Political Economics II
Spring 2016

Part III: Political Institutions and Economic Policy

Torsten Persson, IIES

Introduction

Effects of political institutions – "constitutional reform" ?

- what happens to economic policy?
- traditionally, neglected between economics and political science

Research in last ten-fifteen years

- new theory developed (applying tools you have seen already)
  and used as stepping stone for empirical analysis

Goal of Lectures 7-8

- give selective introduction to research results and methods
- underlying theoretical models and predictions (Lecture 7)
- empirical measurement, strategy and results (Lecture 8)

Reforms of which constitutional features?

two reasonably fundamental features, where theory exists

electoral rules: "majoritarian" vs. "proportional" (and finer detail)

forms of government: "presidential" vs. "parliamentary"

Which policies?

again, follow theory: fiscal policy (rent extraction)

large variation in observed outcomes, even among similar countries

Domain of application?

rules and policy outcomes at national level

parallel literature about sub-national institutions and policy

Background and Theoretical Predictions

today, we begin with some theory and its predictions

Agenda

A. A public-finance problem
B. Electoral rules and political competition
C. Forms of government and legislative bargaining
D. Summary of predictions

A. A public-finance problem

Groups of voters

homogenous (continuum) within $J = 1, 2, 3$, equal size, $N^J = 1$

Private preferences over economic outcomes

$$w^J = c^J + H(g) = y + (1 - \tau) + f^J + H(g) \tag{1}$$

policies both targeted (f^J) and non-targeted (τ, g) – i.e., combine the earlier size and composition of government models

Policy

$\mathbf{q} = [\tau, g, \{f^J\}, r]$

$$\Sigma_J f^J + g + r = 3\tau \tag{2}$$

r endogenous rent extracted by politicians

Rich, three-dimensional conflict over $\mathbf{q}$

among voters: over $\{f^J\}$

between voters and politicians: over τ, g vs. r

among politicians: over r

Basic hypothesis

resolution of these conflicting interests over $\mathbf{q}$
depends systematically on constitutional rules

cf. IO/micro theory, where resolution of conflicting interests
between firms and consumers depend on market regulation rules

B. Electoral rules and political competition

Electoral rules: two basic aspects

(i) electoral formula: vote shares → seat shares?)
winner takes all with plurality rule – proportionality with PR

(ii) district magnitude: # of lawmakers elected in average district?
typically low (often 1) with plurality rule – higher with PR

a third is ballot structure: voting for individuals or party lists?
are lists are open or closed – can preference votes be cast?

(i) and (ii) quite strongly correlated in real-world systems,
but not perfectly so

1. Adapt probabilistic voting model of Lecture 1

Traditional electoral competition

two parties (candidates), $C = A, B$

commitment to platforms before election, here $\mathbf{q}_C$ (r_C implicit)

opportunistic objectives, maximize

$$E(v_C) = p_C(R + \gamma r_C) \tag{3}$$

R exogenous ego-rent, p_C probability C wins the election,
$(1 - \gamma)$ transaction cost

Candidate preferences

i in $J = 1, 2, 3$ votes for A if

$$W^J(\mathbf{q}_A) > W^J(\mathbf{q}_B) + \sigma^{i,J} + \delta \tag{4}$$

Individual preferences: assume, not quite as before

$\sigma^{i,J} \lesseqgtr 0$ have group-specific distribution

uniform on $[-\frac{1}{2\phi^J} + \overline{\sigma}^J, \frac{1}{2\phi^J} + \overline{\sigma}^J]$

$\phi^2 > \phi^1, \phi^3, \ \overline{\sigma}^1 < \overline{\sigma}^2 = 0 < \overline{\sigma}^3, \ \overline{\sigma}^1\phi^1 + \overline{\sigma}^3\phi^3 = 0$

generate group biases for each party + convenient normalization

Aggregate popularity: assume (as before)

$\delta \lesseqgtr 0$ uniform on $[-\frac{1}{2\psi}, \frac{1}{2\psi}]$

Timing

parties know $\{\phi^J, \overline{\sigma}^J\}$ and ψ, but not δ, when set $\mathbf{q}_A$

Swing voters

in group J indifferent between A and B

$$\sigma^J = W^J(\mathbf{q}_A) - W^J(\mathbf{q}_B) - \delta \tag{5}$$

Vote share of A in J

all voters with $\sigma^{i,J} \leq \sigma^J$ cast ballot for A

$$\pi_A^J = \phi^J(\sigma^J - \overline{\sigma}^J + \frac{1}{2\phi^J}) = \frac{1}{2} + \phi^J(\sigma^J - \overline{\sigma}^J) \tag{6}$$

depends on policy via (5)

How do parties evaluate policy?

q affects $E\{\pi_A^J\}$ via identity of swing voters in each J
parties trade off votes against votes (choice of $\tau, g, \{f^J\}$)
and votes against rents (choice of r)

But ...

how p_A varies with $E\{\pi_A^J\}$ depends on electoral rule

2. PR elections (or large district magnitude)

Probability of winning

3 equal-sized districts J, same as groups: f^J regional transfer

enough seats and PR $\Rightarrow$ A seat share is $s_A^J = \pi_A^J$

winning requires 50% of national vote $\pi_A \equiv \frac{1}{3}\Sigma\pi_A^J \geq \frac{1}{2}$

$$p_A = \text{Prob}_\delta[\pi_A \geq \frac{1}{2}] = \frac{1}{2} + \frac{\psi}{3\phi}[\Sigma\phi^J(W^J(\mathbf{q}_A) - W^J(\mathbf{q}_B))] \quad (7)$$

by (5)-(6), distributional assumptions, and $\phi \equiv \frac{1}{3}\Sigma\phi^J$

(alternative interpretation: plurality rule in 1 national district)

Equilibrium – establish a benchmark for comparison

each party chooses identical policy platforms $\mathbf{q}_A = \mathbf{q}_B$

characterize equilibrium via four trade-offs associated with FOC for each policy instrument

(i) *Redistributive transfers:* $\{f^J\}$, votes in different groups?

chase most responsive voters: $\phi^2 > \phi^1, \phi^3$

complementary slackness: only $f^2 > 0$

hinges on quasi-linear utility (and uniform distribution)

(ii) *Size of government:* f^2 vs. τ, votes in different groups?

trade off group 2 votes against votes in all groups: because

$$\phi^2 \cdot 1 > \sum_J \phi^J \cdot \frac{1}{3} = \phi$$

$\tau = 1$ optimal (with distortionary taxes $\tau < 1$), as $\phi^2 > \phi$

(iii) *Public goods:* f^2 vs. g, votes in different groups?

similar trade-off pins down supply of g

$$\phi^2 \cdot 1 = \sum_J \phi^J \cdot H_g(g)$$

note underprovision $H_g = \frac{\phi^2}{3\phi} > \frac{1}{3}$ (utilitarian optimum)

more pronounced if ϕ^2 higher (more group-2 swing voters)

(iv) *Rents:* r, rents vs. votes (via lower f^2)?

equilibrium rents may well be positive

$$\frac{\gamma}{2} = p_A\gamma \leqslant -[R + \gamma r] \cdot \frac{dp_A}{dr} = [R + \gamma r] \cdot \frac{\psi\phi^2}{3\phi} \qquad [r \geqslant 0]$$

($p_A = \frac{1}{2}$ in equilibrium), higher if $\frac{\phi^2}{\phi}, \psi, R, (1-\gamma)$ lower

3. Plurality elections (or small district magnitude)

Probability of winning

again (now), 3 districts $J = 1, 2, 3$, same as groups

by plurality, winner takes all, $s_A^J = 1$ if $\pi_A^J \geq \frac{1}{2}$

if $-\overline{\sigma}^1$ and $\overline{\sigma}^3$ "large enough" $\Rightarrow$

A and B win their "safe districts", 1 and 3, for sure

Equilibrium

A, B only compete for $J = 2$, where most swing voters

$$p_A = \text{Prob}_\delta[\pi_A^2 \geq \frac{1}{2}] = \frac{1}{2} + \psi[W^2(\mathbf{q}_A) - W^2(\mathbf{q}_B)]$$

(i) *Redistributive transfers*

A,B incentives propose more f^2 than under PR

benefits the same, costs lower (ignored in district 1 and 3)

(ii) *Taxes*

still optimal to set $\tau = 1$

(iii) *Public goods*

further underprovision of g than under PR, since

$$\phi^2 \cdot 1 = \phi^2 \cdot H_g(g)$$

internalize benefits only for more narrow group of voters

$$H_g(g) = 1 > \frac{\phi^2}{3\phi} = \frac{\phi^2}{\Sigma\phi^J}$$

(iv) *Rents*

$$\frac{\gamma}{2} = p\gamma \leqslant -[R + \gamma r] \cdot \frac{dp_A}{dr} = [R + \gamma r] \cdot \psi$$

competition now more fierce, i.e., $-\frac{dp_A}{dr}$ higher as $\psi > \psi\frac{\phi^2}{3\phi}$

punishment for inefficiency larger than under PR

4. Comparative politics results

Plurality rule vs. PR (or small vs. large district magnitude)

focus competition to districts with many swing voters $\Rightarrow$

- more targeted redistribution (higher, concentrated f)
- less public goods (lower g)
- less rents (lower r)
- same size of government (same τ)

Compare to predictions in other models

P-T (2000, Ch. 9)
result on g vs. f, and τ the same, result on r opposite

L-P (2001) result on g vs. f the same

M-P-R (2002)
result on g vs. f the same, τ lower with majoritarian elections

C. Forms of government and legislative bargaining

Parliamentary and presidential regimes

differ both in executive-legislature relations and rules for legislation

Confidence requirements of legislature

parliamentary: executive accountable to legislature

presidential: no confidence requirement – as the executive (president) typically directly elected by voters

Separation of powers over legislation

parliamentary: powers often concentrated in cabinet

presidential: powers often separated across offices and legislators

1. Combine retrospective voting and legislative bargaining

Use earlier building blocks

policy problem in **A**, legislative-bargaining model in Lecture **4**, retrospective-voting/political-agency model in Lecture **4**

Legislators

3 incumbents J with opportunistic objectives

$$E(v_J) = \gamma r_J + p_J R,$$

Voters

in each district J, use retrospective voting rules

$$p_J = \begin{cases} 1 & \text{if } W^J(\mathbf{q}) \geq \varpi^J \\ 0 & \text{otherwise} \end{cases}$$

coordination within, but not across, districts

2. A simple legislature

Unrealistic warm up

neither confidence requirement, nor separation of powers

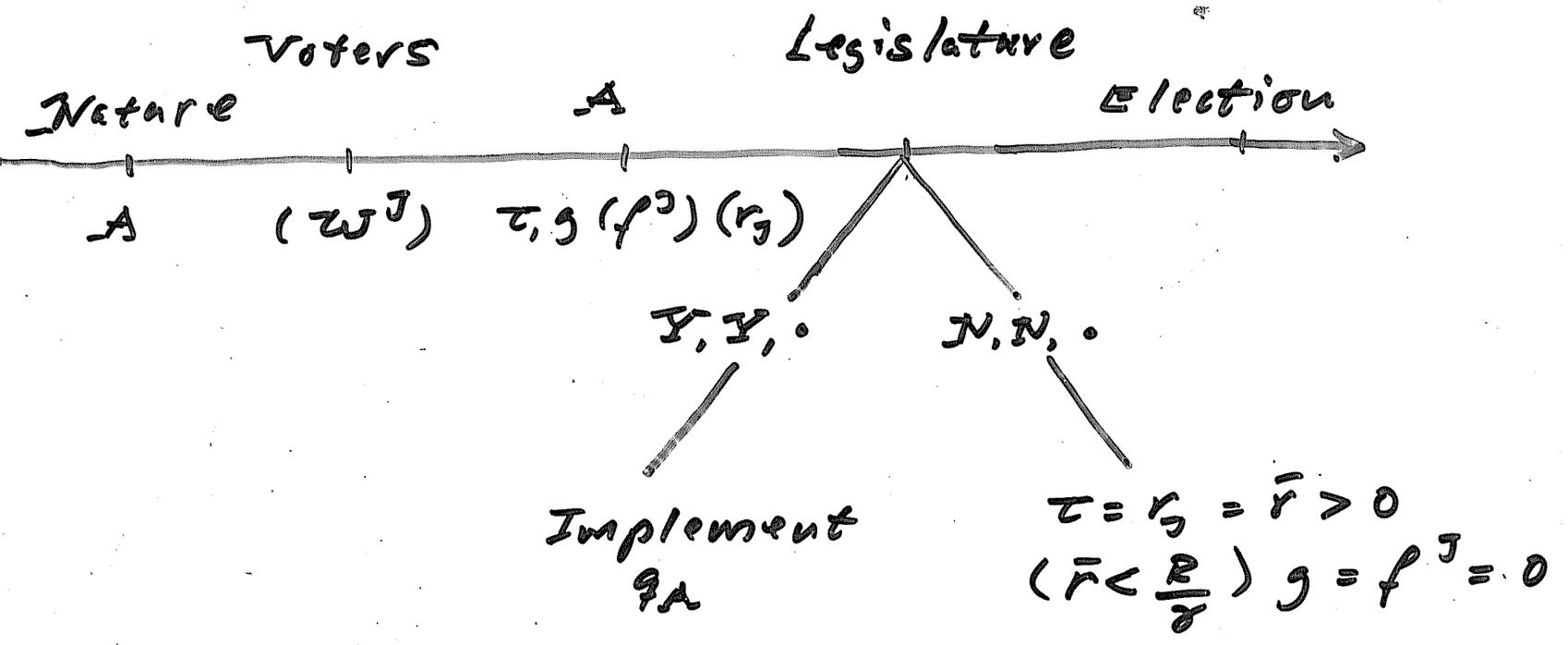
Timing

see *Figure 1*

Equilibrium conditions

(i) $\mathbf{q}(\boldsymbol{\varpi})$ dominates $\overline{\mathbf{q}}$ for at least one legislator $J \neq A$, any $\boldsymbol{\varpi}$
(ii) $\mathbf{q}(\boldsymbol{\varpi})$ optimal for A, any $\boldsymbol{\varpi}$, given (i)
(iii) ϖ^J optimal for voters in J, given (i), (ii) and $\boldsymbol{\varpi}^{-J}$

Figure 1


Nature
Voters
A
Legislature
Election
A
(w^J)
$\tau, g(f^J)(r_J)$
Y, Y, .
N, N, .
Implement g_A
$\tau = r_J = \bar{r} > 0$
$(\bar{r} < \frac{R}{2})$ $g = f^J = 0$

Equilibrium policy

$$\begin{aligned} r^{\mathrm{S}} &= r_A = 3 - \overline{r} - \frac{R}{\gamma} \\ \tau^{\mathrm{S}} &= 1 \\ f^{\mathrm{S}} &= f^A = \overline{r} + \frac{R}{\gamma} - g \\ H_g(g^{\mathrm{S}}) &= 1 \end{aligned} \tag{8}$$

Sketch of proof

consider arbitrary $\boldsymbol{\varpi} = (\varpi^A, \varpi^I, \varpi^K) > y + 1 - \overline{r}$
and A seeking re-election for herself and only one more J

Incentive constraint

joint "Leviathan" deviation should not pay

$$\gamma r + R \geq \gamma(3 - \overline{r}) \tag{9}$$

as re-election good enough for X $(R > \gamma\overline{r}) \Rightarrow r_J = 0, J \neq A$

Optimal behavior for A

pick $I = X$ if $\varpi^I < \varpi^K$, assume this is the case

Max $r_A \Rightarrow \tau = 1$, $f^K = 0$,
$W^I(\mathbf{q}(\boldsymbol{\varpi})) = \varpi^I$, $W^A(\mathbf{q}(\boldsymbol{\varpi})) = \varpi^A$

nail lawmaker I and voters in A, I, K to lowest possible payoff

Utility of voters

$$W^A(\mathbf{q}(\boldsymbol{\varpi})) = y + f^A + H(g) = (\varpi^A - H(g)) + H(g)$$
$$W^I(\mathbf{q}(\boldsymbol{\varpi})) = y + f^I + H(g) = (\varpi^I - H(g)) + H(g)$$
$$W^K(\mathbf{q}(\boldsymbol{\varpi})) = y + H(g)$$

W^I and W^K discontinuous at $\varpi^I = \varpi^K$ as long as $f^I > 0$
voters in I and K compete for any policy favors $\Rightarrow f^I = f^K = 0$

only lawmaker A has a binding re-election constraint

Voters in A

set ϖ^A equal to solution of

$$\text{Max}_{\tau, f^A, g}[y + (1 - \tau) + f^A + H(g)] \ ,$$

$$\text{s t} \quad f^A + g \leq \overline{r} + \frac{R}{\gamma} + 3(\tau - 1)$$

where constraint combines (2) and (9)

Optimal policy

voters in A agree with legislator A on $\tau = 1$

want to set $H_g(1) = 1$, trade off g and f^A one for one

induce legislator A to give up minimum rents

Three political failures (from the voters' horizon)

powerful A gets large benefits for her district

waste via positive equilibrium rents

underprovision of public goods

3. Presidential-congressional regime

Crucial features

separation of powers, but no confidence requirement (cf. US)

Timing

Figure 2

A_τ, A_g control different instruments, sequential decisions

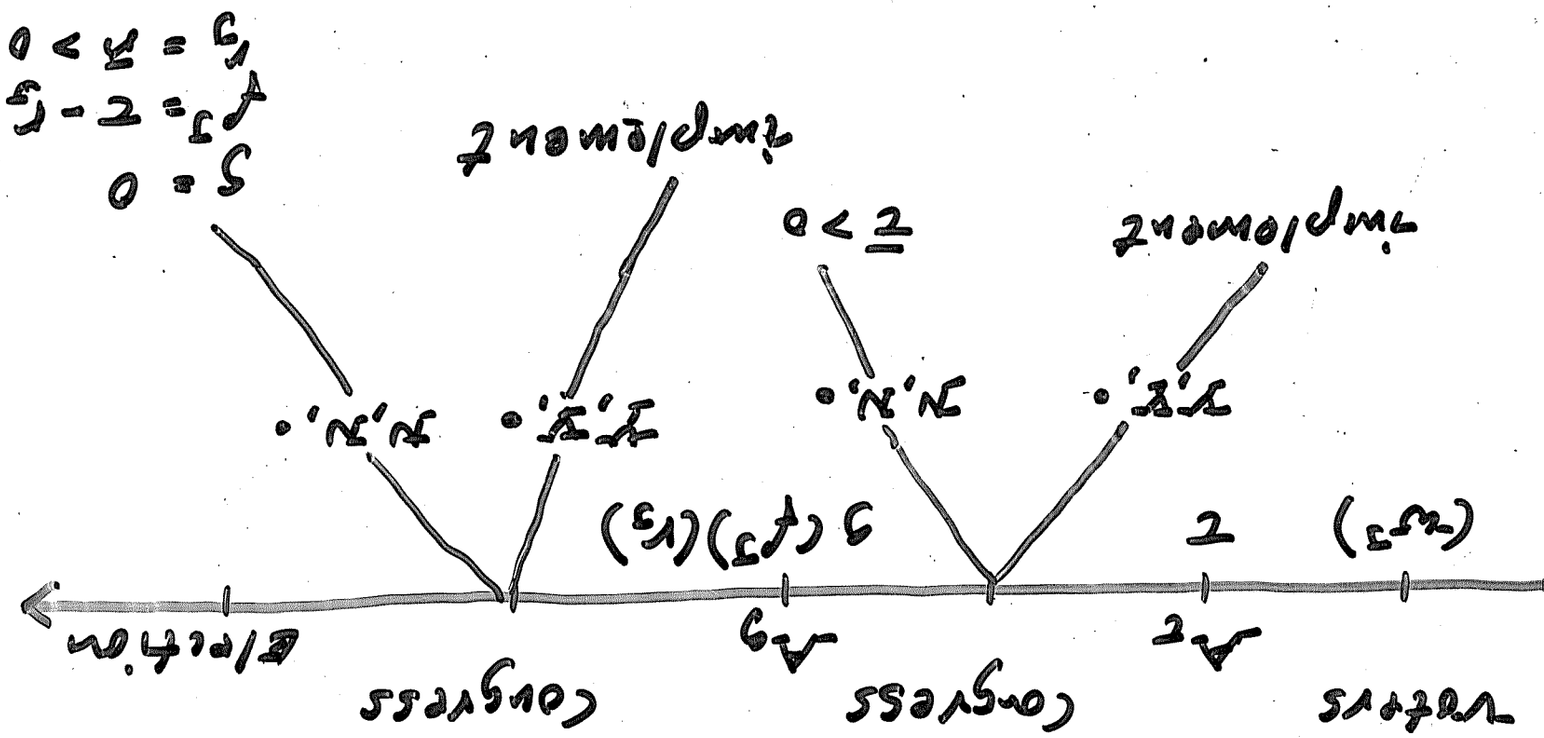

vectors
congress
congress
election
(w^3)
z
λ_2
λ_1
$3\,(r^3)(v_3)$
$z > 0$
timpleworst
timpleworst
$s = 0$
$r^3 = z - v_3$
$v_3 = \lambda > 0$

Figure 2

Equilibrium policy

$$\begin{aligned} r^{\mathrm{Pr}} &= 0 \\ \tau^{\mathrm{Pr}} &= \in [\frac{g}{3}, \frac{1}{3}(\overline{r} + \frac{R}{\gamma})] \\ f^{\mathrm{Pr}} &= f^{A_g} \in [0, \overline{r} + \frac{R}{\gamma} - g] \\ H_g(g^{\mathrm{Pr}}) &= 1 \end{aligned} \tag{10}$$

only give intuition for these results

Spending stage

(i) competition among voters in $J \neq A_g$ drives transfers to 0

(ii) A_g can please her voters, given available τ,
$\Rightarrow$ trade off f^{A_g} and g one for one, so underprovision of g

(iii) as τ is given, voters in A_g can insist on $r = r_A = 0$

Taxation stage

(i) A_τ and voters in A_τ not residual claimants on τ

(ii) may have to give up some revenue to voters in A_g
(so have multiple equilibria)

Overall conclusion

separation of powers and lack of stable majority $\Rightarrow$
f is targeted to minority,
get small broad program g,
r and τ are kept minimal

4. Parliamentary regime

Crucial features

concentration of powers to cabinet

confidence requirement: powers are maintained
only if the cabinet survives (cf. UK)

Timing

Figure 3

Incentives for stable majority

veto right for government members A_τ, A_g, costly to exercise

confidence vote followed by a government crisis
this creates "legislative cohesion"

Figure 3

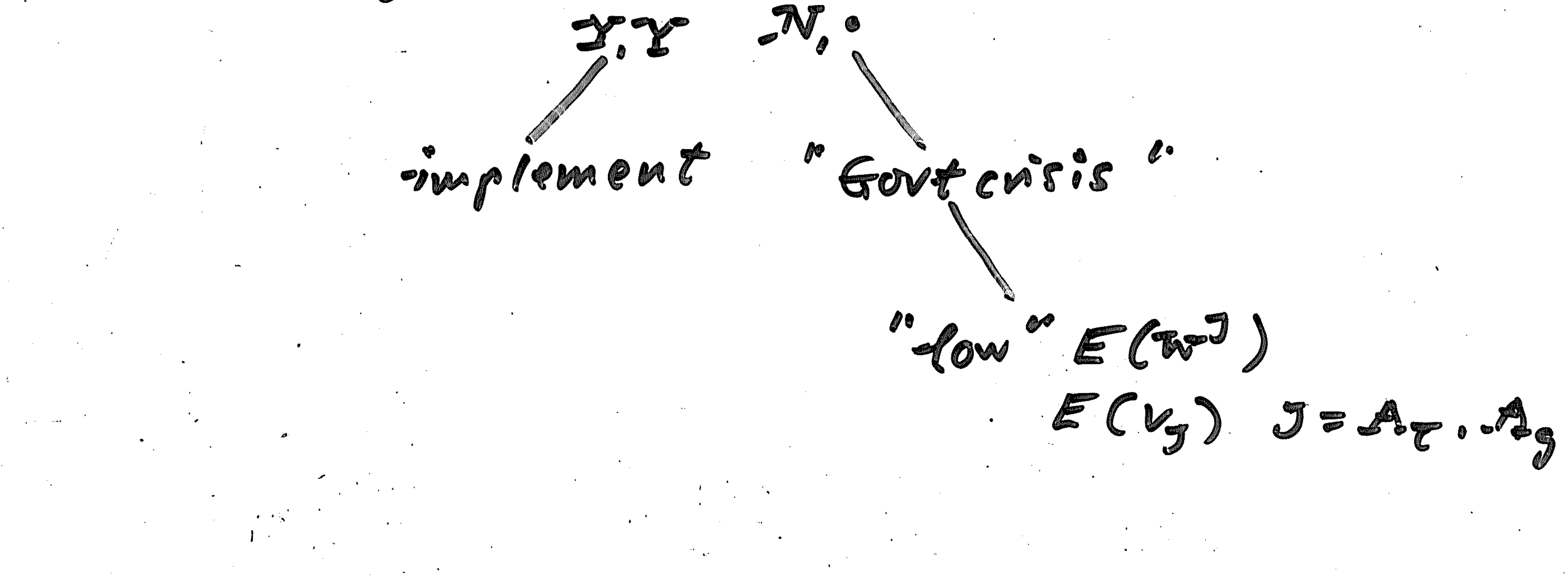

Voters
Governmat
Election
A_τ
A_g
(w^J)
τ
$g, (f^J), (r_J)$
Y, Y
N,
implement
"Govt crisis"
"low" $E(w^J)$
$E(v_J)$ $J = A_\tau \ldots A_g$

Table 2.1
Constitutions and economic policy
Theoretical predictions

Policy outcome	Electoral rules Majoritarian vs. proportional	Form of government Presidential vs. parliamentary
Overall size of government	– / ?	–
Composition: broad vs. narrow programs	–	–
Rent extraction	+ / –	–

Equilibrium policy

$$\begin{aligned} r^{\text{Pa}} &= 3 - \frac{2R}{\gamma} \\ \tau^{\text{Pa}} &= 1 \\ f^{\text{Pa}} &= f^{A_g} + f^{A_\tau} > 0 \\ \frac{1}{2} &\leq H_g(g^{\text{Pa}}) < 1 \end{aligned} \tag{11}$$

again only give intuition

Spending

(i) cost of break-up: both groups of voters backing government have some bargaining power

(ii) benefits of g internalized by majority, this gives less under-provision than in **2** and **3**

(iii) outcome jointly optimal for $J = A_\tau, A_g$ but can split benefits in different ways (multiple equilibria)

Rents

r high, as politicians can collude, absent separation of powers

Taxes

voters backing government want high τ as they,
and their legislators, are residual claimants on revenue

5. Comparative politics results

Presidential vs. parliamentary regimes

no confidence requirement, but separation of powers $\Rightarrow$

- redistribution more targeted (f to one group)
- less public goods (lower g)
- less rents (lower r)
- smaller size of government (lower τ)

Compare to predictions in other work

C-K (2011) on US cities find analogous result for τ

D. Summary of predictions

Theoretically predicted effects

size of government spending

composition into broad vs. targeted programs

(rents extracted by politicians)

$+$ or $-$ in *Table 2.1*

How take these to data? Do the predictions hold up?

those are the topics in last lecture

Table 2.1
Constitutions and economic policy
Theoretical predictions

Policy outcome	Electoral rules Majoritarian vs. proportional	Form of government Presidential vs. parliamentary
Overall size of government	– / ?	–
Composition: broad vs. narrow programs	–	–
Rent extraction	+ / –	–

Empirical Strategy and Results

Introduction

How do we test predictions in last lecture

as summarized in Table 2.1

Describe data, methods, and selected results

Agenda

A. Data and their properties
B. Statistical concerns
C. Identification: effects on government size
D. Tests of other predictions

A. Data and their properties

Sample selection

- given questions: most interesting variation across countries
- limit sample to democracies (in retrospect, a mistake!)
- treat democracy status as random
- generous definition of democracy, test if matters

Two data sets

- average annual obs. in 1990s for 85 democracies (cross section)
- annual obs. in 1960-98 for 60 democracies (panel)

Policy and performance measures

for each question in *Table 2.1*
describe in context

Country characteristics

many socioeconomic, historical, geographical, cultural variables
covariates with policy or constitutional rules

Table 2.1
Constitutions and economic policy
Theoretical predictions

Policy outcome	Electoral rules Majoritarian vs. proportional	Form of government Presidential vs. parliamentary
Overall size of government	– / ?	–
Composition: broad vs. narrow programs	–	–
Rent extraction	+ / –	–

Measures of electoral rules

binary indicator, MAJ, for electoral formula
is only plurality rule used in election of lower house?
yes: $MAJ = 1$, no: $MAJ = 0$

Measures of forms of government

binary indicator, $PRES$, for confidence requirement
is the executive independent of confidence in legislature?
yes: $PRES = 1$, no: $PRES = 0$

no measure of separation of powers

Main characteristics

two features important for empirical strategy

(i) Constitutional inertia

deep reforms rare events: panel has no switch in *PRES*

five in *MAJ* – ten if count mixed systems as well
and several more in detailed measures of electoral rules

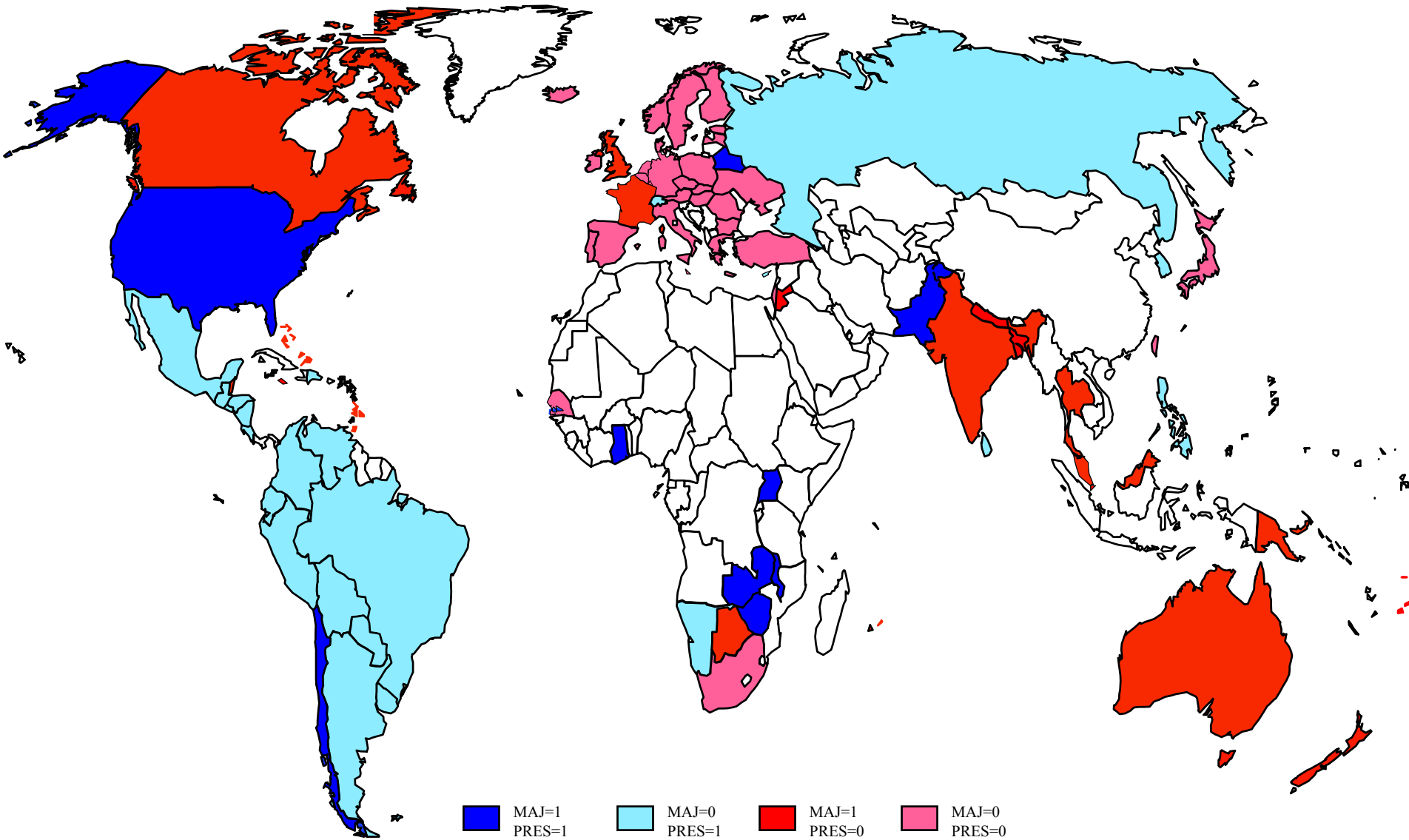
must estimate constitutional effects from cross-country variation

(ii) Non-random selection of constitutional rules

reflect history, geography and culture *Figure 4.1*

must be careful in inference

Constitutional Atlas 1998



B. Statistical concerns

Causal effects or statistical correlations?

- serious pitfalls in inference from cross-sectional data
- try to address by applying alternative methods

What form of simultaneity

- 'reverse causation' perhaps not major issue (inertia over 40 years)
- 'omitted variables' more serious problem

How avoid confounding constitutions and other policy determinants?

- hold constant observables which are correlated with policy, outcomes as well as constitution selection $\Rightarrow$
- use regression with many controls

How about remaining unobservables?

lacking imagination or difficulties in measurement can lead to bias

isolate exogenous variation $\Rightarrow$ IV-estimation

clean estimates from selection bias $\Rightarrow$ Heckman-style adjustment

How about unwarranted extrapolation for heterogenous groups?

allow for non-linear relations in data and give higher weight to 'local' comparisons $\Rightarrow$

estimate with matching methods

Illustrate in the context of our problem

use alternative methods to estimate same parameter and discuss specific identifying assumptions in theory and practice

apply to constitutional effects on the size of government

C. Identification: effects on government size

Details of empirical strategy for cross-sectional data

what's the parameter of interest?
under what assumptions can we estimate it, given concerns in **B**?
what results do we obtain?

Use size of government as specific example

1. Overall question and problem

Parameter of interest

what is direct effect of constitutional reform
in country selected at random (ATE)?

How does rule $S = 1$, vs. $S = 0$, affect policy outcome Y?

$$\alpha \equiv \mathsf{E}(Y^1 - Y^0) = \mathsf{E}\{E(Y^1 \mid \mathbf{X})\} - \mathsf{E}\{E(Y^0 \mid \mathbf{X})\} \qquad (14)$$

where last equality relies on law of iterated expectations when E and E refer to unobservables and observables, respectively

More concretely

how does switch *PRES* or *MAJ* from 0 to 1 affect central government spending (revenue) as % of GDP?

Problem with observational data

need country-i outcome in two potential states: Y_i^1, Y_i^0

observe only $Y_i = S_i Y_i^1 + (1 - S_i) Y_i^0$

other outcome is *unobserved counterfactual*

Pose problem differently

condition on S rather than on $\mathbf{X}$, rewrite (14) as

$$\alpha = \mathsf{E}(Y^1 - Y^0) = P \cdot [E(Y^1 \mid S = 1) - E(Y^0 \mid S = 1)] \quad (15)$$
$$+ (1 - P) \cdot [E(Y^1 \mid S = 0) - E(Y^0 \mid S = 0)]$$

α sums effects (ATN and ATT) for countries in states $S = 0, 1$ which have probabilities $1 - P$ and P, respectively

(Hypothetical) Experimental data

safe assume e.g., $\mathsf{E}(Y^0 \mid S = 1) = \mathsf{E}\{E(Y^0 \mid \mathbf{X}, S = 1)\}$
$= \mathsf{E}\{E(Y^0 \mid \mathbf{X}, S = 0)\} = \mathsf{E}(Y^0 \mid S = 0)$

because both unobserved and observed features balanced across $S = 0, 1$ – by randomization

Yet another way to pose the problem

consider general data-generating process

$$Y_i^S = F^S(\mathbf{X}_i) + \varepsilon_i^S, \quad S = 0, 1$$

$$S_i = \begin{cases} 1 \text{ as } G(\mathbf{W}_i, \mathbf{Z}_i) + \eta_i \geq 0 \\ 0 \text{ otherwise} \end{cases} \tag{16}$$

where $\mathbf{W}$ is subset of $\mathbf{X}$, while $\mathbf{Z}$ is not

Identification of α

need strong explicit or implicit assumptions about $F^S(\mathbf{X}), \varepsilon^S, \eta$ to tackle the statistical problems mentioned in **B**

2. Linear regression estimates

(i) *Recursivity*: $\text{Cov}(\varepsilon^S, \eta) = 0$

or, equivalently, *conditional independence*

$$E(Y^1 \mid \mathbf{X}, S = 1) = E(Y^1 \mid \mathbf{X}, S = 0) = E(Y^1 \mid \mathbf{X})$$

selection is random, after controlling for $\mathbf{X}$

$\mathbf{X}$ income, age and quality of democracy, openness, demographics, indicators for federal structure, OECD, continents, colonial history

(ii) *Linearity*

$$Y_i^S = F^S(\mathbf{X}_i) + \varepsilon_i^S = \alpha^S + \beta \mathbf{X}_i + \varepsilon_i^S$$

By (i) and (ii),

$$\alpha \ = \alpha^1 - \alpha^0 = \mathsf{E}\{E(Y^1 - Y^0 \mid \mathbf{X})\}$$

causal effect estimated by coefficient on S in OLS of Y on $\mathbf{X}, S$

$$Y_i = \alpha^0 + \alpha S_i + \boldsymbol{\beta}\mathbf{X}_i + e_i$$

where $e_i = \varepsilon_i^0 + S_i(\varepsilon_i^1 - \varepsilon_i^0)$

Results in *Table 6.1*

Table 6.1
Size of government and constitutions
Simple regression estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. var.	*CGEXP*	*CGEXP*	*CGEXP*	*CGREV*	*CGEXP*	*CGEXP*	*CGEXP*
PRES	-6.08	-5.29		-5.17	-8.29	-3.46	-7.49
	(1.97)***	(1.92)***		(2.44)**	(2.72)***	(3.88)	(2.72)***
MAJ	-3.29	-5.74		-3.03	-5.59	-2.93	-4.81
	(1.73)*	(1.95)***		(1.85)	(2.68)**	(3.09)	(2.75)*
PROPRES			-7.08				
			(2.70)**				
MAJPAR			-7.30				
			(3.02)**				
MAJPRES			-10.36				
			(2.70)***				
Continents	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Colonies	No	Yes	Yes	Yes	Yes	Yes	Yes
Sample	90s, broad	90s, broad	90s, broad	90s, broad	90s, narrow	60-90s, broad	90s,obs as(6)
Obs.	80	80	80	76	62	60	60
Adj. R2	0.58	0.63	0.63	0.58	0.60	0.54	0.63

Summary

$PRES = 1 \Rightarrow$ spending more than 5% of GDP smaller

$MAJ = 1 \Rightarrow$ slightly smaller effect

reflect more rapid growth of government, in 1965-85

constitutional effects appear additive

Is identification convincing ?

have we included all the relevant variables in $\mathbf{X}$ to rule out 'history and culture' determining both S and Y ?

can we trust there are no interaction effects or other non-linearities?

address these in turn in sections **3** and **4**

3. Relax conditional independence

Can we rule out selection on unobservables ?

i.e., is $\text{Cov}(e, \eta) \neq 0$?

Two prospective sources of selection bias

$$p\lim(\hat{\alpha}) = \alpha + \mathsf{E}(\varepsilon^1 \mid S = 1) - \mathsf{E}(\varepsilon^0 \mid S = 0) =$$
$$\alpha + \{\mathsf{E}(\varepsilon^0 \mid S = 1) - \mathsf{E}(\varepsilon^0 \mid S = 0)\} + \mathsf{E}(\varepsilon^1 - \varepsilon^0 \mid S = 1)$$

conventional omitted variables: ε^0 and S correlated

heterogenous constitutional effects: $\varepsilon^1 - \varepsilon^0$ and S correlated

How relax conditional independence?

a. use instrumental variables to isolate exogenous variation in S

b. adjust estimates, Heckman-style, for correlation e, η

a. Instrumental variables

Identifying assumptions

how find $\mathbf{Z}$ such that $\text{Cov}(\mathbf{Z}, S) \neq 0$ and $\text{Cov}(\mathbf{Z}, e) = 0$?

timing of latest constitutional reform (3 indicator variables)
historical waves of reform (hold constant age of democracy)

latitude, fractions speaking English, European language
geographic and cultural 'distance' to old democratic institutions

Relevant?

constitutional timing: yes, weakly

distance measures: yes, definitely

Exogenous?

constitutional timing: yes, a priori
distance measures: less certain

can test over-identifying restrictions, but low power

Apply to our data – results *Table 6.2*

Compare to earlier estimates in *Table 6.1*

- point estimates agree pretty well
- note parsimonious 1st stage (weak instruments)
- standard errors grow as 2nd stage is richer
- can't reject over-identifying assumptions

Table 6.2
Size of government and constitutions
Heckman and Instrumental Variables estimates

	(3)	(4)
Dep. var.	*CGEXP*	*CGEXP*
PRES	-8.65	-4.50
	(3.63)**	(3.89)
MAJ	-3.90	-5.12
	(3.46)	(3.61)
Conts & Cols	No	*COL_UKA, LAAM*
Sample	90s, broad	90s, broad
Endogenous	*PRES*	*PRES*
selection	*MAJ*	*MAJ*
Estimation	2SLS	2SLS
rho		
Chi-2: over-id	4.64	3.61
Adj. R2	0.59	0.60
Obs.	75	75

b. Heckman-style adjustment

Well-known idea

estimate selection equation (probit or logit) corresponding to

$$S_i = \begin{cases} 1 \text{ as } G(\mathbf{W}_i, \mathbf{Z}_i) + \eta_i \geq 0 \\ 0 \text{ otherwise} \end{cases}$$

correct estimates of α for remaining correlation $\rho = \text{Corr}(\eta, e)$

Identifying assumptions

same exclusion restrictions on $\mathbf{Z}$, as in **a**
to avoid relying (only) on functional form

Apply to our data – results *Table 6.2*

Compare to estimates in *Tables 6.1, 6.2*

point estimates, if anything, more negative

estimated $\rho > 0$ (for *PRES*) $\Rightarrow$ OLS had positive bias

could allow joint selection of *MAJ* and *PRES*, or
separate distributions for $\varepsilon^1, \varepsilon^0$ (heterogenous treatment effect)

Table 6.2
Size of government and constitutions
Heckman and Instrumental Variables estimates

	(1)	(2)	(3)	(4)
Dep. var.	*CGEXP*	*CGEXP*	*CGEXP*	*CGEXP*
PRES	-10.50	-5.37	-8.65	-4.50
	(3.98)***	(2.19)**	(3.63)**	(3.89)
MAJ	-5.69	-4.92	-3.90	-5.12
	(1.86)***	(2.57)*	(3.46)	(3.61)
Conts & Cols	Yes	Yes	No	*COL_UKA, LAAM*
Sample	90s, broad	90s, broad	90s, broad	90s, broad
Endogenous selection	*PRES*	*MAJ*	*PRES* *MAJ*	*PRES* *MAJ*
Estimation	Heckman 2-step	Heckman 2-step	2SLS	2SLS
rho	0.64	-0.02		
Chi-2: over-id			4.64	3.61
Adj. R2			0.59	0.60
Obs.	75	75	75	75

4. Relax linearity

Many reasons believe $Y^S = F^S(\mathbf{X})$ non-linear

no great concern if $\mathbf{X}^1$ and $\mathbf{X}^0$ have similar distribution

if not, specification bias – selection on observables – can be severe

Are $\mathbf{X}^1$ and $\mathbf{X}^0$ similar ?

test $E(\mathbf{X}^1) = E(\mathbf{X}^0)$ suggests not, cf. *Table 5.3*

PRES: reject in 7 cases out of 9

MAJ: reject in 4 cases out of 9

How take care of prospective specification bias ?

parsimonious assumption on functional form

rely on 'local' comparisons

Table 5.3
Balancing property
Equal-means tests for different constitutional groups

	Whole sample
MAJ=1 vs. MAJ=0	
LYP	0.04
PROP65	0.01
GASTIL	0.08
FEDERAL	0.93
COL_UKA	0.00
LAAM	0.34
TRADE	0.44
PROT80	0.94
CATHO80	0.00
PRES=1 vs. PRES=0	
LYP	0.00
PROP65	0.00
GASTIL	0.00
FEDERAL	0.22
COL_UKA	0.44
LAAM	0.00
TRADE	0.01
PROT80	0.03
CATHO80	0.00

Probabilities of falsely rejecting the hypothesis of equal means across constitutional groups under the hypothesis of equal variances.

Central idea in matching: mimic experimental measurement

split data in 'treated' and 'controls'

counterfactual for treated: controls with similar $\mathbf{X}_i$

estimate treatment effect of S on Y non-parametrically

Difficulty

too data-hungry, dimension of $\mathbf{X}$ large

Resolution

match on propensity score, rather than directly on $\mathbf{X}$

$p_i = p(\mathbf{X}_i) = \text{Prob}[S = 1 \mid \mathbf{X}_i]$

Identification

(i) conditional independence of Y given $\mathbf{X}$

(ii) common-support condition: $0 < p(\mathbf{X}_i) < 1$, all $\mathbf{X}_i$

can rewrite (15) as

$$\alpha = P \cdot \mathsf{E}\{[E(Y^1 \mid p) - E(Y^0 \mid p)] \mid S = 1\} +$$
$$(1 - P) \cdot \mathsf{E}\{[E(Y^1 \mid 1 - p) - E(Y^0 \mid 1 - p)] \mid S = 0\} \qquad (17)$$

i.e., condition on p and thus indirectly, not directly, on X

Number of practical questions in evaluating α

1. How estimate $p(\mathbf{X}_i)$?

   simple logit (or probit)

   specification of $\mathbf{X}$ reflect concern for (trade-off between)
   conditional-independence and common-support conditions

   highest t-statistics in equal-means tests and regressions

2. Does matching indeed balance the observations?

   Equal-means, for same $\mathbf{X}$ as before, in three strata for p

   PRES: now reject only in 2 out of 27 cases (cf. *Table 5.3*)

   MAJ: again, reject in 2 out 27 cases

3. How estimate P ?

   relative sample frequency of $S = 1$

Table 5.3
Balancing property
Equal-means tests for different constitutional groups

	Whole sample	$p < 0.33$	$0.33 < p < 0.67$	$0.67 < p$
MAJ=1 vs. MAJ=0				
LYP	0.04	0.04	0.62	0.21
PROP65	0.01	0.32	0.90	0.04
GASTIL	0.08	0.33	0.55	0.37
FEDERAL	0.93	0.79	0.57	0.48
COL_UKA	0.00	0.69	0.42	0.35
LAAM	0.34	0.27	0.39	0.17
TRADE	0.44	0.13	0.93	0.31
PROT80	0.94	0.56	0.75	0.37
CATHO80	0.00	0.11	0.46	0.83
PRES=1 vs. PRES=0				
LYP	0.00	0.87	0.01	0.54
PROP65	0.00	0.34	0.39	0.86
GASTIL	0.00	0.59	0.22	0.71
FEDERAL	0.22	0.07	0.30	0.27
COL_UKA	0.44	0.88	0.56	0.83
LAAM	0.00	0.53	0.23	0.22
TRADE	0.01	0.33	0.34	0.40
PROT80	0.03	0.65	0.60	0.22
CATHO80	0.00	0.28	0.24	0.02

Probabilities of falsely rejecting the hypothesis of equal means across constitutional groups under the hypothesis of equal variances.
Strata defined on the common support of propensity scores, p, estimated by logit regressions including: *LYP, PROP65, GASTIL, FEDERAL, COL_UKA, LAAM.*

4. How estimate $\mathsf{E}\{E(Y^1 \mid p) \mid S = 1\}$?

   sample mean among "treated"

5. How estimate $\mathsf{E}\{E(Y^0 \mid p) \mid S = 1\}$?

   which controls matched with given p_i among treated ?

   (i) nearest-neighbor: the (one) control with closest p_i
   produces (mostly) natural matches, cf. *Table 5.2*

   (ii) stratification: arithmetic mean, all controls in same p interval

   (iii) kernel: geometric mean, all controls in radius of p_i

6. How impose common-support condition (comparability) ?

   compute 3-5 only for overlapping support of p_i

Table 5.2
Estimated propensity scores

(a) Majoritarian elections

Country	*PSCORE*	*MAJ*	Country	*PSCORE*	*MAJ*
Uruguay	**0.052**	**0**	Nepal	0.337	1
Sweden	**0.070**	**0**	South Korea	0.355	0
Greece	**0.073**	**0**	Bangladesh	0.371	1
Bulgaria	**0.075**	**0**	Philippines	0.377	1
Italy	**0.077**	**0**	Namibia	0.419	0
UK	0.078	1	Barbados	0.496	1
Romania	0.083	0	New Zeland	0.568	1
Peru	0.084	0	Jamaica	0.582	1
Belgium	0.090	0	Ireland	0.617	0
Norway	0.090	0	Canada	0.641	1
France	0.093	1	Singapore	0.659	1
Spain	0.095	0	Israel	0.673	0
Latvia	0.101	0	Sri Lanka	0.674	0
Portugal	0.104	0	Trinidad&Tobago	0.694	1
Denmark	0.105	0	Australia	0.735	1
Hungary	0.106	0	South Africa	0.757	0
Japan	0.108	1	Cyprus (G)	0.759	0
Colombia	0.112	0	Malta	0.760	0
Estonia	0.114	0	Bahamas	0.763	1
Guatemala	0.115	0	Pakistan	0.781	1
Czech Republic	0.126	0	Uganda	0.790	1
Luxembourg	0.127	0	Gambia	0.794	1
Chile	0.128	1	Ghana	0.797	1
Argentina	0.132	0	Zimbabwe	0.808	1
Finland	0.132	0	Belize	0.812	1
Paraguay	0.133	0	Fiji	0.828	0
Slovak Republic	0.141	0	**Malawi**	**0.831**	**1**
Nicaragua	0.148	0	**St. Vincent&Granada**	**0.856**	**1**
Dominican Republic	0.152	0	**Zambia**	**0.856**	**1**
Netherlands	0.153	0	**Malaysia**	**0.857**	**1**
Ecuador	0.157	0	**Mauritius**	**0.873**	**1**
Germany	0.160	0	**India**	**0.886**	**1**
Russia	0.161	0	**Papua New Guina**	**0.904**	**1**
Poland	0.177	0	**Botswana**	**0.924**	**1**
Bolivia	0.181	0			
Honduras	0.185	0			
Mexico	0.194	0			
Austria	0.199	0			
Iceland	0.212	0			
Switzerland	0.214	0			
Turkey	0.220	0			
Brazil	0.230	0			
Costa Rica	0.240	0			
El Salvador	0.258	0			
Thailand	0.264	1			
Venezuela	0.292	0			
USA	0.297	1			
Senegal	0.320	0			

PSCORE is the predicted value of a logit regression of *MAJ* on *LYP, PROP65, FEDERAL, GASTIL, LAAM, COL_UKA*
Boldface observations are discarded to impose common support.

Apply to our data – results *Table 6.3*

Compare to earlier estimates in *Tables 6.1 - 6.2*

- point estimates agree
- based on fewer observations on common support
- higher standard errors: trade off less bias against less efficiency

Summary: estimated constitutional effects

- consistent with theory
- larger government, by $\sim 5\%$ of GDP, of parliamentary democracy
- by about same amount of proportional democracy

Table 6.3
Size of government and constitutions
Matching estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.	*CGEXP*	*CGEXP*	*CGEXP*	*CGEXP*	*CGEXP*	*CGEXP*
PRES	-7.30	-7.91	-5.87	-7.92	-2.54	-4.00
	(2.30)***	(2.90)***	(4.93)	(5.11)	(2.30)	(3.45)
MAJ	-5.76	-6.55	-4.87	-4.08	-6.59	-8.81
	(2.94)*	(2.82)**	(3.65)	(4.16)	(3.06)**	(3.15)***
Estimation	Kernel	Kernel	Strat	Strat	Nearest	Nearest
Sample	90s, broad	90s, broad	90s, broad	90s, broad	90s, broad	90s, broad
Logit Specif.	1	2	1	2	1	2
Obs. on common	65 *PRES*	40 *PRES*	65 *PRES*	40 *PRES*	65 *PRES*	40 *PRES*
support	67 *MAJ*	5 7 *MAJ*	67 *MAJ*	57 *MAJ*	67 *MAJ*	57 *MAJ*

5. Summary: size of government

Estimated constitutional effects

consistent with theory

larger government, by $\sim 5\%$ of GDP, of parliamentary democracy

by about same amount of proportional democracy

D. Tests of other predictions

Quick description of results

1. Composition of government

Measurement

Y welfare-state spending (social transfers), as % of GDP

$\mathbf{X}$ same as for overall spending

Estimation

same battery of methods as in **C**

Summary of findings

partly consistent with theory

2-3% of GDP higher in (good, old) proportional democracies

2-3% of GDP higher in (good, old) parliamentary democracies

2. Political rents and corruption

Measurement

Y perception indexes for corruption,
and government (in)effectiveness, 0-10 scale
$\mathbf{X}$ dozen covariates with corruption, suggested by earlier studies

Here can also exploit (piecemeal) electoral reforms

district magnitude & ballot structure, as well as electoral formula
cross-section and panel (fixed–effect) estimates agree

Summary of findings: electoral rules

larger electoral districts: less corruption (and inefficiency)

more list voting: more corruption (and inefficiency)

quantitatively important effects

binary indicator: no robust effect

Summary of findings: form of government

less corruption in presidential regimes?

perhaps, but only in better democracies

3. Extensions – examples of more recent research

(i) Deeper theoretical and empirical analysis
mechanism for higher government spending under PR?
incentives for politicians vs. indirect effect via party system

(ii) Wider scope of empirical work
systematic effects beyond fiscal policy (and corruption)?
trade policy, regulatory policy, economic performance

(iii) More extensive data sets with
sharper identification from time variation around reforms?
exploit switches in and out of democracy