ denote the associated level of indirect utility; i.e.,

$$\begin{aligned} v(t, a) &= (1 - t) a l^*(t; a) - \frac{l^*(t; a)^{(1+\frac{1}{\varepsilon})}}{1 + \varepsilon} \\ &= \frac{\varepsilon^{\varepsilon}((1 - t)a)^{\varepsilon+1}}{\varepsilon + 1} \end{aligned}$$

The uniform transfer when the tax rate is t is therefore

$$T(t) = \int_{a_{\mathsf{min}}}^{a_{\mathsf{max}}} t(1 - t)^{\varepsilon} \varepsilon^{\varepsilon} a^{\varepsilon+1} dF(a).$$

Meltzer and Richards assume that the equilibrium tax rate t is that preferred by the median voter.

They had in mind that the tax rate would be determined by two party Downsian competition.

The preferred tax rate of an individual of ability a is

$$t^*(a) = \arg\max_{t \geq 0} T(t) + v(t, a).$$

It can be shown that $t^*(a) = 0$ if $a \geq a_{mean}$ and that $t^*(a) > 0$ if $a < a_{mean}$.

Moreover, $t^*(a)$ is decreasing in a when $a < a_{mean}$.

The median voter will therefore be the voter of median ability and the equilibrium tax rate is $t^*(a_{med})$.

Assuming that the ability distribution is such that the median income is below the mean (which is true in reality), the result will be a positive level of redistribution (i.e., $t^*(a_{med}) > 0$).

However, the result will not be a full equalization of incomes, since $t^*(a_{med}) < 1$.

This reflects the fact that taxation reduces work effort and thus reduces the size of the pie.

The level of redistribution will be lower the higher the elasticity of labor supply and the greater the difference between the mean and median ability.

The model therefore provides a simple theory of the extent of redistribution in a democracy.

The socially optimal tax rate in this economy depends on the social welfare function.

With a Utilitarian social welfare function, the optimal tax rate is 0.

This is because citizen utility is linear in consumption and so there is no aggregate gain from transferring income from high to low earners.

Thus, there is no benefit from taxation but there is a cost since taxation is distortionary.

With a Rawlsian social welfare function, the optimal tax rate is $t^*(a_{\text{min}})$ which is higher than the equilibrium tax rate.

It follows that there exists a parameter α of the Isoelastic social welfare function at which the equilibrium and socially optimal tax rates coincide.

The main problem with the model is that it requires redistribution to be in the form of a uniform transfer.

In reality, governments can target transfers to different groups (seniors, farmers, auto workers, Floridians, etc)

The difficulty with extending the model to allow for richer tax/transfer systems is that it makes the policy space multidimensional.

With a multidimensional policy space, there does not exist a Condorcet winner.

Dealing with this technical problem requires a more sophisticated approach.

Lindbeck and Weibull Model

Lindbeck and Weibull were also interested in political income redistribution, but wanted a model which allowed for group-specific transfers.

There are n voters indexed by $i \in \{1, .., n\} = I$

Each voter i has exogenous income ω_i

The voters belong to different groups.

Formally, voters may be partitioned into m subsets $(I_1, ..., I_m)$ where $m \in \{2, .., n\}$

The subset of voters in I_k is referred to as *group k.*

Let $n_k = \# I_k$ and let $k(i)$ be the group to which voter i belongs.

Groups can be ethnicities, ages, professions, or locations.

A *policy* is a vector $p = (p_1, ..., p_m)$ where p_k is a transfer given to each member of group k.

To be *feasible*, the policy must be such that $\sum_k n_k p_k = 0$ and for all groups k, $\omega_i + p_k > 0$ for all voters i in group k.

Let P denote the set of feasible policies.

Lindbeck and Weibull assumed that policies were determined by Downsian candidate competition.

There are two political parties, indexed by $J \in \{A, B\}$

Each party J must choose a policy $p^J = (p_1^J, ..., p_m^J) \in P$

Voter i obtains consumption $\omega_i + p_{k(i)}^J$ if party J wins

Each voter has a utility function defined over consumption levels $\nu : \Re_+ \to \Re$.

It is assumed that ν is increasing, strictly concave and satisfies the boundary conditions $\lim_{x\to 0} \nu'(x) = \infty$ and $\lim_{x\to\infty} \nu'(x) = 0$.

Thus, we do have diminishing marginal utility of consumption in this model.

To deal with the problem of non-existence of Condorcet winners, Lindbeck and Weibull assume that the two parties have some exogenous differences in policy preferences over "non-pliable" issues (such as abortion, gay marriage, family values, guns, etc).

Moreover, parties are uncertain about voter preferences over these exogenous differences.

This leads to a model in which voting is probabilistic.

Formally, voter i's payoff if party J wins is $\nu(\omega_i + p^J_{k(i)}) + i_J$ where i_A and i_B represent the payoffs that i derives from the exogenous aspects of A and B

Voter i votes for party A if $\nu(\omega_i + p^A_{k(i)}) - \nu(\omega_i + p^B_{k(i)}) > i_B - i_A$

The parties are assumed not to observe $i_B - i_A$ - they only know that $i_B - i_A$ is the realization of a random variable with smooth CDF F_i with density f_i.

Thus, from the parties' perspective, the probability that voter i votes for party A is

$$\pi_i(p^A, p^B) = F_i(\nu(\omega_i + p^A_{k(i)}) - \nu(\omega_i + p^B_{k(i)}))$$

The expected number of votes for party A is $\sum_i \pi_i(p^A, p^B)$

Each party seeks to maximize his expected votes.

This is not quite the same thing as maximizing the probability of winning, but the two objectives converge as $n \to \infty$.

Equilibrium and results

Each party J simultaneously chooses its policy platform $(p_1^J, ..., p_m^J) \in P$

Each voter votes probabilistically as described above

A pair of policy platforms (p^A, p^B) is an *equilibrium* if (i) $p^A \in \arg\max_{p \in P} \sum_i \pi_i(p, p^B)$ and (ii) $p^B \in \arg\min_{p \in P} \sum_i \pi_i(p^A, p)$.

Proposition *If* (p^A, p^B) *is an equilibrium, then* $p^A = p^B = p^*$ *and there exists* $\lambda > 0$ *such that for all* $k = 1, ..., m$

$$\sum_{i \in I_k} \nu'(\omega_i + p_k^*) f_i(0) = \lambda n_k$$

Proof The Lagrangian for party A's problem is

$$\mathcal{L} = \sum_{k=1}^{m} \sum_{i \in I_k} F_i(\nu(\omega_i + p_k) - \nu(\omega_i + p_k^B)) - \lambda_A \sum_{k=1}^{m} n_k p_k$$

where λ_A is the Lagrange multiplier.

The first order conditions are for all groups k

$$\sum_{i \in I_k} \nu'(\omega_i + p_k^A) f_i(\nu(\omega_i + p_k^A) - \nu(\omega_i + p_k^B)) = \lambda_A n_k$$

Similarly, the first order conditions for party B's problem imply

$$\sum_{i \in I_k} \nu'(\omega_i + p_k^B) f_i(\nu(\omega_i + p_k^A) - \nu(\omega_i + p_k^B)) = \lambda_B n_k$$

It follows that for each group k

$$\begin{aligned} \rho_k &\equiv \frac{\sum_{i \in I_k} \nu'(\omega_i + p_k^A) f_i(\nu(\omega_i + p_k^A) - \nu(\omega_i + p_k^B))}{\sum_{i \in I_k} \nu'(\omega_i + p_k^B) f_i(\nu(\omega_i + p_k^A) - \nu(\omega_i + p_k^B))} \\ &= \frac{\lambda_A}{\lambda_B} \end{aligned}$$

This implies that $p_k^A = p_k^B$ for all k.

For otherwise, there would exist groups k and h such that $p_k^A - p_k^B > 0 > p_h^A - p_h^B$.

But this would mean that $\rho_k < \rho_h$

Thus, $p^A = p^B = p^*$ where

$$\sum_{i \in I_k} f_i(0)\nu'(\omega_i + p_k^*) = \lambda_A n_k.$$

■

To see the implications of this Proposition suppose first that $f_1(0) = ... = f_n(0)$.

Intuitively, this corresponds to the idea that all voters are expected to have the same preferences over the parties' exogenous differences.

Then, $p^A = p^B = p^*$ and there exists $\lambda > 0$ such that for all $k = 1, ..., m$

$$\sum_{i \in I_k} \nu'(\omega_i + p_k^*) = \lambda n_k$$

Thus, the politically determined distribution of income maximizes the Utilitarian social welfare function $\sum_{k=1}^{m} \sum_{i \in I_k} \nu(\omega_i + p_k)$!

This suggests that politically income redistribution will coincide with socially optimal redistribution.

Given the diminishing marginal utility of consumption assumption, political income redistribution will involve transferring income from richer to poorer groups.

The intuition is that poor people value consumption more than rich and are thus more likely to be convinced to vote for a party by a promise of more redistribution.

However, this conclusion rests heavily on the assumption that all voters have the same preferences over the parties' exogenous differences.

Suppose instead that (i) $f_i(0) = f^k(0)$ for all $i \in I_k$ and (ii) for all k there exists α_k such that $f^k(x) = f(x + \alpha_k)$ where f is symmetric around 0 and unimodal

The parameter α_k can be interpreted as the expected bias of group k in favor of party A.

If $\alpha_k = 0$ group k is unbiased; if $\alpha_k > 0$ group k is biased in favor of party A and if $\alpha_k < 0$ group k is biased in favor of party B

The proposition implies that $p^A = p^B = p^*$ and there exists $\lambda > 0$ such that for all $k = 1, ..., m$

$$\frac{\sum_{i \in I_k} \nu'(\omega_i + p_k^*)}{n_k} = \lambda / f(\alpha_k)$$

The implication of this is that groups who are unbiased get greater transfers or, in other words, political income redistribution is targeted to groups who are swing voters.

Obviously, this will not be socially optimal.

Discussion

The Lindbeck-Weibull model has been quite influential.

Dixit and Londregan (1998) extend the model to allow the two parties to have different preferences over the distribution of income.

Swing voters still have an advantage, but now there is ideologically motivated redistribution as well as tactical redistribution.

There have been a number of attempts to test the predictions of the model empirically using location as the determinant of group identity.

The question is whether states with more swing voters get more federal money.

There is actually not much evidence for this.

Work on understanding political income redistribution continues.

III.3 Politics and Efficiency

Will policies emerging from the political process be Pareto efficient?

This is an interesting and controversial question (see Wittman (1989) for general discussion).

All the models we have looked at in Part I of the class suggest that the answer is yes.

In all these models, the implemented policies are Pareto efficient.

For example, if the implemented policy maximizes the utility of the median voter then it must be Pareto efficient because any other policy would make the median voter worse off.

Even if extremist candidates are elected as in the citizen-candidate model with strategic voting, policies are still (ex post) Pareto efficient because they maximize the utility of those citizens with the extremist preferences.

Nonetheless, there are a number of types of inefficiency that have been identified.

We focus primarily on inefficiencies stemming from reputational concerns and lack of commitment.

III.3.i Reputational inefficiencies

The basic idea is that incumbent politicians want to preserve good reputations with voters and this sometimes leads them to do inefficient things.

The underlying framework for thinking through these inefficiencies is the political agency model.

The simplest inefficiency of this type is pandering.

This occurs when politicians choose policies they know are inefficient just because voters think they are good policies.

Pandering can arise when politicians have more information about what the best policies are than do voters.

I will go through a simple political agency model of pandering which is simplified version of that used by Maskin and Tirole in a (2004) paper in *American Economic Review*.

For a related model see Canes-Wrone, Shotts, and Herron (2001).

A Model of Pandering

There are two time periods, period 1 and period 2.

There are two politicians, an incumbent and a challenger, and a representative voter.

The incumbent holds office in period 1, but at the beginning of period 2 faces an election against the challenger to determine who holds office in period 2.

In each period, the politician who holds office has to choose policy a or policy b.

The voter does not know which option is best for him.

In each period, the voter believes that the optimal policy for him is a with probability $p > 1/2$.

Thus, the voter believes that the optimal policy for him is more likely to be a.

The voter's payoff from the optimal policy is 1 and from the wrong policy is 0.

Importantly, we assume that the voter does not experience his payoff from the period 1 policy choice until period 2.

Each politician can be *congruent* or *dissonant*.

A congruent politician shares the voter's policy preferences.

A dissonant politician has the exact opposite preferences.

When in office, each type of politician (i.e., congruent or dissonant) gets a policy-related payoff of 1 from choosing his preferred policy and 0 from the other policy.

The politician in office is assumed to have more information than the voter.

To capture this as simply as possible we assume that in each period the politician in office is assumed to know perfectly which option is best for the voter (and himself).

Each type of politican also gets a non-policy related payoff r when in office.

When not in office, politicians get a payoff of 0.

Politicians discount period 2 payoffs at rate $\delta < 1$.

The voter cannot directly observe politicians' types.

At the beginning of period 1, he believes the incumbent is congruent with probability π_I.

The voter is assumed to observe the incumbent's period 1 policy choice and to update his beliefs about the incumbent's type rationally.

Recall that the voter does not observe his payoff from the period 1 policy choice until period 2.

Game

The game played between the politicians and the voter is as follows.

In period 1, the incumbent observes what policy is best for the voter and chooses the policy a or b.

At the beginning of period 2, the election is held and the voter chooses whether to re-elect the incumbent or elect the challenger.

In period 2, the winning politician observes what policy is best for the voter and chooses the policy a or b.

Solving the game

In period 2, whichever politician is holding office will simply choose his preferred policy - there is no gain to doing anything else.

Accordingly, at election time, the voter wants to elect the candidate most likely to be congruent.

Suppose that at the time of the election he believes the challenger to be congruent with probability π_C and the incumbent to be congruent with probability π_I'.

Then the voter will re-elect the incumbent if and only if $\pi_I' \geq \pi_C$.

The key issue is how the voter forms the belief π_I'.

Intuitively, it should depend on the voter's initial belief π_I and also on his period 1 policy choice.

While there are a number of possibilities, we will focus on the pandering equilibrium.

At the outset, we will make the assumption that $\pi_I = \pi_C$, so that the voter has the same initial beliefs about the challenger as he does about the incumbent.

In the *pandering equilibrium* both types of incumbent choose policy a irrespective of whether it is the best option for them or the voter.

Then the voter learns nothing about the incumbent's type from the period 1 policy choice if he chooses a and thus has no reason to change his prior belief.

What would the voter think if the incumbent were to choose b?

This is not so clear, because even a congruent politician could prefer b if this were the optimal policy for the voter.

Nonetheless, since the voter knows that policy a is more likely to be the optimal policy for him, in expectation he knows that dissonant politicians are more likely to prefer b.

Thus, it is natural to assume that the voter would revise his belief that the incumbent was congruent downwards if he observed him choosing b.

Under this assumption

$$\pi_I' = \begin{cases} \pi_I \text{ if } a \\ \pi \text{ if } b \end{cases}$$

where $\pi < \pi_I$

Given the assumption that $\pi_I = \pi_C$, it follows that the voter will re-elect the incumbent if he observes policy a chosen in period 1.

If he observes policy b chosen in period 1, the voter will elect the challenger.

In order for this to be an equilibrium, we need to check that both types of incumbent will play according to the proposed equilibrium.

We begin with the Congruent Incumbent.

If the optimal policy is a there is nothing to check, since the congruent incumbent not only prefers a but it also gets him re-elected.

Suppose the optimal policy is b.

In this case, if the congruent incumbent chooses a his payoff in period 1 is just r, but he gets re-elected and obtains a payoff in period 2 of $1 + r$.

His discounted expected payoff from choosing a is therefore

$$r + \delta[1 + r]$$

If the congruent incumbent chooses b his payoff in period 1 is $1 + r$, but he does not get re-elected and obtains a payoff in period 2 of 0.

His discounted expected payoff from choosing b is therefore

$$1 + r$$

He will choose a if

$$\delta \geq \frac{1}{1 + r}.$$

We now turn to the Dissonant Incumbent.

If the optimal policy (for the voter) is b there is nothing to check, since the dissonant incumbent not only prefers a but it also gets him re-elected.

Suppose the optimal policy (for the voter) is a.

In this case, if the dissonant incumbent chooses a his payoff in period 1 is just r, but he gets re-elected and obtains a payoff in period 2 of $1 + r$.

His discounted expected payoff from choosing a is therefore

$$r + \delta[1 + r]$$

If the congruent incumbent chooses b his payoff in period 1 is $1 + r$, but he does not get re-elected and obtains a payoff in period 2 of 0.

His discounted expected payoff from choosing b is therefore

$$1 + r$$

He will choose a if

$$\delta \geq \frac{1}{1 + r}.$$

We have therefore proved the following proposition.

Proposition *There exists a pandering equilibrium if* $\delta \geq 1/(1+r)$.

Discussion

To understand the nature of inefficiency here, consider the case in which the incumbent is congruent and knows that the optimal policy for the voter is b.

Even though both he and the voter prefer policy b, he chooses a.

He therefore makes a Pareto inefficient policy choice.

This is a political failure.

The logic of the political failure is that the incumbent chooses a to preserve his reputation with the voter.

Choosing b would damage his re-election prospects.

Disguised transfers

Another important example of an inefficiency arising from reputational concerns is the use of disguised transfers.

These arise when politicians try to hide the fact that they are redistributing to a particular group by using inefficient but sneaky forms of redistribution.

Examples would be a mayor building a convention center because he wanted to help out his buddies in the construction industry or legislators ordering a new fighter plane because they wanted to help out campaign contributors in the defense industry.

Such decisions can be Pareto inefficient if the projects are so unnecessary that citizens would be better off if the politicians canceled the project and just gave their contributors a transfer equal to the profits.

Coate and Morris Model

Coate and Morris (1995) illustrate the use of such inefficient disguised transfer mechanisms in a two period political agency model.

There is a single representative citizen, a special interest (e.g., a construction company), an incumbent politician, and a challenger.

The incumbent politician is in office at the beginning of period 1 and faces an election against the challenger at the beginning of period 2.

In each period, the citizen's utility is

$$u_C = y_C - t + B$$

where y_C is exogenous income; t are taxes; and B are benefits from public projects.

In each period, the special interest's utility is

$$u_s = R + T$$

where R is income derived from public projects and T is a cash transfer.

In period 1, the incumbent chooses a policy (T, D).

$T \in \Re_+$ is a cash transfer to the special interest and $D \in \{P, N\}$ is a discrete public project decision (think of this as a construction project).

$D = P$ means the project is undertaken, $D = N$ means it is not.

In period 2, the winner of the election just chooses a policy T.

The public project provides income R_s to the special interest and has a tax cost to the citizen C.

The benefits of the project to the citizen are $B \in \{B_L, B_H\}$ where $0 < B_L < B_H$.

Let $\theta = \Pr\{B = B_H\}$ and assume that $\theta \in \{\theta_0, \theta_1\}$ where $0 < \theta_0 < \theta_1 < 1$.

The expected net benefits of the project to the citizen given θ are

$$\Delta(\theta) = \theta B_H + (1 - \theta) B_L - C$$

Assumption 1 *(i)* $\Delta(\theta_1) > 0$ *and (ii)* $\Delta(\theta_0) < -R_s$.

Thus, when $\theta = \theta_1$, the project is an expected Pareto improvement.

When $\theta = \theta_0$, the project is an inefficient way of redistributing to the special interest.

Assume that, as far as the citizen knows, $\Pr\{\theta = \theta_1\} = \pi$.

Each politican can be either *good* or *bad*.

Each type receives a per-period payoff 0 when not in office.

When in office, a good politician receives a payoff $v_g(u_c - y_c)$ where $v_g'(\cdot) > 0$ and $v_g(0) > 0$

A bad politician receives $v_b(u_c - y_c, u_s)$ where $\partial v_b / \partial u_c > 0$ and $\partial v_b / \partial u_s > 0$ and $v_b(0, 0) > 0$.

Thus bad politicians wish to help the special interest.

Assumption 2 $v_b(\Delta(\theta_0), R_s) > v_b(0, 0)$

This says that bad politicians prefer introducing the inefficient project than doing nothing.

The citizen faces two types of uncertainty:

(i) *Policy uncertainty* - he does not observe the realization of θ whereas the incumbent politician does.

Note that he cannot infer θ ex post since even a good project may fail.

(ii) *Politician uncertainty* - he does not observe the type of the incumbent or the challenger.

Let λ_I be his prior that the incumbent is good and let λ_C be his prior that the challenger is good.

λ_I is a given, while λ_C is realized at the time of the election - ex ante, it is draw from the CDF $G(\lambda)$

Game and equilibrium

The game played between the incumbent, challenger and voter has the following timing.

1. The quality of project $\theta \in \{\theta_0, \theta_1\}$ is chosen - this is observed only by the incumbent.

2. The incumbent chooses $(T, D) \in \Re_+ \times \{P, N\}$ - this is observed by the citizen.

3. If $D = P$, the benefits from the project $B \in \{B_L, B_H\}$ are realized and this is observed by the citizen.

4. The reputation of the challenger λ_C is determined.

5. Knowing λ_C and incumbent's first period record (T, D, B) the citizen decides whether to re-elect the incumbent.

6. The winner of the election chooses second period T.

A strategy for the incumbent consists of a first period policy choice conditional on his type and θ and a second period transfer conditional on his type.

A strategy for the challenger consists of a second period transfer conditional on his type.

A strategy for the citizen is a rule specify the probability that he will re-elect the incumbent conditional on (T, D, B) and λ_C.

We also need to specify the citizen's beliefs concerning the probability that the incumbent is good which we denote $\mu(T, D, B)$.

An equilibrium consists of strategies for all the players that are optimal and beliefs for the citizen which are consistent with the strategy of the incumbent of this game.

Main result

Proposition *Under appropriate assumptions, there exists* $\widehat{\lambda}$ *such that if* $\lambda_I > \widehat{\lambda}$*, in any equilibrium a bad incumbent always chooses* $(0, P)$ *and a good incumbent chooses* $(0, N)$ *when* $\theta = \theta_0$ *and* $(0, P)$ *when* $\theta = \theta_1$.

Thus, when $\theta = \theta_0$ the bad incumbent redistributes inefficiently.

The logic is that the reputational penalty for choosing the project is less severe than that for choosing cash transfers.

This is because the good incumbent chooses the project when it is a good idea and the voter cannot observe whether it is a good idea.

The project is therefore a disguised transfer.

Note that for a public project to be a disguised transfer it must have four key features.

(i) It benefits a special interest.

(ii) It may or may not benefit the rest of society.

(iii) Citizens have less information on the project's expected benefits than do politicians.

(iv) Citizens cannot infer the project quality ex post.

Many public projects have these features.

Coate and Morris also show that for the argument to make sense, voters must be uncertain of the incumbent's desire to redistribute to the special interest.

If all incumbents are bad, then there is nothing to hide and voters penalize inefficient redistribution.

In essence, the use of disguised transfers arises from the desire of incumbents to avoid losing their reputations as being on the side of taxpayers (as opposed to special interests).

Discussion

The above examples illustrate how reputational concerns can lead to inefficiencies.

There are by now many papers with a similar theme.

Two other nice examples are Groseclose and McCarty (2001) and Majumdar and Mukand (2004).

III.3.ii Commitment problems

The literature has identified a number of reasons why public policy decisions with dynamic consequences (such as budget deficits or public investments) may be inefficient.

The fundamental factor underlying these inefficiencies is lack of commitment - there is no mechanism by which future leaders can commit to current policy-makers what their policy choices will be.

We will illustrate these problems with two examples.

The first example illustrates why a reform with short run costs but long run benefits is not undertaken.

The second example shows why debt levels can be too high.

Reform example

This example is based on Besley and Coate (1998).

Consider a society of N citizens divided equally into two groups, A and B.

Suppose that there are two time periods, period 1 and period 2.

In period 1, the society can undertake a reform.

The reform imposes costs on members of group A in period 1, but benefits all citizens in period 2.

Good examples might be the removal of a trade barrier protecting group A or a subsidy benefiting group A.

The per-citizen cost to members of group A in period 1 is c and the per-citizen benefit to all citizens in period 2 is b.

Let δ denote citizens' discount rate, and assume that the discounted aggregate benefits of the reform exceed the aggregate costs.

Formally, the condition is

$$\delta N b > c\frac{N}{2}.$$

Also assume that that the reform does not directly benefit members of group A.

Formally, this condition is that

$$\delta b < c.$$

However, assume that the government can levy a tax on the period 2 benefits of the reform and compensate members of group A by redistributing the revenues to them.

If the tax is t, revenues are

$$R(t) = tNb - C(t)$$

where $C(t)$ represent collection costs.

Let

$$t^* = \arg\max b(1-t) + \frac{R(t)}{N/2}$$

denote the tax rate that maximizes the after tax income of group A members.

Assume that collection costs are such that $t^* \in (0, 1)$.

If members of group A can be compensated for undertaking the reform, it must be that

$$\delta \left[b(1-t^*) + \frac{R(t^*)}{N/2} \right] > c.$$

We assume this inequality holds in what follows.

Political decision-making

Consider a citizen-candidate model of political decision-making where a single citizen is elected to be policy-maker.

Suppose that in period 1, a member of group A has been elected and has to decide whether to undertake the reform.

At the beginning of period 2 there will be an election to pick the policy-maker for period 2.

The period 2 policy-maker has to choose the tax on the benefits of the reform if it is enacted.

The period 1 policy-maker has to figure out what will happen in period 2 in order to decide whether to undertake the reform.

Assuming that the cost of running is not too high, if the reform is enacted, there will be an equilibrium in which one member of each group enters the race.

All citizens will vote for the candidate from their group and the result will be a tie.

Each candidate will therefore win with probability $1/2$.

If the group A candidate is elected, he will choose a tax t^*.

If the group B candidate is elected, he will choose a tax 0.

This is because in period 2, group B members obtain a payoff $b(1-t)$ which is decreasing in t.

If the reform is not enacted, there is nothing for the period 2 policy-maker to do (with respect to the reform) so it does not matter who wins.

The group A policy-maker's expected discounted payoff if he undertakes the reform is

$$-c + \delta \left[\frac{1}{2} \left(b(1 - t^*) + \frac{R(t^*)}{N/2} \right) + \frac{1}{2} b \right]$$

If he does not undertake the reform, he just gets a payoff of zero.

He will not undertake the reform if

$$\delta \left[\frac{1}{2} \left(b(1 - t^*) + \frac{R(t^*)}{N/2} \right) + \frac{1}{2} b \right] < c$$

He will undertake the reform otherwise.

If the reform is undertaken, in period 2 the tax on benefits will be t^* with probability $1/2$ and 0 with probability $1/2$.

Political failure

If the reform is not undertaken, the politically determined policy choice will be Pareto inefficient.

The policy choice of no reform is *Pareto dominated* by the policy choice that involves the reform being undertaken in period 1 and, in period 2, the benefits being taxed at rate t^* and the revenues given to group A members.

Political decision-making in this example therefore gives rise to *political failure*.

The intuition underlying the political failure is as follows.

In order to make the reform Pareto improving, Group A members must be compensated for incurring the short run costs of the reform from the future benefits.

However, Group A members realize that when compensation can be paid they may no longer hold power.

They understand that if Group B members control policy they will not receive compensation.

This is because by then the reform has been implemented and Group B members have no incentive to compensate Group A.

Because of this concern, Group A members do not introduce the reform in the first place.

In order for the story to hold together, it must be that the group who currently holds power must anticipate losing that power in the future.

To resolve this inefficiency, Group B needs to commit to Group A that it will compensate them when the benefits of the reform are realized if it holds power.

Unfortunately, such future commitments are difficult in the political process.

This is because the future decision-makers are not even selected yet.

Debt example

This example is based on Alesina and Tabellini (1990).

Consider a legislature which consists of representatives from two Parties D and R.

There are two time periods, period 1 and period 2.

The legislature uses revenues to finance projects.

There are two types of projects: D projects and R projects.

Let P_{Dt} and P_{Rt} denote spending on D projects and R projects in period $t \in \{1, 2\}$.

D party representatives only care about spending on D projects and R party representatives only care about spending on R projects.

Specifically, in each period t, D representatives have payoffs $V_D(P_{Dt}, P_{Rt}) = P_{Dt}$ and R representatives have payoffs $V_R(P_{Dt}, P_{Rt}) = P_{Rt}$.

Representatives have discount rate δ.

In each period, the legislature receives a fixed amount of tax revenue T.

In period 1, the legislature can also obtain extra revenues by borrowing.

If it borrows an amount B in period 1 it must repay an amount $C(B)$ in period 2 where $C'(B) > 0$ and $C''(B) > 0$.

Assume that $1/\delta < C'(0) < 2/\delta$.

Political decision-making

Suppose that in period 1 the D party has a majority of seats in the legislature.

In period 2, there is a probability $1/2$ that the D party will have a majority of seats and a probability $1/2$ that the R party will have a majority of seats.

Obviously in period 1, the D party will devote all its revenues to spending on D projects.

The only question is how much it should borrow.

To understand this, we need to figure out what will happen in period 2.

If the D party borrows B in period 1, there will be revenues of $T - C(B)$ in period 2.

If the D party has a majority of seats, it will spend these revenues on D projects so that $(P_{D2}, P_{R2}) = (T - C(B), 0)$

If the R party has a majority of seats, it will spend these revenues on R projects so that $(P_{D2}, P_{R2}) = (0, T - C(B))$

If the representatives have discount rate δ, the optimal choice of borrowing for the D party solves the problem

$$\max T + B + \delta \left[\frac{1}{2}(T - C(B))\right]$$

The optimal debt level B^* satisfies the first order condition

$$1 - \frac{\delta}{2} C'(B) = 0$$

Given what we have assumed about $C'(0)$, we know that $B^* > 0$.

Thus, the D party will borrow B^* and spend revenues on D projects so that $(P_{D1}, P_{R1}) = (T + B^*, 0)$.

The period 2 spending levels will be $(P_{D2}, P_{R2}) = (T - C(B^*), 0)$ with probability 1/2 and $(P_{D2}, P_{R2}) = (0, T - C(B^*))$ with probability 1/2.

Political failure

The politically determined policy choices are Pareto inefficient.

Both groups would be better off under the following policies: in period 1, $(P_{D1}, P_{R1}) = (T, 0)$ and, in period 2, $(P_{D2}, P_{R2}) = (T, 0)$ with probability 1/2 and $(P_{D2}, P_{R2}) = (B^*/\delta, T - B^*/\delta)$ with probability 1/2.

To demonstrate that D party representatives are better off, we need to show that

$$T + \delta \left[\frac{1}{2} T + \frac{1}{2} B^* / \delta \right] > T + B^* + \delta \left[\frac{1}{2} (T - C(B^*)) \right]$$

This is equivalent to

$$\delta C(B^*) > B^*.$$

To demonstrate that R party representatives are better off, we need to show that

$$T - B^*/\delta > T - C(B^*)$$

This is equivalent to

$$\delta C(B^*) > B^*.$$

To see that the key inequality holds, note that

$$\begin{aligned} \delta C(B^*) &= \delta \int_0^{B^*} C'(B) dB \\ &> \delta C'(0) B^* \\ &> B^* \end{aligned}$$

where the first inequality follows from the assumption that $C''(B) > 0$ and the second inequality follows from the assumption that $C'(0) > 1/\delta$.

Political decision-making in this example again gives rise to *political failure*.

The logic of the political failure is as follows.

Party D borrows even though it is costly because it can control how the borrowed resources are utilitized in period 1.

It may not control the legislature in period 2 in which case Party R will control how resources are allocated.

Essentially, borrowing allows Party D to take resources away from Party R.

To resolve this inefficiency, Party R's period 2 legislators would have to promise to spend B^*/δ on Party D's projects if they were to hold power and if Party D had not borrowed.

Again, such future commitments are difficult in the political process.

This is because the future legislators are not even selected yet.

Discussion

There are many other papers exploring political failures arising from lack of commitment in a variety of dynamic contexts.

See Acemoglu (2003), Acemoglu and Robinson (2001), Azzimonti (2011), Battaglini and Coate (2008), and Hassler et al (2003).

III.3.iii Political risk

So far, our discussion of inefficiency has focused on ex post Pareto inefficiency; i.e., can the policy that is actually chosen be Pareto dominated?

The literature has also looked at ex ante Pareto inefficiency which takes into account the costs of any risk that arises in political outcomes.

For example, in the model of candidate competition with policy-motivated candidates or the citizen candidate model, risk arises because candidates choose different policies and it is uncertain which candidate will win.

If individuals are risk averse, this uncertainty will be costly.

On the other hand, if there is aggregate uncertainty in citizens' preferences, the fact that candidates have different positions provides an option value to society.

This raises the question of whether candidates will be optimally differentiated.

On this issue see Bernhardt, Duggan, and Squintani (2009).

III.4 Implications for the Role of Government

There is no good reason to expect policies emerging from the political process to produce policies that are socially optimal according to commonly used social welfare functions.

The issue of the Pareto efficiency of politically determined policies is a little more subtle, but we have identified solid reasons why even this will not be achieved.

In terms of what all this means, there are three points to note.

Point 1

The first point to note is that the prescriptions of normative theories of government can not simply be interpreted as positive predictions.

We need to study politics if we want to explain policy choices.

This is comforting, but not particularly surprising.

Point 2

The second point to note is that we need to rethink our standard justifications for government intervention.

To illustrate, consider the standard analysis of intervention to deal with an externality.

Figure 1 depicts the textbook analysis of a polluting industry.

The good is produced at constant cost c and the market equilibrium is Q_m units.

The textbook approach recommends intervention using a corrective tax z on the grounds that aggregate welfare would be higher if the tax were set at z^*, the aggregate utility maximizing level.

However, as we have argued, there is no reason to expect the political process to set the tax at the aggregate utility maximizing level.

The political economist would argue that intervention should be recommended only if aggregate surplus at the *politically determined level of the tax* is higher than at the market equilibrium.

There is no guarantee that this would be the case.

Suppose, for example, that a pro-environment lobby group would pressure policy makers to set a tax equal to $\widehat{z}$.

Then the intervention would actually reduce aggregate utility.

Acccordingly, the economic policy analyst should anticipate the politically determined level and take this into account.

This leads to more conservative policy advice because political determination of the new policy is a constraint on policy choices that makes intervention less attractive.

Point 3

Taking into account political distortions can also justify imposing fiscal restraints on government.

Fiscal restraints include limitations on the types and levels of taxes, balanced budget requirements, spending limits, etc.

The idea is that such restraints can be written into a society's constitution and are therefore hard for politicians to overcome.

In the debt example, for example, it would be good idea to require the government to run a balanced budget.

This would restore efficiency.

We will come back to look at fiscal restraints in more detail later in the class.

IV Contemporary Political Issues

We now apply some of the tools we have learned to think about some current political issues.

We begin with districting.

IV.1 Districting

Districting refers to the process by which political districts are redrawn to reflect population changes.

For example, suppose there are n seats in the legislature and that legislators are elected by n distinct political districts.

Each district is required to have approximately the same population.

Then if the population changes, the districts have to be withdrawn: districting refers to that process.

For U.S. house elections, districting is done at the state level.

The relative size of a state determines the size of its congressional delegation.

The state then divides up the population into political districts that will elect these representatives.

Districting must also be done for the election of state legislatures, county legislatures and city councils.

Indeed, districting is an essential part of any political system with geographic based (as opposed to at large) representation.

Districting in the U.S. is typically done on a ten year cycle following the dicennial census.

In the U.S., districting is often a highly controversial process.

There are two main issues: *partisan* districting and *racial* districting.

With partisan districting, districters design districts with an eye towards increasing the representation of legislators from a particular party or ideology.

This process is also known as *gerrymandering* after Governor Elbridge Gerry of MA who designed a salamander-shaped electoral district in the nineteenth century.

To illustrate, consider a legislature with 10 seats and two parties Democrats and Republicans.

Suppose that 50% of the population are Democrats and 50% are Republicans.

Then, while the aggregate vote share for the Democrats will be 50%, they can get anywhere between

1 and 9 seats depending on how the districts are created.

For example, they can get 9 seats by creating 9 districts that are 51% Democrat.

Relatedly, districting can also be designed to *protect incumbents*.

Thus, the idea is that districters conspire to design districts to maximize the chances that incumbents are reelected.

For example, by adding more Republicans to the districts of Republican incumbents.

This is called a *bipartisan gerrymander*.

With racial districting, redistricters design districts with an eye towards minimizing the political influence of a particular ethnic group - typically, African Americans.

The topic of districting has attracted a great deal of attention from political scientists and political economists.

It is a subject that is historically rich, the underlying theoretical issues are interesting, and the topic is practical and policy relevant.

We will focus on the literature on partisan districting.

We discuss three topics: seat-vote curves; optimal partisan gerrymandering; and socially optimal districting.

IV.1.i Seat-vote curves

The standard way of thinking about districting plans is in terms of the seat-vote curves they generate.

The seat-vote curve gives the relationship between one party's seats in the legislature and its aggregate vote.

Formally, it is represented by a function $S(V)$ where $V \in [0,1]$ is the aggregate fraction of votes received by (say) the Democrats and $S \in [0,1]$ is the fraction of seats in the legislature that they hold.

For example, in a *PR system*, $S(V) = V$ and in a *winner take all system*

$$S(V) = \begin{cases} 1 & \text{if } V > 1/2 \\ 0 & \text{if } V < 1/2 \end{cases}$$

It should be clear intuitively that different districtings will give rise to very different seat-vote curves.

In our earlier example with a 10 seat legislature, V was 0.5 while $S(0.5)$ could be any number in $\{0.1, ..., 0.9\}$.

Key properties of seat-vote curves are *partisan bias* and *responsiveness*.

A seat vote curve exhibits *partisan symmetry* if $S(V) = 1 - S(1 - V)$ for all V.

Intuitively, what this means is that the Democrat seat share when it has a particular fraction of the votes is the same as the Republican seat share when it has that same fraction of votes.

A seat-vote curve exhibits *partisan bias* if it deviates from partisan symmetry in a systematic way.

The simplest way of measuring this bias is just to look at the difference $S(1/2) - 1/2$.

If this difference is postive, the bias favors Democrats and if it is negative, the bias favors Republicans.

The *responsiveness* of a seat-vote curve is measured by the proportionate change in seat share following an increase in vote share; that is,

$$\frac{S(V + \Delta V) - S(V)}{\Delta V}.$$

If the seat-vote curve is differentiable, then its responsiveness at V is measured by the derivative $S'(V)$.

In the case of a *linear seat-vote curve*, partisan bias and responsiveness can be defined unambiguously.

The linear seat-vote curve can be written as

$$S(V) = \frac{1}{2} + b + r(V - \frac{1}{2}),$$

where the parameter b measures partisan bias and r measures responsiveness.

The political science literature has devised ever more sophisticated methods for estimating seat-vote curves.

Sample references are Tufte (1973), King (1989), and Gelman and King (1994).

A standard exercise is to estimate the seat-vote curves for a state legislature associated with particular districting plans (recall that districting is changed every 10 years).

The analyst then studies how partisan bias and responsiveness change as the districting plan changes.

It is also common to study how the changes in bias and responsiveness depend on the process by which districting is done at the state level.

Some states have partisan districting processes; others bipartisan districting processes; and still others non-partisan districting processes.

A standard finding is that partisan districting increases partisan bias and responsiveness; bipartisan districting reduces bias and responsiveness; non-partisan districting increases responsiveness and has no effect on bias.

IV.1.ii Optimal partisan gerrymandering

How should a party that controls the districting process choose a districting plan to maximize the probability of controlling the legislature?

This is a classic question in theoretical political science.

Lets work through a simple model to illustrate the main principles underlying an optimal partisan gerrymander.

The model is taken from Coate and Knight (2007).

A simple districting model

Consider a community in which there are three types of voters: Democrats, Independents, and Republicans.

Democrats and Republicans have ideologies 0 and 1, respectively.

Independents have ideologies that are uniformly distributed on the interval $[i_m - \tau, i_m + \tau]$ where $\tau > 0$ and i_m is the realization of a random variable uniformly distributed on $[1/2 - \varepsilon, 1/2 + \varepsilon]$, where $\varepsilon \in (0, \tau)$ and $\varepsilon + \tau < 1/2$.

This is similar to the set up we used in Section I.2.ii - i_m is the ideology of the median Independent voter.

The fractions of voters who are Democrats, Independents, and Republicans are π_D, π_I, and π_R.

Policy choices are made by a 3-seat legislature (we assume 3 just for tractability).

The policy outcomes chosen by the legislature depend upon the ideology of the median legislator.

Voters have quadratic preferences so that if the median legislator has ideology i', a voter with ideology i obtains a payoff given by $-(i - i')^2$.

There are two political parties: the Democrats and Republicans.

The Democrat Party consists of Democrats and the Republican Party consists of Republicans.

Legislators are all affiliated with one or the other party.

To select legislators, the state is divided into 3 equally sized districts indexed by $j \in \{1, 2, 3\}$

Each district then elects a legislator.

Candidates are put forward by the two political parties.

The Democrat Party puts up Democrats and the Republican Party puts up Republicans.

The citizen-candidate model applies so legislators just follow their true ideologies when elected - Democrats 0 and Republicans 1.

Elections are held simultaneously in each of the 3 districts and the candidate with the most votes wins.

In each district, voters vote sincerely for the representative whose ideology is closest to their own.

Districtings

A *districting* is a division of the population into 3 districts.

Formally, a districting is described by $(\pi_D(j), \pi_I(j), \pi_R(j))_{j=1}^3$ where $\pi_D(j)$ represents the fraction of Democrats in district j, $\pi_R(j)$ the fraction of Republicans, and $\pi_I(j)$ the fraction of Independents.

Assume the Party doing the districting knows the group membership of citizens and faces no geographic constraints in terms of how it can group citizens.

Thus, any districting $(\pi_D(j), \pi_I(j), \pi_R(j))_{j=1}^3$ such that the average fractions of voter types equal the actual is feasible.

For Democrats, this requires that

$$\frac{\pi_D(1) + \pi_D(2) + \pi_D(3)}{3} = \pi_D.$$

Similarly, for Independents and Republicans.

Seat-vote curves

Any particular districting implies a relationship between the community-wide fraction of Democrat votes and the number of seats Democrats obtain.

If the median Independent has ideology i_m, the fraction $V(j; i_m)$ of voters in district j voting for the Democrat is

$$V(j; i_m) = \pi_D(j) + \pi_I(j)[\frac{1/2 - (i_m - \tau)}{2\tau}].$$

Let $V(i_m)$ denote the community-wide fraction of voters voting Democrat; i.e.,

$$V(i_m) = \pi_D + \pi_I[\frac{1/2 - (i_m - \tau)}{2\tau}],$$

and let V_{min} and V_{max} denote, respectively, the statewide maximum and minimum Democrat vote shares.

We have that

$$V_{\text{min}} = \pi_D + \pi_I[\frac{\tau - \varepsilon}{2\tau}]$$

and

$$V_{\max} = \pi_D + \pi_I[\frac{\tau + \varepsilon}{2\tau}]$$

We can associate with any given community-wide Democratic vote share $V \in [V_{\min}, V_{\max}]$, the ideology of the median Independent $i_m(V)$ that would generate this vote share.

This is given by

$$i_m(V) = \frac{1}{2} + \tau[\frac{\pi_I + 2\pi_D - 2V}{\pi_I}].$$

It follows that if the Democratic vote share is V, district j elects a Democrat if

$$V(j; i_m(V)) \geq 1/2,$$

which turns out to be equivalent to

$$V \geq V^*(j) \equiv \pi_D + \pi_I[\frac{1/2 - \pi_D(j)}{\pi_I(j)}],$$

where $V^*(j)$ is the community-wide *vote threshold* above which district j elects a Democrat.

District j is a *safe Democrat* (*safe Republican*) seat if $V^*(j) \leq V_{\text{min}}$ ($V^*(j) \geq V_{\text{max}}$).

A seat which is not safe is *competitive*.

By relabelling as appropriate, order the districts so that $V^*(1) \leq V^*(2) \leq V^*(3)$.

The Democrats will get 0 seats if $V < V^*(1)$, 1 if $V \in (V^*(1), V^*(2))$, 2 if $V \in (V^*(2), V^*(3))$, and 3 if $V > V^*(3)$.

This relationship between votes and seats defines the seat-vote curve.

In terms of policy outcomes, the median legislator will be a Republican if $V < V^*(2)$ and a Democrat if $V > V^*(2)$.

Partisan gerrymandering

Suppose that the Democrats control the districting authority and want to come up with a districting plan that maximizes the probability that the Democrats have a majority of seats.

Clearly, they needs to choose a districting that maximizes the probability that $V > V^*(2)$.

This involves making $V^*(2)$ as small as possible.

To make things interesting assume that $\pi_R > 1/3$.

The Democrats do not care who wins district 3.

Thus, it is obvious that the solution will involve packing district 3 with Republicans; i.e., $(\pi_D(3), \pi_I(3), \pi_R(3)) = (0, 0, 1)$.

For districts 1 and 2, $(\pi_D(1), \pi_I(1))$ and $(\pi_D(2), \pi_I(2))$ solve the problem

$$\begin{gathered} \min V^*(2) \\ s.t.\ V^*(1) \leq V^*(2) \\ \frac{\pi_D(1)+\pi_D(2)}{3} = \pi_D \\ \frac{\pi_I(1)+\pi_I(2)}{3} = \pi_I \end{gathered}$$

This looks complicated but it is not really.

Since there is no gain to making District 1 any easier to win than District 2, it is clear that

$$V^*(1) = V^*(2).$$

This goal can be achieved by setting $(\pi_D(1), \pi_I(1))$ equal to $(\pi_D(2), \pi_I(2))$.

But then from the constraints we get that $\pi_D(1) = \pi_D(2) = \frac{3}{2}\pi_D$ and that $\pi_I(1) = \pi_I(2) = \frac{3}{2}\pi_I$.

It follows that

$$\begin{aligned}(\pi_D(1), \pi_I(1), \pi_R(1)) &= (\pi_D(2), \pi_I(2), \pi_R(2)) \\ &= (\frac{3}{2}\pi_D, \frac{3}{2}\pi_I, \frac{3}{2}\pi_R - \frac{1}{2})\end{aligned}$$

Notice that the Republicans who cannot be fit into District 3 are spread evenly over the other districts.

This strategy for dealing with the opposition voters is sometimes referred to as *cracking* and *packing*.

You pack a minority of districts with the opposition and then spread the remainder smoothly over the majority districts.

Notice that under the optimal districting plan

$$\begin{aligned}V^*(2) &= \pi_D + \pi_I[\frac{1/2 - \pi_D(2)}{\pi_I(2)}] \\ &= \frac{1}{3}.\end{aligned}$$

It follows that the Democrats will hold a majority of seats with probability 1 if $V_{\text{min}} > 1/3$

Using the expression for V_{min}, this condition amounts to

$$\pi_D + \pi_I[\frac{\tau - \varepsilon}{2\tau}] > 1/3$$

If this condition is not satisfied, the probability that the Democrats will hold a majority of seats is the probability that $V \geq 1/3$.

This probability is given by

$$\frac{V_{\text{max}} - 1/3}{V_{\text{max}} - V_{\text{min}}} = \frac{\pi_D + \pi_I[\frac{\tau+\varepsilon}{2\tau}] - 1/3}{\pi_I\left(\frac{\varepsilon}{\tau}\right)}.$$

Note that the seat-vote curve generated by this districting plan displays extreme partisan bias in favor of Democrats since they get more than $1/2$ the seats with only $1/3$ of the votes.

For more on optimal partisan gerrymandering see Friedman and Holden (2008).

See also Gul and Pesendorfer (2010) for an interesting analysis of what happens when the two parties control different states and choose their districting strategies in competitive fashion.

IV.1.iii Socially optimal districting

The previous exercise explores how a party should choose a districting plan to maximize political advantage.

A more public-spirited question is this: how should a districting authority that controls the districting process choose a districting plan?

The way this is usually approached is in terms of the shapes of the districts: for example, a good districting plan should have appropriately "compact" districts.

There are alternative ways of measuring the compactness of political districts - see, for example, Fryer and Holden (in press).

A more economic way to think about it is to ask what districting plan will maximize social welfare as measured by a social welfare function?

We can consider this problem using the same model we just used to study optimal partisan gerrymandering.

If the Democrats have a majority of seats in the legislature and the median Independent has ideology i_m, social welfare is

$$W_D(i_m) = -[\pi_R + \pi_I \int_{i_m-\tau}^{i_m+\tau} i^2 \frac{di}{2\tau}]$$

If the Republicans have a majority of seats in the legislature and the median Independent has ideology i_m, social welfare is

$$W_R(i_m) = -[\pi_D + \pi_I \int_{i_m-\tau}^{i_m+\tau} (1-i)^2 \frac{di}{2\tau}]$$

We know that if the Democratic vote share is V, the Republicans will have a majority of seats if $V < V^*(2)$ and Democrats will have a majority of seats if $V > V^*(2)$.

Thus, social welfare just depends on $V^*(2)$ which is the critical vote threshold at which District 2 flips from Republican to Democrat.

Recall that when the Democratic vote share is V, the median Independent has ideology $i_m(V)$.

Thus, when the critical vote threshold is $V^*(2)$, expected aggregate utility is given by

$$\int_{V_{\mathsf{min}}}^{V^*(2)} W_R(i_m(V)) \frac{dV}{V_{\mathsf{max}} - V_{\mathsf{min}}} + \int_{V^*(2)}^{V_{\mathsf{max}}} W_D(i_m(V)) \frac{dV}{V_{\mathsf{max}} - V_{\mathsf{min}}}.$$

Maximizing this expression with respect to $V^*(2)$, we obtain a first order condition

$$W_R(i_m(V^*(2))) - W_D(i_m(V^*(2))) = 0$$

Thus, the socially optimal districting must be such that

$$\begin{aligned} & \pi_R + \pi_I \int_{i_m(V^*(2))-\tau}^{i_m(V^*(2))+\tau} i^2 \frac{di}{2\tau} \\ = \;& \pi_D + \pi_I \int_{i_m(V^*(2))-\tau}^{i_m(V^*(2))+\tau} (1-i)^2 \frac{di}{2\tau} \end{aligned}$$

This can be simplified to

$$1 - 2i_m(V^*(2)) = \frac{\pi_R - \pi_D}{\pi_I}$$

Now using the expression for $i_m(V)$, we obtain

$$V^*(2) = \frac{1}{2} + (\frac{1-2\tau}{4\tau})(\pi_R - \pi_D)$$

We conclude that the districting authority should choose a districting plan such that

$$V^*(1) < V^*(2) < V^*(3)$$

and

$$V^*(2) = \frac{1}{2} + (\frac{1-2\tau}{4\tau})(\pi_R - \pi_D).$$

There are many ways this can be done, so that there are many different "optimal districting plans".

The important point to note is that $V^*(2) \neq \frac{1}{2}$.

The seat-vote curve associated with the optimal districting plans is biased towards the Party with the largest partisan base.

This reflects the fact that partisans (i.e., Republicans and Democrats) care more about who is in office than do Independents

Thus, the main lesson from this analysis is that partisan bias is not necessarily a bad thing!

Coate and Knight (2007)

Coate and Knight (2007) study socially optimal districting under the assumption that the policy outcomes chosen by the legislature depend upon the *average* ideology of the legislators rather than the *median* ideology.

They do this for two reasons:

i) to square with the political science literature which focuses on the global properties of the seat vote curve rather than its behavior in the neighborhood in which $S(V) = 1/2$.

ii) because there is empirical evidence from the U.S. states that policy outcomes are sensitive to the fraction of seats controlled by each party.

The districting problem is more involved under this assumption because social welfare depends on the critical vote shares of all the districts rather than just the median district.

They also compare empirically the actual seat-vote curves from the U.S. states with those that would be generated by socially optimal districtings.

This allows them to evaluate quantitatively the gains from socially optimal districting.

Besley and Preston (2007)

Our analysis of socially optimal districting also embodies a very spartan conception of electoral competition - parties just put up candidates representing their ideology and voters vote.

Parties and candidates therefore have no strategic choices to make.

Intuitively, however, one might imagine that districting would influence the type of candidates parties put up and/or the platforms that parties run on.

In this case, an optimal districting would have to take into account the implications for strategic choices.

Besley and Preston (2007) is a very interesting study using U.K. data which relates to this general issue.

They argue first theoretically that the degree of partisan bias in districting should impact the platforms that parties run on.

They then investigate this prediction empirically using data on U.K. local governments.

Valence and districting

Our analysis of socially optimal districting assumes that all districting determines is the ideology of the legislature.

One might also think of districting being important in determining the average valence characteristics (competence, honesty, etc) of legislators.

Imagine that candidates differ in both their ideology and quality as measured by $q \in [0, q_{\text{max}}]$.

Assume that if the median legislator has ideology i' and the average legislator quality is q, a citizen with ideology i experiences a payoff given by $q - \beta(i - i')^2$.

Assume as before that Democratic candidates have ideology 0 and Republican candidates have ideology 1.

Let q_j^K denote the quality of Party K's candidate ($K \in \{D, R\}$) in district j.

Assume that q_j^K is the realization of a random variable with support $[0, q_{\text{max}}]$ and CDF $F(q)$.

Further assume that $q_{\text{max}} < \beta$ implying that partisan voters vote on the basis of ideology.

The key point to note about this environment is that expected quality of the winning candidate is increasing in the competitiveness of the district.

This is because Independent voters vote more on the basis of quality and partisan voters vote on the basis of ideology.

If the partisan votes cancel each other the race is more likely to be determined by quality.

This gives a reason to make districts competitive.

It would be interesting (but hard) to solve for socially optimal districting plans in this environment.

Polarization and districting

It is common to argue that districting is responsible for the growing polarization that is observed in American federal politics.

The intuitive idea is that there are too many safe districts and that those districts elect more extreme candidates.

The safe districts are created by bipartisan gerrymandering which seeks to help out incumbents.

To capture this idea we would need a model where the ideology of the candidates that the two parties run depends on the characteristics of the district.

We would also need a model of legislative decision-making that captured the difficulties that polarization appears to cause.

This may be possible to do, but is much more complicated.

It should also be noted that the idea that districting causes polarization is not shared by most specialists on American politics.

The reason is because the Senate has also become more polarized over time and Senators represent states and not districts.

In any case, there is probably scope for work on this issue.

IV.2 Citizens' Initiatives

In more than half of U.S. States, the constitution allows citizens to place legislation directly on the ballot for voters to approve.

Such a piece of legislation is known as a *citizens' initiative*.

Examples of recent citizens' initatives include proposals to legalize marijuana (in Alaska); create voucher programs (in California); prohibit same-sex marriages (numerous states); eliminate affirmative action in public universities (California and Texas); require background checks at gun shows (Colorado); increase minimum wage (Florida).

To get an initiative placed on the ballot, it is necessary to present a petition signed by some fraction of the registered voters in the state in question.

The fraction varies across states, but ranges between 2-15% or registered voters.

Initiatives were first introduced in South Dakota in 1898 and were associated with the populist movement in the late nineteenth century.

Most states which have them introduced them in the early 1900s.

There were 360+ statewide initiatives over the 10 year period from 1995-2004 - over 30% in California, Washington and Oregon.

A large amount of money is spent trying to get initiatives passed or blocked.

In 2004, gambling interests spent $90 million in California on two ballot measures alone.

There is considerable debate about the role of initiatives - in particular, are they a good idea?

There is a large academic literature on the topic (see Matsusaka's papers on the reading list for discussion).

This literature seeks to understand what impact initiatives have on policy choices and whether they are a good idea.

It is popularly believed that permitting citizens' initiatives improves the "congruence" between citizens' preferences and policy outcomes across the spectrum of issues on which initiatives may be brought.

This premise is accepted by both advocates and opponents of initiatives, with the main debate focusing on the desirability of allowing voters' opinions more weight.

This premise receives backing from a variety of academic sources.

One nice example is Elizabeth Gerber's 1996 paper "Legislative Response to the Threat of Popular Initiatives".

She begins with a simple game theoretic model which illustrates why initiatives might move policy closer to the median voter's ideal and then tests the idea empirically.

We will begin our discussion of initiatives by going through Gerber's theory and evidence.

We will then discuss other evidence and relevant theoretical ideas.

Finally, we will come back to the question of whether initiatives are good or bad.

IV.2.i Gerber's paper

Gerber's model

There are three players: an incumbent politician p; an interest group i; and a representative voter v.

There is a one dimensional policy $z \in [0, 1]$

The preferences of the three players are

$$\begin{aligned} u_p(z) &= -|z - x_p| \\ u_i(z, m) &= \gamma m - |z - x_i| \\ u_v(z) &= -|z - x_v| \end{aligned}$$

where m is money and $\gamma > 0$.

Thus, players have *distance preferences*, with ideal points x_p, x_v, and x_i.

It is assumed that $x_p < x_v < x_i$, so that the politician prefers a lower level of the policy, than does the voter and the interest group prefers a higher level.

Game

The interaction between the players has three stages.

Stage 1. The politician selects a policy level z_p.

Stage 2. The interest group decides whether to propose an initiative - it costs the interest group C to propose.

This cost represents the cost of collecting the required signatures on the petition.

Let z_i denote the proposed policy level in the initiative.

Stage 3. If no initiative is proposed, the policy outcome is simply z_p.

If an initiative is proposed, the voter chooses whether to approve it.

If he does approve it, the policy outcome is z_i, and if he does not, the policy outcome is z_p.

Equilibrium

A *strategy* for the politician is simply a policy choice z_p.

A *strategy* for the interest group is a rule which tells us, as a function of the politician's choice z_p, whether or not it will propose an initiative, and, if so, the proposed policy level z_i.

A *strategy* for the voter is a rule which tells us whether or not he will approve the initiative as a function of the politician choice z_p and the interest group's proposal z_i.

An *equilibrium* consists of strategies for each of the three players that are optimal for each player given the other players' strategies.

Solving for the equilibrium

1. Voter strategy

Working backwards, first consider what the voter will do if there is an initiative.

The voter will approve the initiative if $-|z_p - x_v| \leq -|z_i - x_v|$

Given our assumptions about preferences, it is natural to assume that if there is an initiative, the proposed policy levels will be such that $z_p < x_v < z_i$.

Thus, we will assume this for now and then later we can verify that the equilibrium indeed satisfies this.

It follows that the voter will approve the intiative if $z_i - x_v \leq x_v - z_p$ or equivalently if $x_v \geq (z_i + z_p)/2$.

2. Interest group strategy

Now consider what the interest group will do.

The interest group will anticipate the voter's behavior and, if it does propose an initiative, it will propose the highest policy level that will be approved by the voter (assuming $z_i < x_i$).

This policy level is $z_i = 2x_v - z_p$.

It will propose an initiative if

$$-\gamma C - |(2x_v - z_p) - x_i| > - |z_p - x_i| \, .$$

This is equivalent to

$$(x_i - z_p) - (x_i - (2x_v - z_p)) > \gamma C$$

which is in turn equivalent to

$$(x_v - z_p) > \gamma C/2.$$

3. Politician strategy

Finally, consider what the politician will do.

When choosing z_p, the politician will anticipate the interest group's behavior.

In particular, it will make sure that $(x_v - z_p) \leq \gamma C/2$ so that no initiative is proposed.

It follows that

$$z_p = \begin{cases} x_p & \text{if } (x_v - x_p) \leq \gamma C/2 \\ x_v - \gamma C/2 & \text{if } (x_v - x_p) > \gamma C/2 \end{cases} \quad (1)$$

Equilibrium outcome

The outcome of the game is that the politician will choose z_p as described in (2) and no initiative will be proposed.

Assuming that $C < 2\left(x_v - x_p\right)/\gamma$, the *threat* of the initiative will bring the politician's choice closer to the voter's ideal point.

The initiative therefore has an effect on policy even though it is not actually used.

If we extend the model to incorporate different beliefs on the part of the politician and interest group on the location of the voter's ideal point (i.e., x_p), we can have initiatives actually being used.

Gerber's empirical work

Gerber tests the theoretical idea using data on parental consent notification laws for abortion.

As of 1990, 24 states required parental consent notification for women aged 18 or younger.

All of these laws were passed by state legislatures as opposed to initiatives.

These consent requirements are a good policy to look at because:

(i) states have primary jurisdiction

(ii) legislatures play a role

(iii) survey data is available state-by-state on voters' preferences on the topic.

Gerber runs regressions of the form:

$$\text{Pr}(policy_j = consent) = \frac{\exp(\beta_0 + \beta_1 V_j + \beta_2 V_j I_j + \beta_3 \cdot X_j)}{1 + \exp(\beta_0 + \beta_1 V_j + \beta_2 V_j I_j + \beta_3 \cdot X_j)}$$

where V_j is a measure of the median voter's preference in state j; X_j is a vector of state characteristics; and I_j is a state dummy variable for whether the state has the initiative.

She finds that β_2 is positive and significant.

This suggests that policy is closer to voters' preferences in states with the inititiative.

She does the same thing for the death penalty with similar results.

IV.2.ii Other evidence and related theory

Evidence

Evidence comparing fiscal policies in states with and without initiatives is also supportive of the general idea that the initiative improves the "congruence" between citizens' preferences and policy outcomes.

Matsusaka in a 2005 paper in the *Journal of Economic Perspectives* reviews evidence on the effect of the initiative on taxes and spending.

More than 10 studies have found that initiative states spent and taxed less than noninitiative states beginning around the mid 1970s, controlling for demographic and political factors.

The estimates imply that the initiative cut the combined spending of state and local governments by about 5% and cut state government spending by over 10%.

Initiative states also tend to decentralize spending from state to local governments and shift revenue out of broad-based taxes into user fees and charges for services.

Surveys of voter preferences suggest that these moves are in the direction in which the majority of voters prefer.

Theory

In light of all this evidence, a natural question to ask is why elections are not by themselves sufficient to bring policies into line with what the majority of voters want.

For example, Gerber's model assumes that the incumbent politician's preferences are different from the voter's.

One may well ask why the voters would have elected such an incumbent in the first place or why they cannot just toss the incumbent out of office if he chooses to ignore their preferences.

In a sense, we have already answered this with those models of elections reviewed in Section I.2 which do not predict median voter outcomes.

However, there are some additional points that can be made on this which are particularly relevant for thinking about the impact of initiatives.

1) Corruption

Elections may provide voters with very little control when incumbent politicians are receiving some financial reward in return for ignoring voters' preferences.

When initiatives were first introduced there was a great deal of concern about corruption at the state level.

Legislators were literally being bribed to do the bidding of corporations.

This problem could not be resolved via elections because any candidate who was elected would succumb to the same temptations.

To see this formally, imagine a citizen-candidate model of elections followed by a common agency model of influence with a single interest group.

If all potential candidates are willing to trade policy for money, there is no way to prevent the interest group from influencing policy via elections.

The only way round this problem is to take legislation out of the control of state legislatures.

2) The multi-dimensional nature of political competition

While most of our models of elections have ignored this point, it is important to note that when voters are voting on candidates, they are typically voting on a bundle of different policy positions.

Even in very competitive elections with no corruption, there is no reason to suppose that a candidate has to agree with the majority of voters on every issue to get elected.

Many issues are not salient in elections and candidates can hold non-majoritarian positions with no electoral penalty.

For example, trade policy: most voters tend to be very protectionist, but elected leaders in both parties tend to promote free trade.

The reason is that very few people are actually voting on this issue.

Other issues are only salient for those with the minority view.

For example, affirmative action and gun control.

In this case, the electoral incentive is for candidates not to champion the majority view.

Because the minority votes on the issue and the majority votes on other issues.

In these cases, if the voters have access to initiatives and the cost of placing an initiative is not too large, congruence on an issue can be restored.

The key feature of initiatives is that they allow voters to *unbundle* issues and force a direct vote on that issue alone.

Besley and Coate (2008) try to formalize this argument in a model in which two parties compete by selecting citizen-candidates and the winning candidate must choose two distinct policies.

They show that allowing citizens' initiatives on one of the issues improves congruence on this issue relative to candidate elections alone.

IV.2.iii Are initiatives good or bad?

In general it is difficult to say whether initiatives are good or bad since there is nothing necessarily good or bad from a social welfare perspectiive about having congruence on an issue-by-issue basis.

In particular, as we have argued Section III.2, it could be the case that on a specific issue the minority feels much more intensely about the issue (for example, gay marriage) than the majority and social welfare would be maximized by respecting the minority's preference.

An additional criticism - usually from elite groups - concerns the lack of voter information.

The reason for representative democracy is so that we can delegate decisions to a few professionals who have the time to find out the right thing to do.

Voters are rationally ignorant about issues.

Thus, majoritarian choices on complex initiatives may be misguided.

This is particularly so if monied interests can fool voters to voting against their interests through slick advertising campaigns, etc.

Advocates of initiatives argue that while indeed voters do not understand all the details of every initiative, they do not need to.

They can take cues from the positions of interest groups, newspaper endorsements, etc.

There is quite a bit of evidence in support of this idea as we discussed in Section I.1.

A final criticism is that initiatives take a huge amount of resources.

First, all the signatures have to be collected to get an issue on the ballot and second, in the campiagn stage, a huge amount of resources are spent on advertising.

These expenditures could be seen as wasteful - dissipating any benefit from aligning policy more closely with majoritarian preferences.

IV.3 Campaign Finance Policy

How should campaigns be financed?

Should private campaign contributions be regulated?

Should public money be used and, if so, how should it be allocated to candidates?

These are interesting questions on which there is little consensus.

Some commentators argue that all that is necessary is to have disclosure requirements which require candidates to publicly announce who is giving them money and in what amounts.

Others argue for full public financing whereby taxpayer dollars are allocated to qualifying candidates.

Still others argue for a voucher scheme whereby all citizens receive a small budget (say, \$100) that they can allocate to political candidates.

IV.3.i Background

Recall from Section I.5 that campaign contributions to candidates for federal office are made by individual citizens and Political Action Committees (PACs).

Interest groups, unions, and corporations contribute through PACs.

The amount that individuals and PACS can give to any one candidate is limited.

The individual limit per candidate is $2500 per election and the PAC limit per candidate is $5000 per election.

There are also limits on how much individuals and PACs can give to national parties and to other PACs.

There are no limits on the total amount of campaign spending that candidates can do - such limits have been ruled unconstitutional by the Supreme Court.

There are disclosure requirements for candidates and parties.

They must identify the PACs that provide them contributions, and the names, occupations, employers, and addresses of all individual contributors who give them more than $200.

They must also disclose exactly how they spend the money they get.

Since the Supreme Court's 2010 Citizens United vs the FEC decision, interest groups, unions, and corporations are free to spend whatever they want on political advertisements.

Thus, for example, they can run ads telling people their opinions about particular candidates.

This expenditure is treated differently from contributions to candidates and is protected as free speech.

At the federal level, public funding is limited to subsidies for presidential candidates.

These subsidies are available for both primaries and the general election.

However, they come with a catch: in return for receiving the taxpayer dollars, candidates must agree to limit their spending.

In the 2008 presidential election, most of the major candidates opted out of the public financing scheme for the primaries.

Moreover, Barack Obama opted out of the scheme for the general election as well.

The public financing scheme is funded by a $3 tax check off on individual tax returns.

In 2006 fewer than 8% of taxpayers checked this box (even though it means no increase in their tax bills), leaving the public financing fund a little short of funds.

Races for non-federal offices are governed by state and local law.

Over half the states allow some level of corporate and union contributions.

Some states have limits on contributions that are more stringent than national limits.

Six states (IL, MO, NM, OR, UT, VA) have no limits at all.

Public financing schemes are more common at the state and local level.

One method is the *Clean Money, Clean Elections* program.

To qualify for receiving money, candidates must first raise a target amount privately in small contributions.

Once they have done this, they receive an amount from the program.

In return, they are not allowed to accept further private contributions or to use their own money.

However, if they are running against a candidate who is not participating in the scheme and outspending them, they are eligible to receive additional funds from the program.

This program is currently used in Arizona and Maine.

It was proposed for California in a citizens' initiative, but the initiative failed.

In addition, the constitutionality of that part of the program that provides additional funds is problematic and in June 2011 the Supreme Court struck down this part of the program.

The ruling will likely be bad for the future of the program.

If additional funds are not granted, candidates may well opt out of the program to avoid being outspent.

This is what we have seen recently for the presidential public financing scheme.

IV.3.ii Economic analysis of campaign finance policy

From an economic perspective, a basic question is how well an unregulated system of campaign giving will perform from a social welfare perspective.

This depends critically on what we think the resources spent by candidates are used for.

As we discussed in Section I.5, there are two main ways to think about campaign spending.

First, spending buys the votes of “noise voters”.

Recall that noise voters are voters who are voting for random non-policy related reasons.

Campaign advertising is simply assumed to attract these voters.

Second, spending allows candidates to provide information about their positions and qualifications to uninformed voters.

In this view, voters are rational and update their beliefs based on the information they are provided by campaigns.

1. The noise-voter perspective

Under this perspective, campaign spending tends to either distort policy in an undesirable way or to just be wasteful rent-seeking with no policy consequences.

Limiting contributions therefore tends to be a good idea.

We can illustrate the implications of the perspective with Grossman and Helpman's (1996) analysis which embeds the common agency model within a theory of electoral competition.

The model assumes that campaign contributions "buy" the votes of noise voters and endogenizes the level of contributions via the common agency model.

It is a useful model to know about independent of its implications for campaign finance policy.

The voting population is divided into J groups indexed by $j \in \{1, ..., J\}$.

The fraction of the population in group j is N_j.

There are two candidates, indexed by $k \in \{A, B\}$.

Following Lindbeck-Weibull, candidates compete by staking out positions on "pliable" policies and have fixed positions with respect to "non-pliable" policies.

Let p^k denote candidate k's positions on the pliable polices and let P be the set of possible positions.

Voters are divided into two types - *rational* and *noise* - the fraction of rational voters is μ.

$I \leq J$ of the citizen groups are represented by interest groups - let $j(i)$ denote the group represented by interest group $i \in \{1, ..., I\}$

The interest groups give the candidates contributions that are used to buy the votes of the noise voters.

Behavior of rational voters

If candidate k is elected, rational voter i in group j obtains a payoff

$$u_j(p^k) + v_{ij}^k$$

where v_{ij}^k is an idiosyncratic shock reflecting i's assessment of k's positions on the non-pliable issues

Letting $v_{ij} = v_{ij}^B - v_{ij}^A$, the voter votes for candidate A if and only if

$$v_{ij} \leq u_j(p^A) - u_j(p^B)$$

It is assumed that v_{ij} is uniformly distributed on the interval

$$\left[\frac{b}{f} - \frac{1}{2f}, \frac{b}{f} + \frac{1}{2f}\right].$$

Note that the higher the value of b, the more voters lean toward candidate B with respect to his positions on the non-pliable issues.

It is assumed that the value of b is ex ante uncertain, being the realization of a random variable with CDF F_b.

For a given realization of b, the fraction of rational voters in group j voting for candidate A is

$$s_j^r = \frac{1}{2} - b + f[u_j(p^A) - u_j(p^B)]$$

The fraction of rational voters voting for candidate A is

$$s^r = \sum_j N_j s_j^r = \frac{1}{2} - b + f[u(p^A) - u(p^B)]$$

where $u(p) = \sum_j N_j u_j(p)$.

Behavior of noise voters

Noise voters are not impacted by candidates positions on the pliable policies but are swayed by campaign spending.

Letting C^k denote the spending by candidate k, the fraction of noise voters in group j voting for candidate A is

$$s_j^n = \frac{1}{2} - b + e[C^A - C^B]$$

where $e > 0$

The fraction of noise voters voting for candidate A is therefore

$$s^n = \sum_j N_j s_j^n = \frac{1}{2} - b + e[C^A - C^B]$$

Winning probabilities

The fraction of all voters voting for candidate A given a realization of b is

$$\begin{aligned} s &= \frac{1}{2} - b + \mu f[u(p^A) - u(p^B)] \\ &\quad +(1-\mu)e[C^A - C^B] \end{aligned}$$

The candidates choose their positions on the pliable issues p^A and p^B *before* the realization of b.

Thus, the probability that candidate A wins given p^A and p^B and C^A and C^B is

$$\Pr\{s \geq 1/2\} = F_b(G^A - G^B) \qquad (1)$$

where

$$G^k = \mu f u(p^k) + (1-\mu)eC^k.$$

Interest group behavior

Let C_i^k denote interest group i's contribution to candidate k.

Interest group i's objective function is

$$N_{j(i)}[F_b(\cdot)u_{j(i)}(p^A)+(1-F_b(\cdot))u_{j(i)}(p^B)]-C_i^A-C_i^B$$

Following the common agency model, each interest group i offers each candidate k a platform contingent contribution schedule $C_i^k(p^k)$.

Thus interest groups can make contributions to *both* candidates.

We are ruling out, by assumption, the possibility of schedules of the form $C_i^k(p^A, p^B)$

Candidate behavior

Candidate k chooses his positions on the pliable issues to maximize his probability of winning, taking into account the contribution schedules of the interest groups.

From (1), this is equivalent to choosing p^k to maximize G^k.

Let $(C_1^k(\cdot), ..., C_I^k(\cdot))$ denote the I interest groups' contribution schedules to candidate k.

Given $(C_1^k(\cdot), ..., C_I^k(\cdot))$, candidate will choose a policy from the set

$$\begin{aligned} & M^k((C_1^k(\cdot), ..., C_I^k(\cdot))) \\ = \; & \arg\max_{p^k \in P} [\mu f u(p^k) + (1-\mu) e \sum_i C_i^k(p^k)] \end{aligned}$$

Note here that candidate k's optimal choice depends only upon the contributions offered to him.

Equilibrium

An equilibrium consists of contribution schedules $(C_1^A(\cdot), ..., C_I^A(\cdot))$ and $(C_1^B(\cdot), ..., C_I^B(\cdot))$ and position choices p^A and p^B which satisfy the following conditions;

(i) each candidate is choosing positions optimally; i.e., $p^k \in M^k((C_1^k(\cdot), ..., C_I^k(\cdot)))$

(ii) no interest group, taking as given the contribution schedules of the other interest groups, can change its contribution schedule in such a way as to induce a preferred expected policy outcome.

Results

Suppose first that there are no interest groups ($I = 0$) or, equivalently, that contributions were banned: then,

$$p^A = p^B = p^* = \arg\max_{p \in P} u(p)$$

Thus, both candidates choose the policy positions that maximize aggregate utility $\sum_j N_j u_j(p)$

This result is similar to that obtained by Lindbeck and Weibull.

The result suggests that anything that contributions might do would be bad for aggregate policy utility.

Moreover, contributions also involve wasteful campaign advertising, which makes things even worse.

Next suppose that there is a single interest group ($I = 1$)

With no contributions, candidate k would simply choose the policy p^* and G^k would equal $\mu f u(p^*)$

Thus, if the interest group wishes to get candidate k to choose policy p^k, it must provide a contribution C_1^k such that

$$(1 - \mu) e C_1^k \geq \mu f[u(p^*) - u(p^k)] \qquad (2)$$

That is, the contribution must allow the purchase of a sufficient number of noise voters to offset the loss of rational voters.

If (2) holds exactly, then the interest group is only giving sufficient contributions to get candidate k to choose the policy p^k.

If it holds strictly, then the interest group is giving strictly more than it needs.

In the former case, the interest group has only an *influence motive* in giving to candidate k.

In the latter case, it also has an *electoral motive*; that is, it is giving additional money to help candidate k to win.

We can formally represent the interest group's problem as:

$$\begin{aligned}
& & \max N_{j(1)}[F_b(\Delta)u_{j(1)}(p^A) \\
& & +(1-F_b(\Delta))u_{j(1)}(p^B)] - C_1^A - C_1^B \\
s.t.\ (2) \text{ for } k & \in & \{A, B\} \\
& & and \\
\Delta & = & \mu f[u(p^A) - u(p^B)] \\
& & +(1-\mu)e[C_1^A - C_1^B]
\end{aligned}$$

If there is only an influence motive then we can solve out for C_1^A and C_1^B and conclude that

$$p^A = \arg\max\{F_b(0)N_{j(1)}u_{j(1)}(p) + \frac{\mu f}{(1-\mu)e}u(p)\} \quad (3)$$

and

$$p^B = \text{arg max}\{(1-F_b(0))N_{j(1)}u_{j(1)}(p)+\frac{\mu f}{(1-\mu)e}u(p)\} \quad (4)$$

In the symmetric case in which $F_b(0) = 1/2$, each candidate chooses the same platform and the equilibrium policy choice maximizes a weighted sum of aggregate and interest group utility, which is similar to the common agency model.

Thus, campaign contributions bias policy towards the interest group's preferences.

The weight on the interest group's utility is higher the larger the fraction of noise voters and the more responsive are their votes to campaign spending.

In the asymmetric case in which $F_b(0) > 1/2$, the interest group gives more to the more popular candidate (A) and the popular candidate's policies are more distorted towards the interest group.

If the difference in the two candidate's policies is sufficiently large, the interest group may wish to give additional money to help candidate A win.

A sufficient condition for this is that

$$(1-\mu)eF_b'(0)N_{j(1)}[u_{j(1)}(p^A) - u_{j(1)}(p^B)] > 1$$

where p^A and p^B satisfy (3) and (4).

However, the interest group will still give some contributions to candidate B.

Finally, suppose there are multiple interest groups.

While the model gets fairly intractable, it does yield three insights.

First, there is the possibility of multiple equilibria - some with candidate A getting the most contributions and some with candidate B.

The expectation that a candidate will attract most of the funds is self-fulfilling.

Second, it is unlikely that the electoral motive will be operative for more than one interest group.

This reflects a free-rider problem in contributing with an electoral motive.

Third, when (i) only the influence motive is operational for all interest groups; (ii) all the interest groups employ differentiable contribution schedules; and (iii) $u_j(p)$ is strictly concave for all j, then we have that

$$\begin{aligned} p^A &= \arg\max\{F_b(G^A - G^B)\sum_i N_{j(i)} u_{j(i)}(p) \\ &\quad + \frac{\mu f}{(1-\mu)e} u(p)\} \end{aligned}$$

and

$$\begin{aligned} p^B &= \arg\max\{(1 - F_b(G^A - G^B))\sum_i N_{j(i)} u_{j(i)}(p) \\ &\quad + \frac{\mu f}{(1-\mu)e} u(p)\} \end{aligned}$$

Thus, again, the platforms maximize a weighted sum of interest group and aggregate utility.

The indeterminacy in the model is with respect to $G^A - G^B$ which determines the advantage of each candidate.

Note that if interest groups are representative of the population $\sum_i N_{j(i)} u_{j(i)}(p)$ will be similar to $u(p)$, and campaign spending will not distort policy.

However, the campaign spending is itself completely wasteful, so it is still optimal to ban it.

2. The informative campaign spending perspective

The Grossman-Helpman model generates some useful insights.

However, the noise-voter perspective seems overly simplistic and cynical.

In reality, it does seem that campaign spending plays a role in allowing candidates to provide information to voters about their (and their opponents') positions on the issues; qualifications for office; and plans for action.

Such information could lead voters to elect better candidates which could lead to aggregate social benefits.

Unfortunately, developing the implications of this informative campaign spending view is more complicated, since we have to model how information is transmitted during the campaign.

I will describe one way of modelling this and discuss the results that the model yields.

An informative campaign spending model

We use the same basic set-up that we have used several times before.

There are three types of voters: Democrats, Independents, and Republicans.

Democrats and Republicans have ideologies 0 and 1, respectively.

Independents have ideologies that are uniformly distributed on the interval $[i_m - \tau, i_m + \tau]$ where $\tau > 0$ and i_m is the realization of a random variable uniformly distributed on $[1/2 - \varepsilon, 1/2 + \varepsilon]$, where $\varepsilon \in (0, \tau)$ and $\varepsilon + \tau < 1/2$.

The size of the population is n and there are n_I Independents.

The non-Independents are equally divided between Democrats and Republicans, so the model is symmetric.

The community is electing a representative.

Candidates are put forward by two political parties: Party D, comprised of Democrats and Party R, comprised of Republicans.

Each party selects a candidate of its own ideology but may select either a “qualified” ($q = 1$) or an “unqualified” ($q = 0$) candidate.

The intuitive idea is that finding qualified candidates is difficult, and a party may not be able to find one.

The probability that a party can find a qualified candidate is σ.

The payoff enjoyed by a citizen with ideology i from having a leader of ideology i' and qualifications q is given by $\delta q - \beta \left| i - i' \right|$.

$\delta < \beta$, so that Democrats and Republicans prefer a candidate of their own ideology, even when he is unqualified.

Independents do not know whether each party's candidate is qualified or unqualified.

Candidates can convey information concerning their candidates' qualifications via advertising.

To simplify, assume candidates can only advertise their own candidates' characteristics, ruling out negative advertising.

Thus only qualified candidates can benefit from advertising.

If a candidate spends an amount C, its message reaches a fraction $\lambda(C)$ of the population, where the function λ is such that $\lambda(0) = 0$, $\lambda'(C) > 0$, and $\lambda''(C) < 0$.

Candidates' advertising is financed by campaign contributions provided by two interest groups - a Democrat group that contributes to Party D's candidate and a Republican group that contributes to Party R's.

Each group constitutes a fraction γ of the population.

Interest groups' objectives are to maximize the expected payoff of their members.

Game

The timing of the model is as follows:

(i) Parties select candidates.

(ii) If their party's candidate is qualified, interest groups decide how much campaign contributions to give.

(iii) Candidates use their contributions to finance advertising.

(iii) Voters, having possibly observed one or both candidates' advertisements, update their beliefs about candidates' qualifications and vote.

Equilibrium

Equilibrium in this model will consist of (i) campaign contribution levels C_D and C_R that the two interest groups make if their party's candidate is qualified, and (ii) a description of how voters behave.

The campaign contributions must be optimal for the interest groups given how they expect voters behave and voters behavior must be appropriately rational.

Given the symmetry of the model, we will study the symmetric equilibrium in which $C_D = C_R = C$.

Democrat and Republican voters always vote for the candidate put forward by their party, so their behavior is straightforward.

Independents may vote for either party's candidate.

Independents behavior is straightforward when they have either seen both candidates' advertisements or seen nothing from either candidate: they just vote for the Democrat if their ideology is less than $1/2$ and the Republican otherwise.

The interesting case is when they have seen an advertisement from one candidate (say the Democrat) and not the other.

Then an Independent with ideology i will vote for the Democrat if

$$\delta - \beta i \geq \rho\delta - \beta(1 - i)$$

where ρ is the probability the voter assigns to the unadvertised Republican candidate being qualified.

Rearranging the above inequality, we see that an Independent with ideology i will vote for the Democrat if

$$i \leq \frac{\delta(1 - \rho)}{2\beta} + \frac{1}{2}.$$

If he had not seen the advertisement, he would vote for the Democrat if $i \leq 1/2$.

Thus, advertising leads the advertised candidate to get an extra slice of voters.

This is what gives the advertised candidate an advantage.

The belief ρ is determined as part of the equilibrium and will take into account how likely it is that a voter will miss seeing an advertisement if a candidate is qualified.

In equilibrium, Bayesian updating implies

$$\rho = \frac{\sigma[1 - \lambda(C)]}{\sigma[1 - \lambda(C)] + (1 - \sigma)}$$

Results

In equilibrium, under appropriate assumptions, interest groups will spend on behalf of their qualified candidates.

As a result of this spending, qualified candidates defeat unqualified opponents with a probability greater than $1/2$.

If there were no campaign spending, qualified candidates would have no advantage over unqualified opponents.

Since all citizens benefit from qualified candidates being more likely to hold office, campaign spending has a social benefit.

In general, the level of campaign spending on behalf of qualified candidates can be smaller or larger than the level that maximizes a utilitarian social welfare function.

There are two offsetting forces.

On the one hand, there remains the force arising in the Grossman-Helpman model - the two groups spending is purely wasteful when the two parties' candidates have the same qualifications.

This force leads campaign spending to be too high.

On the other hand, when choosing how much to spend on behalf of qualified candidates, interest groups only take into account the benefits to their members of their candidate winning.

This force leads campaign spending to be too low.

Formally, it can be shown that as σ gets close to 1 there is necessarily too much campaign spending.

Intuitively, if both parties' candidates are highly likely to be qualified, there is little chance that we will be in the situation in which campaign spending can be socially beneficial; i.e., one candidate qualified and the other unqualified.

Similarly, as δ gets close to zero, there is too much campaign spending.

This is because there is little gain from having a qualified candidate in office rather than an unqualified candidate.

Campaign spending is more likely to be underprovided, the larger is δ, the larger is n_I, and the smaller is γ.

All in all, the model provides a logical foundation for believing that unregulated campaign spending may be too low.

It therefore provides a possible justification for policies that subsidize campaign giving, like tax deductions.

Favors and public financing

The above model assumes that campaign contributions are provided by interest groups for purely ideological reasons.

Many of the arguments in favor of limits or public financing are worried about the quid-pro-quo of candidates providing policy favors for their contributors once they are elected.

Policy favors are things like special tax breaks or regulatory exemptions which benefit interest groups but are paid for by regular citizens.

The aggregate cost of these favors exceeds the benefits they provide, so that they are a net loss to society.

As we saw in Section I.5, the evidence of a quid-pro-quo is not particularly strong.

On the other hand, giving to federal candidates is already heavily regulated so who can say what would happen without such regulations.

It seems highly likely that quid-pro-quo would arise in an unregulated system.

Thus, it is interesting to think about what might happen in an unregulated system with policy favors.

Coate (2004) argues that in such a world campaign finance policy, in the form of contribution limits and public matching grants, can be Pareto improving.

To see the argument, we just need to modify the above model to allow policy favors.

Specifically, assume that after selection, each party's candidate (if qualified) requests a contribution from his interest group and that, to obtain a larger contribution, a candidate can offer to implement favors if elected.

When a candidate implements a level of favors f each interest group member enjoys a monetary benefit $b(f)$ at a cost f to every citizen, where $b(0) = 0$ and b is increasing and strictly concave.

In addition, $\gamma b(f) \leq f$ for all f (so favors create no aggregate benefits) and $b'(\delta) > 1$.

Each interest group agrees to a candidate's request only if it benefits it to do so.

Also so we can vary candidate incentives to offer favors, assume candidates obtain a non-policy related payoff $r > 0$ from holding office.

Higher r makes candidates more power hungry and hence more willing to grant favors.

The timing of the extended model is as follows:

(i) Parties select candidates.

(ii) If qualified, candidates request contributions from their interest groups and promise favors in exchange.

(iii) Candidates use their contributions to finance advertising.

(iv) Voters, having possibly observed one or both candidates' advertisements, update their beliefs about candidates' qualifications and vote.

Voters do not observe the amount of favors promised, but do have rational expectations.

(v) The winning candidate chooses policy and implements favors.

Equilibrium

A symmetric equilibrium in this model will consist of (i) a description of the contribution-favor package (C, f) qualified candidates offer to their interest groups, and (ii) a description of how voters behave.

The campaign contributions must be optimal for the candidates given how they expect voters behave and the constraint of interest group acceptance.

Voters behavior must be appropriately rational.

Again, Democrat and Republican voters always vote for the candidate put forward by their party, so their behavior is straightforward.

Independents behavior is straightforward when they have either seen both candidates' advertisements or seen nothing from either candidate: they just vote for the Democrat if their ideology is less than $1/2$ and the Republican otherwise.

The interesting case is again when they have seen an advertisement from one candidate (say the Democrat) and not the other.

Then an Independent with ideology i will vote for the Democrat if

$$(\delta - f) - \beta i \geq \rho\,(\delta - f) - \beta(1 - i)$$

where ρ is the probability the voter assigns to an unadvertised candidate being qualified.

Notice that this calculation recognizes that qualified candidates, if elected, will implement favors for their interest groups.

Rearranging the above inequality, we see that an Independent with ideology i will vote for the Democrat if

$$i \leq \frac{(\delta - f)(1 - \rho)}{2\beta} + \frac{1}{2}.$$

If he had not seen the advertisement, he would vote for the Democrat if $i \leq 1/2$

Thus, advertising leads the advertised candidate to get an extra slice of voters, but higher favors reduces the size of this slice.

Letting

$$\xi \equiv \frac{(\delta - f)(1-\rho)}{2\beta}$$

we can summarize voter behavior by ξ.

As before, in equilibrium, the belief ρ is given by

$$\rho = \frac{\sigma[1-\lambda(C)]}{\sigma[1-\lambda(C)] + (1-\sigma)}.$$

Results

What would happen if contributions were unrestricted?

Proposition 1 *Under appropriate assumptions, in any equilibrium with unrestricted contributions* (C, f, ξ), *qualified candidates offer to implement favors for their interest groups to extract larger contributions. The contributions they receive allow them to defeat unqualified opponents with a probability* $\pi = \frac{1}{2} + \frac{\xi}{2\varepsilon}\lambda(C)$. *The level of favors promised must be less than the gain from having a qualified candidate (i.e.,* $f < \delta$*).*

The problem with "favor-finance" is that it makes campaign advertising less effective.

This means that the probability that qualified candidates defeat their unqualified opponents is not as high as it could be, given the resources expended on advertising.

This difficulty is dramatically illustrated in the limiting case when candidates are infinitely power-hungry.

Proposition 2: *For all r, let $(C(r), f(r), \xi(r))$ be the equilibrium that would arise with no limits when non-policy related office holding benefits are r. Then,*

$$\lim_{r \to \infty} (C(r), f(r), \xi(r)) = (\frac{\gamma(b(\delta) - \delta)}{2}, \delta, 0).$$

The proposition implies that the equilibrium probability that a qualified candidate defeats an unqualified one tends to $1/2$ as candidates become more power-hungry; i.e., $\lim_{r \to \infty} \pi(r) = 1/2$.

Thus, campaign advertising is totally ineffective in the limit!

Now lets introduce campaign finance policy.

Consider policies that impose a limit l on the amount an interest group can contribute and provide matching public financing at rate s (i.e., if a candidate raises C in private contributions he gets sC).

Assume public financing is funded by a head tax T.

Consider first what can be achieved with a pure contribution limit policy ($s = 0$).

Fact *If imposing a limit reduces the level of favors and does not appreciably change the probability that a qualified candidate defeats an unqualified one, then it makes all types of citizens strictly better off.*

The reason why interest groups are better off is that all their gains from favors are paid for up front with contributions.

Combining this result with Proposition 2 enables us to establish:

Proposition 3 *If candidates are sufficiently power-hungry (i.e.,* r *is sufficiently large), banning contributions will create a Pareto improvement.*

Intuition: if contributions are banned entirely then no favors will be promised and the probability that a qualified candidate defeats an unqualified one is $1/2$.

With less power-hungry candidates, *banning* contributions could lead to a significant reduction in the probability that qualified candidates defeat unqualified ones.

However, *limiting* contributions need not appreciably reduce the probability that qualified candidates win.

Limiting contributions reduces the level of favors and this may increase advertising effectiveness.

This increase in effectiveness may compensate for the reduction in the level of advertising.

Unfortunately, while all this is possible, it is difficult to find sufficient conditions under which there exists a Pareto improving limit.

With public financing, public funds can substitute for the reduction in spending caused by limits.

Proposition 4 *Suppose that the Assumptions of Proposition 1 are satisfied and let* (C^*, f^*, ξ^*) *be an equilibrium with unrestricted contributions. Then there exists a campaign finance policy* (l, s) *and an equilibrium* (C, f, ξ) *under* (l, s) *such that* (C, f, ξ) *Pareto dominates* (C^*, f^*, ξ^*).

Thus, there always exists a Pareto improving campaign finance policy!

It is interesting to note that the campaign finance policies that the model suggests can be Pareto improving, closely resemble the *Clean Money, Clean Elections* program discussed earlier.

Indirectly informative advertising

The above model assumes that advertising directly informs voters of candidate characteristics.

Another branch of literature looks at indirectly informative advertising.

According to this perspective, it is the fact that a candidate has been given contributions that is the signal - not the messages per se.

When advertising is indirectly informative, banning contributions may improve aggregate voter welfare as shown by Prat (2003).

Basic trade off is between losses of information and benefits from reduced policy distortion.

However, the indirectly informative view does not appear to support public financing or contribution limits.

In many respects, neither the directly or the indirectly informative advertising models are very satisfactory.

There is a need for a better model of advertising incorporating the idea of persuasion - see DellaVigna and Gentzkow (2010) for more discussion.

Other relevant work

Houser and Stratmann (2008) study experimentally a game based on Coate (2004).

They provide experimental evidence for the idea that campaign advertising financed by favors is less effective than advertising financed by clean money.

Ashworth (2006) develops a model very similar to Coate (2004), but relaxes the assumption of candidate symmetry.

He notes that incumbent candidates have their advantage reinforced because challengers must promise more favors to get money for campaigning since they are less likely to be elected.

This means that challenger advertising is less effective and helps explain the extent of incumbent bias.

Vanberg (2008) explores theoretically and empirically a different argument for campaign contribution limits - the *equalization argument*.

The argument is that contribution limits are desirable because they equalize the influence of different donors and therefore cause a candidate's campaign resources to better reflect his public support.

Vanberg uses a theoretical model to identify the assumptions under which this argument makes sense and then analyzes whether these assumptions are satisfied empirically.

Prat, Puglisi, and Snyder (2010) study the extent to which private campaign contributions are targeted to better candidates.

They are interested in analyzing the common critique of public funding that funds allocated publicly will not get to the right candidates.

This critique is dealt with in practice in the Clean Money schemes by giving public money only to candidates with the required number of small contributions.

Prat, Puglisi, and Snyder provide some support for this practice by showing that the sum of small contributions (as opposed to the total sum of contributions) is strongly correlated with legislator effectiveness.

IV.4 Fiscal Policy and Fiscal Restraints

In public choice theory it is common to distinguish between the political and the fiscal constitution.

The political constitution specifies the electoral rules which govern how representatives are elected.

The fiscal constitution constrains the policies representatives may choose.

For example, the fiscal constitution could require that representatives run a balanced budget; that revenue by raised only from property taxes; that taxes cannot be person-specific; etc.

The fiscal constitution might also specify procedures by which fiscal decisions must be made; for example, raising taxes requires a super-majority vote in the legislature.

The idea of a fiscal constitution is to appropriately constrain politicians' fiscal choices and reflects political economy concerns.

Obviously, if government were benevolent no such constraints would be desirable.

In the public choice literature, the case for having a fiscal constitution has been made by James Buchanan and his co-authors.

Buchanan has also written a number of papers analyzing what principles should guide the choice of fiscal constitution.

Brennan and Buchanan (1977) is the best known example.

Brennan and Buchanan

This paper employs a very simple model of government - the so-called Leviathan model.

According to this model, the government sets taxes to maximize revenue.

It spends a fixed fraction of this revenue on public goods and services and the remainder is wasted (the idea is that the government consumes it somehow in unproductive ways).

The question addressed by Brennan and Buchanan, is that if this is the right model of government, what should it be allowed to tax?

The key problem is to allow the government to raise sufficient revenue as to provide valued public goods and services, but to not allow it to extract too much revenue.

The main point made by Brennan and Buchanan is that all the lessons of traditional public finance are turned on their head.

For example, taxing goods that are in inelastic demand is recommended by optimal tax theory because this minimizes distortions.

But allowing a Leviathan government to tax a good in inelastic demand would be a disaster - it would raise too much revenue.

Similarly, while traditional public finance recommends government have a comprehensive tax base (i.e., income, commodity, wealth, etc) to minimize distortions, a narrow base (e.g., only income) is desirable with a Leviathan government.

Subsequent literature

The subsequent literature on fiscal restraints has two strands.

One strand is devoted to the empirical question of whether the fiscal restraints that are used in practice actually have any effect.

For example, Poterba (1996) discusses the impact of balanced budget rules which are common at the state level in the U.S..

The key question is whether legislators can circumvent these rules via accounting gimmicks and the like.

The literature tends to finds that rules do matter: Poterba shows that states with more stringent restraints are quicker to reduce spending and/or raise taxes in response to negative revenue shocks than those without.

Knight (2000) looks at the impact of supermajority requirement for tax increases which are found in some states.

He argues that such requirements have significantly reduced taxes.

The second strand of work is devoted to the basic theoretical question of whether, assuming that they can be enforced and will not be circumvented, particular fiscal restraints are desirable.

Besley and Smart (2007) provide an analysis of balanced budget rules and other fiscal restraints in the context of a two period political agency model.

The key issue in their analysis is how having fiscal restraints influences the flow of information to citizens concerning the characteristics of their politicians.

Azzimonti, Battaglini, and Coate (2008) provide a theoretical analysis of the case for a balanced budget rule in the context of Battaglini and Coate's (2008) political economy model of fiscal policy.

They also calibrate the model to the U.S. economy to assess the case for a balanced budget rule for the federal government.

This is motivated by the recurring debate in American politics concerning the desirability of amending the U.S. constitution to require that the federal government operate under such a rule.

We will review this analysis, beginning with a review of the Battaglini and Coate model.

Battaglini and Coate model

This combines a "tax smoothing model" of fiscal policy with a legislative bargaining model of policy determination.

It incorporates the friction that legislators can distribute pork back to their districts and analyzes how this distorts fiscal policy.

The key contribution is to shed light on how a standard feature of legislative policy-making (pork-barrel politics) distorts macro fiscal policies.

The economy

A continuum of infinitely-lived citizens live in n identical districts indexed by $i = 1, ..., n$.

The size of the population in each district is normalized to be one.

There is a single (nonstorable) consumption good, denoted by z, that is produced using a single factor, labor, denoted by l, with the linear technology $z = wl$.

There is also a public good, denoted by g, that can be produced from the consumption good according to the technology $g = z/p$.

Consumers consume the consumption good, benefit from the public good, and supply labor.

Each consumer's per period utility function is

$$z + Ag^{\alpha} - l^{(1+\frac{1}{\varepsilon})}/(1+\varepsilon),$$

where $\alpha \in (0, 1)$ and $\varepsilon > 0$.

The parameter A measures the value of the public good.

Citizens discount future per period utilities at rate δ.

The value of the public good varies across periods in a random way, reflecting shocks to the society such as wars and natural disasters.

Specifically, in each period, A is the realization of a random variable with range $[A_0, A_1]$ (where $0 < A_0 < A_1$) and cumulative distribution function $G(A)$.

There is a competitive labor market and competitive production of the public good.

Thus, the wage rate is equal to w and the price of the public good is p.

There is also a market in risk-free one period bonds.

The assumption of a constant marginal utility of consumption implies that the equilibrium interest rate on these bonds must be $\rho = 1/\delta - 1$.

At this interest rate, consumers will be indifferent as to their allocation of consumption across time.

Government

The public good is provided by the government.

The government can raise revenue by levying a proportional tax on labor income.

It can also borrow and lend in the bond market by selling and buying bonds.

Revenues can not only be used to finance the provision of the public good but can also be used to finance a district specific transfers (non-distortionary “pork-barrel” spending).

At the beginning of each period, the government inherits a level of public debt b from the previous government.

This must be repaid with interest, which costs $(1+\rho)b$.

Government policy in the period is described by a $n+$ 3-tuple

$$\{\tau, g, b', s_1, ..., s_n\},$$

where τ is the income tax rate; g is the amount of the public good provided; b' is the proposed new level of public debt; and s_i is the proposed transfer to district i's residents.

When b' is negative, the government is buying bonds.

In a period in which government policy is $\{\tau, g, b', s_1, ..., s_n\}$, each consumer will supply an amount of labor

$$l^*(w(1-\tau)) = \arg\max_l \{w(1-\tau)l - \frac{l^{(1+1/\varepsilon)}}{\varepsilon + 1}\}.$$

It is straightforward to show that $l^*(w(1-\tau)) = (\varepsilon w(1-\tau))^{\varepsilon}$, so that ε is the elasticity of labor supply.

A consumer who simply consumes his net of tax earnings and his transfer will obtains a per period utility of $u(w(1-\tau), g; A) + s_i$ where

$$u(w(1-\tau), g; A) = \frac{\varepsilon^{\varepsilon}(w(1-\tau))^{\varepsilon+1}}{\varepsilon+1} + Ag^{\alpha}.$$

Since consumers are indifferent as to their allocation of consumption across time, their lifetime expected utility will equal that which they would obtain if they simply consumed their net earnings and transfer each period plus the value of their initial bond holdings.

Feasibility constraints

Government policies must satisfy three feasibility constraints.

The first is that revenues must be sufficient to cover expenditures.

To see what this implies, consider a period in which the initial level of government debt is b and the policy choice is $\{\tau, g, b', s_1, ..., s_n\}$

Expenditure on public goods and debt repayment is $pg + (1 + \rho)b$.

Tax revenues are

$$R(\tau) = n\tau w l^*(w(1 - \tau)) = n\tau w(\varepsilon w(1 - \tau))^{\varepsilon}$$

and total revenues are $R(\tau) + b'$.

Letting the *net of transfer surplus* (i.e., the difference between revenues and spending on public goods and debt repayment) be denoted by

$$B(\tau, g, b'; b) = R(\tau) - pg + b' - (1 + \rho)b,$$

the constraint requires that $B(\tau, g, b'; b) \geq \sum_i s_i$.

The second constraint is that the transfers must be non-negative (i.e., $s_i \geq 0$).

This rules out financing public spending via lump sum taxation.

The final constraint is that the amount of government borrowing must be feasible.

In particular, there is an upper limit on the amount the government can borrow given by $b_1 = \max_\tau R(\tau)/\rho$.

Political decision-making

Public decisions are made by a legislature consisting of representatives from each of the n districts.

One citizen from each district is selected to be that district's representative.

Since all citizens are the same, the identity of the representative is immaterial and hence the selection process can be ignored.

The legislature meets at the beginning of each period.

The affirmative votes of $q < n$ representatives are required to enact any legislation.

One of the legislators is randomly selected to make the first policy proposal, with each representative having an equal chance of being recognized.

If the first proposer's plan is accepted by q legislators, then it is implemented and the legislature adjourns until the beginning of the next period.

At that time, the legislature meets again with the difference being that there is new initial level of public debt and a new realization of the value of public goods.

If, on the other hand, the first proposal is not accepted, another legislator is chosen to make a proposal.

There are $T \geq 2$ such proposal rounds, each of which takes a negligible amount of time.

If the process continues until proposal round T, and the proposal made at that stage is rejected, then a legislator is appointed to choose a default policy that treats districts uniformly.

Political equilibrium

We look for a symmetric Markov-perfect equilibrium.

In this type of equilibrium, any representative selected to propose at round $r \in \{1, ..., T\}$ of the meeting at some time t makes the same proposal and this depends only on the current level of public debt (b), the value of the public good (A), and the bargaining round (r).

We assume that legislators vote for a proposal if they prefer it (weakly) to continuing on to the next proposal round.

We focus, without loss of generality, on equilibria in which at each round, proposals are immediately accepted by at least q legislators, so that on the equilibrium path, no meeting lasts more than one proposal round.

Accordingly, the policies that are actually implemented in equilibrium are those proposed in the first round.

Equilibrium policies

To understand equilibrium behavior note that to get support for his proposal, the proposer must obtain the votes of $q-1$ other representatives.

Given that utility is transferable, the proposer is effectively making decisions to maximize the utility of q legislators.

It is therefore *as if* a randomly chosen minimum winning coalition (mwc) of q representatives is selected in each period and this coalition chooses a policy choice to maximize its aggregate utility.

In any given state (b, A), there are two possibilities: either the mwc will provide pork to the districts of its members or it will not.

Providing pork requires reducing public good spending or increasing taxation in the present or the future (if financed by issuing additional debt).

When b and/or A are sufficiently high, the marginal benefit of spending on the public good and the marginal cost of increasing taxation may be too high to make this attractive.

In this case, the mwc will not provide pork and the outcome will be *as if* it is maximizing the utility of the legislature as a whole.

If the mwc does provide pork, it will choose a tax rate-public good-public debt triple that maximizes coalition aggregate utility under the assumption that they share the net of transfer surplus.

Thus, (τ, g, b') solves the problem:

$$\max u(w(1-\tau), g; A) + \frac{B(\tau,g,b';b)}{q} + \delta E v(b', A')$$
$$s.t. \quad b' \leq b_1.$$

where v is the continuation value function.

The optimal policy is $(\tau^*, g^*(A), b^*)$ where the tax rate τ^* satisfies the condition that

$$\frac{1}{q} = \frac{[\frac{1-\tau^*}{1-\tau^*(1+\varepsilon)}]}{n},$$

the public good level $g^*(A)$ satisfies the condition that

$$\alpha A g^*(A)^{\alpha-1} = \frac{p}{q},$$

and the public debt level b^* satisfies

$$\frac{1}{q} = -\delta E[\frac{\partial v(b^*, A')}{\partial b'}].$$

The first condition says that the benefit of raising taxes in terms of increasing the per-coalition member transfer $(1/q)$ must equal the per-capita cost of the increase in the tax rate.

The second condition says that the per-capita benefit of increasing the public good must equal the per-coalition member reduction in transfers it necessitates.

The third condition says that the benefit of increasing debt in terms of increasing the per-coalition member transfer must equal the per-capita cost of an increase in the debt level.

The mwc will choose pork if the net of transfer surplus at the optimal policy $B(\tau^*, g^*(A), b^*; b)$ is positive.

Otherwise the coalition will provide no pork and its policy choice will then maximize aggregate legislator (and hence citizen) utility.

In this case, the tax rate will exceed τ^*, the public good level will be less than $g^*(A)$, and the public debt level will exceed b^*.

The following result explains how the temptation to distribute pork distorts the political equilibrium:

Proposition 1 *The equilibrium value function* $v(b, A)$ *solves the functional equation*

$$v(b, A) = \max_{(\tau,g,b')} \left\{ \begin{array}{c} u(w(1-\tau), g; A) + \frac{B(\tau,g,b';b)}{n} + \delta E v(b', A') : \\ B(\tau, g, b'; b) \geq 0,\ \tau \geq \tau^*,\ g \leq g^*(A),\ \&\ b' \in [b^*, b_1] \end{array} \right\}$$

and the equilibrium policies $\{\tau(b, A), g(b, A), b'(b, A)\}$ *are the optimal policy functions for this program.*

Note that the planning problem for this model would be just to solve this problem without the lower bound constraints on taxes and debt, and the upper bound constraint on the public good.

Thus, political determination simply amounts to imposing three additional constraints on the planning problem.

This result makes clear exactly how politics distorts the planning solution: taxes can be too high, public good provision too low and borrowing too high.

Given Proposition 1, it is straightforward to characterize the equilibrium policies.

Define the function $A^*(b, b')$ from the equation

$$B(\tau^*, g^*(A), b'; b) = 0.$$

Then, if the state (b, A) is such that $A \leq A^*(b, b^*)$ the tax-public good-debt triple is $(\tau^*, g^*(A), b^*)$ and the mwc shares the net of transfer surplus $B(\tau^*, g^*(A), b^*; b)$.

If $A > A^*(b, b^*)$ the budget constraint binds and no transfers are given.

The tax-debt pair exceeds (τ^*, b^*) and the level of public good is less than $g^*(A)$.

In this case, the solution can be characterized by obtaining the first order conditions for the problem in Proposition 1 with only the budget constraint binding.

It can be shown that the tax rate and debt level are increasing in b and A, while the public good level is increasing in A and decreasing in b.

The determination of b^*

The characterization in Proposition 1 takes as fixed the lower bound on debt b^*.

However, b^* depends on the expected derivative of the value function.

Using Proposition 1, we can show that:

$$-\delta E[\frac{\partial v(b^*,A)}{\partial b'}] =$$
$$[G(A^*(b^*,b^*)) + \int_{A^*(b^*,b^*)}^{A_1} (\frac{1-\tau(b^*,A)}{1-\tau(b^*,A)(1+\varepsilon)})dG(A)]/n.$$

The intuition is this: in the event that $A \leq A^*(b^*,b^*)$ in the next period, increasing debt will reduce pork by an equal amount since that is the marginal use of resources.

By contrast, in the event that $A > A^*(b,b^*)$, there is no pork, so reducing debt means increasing taxes and $\frac{1-\tau}{1-\tau(1+\varepsilon)}$ is the marginal cost of taxation when the tax rate is τ.

Observe that since $1/q > 1/n$, $A^*(b^*, b^*)$ must lie strictly between A_0 and A_1.

Intuitively, this means that the debt level b^* must be such that next period's mwc will provide pork with a probability between zero and one.

If $R(\tau^*) > pg^*(A_1)$, this implies that b^* must be positive, so that $pg^*(A) + \rho b^*$ exceeds $R(\tau^*)$ for some realizations.

If $R(\tau^*) < pg^*(A_0)$, this implies that b^* must be negative, so that $R(\tau^*) - \rho b^*$ exceeds $pg^*(A)$ for some realizations.

The key determinant of the magnitude of b^* is the size of the tax base as measured by $R(\tau^*)$ relative to the public good needs of the economy as measured by $pg^*(A)$.

Equilibrium dynamics

The long run behavior of fiscal policies in the political equilibrium is summarized in the following proposition:

Proposition 2 *The equilibrium debt distribution converges to a unique, non-degenerate invariant distribution whose support is a subset of* $[b^*, b_1]$. *When the debt level is* b^*, *the tax rate is* τ^*, *the public good level is* $g^*(A)$, *and a minimum winning coalition of districts receive pork. When the debt level exceeds* b^*, *the tax rate exceeds* τ^*, *the public good level is less than* $g^*(A)$, *and no districts receive pork.*

Thus equilibrium fiscal policies fluctuate in the long run in response to shocks in the value of the public good.

Legislative policy-making oscillates between periods of pork-barrel spending and periods of fiscal responsibility.

Periods of pork are brought to an end by high realizations in the value of the public good.

These trigger an increase in debt and taxes to finance higher public good spending and a cessation of pork.

Once in the regime of fiscal responsibility, further high realizations of A trigger further increases in debt and higher taxes.

Pork returns only after a suitable sequence of low realizations of A.

The larger the amount of debt that has been built up, the greater the expected time before pork re-emerges.

It is worth comparing this long run behavior with that arising under the planning solution.

In the planning solution, the government gradually builds up sufficient assets to finance the Samuelson level of the public good in each period with no taxation.

Azzimonti, Battaglini, and Coate

Azzimonti et al address the following question: what would happen if, once debt was at long run equilibrium levels (i.e., $b \in [b^*, b_1]$), a balanced budget rule were to be imposed on the economy?

They model a balanced budget rule as a requirement that tax revenues must always be sufficient to cover spending and the costs of servicing the debt.

If the initial level of debt is b, this requires that

$$R(\tau) \geq pg + \sum_i s_i + \rho b.$$

Given the definition of $B(\tau, g, b'; b)$, such a rule is equivalent to adding, in each period, the feasibility constraint that $b' \leq b$; i.e., that debt cannot increase.

Thus, under a balanced budget rule, next period's feasible debt levels are determined by this period's debt choice.

In particular, if debt is paid down in the current period, that will tighten the debt constraint in the next period.

To state the main result of their paper, let $(\tau_b(A), g_b(A))$ be the tax rate and public good level that solve the static maximization problem

$$\max_{(\tau,g)} \left\{ u(w(1-\tau), g; A) + \frac{B(\tau, g, b, b)}{n} : B(\tau, g, b, b) \geq 0 \right\}.$$

Proposition 3 *Suppose that a balanced budget rule is imposed on the economy when the debt level is in the range $[b^*, b_1)$. Then, debt will converge monotonically to a steady state level b_0 smaller than b^*. At this steady state level b_0, when the value of the public good is less than $A^*(b_0, b_0)$, the tax rate will be τ^*, the public good level will be $g^*(A)$, and a mwc of districts will receive pork. When the value of the public good is greater than $A^*(b_0, b_0)$, the tax rate will be $\tau_{b_0}(A)$, the public good level will be $g_{b_0}(A)$, and no districts will receive pork.*

The intuition behind the debt reduction is that a balanced budget rule, by restricting future policies, increases the expected cost of taxation and increases legislators' incentive to save.

The above proposition provides a reasonably complete picture of how imposing a balanced budget rule will impact fiscal policy.

However, Azzimonti et al are also interested in the impact on citizens' welfare.

The argue that when the rule is first imposed, it will reduce contemporaneous utility.

When A is low, instead of transfers being paid out to the citizens, debt will be being paid down.

When A is high, the increase in taxes and reduction in public goods will be steeper than would be the case if the government could borrow.

Thus, in either case, citizen welfare will be lower.

As debt falls, the picture becomes less clear.

On the one hand, citizens gain from the higher average public spending levels and/or lower taxes resulting from the smaller debt service payments.

On the other hand, public good provision will be less responsive and there will be greater volatility in tax rates.

The inability to run deficits means that the only way to respond to positive shocks in the value of the public good is to raise taxes.

This leads to sharper tax hikes and, since the marginal cost of public funds is higher, public good provision incentives are dampened.

Thus, there is a clear trade-off whose resolution will depend on the parameters.

In their calibrated model, long run welfare is actually increased by 2.85%.

However, when account is taken of the short run transition costs, imposing a balanced budget rule reduces welfare.

Discussion

The balanced budget rule studied by Azzimonti et al differs from the rules actually proposed in Congress.

In particular, these rules can be overrided with a super-majority vote.

Azzimonti et al show that imposing a balanced budget rule with a super-majority override will have no effect on fiscal policy or citizen welfare.

Such a rule will only have an effect if imposed at the foundation of the state before debt has risen to equilibrium levels.

Intuitively, this is because in the Battaglini and Coate model, once debt has reached equilibrium levels, additional debt will be issued only when it is in the interests of *all* legislators to do so, rather than just a minimum winning coalition.

This result reflects the stationarity of the Battaglini and Coate model and would not necessarily apply in a growing economy.

Another feature of the rules actually proposed is that they are automatically suspended in times of war.

If we interpret war as periods in which A is high, having such an automatic suspension rule may be helpful in reducing the costs of a balanced budget rule.

This has yet to be studied.

There is plenty of scope for much more analysis of fiscal restraints in this style (see Bassetto and Sargent (2006) for another example).

IV.7 Electing the President

The president of the U.S. is probably the world's most powerful leader.

The president is elected by a very distinctive process which happens every four years.

At the beginning of an election year, there are presidential primaries which consist of sequential statewide contests that run from January through June.

The first contests are traditionally in Iowa, New Hampshire, and South Carolina.

These primaries are designed to pick the two parties' nominees who are annointed at the parties' national conventions which are held in the late summer.

Over the late summer and early fall, the two nominees campaign for the national election which is held in the first week of November.

The November election is not decided by a direct national vote but by the Electoral College system.

In this system, each state is represented by a slate of "electors" the number of which equals the state's number of senators (i.e., 2) plus the number of its House representatives.

The electors then cast their votes on behalf of their state and the candidate with the most votes wins.

By convention, the electors of all states except Maine and Nebraska cast their vote in favor of the candidate who wins a majority of votes in their state.

In Maine and Nebraska, two electoral votes are determined by the statewide vote, and the remainder by the votes in each congressional district.

Two recent papers study some of the anomalies created by this system.

Knight and Schiff (2010) focus on the idea that the sequential nature of the primary system gives voters in the early states more weight in determining the outcome.

This is a very common complaint about the system.

Stromberg (2008) focuses on the incentives created by the Electoral College in terms of which states receive attention during the campaign.

Both papers are very good and represent examples of frontier research in political economy combining theory and empirics.

We will discuss only the Stromberg paper because of time constraints.

Stromberg's Paper

Stromberg is interested in the long-standing argument that the Electoral College creates incentives for presidential candidates to only pay attention to a handful of key swing states.

To empirically measure the power of these incentives, he focuses on the allocation of presidential and vice-presidential visits during the campaign period following the two parties' nominating conventions.

He develops a model of campaign visits based on the Lindbeck-Weibull model and studies it's performance empirically.

He then uses the model to compare the equilibrium allocation of campaign visits with that which would arise if the president were elected by a direct national vote.

Stromberg's Model

There are two presidential candidates, indexed by D and R.

Both candidates have too make plans for the last I days before the election.

The plans specify how many of the days to use to visit each of S states.

Let d_s^J denote the number of days candidate $J \in \{D, R\}$ visits state $s \in \{1, ..., S\}$.

Each candidate J chooses $(d_1^J, ..., d_S^J)$ subject to the time constraint

$$\sum_s d_s^J \leq I. \tag{2}$$

State s has e_s electoral college votes and these are allocated to the candidate who wins the majority of votes in state s.

Voter i in state s votes for candidate D if

$$u(d_s^D) > u(d_s^R) + R_i + \eta_s + \eta$$

where $u(\cdot)$ is an increasing and strictly concave function.

The function $u(\cdot)$ tells us how voters are influenced by candidate visits - this influence is assumed to be common for all voters.

Let

$$\Delta u_s = u(d_s^D) - u(d_s^R).$$

The term $R_i + \eta_s + \eta$ represents voter i's ideological preference for candidate R over candidate D.

Voters' preferences depend on an idiosyncratic term R_i which is distributed in state s according to the CDF F_s; a statewide shock η_s which is distributed according to the CDF G_s; and a national shock η which is distributed according to the CDF H.

Given values of η_s and η, the fraction of votes candidate D receives in state s is

$$F_s(\Delta u_s - \eta_s - \eta).$$

It is assumed that when planning their schedule of campaign visits, candidates do not know the shocks η_s and η.

Let

$$D_s(\eta_s, \eta) = \begin{cases} 1 \text{ if } F_s(\Delta u_s - \eta_s - \eta) > 1/2 \\ 0 \text{ if } F_s(\Delta u_s - \eta_s - \eta) < 1/2 \end{cases}$$

Then, the probability that candidate D wins is

$$P_{EC} = \Pr\left[\sum_s D_s(\eta_s, \eta) e_s > \frac{1}{2} \sum_s e_s\right] \qquad (3)$$

Candidate D is assumed to choose $(d_1^D, ..., d_S^D)$ to maximize (3) subject to (2) taking as given $(d_1^R, ..., d_S^R)$.

Candidate R is assumed to choose $(d_1^R, ..., d_S^R)$ to minimize (3) subject to (2) taking as given $(d_1^D, ..., d_S^D)$.

Actually maximizing or minimizing (3) is complicated, so Stromberg comes up with a clever way to approximate (3).

He also chooses parametrized functional forms for the CDFs which allow him to solve for the equilibrium in closed form as a function of the parameters.

In equilibrium, the two candidates will choose identical schedules of campaign visits, following the logic of the Lindbeck- Weibull model.

Stromberg then estimates the parameters of the CDFs using past voting data, polls, etc and obtains his equilibrium predictions.

Figure 1 of his paper compares the predictions of his model with data on actual presidential and vice-presidential visits from the 2000 and 2004 campaigns.

The estimates are pretty close to the actual visits.

Direct Vote vs Electoral College

Let the number of voters in state s be denoted V_s.

Then, with a direct national vote, the probability that candidate D wins is

$$P_{DV} = \mathsf{Pr}\left[\sum_s F_s(\Delta u_s - \eta_s - \eta)v_s > \frac{1}{2}\sum_s v_s\right] \quad (4)$$

How would the allocation of campaign visits compare under the two systems?

Figure 6 of his paper deals with this question.

States above 1 on the vertical axis receive more than average visits per capita under the Electoral College system.

States above 1 on the horizontal axis receive more than average visits per capita with a Direct Vote.

States above the 45^o line lose from the reform and states below gain.

States gain or lose attention because of (i) their electoral size per capita and (ii) influence relative to electoral size.

States like RI and MA gain under Direct Vote because, while their average partisanship is high, they do have quite a lot of swing voters.

Their high average partisanship means that they are rarely competitive and thus attract little attention under the Electoral College.

Their large fraction of swing voters make them attractive targets under the Direct Vote.

States like NV and DE lose because they have a heavy endowment of electoral votes relative to popular votes (i.e., high e_s relative to v_s).

In terms of the aggregate distribution, equilibrium visits are much more concentrated under the Electoral College than the Direct Vote.

IV.8 Federalism

In many countries, there are different levels of government.

In the U.S., for example, we have the federal government, state governments, and local governments.

A classic question in public choice concerns the appropriate allocation of policy responsibility to the different government levels.

What policies should the various levels of government be responsible for?

We will review a number of arguments that are relevant for thinking about this question.

IV.8.i Spillovers vs Preference Heterogeneity

Wallace Oates in his classic book *Fiscal Federalism* presented an influential way of thinking about the assignment of policy responsibility between national and local governments.

Assignment of policy responsibility to the local level was desirable if there was significant preference heterogeneity between localities and if there were limited policy spillovers across localities.

To see the argument, consider the following model.

An economy is divided into two geographically distinct districts indexed by $i \in \{1, 2\}$.

Each district has a continuum of citizens with a mass of unity.

There are three goods in the economy; a single private good, x, and two local public goods, g_1 and g_2, each one associated with a particular district.

These local public goods could be thought of as roads or parks.

Each citizen is endowed with some of the private good.

To produce one unit of either of the public goods, requires p units of the private good.

Each citizen in district i is characterized by a public good preference parameter λ.

The preferences of a type λ citizen in district i are

$$x + \lambda[(1-\kappa)\ln g_i + \kappa \ln g_{-i}].$$

The parameter $\kappa \in [0, 1/2]$ indexes the degree of *spillovers*.

When $\kappa = 0$ citizens care only about the public good in their own district, while when $\kappa = 1/2$ they care equally about the public goods in both districts.

The *mean type* in district i is denoted by m_i, where $m_1 \geq m_2$.

We are interested in whether responsibility for choosing the local public goods should be assigned to local or national governments.

With local provision, the level of public good in each district is chosen by the government of that district and public expenditures are financed by a uniform head tax on local residents.

Thus, if district i chooses a public good level g_i, each citizen in district i pays a tax of pg_i.

With national provision, a uniform public good level is chosen by a national government with spending being financed by a uniform head tax on all citizens.

Thus, a public good levels g results in a head tax of pg.

The criterion for comparing the performance of centralized and decentralized systems is aggregate public good surplus.

With public good levels (g_1, g_2), this is

$$\begin{aligned} S(g_1, g_2) &= [m_1(1-\kappa) + m_2\kappa] \ln g_1 \\ &\quad + [m_2(1-\kappa) + m_1\kappa] \ln g_2 - p(g_1 + g_2). \end{aligned}$$

The surplus maximizing public good levels are given by

$$(g_1, g_2) = (\frac{m_1(1-\kappa) + m_2\kappa}{p}, \frac{m_2(1-\kappa) + m_1\kappa}{p}).$$

Local Provision

Assume each district's government maximizes public goods surplus in the district..

Policies are chosen simultaneously and independently.

Accordingly, the expenditure levels in the two districts (g_1^l, g_2^l) will be such that

$g_i^l = \arg\max_{g_i}\{m_i[(1-\kappa)\ln g_i + \kappa \ln g_{-i}^l] - pg_i\},\ i \in \{1, 2\}.$

Taking first order conditions and solving yields:

$$(g_1^l, g_2^l) = (\frac{m_1(1-\kappa)}{p}, \frac{m_2(1-\kappa)}{p}).$$

With spillovers, public goods are under-provided in both districts and this under-provision is increasing in the extent of spillovers.

National Provision

Assume the national government chooses the level that maximizes aggregate public goods surplus.

This level, denoted g^n, satisfies

$$g^n = \arg\max_g \{[m_1 + m_2] \ln g - 2pg\},$$

yielding

$$g^n = \frac{m_1 + m_2}{2p}.$$

When m_1 exceeds m_2, centralization under-provides public goods to district 1 and over-provides them to district 2 except when $\kappa = \frac{1}{2}$.

Local vs National Provision

Proposition *(i) If the districts are identical and spillovers are present, national provision produces a higher level of surplus than does local provision. Absent spillovers, the two systems generate the same level of surplus. (ii) If the districts are not identical, there is a critical value of* κ*, greater than* 0 *but less than* $\frac{1}{2}$*, such that national provision produces a higher level of surplus if and only if* κ *exceeds this critical level.*

Thus, without spillovers, local provision is superior - a result referred to as *Oates' Decentralization Theorem*.

With spillovers and identical districts, national provision is preferred.

With spillovers and non-identical districts, it is necessary to compare the magnitude of the two effects.

National provision is desirable if and only if spillovers are sufficiently large.

IV.8.ii Spillovers vs Preference Heterogeneity Revisited

The trade-off identified by Oates relies critically on the assumption that policies under national provision are uniform across districts.

If the government were permitted to choose different levels of public goods for the two districts, national provision is always better.

The justification for the assumption of uniformity under national provision is not completely clear.

One possibility is informational, the national government does not know what local people want and therefore chooses a "one size fits all" outcome.

Another possibility is that a uniform policy is chosen on fairness grounds.

Empirically, the uniformity assumption seems reasonable for things like environmental standards, but not for spending on roads, parks, etc.

Besley and Coate (2003) compare local and national provision without the uniformity assumption.

They argue that the sharing of the costs of local public spending under national provision will create a conflict of interest between citizens in different localities.

Assuming national spending decisions are made by a legislature of locally elected representatives, this conflict of interest will play out in the legislature.

Thus the performance of national provision will depend upon how the legislature chooses policy.

To make predictions concerning legislative policy-making, they draw on the citizen-candidate model of elections and the legislative bargaining approach to legislative decision-making.

Local Provision

Under local provision, each district elects a single representative from among its members to choose policy.

Representatives are characterized by their public good preferences λ and a representative of the majority preferred type is elected.

The policy determination process has two stages.

First, elections determine which citizens are selected to represent the two districts.

Second, policies are chosen simultaneously by the elected representative in each district.

To faciltate comparison with Oates, assume that the median preference equals the mean preference in each district.

Then, in equilibrium, each district i elects a representative with public good preferences m_i and the policy outcome is

$$(g_1, g_2) = (\frac{m_1(1-\kappa)}{p}, \frac{m_2(1-\kappa)}{p}).$$

This is the same as under the Oates approach.

National Provision

Under national provision, each district elects a representative to the legislature which then chooses policy.

The legislative bargaining approach predicts that policies will be determined by a minimum winning coalition and that the identitity of this minimum winning coalition will be uncertain.

To capture this, assume that if the representatives are of types λ_1 and λ_2, the policy outcome will be $(g_1^1(\lambda_1), g_2^1(\lambda_1))$ with probability $1/2$ and $(g_1^2(\lambda_2), g_2^2(\lambda_2))$ with probability $1/2$ where $(g_1^i(\lambda_i), g_2^i(\lambda_i))$ is the optimal choice of district i's representative; that is,

$$\begin{aligned}(g_1^i(\lambda_i), g_2^i(\lambda_i)) &= \arg\max_{(g_i, g_{-i})} \{\lambda_i[(1-\kappa)\ln g_i + \kappa \ln g_{-i}] \\ &\quad -\frac{p}{2}(g_i + g_{-i})\}.\end{aligned}$$

Note that

$$(g_i^i(\lambda_i), g_{-i}^i(\lambda_i)) = (\frac{2\lambda_i(1-\kappa)}{p}, \frac{2\lambda_i\kappa}{p}),\ i \in \{1, 2\}.$$

Given these policy rules, each citizen prefers a representative of his own type.

Thus, a pair of representative types is majority preferred if and only if it is a median pair; i.e., $(\lambda_1^*, \lambda_2^*) = (m_1, m_2)$.

The policy outcome with national provision will be random, generating $(g_1, g_2) = (\frac{2m_1(1-\kappa)}{p}, \frac{2m_1\kappa}{p})$ with probability $1/2$ and $(g_1, g_2) = (\frac{2m_2\kappa}{p}, \frac{2m_2(1-\kappa)}{p})$ with probability $1/2$.

It follows that there are two drawbacks of national provision: uncertainty and misallocation across districts.

Notice, however, that these problems are attenuated the greater are spillovers and the lower is heterogeneity.

This is because the basic conflict of interest between citizens in the two localities is lower when spillovers are high and preferences are similar.

Local vs National Provision

Proposition *(i) If the districts are identical, there is a critical value of* κ, *strictly greater than* 0 *but less than* $\frac{1}{2}$, *such that national provision produces a higher level of surplus if and only if* κ *exceeds this critical level. (ii) If the districts are not identical, there is a critical value of* κ, *strictly greater than* 0 *but less than* $\frac{1}{2}$, *such that national provision produces a higher level of surplus if and only if* κ *exceeds this critical level. This critical level is higher than that in Oates' approach.*

The bottom line is that the relative performance of local and national provision still depends upon spillovers and differences in tastes for public spending, but for different reasons than suggested by Oates.

IV.8.iii Voting with your Feet

In a setting of mobile households, local provision of public goods and services has a number of important advantages.

First, citizens can seek out jurisdictions with public outputs that are well suited to their tastes - so called “Tiebout sorting” (Tiebout (1956)).

Second, in contrast to the monopolist position of national government, local governments must compete for residents and this competition mitigates government exploitation of their citizens.

Taking a Leviathan view of government, competition can substitute for explicit fiscal restraints on government (Brennan and Buchanan).

Epple and Zelenitz (1981) provide a well-known analysis of this argument.

Consider a geographic area of land size L.

The area can be divided into J identical jurisdictions of size L/J, where $J \geq 1$.

The housing supply function on each unit of land in the area is $h_s(P)$ where P is the price of housing.

These supply functions are assumed exogenous and the suppliers are outside the model.

Jurisdiction j provides G_j units of a public service to each resident.

The cost of a unit of service is c and service provision is financed by a property tax t_j.

The total number of households is N and the number residing in jurisdiction j is N_j.

Each household has an identical utility function $U(x, h, G)$ defined over private consumption x, housing h, and public services G, and an identical income I.

Let

$$h_d(P(1+t), G) = \arg\max U(I - P(1+t)h, h, G)$$

denote the individual household's demand for housing and

$$V(P(1+t), G) = U(I - P(1+t)h_d(\cdot), h_d(\cdot), G)$$

the associated indirect utility function.

The timing of the model is that: (i) each jurisdiction chooses a tax-service level (t_j, G_j), and (ii) housing markets open and each household chooses a jurisdiction and housing level.

Housing prices must be such that for each jurisdiction j supply equals demand; i.e.,

$$N_j h_d(P_j(1+t_j), G_j) = h_s(P_j)L/J.$$

In addition, if $J > 1$, it must be the case that households are indifferent as to which jurisdiction they locate, which implies that

$$V(P_j(1 + t_j), G_j) = V(P_i(1 + t_i), G_i).$$

for any pair of jurisdictions j and i.

In choosing (t_j, G_j), jurisdiction j seeks to maximize the difference between tax revenues and expenditures which is given by

$$N_j \left[t_j P_j h_d(P_j(1 + t_j), G_j) - cG_j \right].$$

The surplus revenues are just consumed by the government and provide no benefits to citizens.

This is similar to Brennan and Buchanan's Leviathan assumption.

National Government

With a single national government ($J = 1$), the government would choose a tax-service pair (t, G) to maximize revenue

$$N\left[tP(\cdot)h_d(P(\cdot)(1+t), G) - cG\right]$$

where $P(\cdot)$ is implicitly defined by the market clearing condition

$$Nh_d(P(1+t), G) = h_s(P)L.$$

This is similar to a monopoly problem and is reasonably straightforward to solve.

The only way the households can avoid paying taxes is by reducing the amount of housing they consume.

The only incentive the government has to provide services is if housing demand is complementary with services.

Optimal tax and public service levels depend on elasticities of housing demand and supply.

Local Governments

With competing local governments ($J \geq 2$), governments have to worry that they will lose residents if they set taxes too high and services too low.

The paper characterizes the symmetric equilibrium (i.e., $(t_j, G_j) = (t, G)$ for all j) with $J \geq 2$ jurisdictions.

Obviously, households are better off with competing local governments than a national government.

However, Epple and Zelenitz (1981) show that, even with a very large number of local governments, jurisdictions are still able to extract a positive surplus.

In particular, they show that for very large J, it is approximately true that the equilibrium tax, public service and housing price (t, G, P) are such that

$$tPh_d = cG + \frac{Ph_d}{\theta}$$

where θ is the elasticity of housing supply (i.e., $\theta = P\partial h_s(P)/\partial P h_s$).

Thus, jurisdictions obtain a surplus from each resident equal to Ph_d/θ.

This surprising result reflects the fact that the government of each jurisdiction has the exclusive right to tax property in its district.

Housing suppliers cannot move their housing supply to another jurisdiction.

Thus, local governments exploit the elasticity of housing supply via taxation.

This results not only in transfers from housing suppliers to local governments, but also higher gross-of-tax housing prices for residents than would arise with benevolent governments.

IV.8.iv Race to the Bottom

An important argument in favor of national rather than local provision is that local control of policies can lead to a *race to the bottom*.

The intuitive idea is that local governments will compete to attract business capital because that will boost wages and employment opportunities for their citizens.

This competition will result in taxes on business that are too low.

Similarly, local competition to get business will lead to environmental standards, workplace regulations, etc that are too permissive.

By contrast, if business taxes and regulations are set nationally, there is less concern about capital mobility.

There is a large literature exploring this idea formally.

The simplest models assume that there is a single mobile factor of production (capital) that is the sole source of revenue for governments that provide a single public service.

Citizens are immobile and government policies are chosen by the median resident in each locality.

In these models, local service provision results in under-provision of public services as local governments compete for mobile capital by lowering taxes.

National service provision therefore generates higher welfare.

If you are interested in this line of argument, you can start with the chapter in W & W by Wildasin and the paper "Economic Competition among Jurisdictions: Efficiency Enhancing or Distortion Inducing" by Oates and Schwab in the *Journal of Public Economics* 1988.