Causal Impact of Democracy on Growth: An Econometrician's Perspective

Shuowen Chen
BU

Victor Chernozhukov
MIT

Iván Fernández-Val
BU

January 4, 2019

Introduction

- Relationship between democracy and economic growth is of long standing interest
- We revisit the panel analysis of Acemoglu, Naidu, Restrepo, Robinson (19, JPE) using modern econometric methods that have been adapted to deal with **high-dimensional nature** of the problem
- Conventional panel estimators suffer from biases arising due to high-dimensionality:
  - Fixed Effects Estimator is biased due to estimation of **many nuisance parameters**;
  - Arellano-Bond Estimator is biased due to the use of **many moments**
- **De-biased estimators** produce **substantially higher estimates of long run effect of democracy on growth**

Dynamic Linear Panel Model

$$Y_{it} = a_i + b_t + D_{it}'\alpha + W_{it}'\beta + \epsilon_{it}, \quad i = 1, ..., N; \quad t = 1, ..., T$$

- Y_{it} is the outcome (log GDP) for a country i in year t
- D_{it} is a vector of treatments (encoding democracy), whose predictive effect α we would like to estimate
- W_{it} is a vector of covariates or controls including a constant and **lags of Y_{it}**
- a_i and b_t are unobserved unit and time effects that can be correlated to D_{it} and W_{it}.
- ϵ_{it} is an error term normalized to have zero mean for each unit that satisfies the **weak sequential exogeneity** condition

$$\epsilon_{it} \perp I_{it}, \quad I_{it} := \{(D_{is}, W_{is}, b_s)_{s=1}^{t}, a_i\}$$

Fixed Effects Approach

- Treat unit and time effects as parameters to be estimated
- Applies OLS in the model:

$$Y_{it} = D_{it}'\alpha + X_{it}'\gamma + \epsilon_{it}, \quad X_{it} := (W_{it}', Q_i', Q_t')'$$

- Q_i is an N-dimensional vector of unit fixed effects and Q_t is a T-dimensional vector of time fixed effects
- FE can be seen as an **exactly identified** GMM estimator with the score function

$$g(Z_i, \alpha, \gamma) = \{(Y_{it} - D_{it}'\alpha - X_{it}'\gamma)M_{it}\}_{t=1}^{T}, \quad M_{it} := (D_{it}', X_{it}')'$$

for $Z_i := \{(Y_{it}, D_{it}', W_{it}')'\}_{t=1}^{T}$

Problems with Fixed Effects Estimator: High-Dimensionality

- Here we estimate at least N nuisance parameters
- Need to rely on T large, formally $T \to \infty$ in order to drive bias b of the estimator to zero under the weak exogeneity condition
- But this alone does not suffice for inference, since bias b is large compared to the stochastic error $O(1/\sqrt{NT})$ of the estimator of α
- Problems with FE with small T are well documented in econometrics and machine learning (e.g., Neyman and Scott, 48; Nickell, 81; Hahn and Kuersteiner, 2011)

AB Approach

- Eliminate a_i by taking differences across time

$$\Delta Y_{it} = \Delta D_{it}'\alpha + \Delta X_{it}'\gamma + \Delta\epsilon_{it}, \quad X_{it} = (W_{it}', Q_t')'$$

- Moment conditions for the variables in differences

$$\Delta\epsilon_{it} \perp M_{it}, \quad M_{it} = [(D_{is}', W_{is}')_{s=1}^{t-1}, Q_t'], \qquad t = 2, \ldots, T$$

- Lead to overidentified GMM with score function

$$g(Z_i, \alpha, \gamma) = \{(\Delta Y_{it} - \Delta D_{it}'\alpha - \Delta X_{it}'\gamma)M_{it}\}_{t=2}^{T}, \quad M_{it} = [(D_{is}', W_{is}')_{s=1}^{t-1}, Q_t']$$

- This is the Arellano-Bond (91) estimator

Problems with AB: High Dimensionality

- Arellano-Bond is consistent under short-panel asymptotics, but it can be biased when T is large due to the many instrument problem
- The number of instruments or moment conditions is

$$m = \dim(g(Z_i, \alpha, \gamma)) = O(T^2)$$

- The bias of the estimator will scale with $m^2 = O(T^4)$ and may not be small compared to the sampling error $O_p(1/\sqrt{NT})$ of the estimator

High Dimensional Asymptotic Approximation

- In the FE approach, the dimension of α is low, but the dimension of γ might be high. Think of this as:

$$p = \dim(\gamma) \to \infty, \quad n \to \infty, \quad \dim(\alpha) = \text{const}.$$

- In the AB approach, the number of moment conditions,

$$m = \dim(g(Z_i, \alpha, \gamma)),$$

can be high, while the dimension of α is low. Think of this as:

$$m \to \infty, \quad n \to \infty, \quad \dim(\alpha) = \text{const}.$$

GMM in High Dimensions

- The approximate normality and consistency results of GMM estimator $\hat{\alpha}$ continue to hold if

$$p^2 \text{ and } m^2 \text{ are small compared to } n,$$

formally

$$(p \vee m)^2/n \to 0 \text{ as } n \to \infty$$

Can interpret as the small bias condition.

- But in the FE approach

$$p^2 = O(N^2 + T^2) \text{ is not small compared to } n = NT$$

- But in the AB approach, if T is large

$$m^2 = O(T^4) \text{ is not small compared to } n = NT$$

GMM in High Dimensions

- To understand the rate condition, focus on the exactly identified case where $p = m$
- An asymptotic second order expansion of $\hat{\alpha}$ around α gives

$$\hat{\alpha} - \alpha = Z_n/\sqrt{n} + b/n + r_n,$$

- Z_n is an asymptotic normal term, $b = O(p)$ is a first order bias term, and r_n is the higher order remainder such as

$$r_n = O_p((p/n)^{3/2} + p^{1/2}/n)$$

- $Z_n/\sqrt{n}$ dominates b/n if

$$\sqrt{n}b/n \to 0, \quad \text{i.e. } p^2/n \to 0,$$

and dominates r_n if

$$\sqrt{n}r_n \to_P 0, \quad \text{i.e. } p^{3/2}/n \to 0$$

Bias Corrections

- The bias is the bottleneck in the expansion, so can do two things:

  a) *Analytical bias correction*: estimate b/n using analytical expressions for the bias and set

  $$\check{\alpha} = \hat{\alpha} - \hat{b}/n.$$

  b) *Split-sample bias correction*: split the sample into two parts, compute the estimator on the two parts $\hat{\alpha}_{(1)}$ and $\hat{\alpha}_{(2)}$, and then set

  $$\check{\alpha} = \hat{\alpha} - (\bar{\alpha} - \hat{\alpha}), \quad \bar{\alpha} = (\hat{\alpha}_{(1)} + \hat{\alpha}_{(2)})/2.$$

  In some cases we can average over many splits to reduce variability

- With the bias correction, the rate requirement for GMM becomes weaker:

  $$(p \vee m)^{3/2}/n \to 0 \text{ as } n \to \infty$$

Why does the sample-splitting method work?

- The key is to split the sample such that the number of nuisance parameters and moment conditions are the same in all the parts
- Then, assuming that the parts are homogenous, the first order biases of $\hat{\alpha}$, $\hat{\alpha}_{(1)}$, and $\hat{\alpha}_{(2)}$ are

$$\frac{b}{n}, \quad \frac{b}{n/2}, \quad \frac{b}{n/2}$$

- The first order bias of $\check{\alpha}$ is

$$2\frac{b}{n} - \left(\frac{1}{2}\left[\frac{b}{n/2}\right] + \frac{1}{2}\left[\frac{b}{n/2}\right]\right) = 0$$

Implementation of Split-Sample Bias Correction

- Need to determine the right partition of the data
- In FE approach: halve the panel along the time series dimension (Dhaene and Jochmans, 15)

  This partition preserves the time series structure and delivers two panels with the same number of unit fixed effects and half the number of observations
- In AB approach: halve the panel along the cross section dimension

  This partition delivers two panels where the number of observations relative to the number of instruments is half of the original panel

  Can average across multiple splits to reduce variability because the cross-sectional ordering of the observations is arbitrary

Analytical Bias Correction for FE Approach

- Bias b is determined by, following Nickell (81) type arguments:

$$Hb = -\frac{1}{T}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\mathrm{E}[D_{it}\epsilon_{is}], \quad H = \frac{1}{NT}\sum_{k=1}^{N}\sum_{t=1}^{T}\mathrm{E}[\widetilde{D}_{it}\widetilde{D}_{it}'],$$

  where $\widetilde{D}_{it}$ is the residual of the linear projection of D_{it} on X_{it}

- $b = O(N)$ because the source of the bias is the estimation of the N unit fixed effects

- There is no bias from time fixed effects because the model is linear

- An estimator of the bias (Hahn and Kuersteiner, 11) can be formed as

$$\hat{H}\hat{b} = -\sum_{t=1}^{T-1}\sum_{s=t+1}^{(t+M)\wedge T}\frac{D_{is}\hat{\epsilon}_{it}}{T-s+t}, \quad \hat{H} = \frac{1}{NT}\sum_{k=1}^{N}\sum_{t=1}^{T}\widetilde{D}_{it}\widetilde{D}_{it}',$$

  where $\hat{\epsilon}_{it}$ is the fixed effect residual and M is such that $M/T \to 0$ and $M \to \infty$ as $T \to \infty$

Democracy and Growth : Acemoglu et al. (19, JPE)

- Extract a balanced panel of 147 countries over the period from 1987 through 2009 from Acemoglu et al. (19, JPE)
- Y_{it} is the logarithm of GDP per capita in 2000 USD as measured by the World Bank for country i at year t
- D_{it} is a democracy indicator that combines information from several sources including Freedom House and Polity IV

  It captures a bundle of institutions that characterize electoral democracies such as free and competitive elections, checks on executive power, an inclusive political process

Summary Statistics

	Mean	SD	Dem $= 1$	Dem $= 0$
Democracy	0.62	0.49	1.00	0.00
Log(GDP)	7.58	1.61	8.09	6.75
Number Obs.	3,381	3,381	2,099	1,282

Empirical Specification

- Dynamic linear panel data model, where we control for unobserved country effects, time effects (e.g. trends) and rich dynamics of GDP using **four lags** of Y_{it}
- The target parameter α measures the instantaneous or short-run effect of a transition to democracy on economic growth
- The permanent or long-run dynamic effect is

$$\alpha/(1 - \sum_{j=1}^{4} \beta_j),$$

where $\beta_1, \ldots, \beta_4$ are the coefficients corresponding to the lags of Y_{it}

Estimators

- FE estimates $p = 170$ parameters with $n = 147 \times 19 = 2,793$ observations, after using the first 4 periods as initial conditions, it **fails small bias** condition:

$$(m \vee p)^2/n \approx 10$$

- Bias corrected FE-ABC and FE-SBC implement the analytical and split-sample bias correction for FE
- AB relies on $m = 632$ instruments and $n = 147 \times 18 = 2,646$ observations, after using the first 5 periods as initial conditions; it **fails small bias condition** because

$$(m \vee p)^2/n \approx 150$$

- Bias-Corrected AB- SBC1 and AB-SBC5 implement the split-sample bias correction for AB with 1 and 5 splits
- Analytical and bootstrap standard errors clustered at the country level

Effect of Democracy on Economic Growth

	Fixed Effects			AB		
	FE	FE-ABC	FE-SBC	AB	AB-SBC1	AB-SBC5
Short-run	1.89	2.27	2.44	3.94	5.22	4.53
(×100)	(0.65)			(1.50)		
	[0.64]	[0.64]	[0.96]	[1.52]	[1.83]	[1.91]
Long-run	16.05	**25.91**	**25.69**	20.97	**26.46**	**25.24**
(×100)	(6.67)			(9.51)		
	[6.63]	[9.31]	[12.12]	[9.38]	[10.72]	[11.29]

Note 1: All the specifications include country and year effects.

Note 2: Clustered standard errors at the country level in parentheses.

Note 3: Bootstrap standard errors in brackets based on 500 replications.

Concluding Remarks

- We revisit the analysis of the causal effect of democracy on economic growth using state of the art econometrics methods that address the high-dimensional aspects of the problem.
- Traditional (FE and Arellano-Bond) estimators might be biased due to high dimensionality (number of estimated nuisance parameters or number of moments too high). Apply bias corrections for high dimensionality based on analytical and split-sample methods.
- Bias corrected estimators produce **substantially higher estimates of long run effect of democracy on growth** (than the estimates without bias correction).
- The state-of-the art methods are simple-to-use and are taught in our first-year econometrics courses, also discussed in our introductory econometrics book draft (available via MIT opencourseware for 14.382).