Political Economics II
Spring 2010

Lecture 8

Part III: Political Institutions and Economic Policy
Empirical Strategy and Results

Torsten Persson

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 and empirical results: constitutional 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 – probably a mistake!
- treat 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

Measures of electoral rules

binary indicators, MAJ, $MIXED$ based on electoral formula

only plurality rule in election of lower house (MAJ =1) or not ($MAJ = 0$)

mix of plurality and PR ($MIXED$ =1)

also finer measures of district magnitude, ballot structure

Measures of forms of government

binary indicator, $PRES$, based on confidence requirement

for executive absent ($PRES$ =1) or present ($PRES$ =0)

no measure of separation of powers

Main characteristics

two features important for empirical strategy

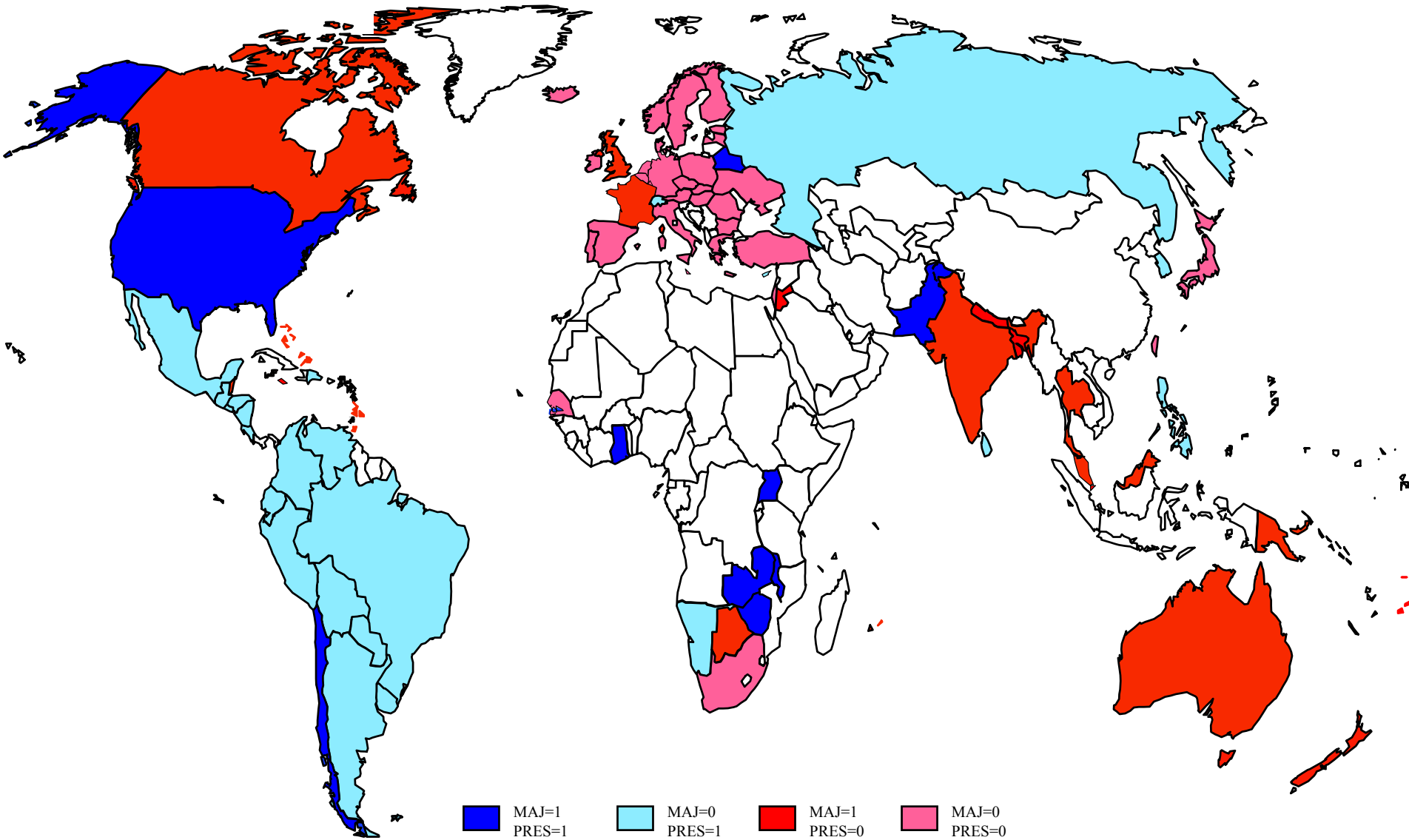
1. Non-random selection of constitutional rules

   reflect history, geography and culture *Figure 4.1*
   must be careful in inference

2. Constitutional inertia

   deep reforms rare events: panel has no switch in *PRES*,
   five in *MAJ*, ten in *MAJ* and *MIXED* together
   (more in detailed measures of electoral rules)
   must estimate constitutional effects from cross-country variation

Constitutional Atlas 1998



B. Statistical concerns

Causal effects or statistical correlations?

- pitfalls in inference from cross-sectional data
- address with alternative methods

Simultaneity

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

Confound constitutions and other policy determinants

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

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

Unwarranted extrapolation for heterogenous groups

allow for non-linear relations in data and

give higher weight to 'local' comparisons $\Rightarrow$

estimation with matching methods

How apply to our problem?

use all these methods to check robustness

discuss specific identifying assumptions in theory and practice

illustration: constitutional effects on the size of government

C. Identification and results: constitutional 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

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 (ATT) for obs. in $S = 0, 1$
with 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

consider general data generating process

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

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

$\mathbf{W}$ is subset of $\mathbf{X}$, $\mathbf{Z}$ 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 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 ?

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-identifyng 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-identifyng 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

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

Number of practical questions in evaluating α

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

   simple logit, or probit

   specification of $\mathbf{X}$ reflect concerns for
   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: reject in 2 out of 27 cases (*Table 5.3*)

   MAJ: again, reject in 2 out 27 cases

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.*

3. How estimate P ?

   relative sample frequency of $S = 1$

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: one 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

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

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*

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

Extensions – examples of 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
sharper identification from time variation around reforms?
exploit switches in and out of democracy