Causal Impact of Democracy on Growth: An Applied Econometrics Perspective

By Shuowen Chen and Victor Chernozhukov and Iván Fernández-Val*

The relationship between democracy and economic growth is of long standing interest in Economics. We revisit the empirical analysis of Acemoglu et al. (forthcoming) using state of the art econometric methods. We consider variations of the GMM Arellano-Bond and fixed effects estimators of a dynamic linear panel data model with country and time fixed effects. We find that both methods produce similar estimates of the short-run and long-run effects of democracy on growth once the GMM estimator is bias-corrected for the many instrument problem and the fixed effect estimator is bias-corrected for the incidental parameter problem. Our estimated effects show that the finding that democracy does cause growth is not sensitive to the econometric methodology.

I. Econometric Methods

A. *The Setting*

We consider the dynamic linear panel data model

$$(1) \qquad Y_{it} = a_i + b_t + D_{it}'\alpha + W_{it}'\beta + \epsilon_{it},$$

where $i = 1, ..., N$ and $t = 1, ..., T$. Here Y_{it} is the outcome for an observational unit i at time t, D_{it} is a vector of variables of interest or treatments, 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 ϵ_{it} is an error term normalized to have zero mean for each unit that satisfies the weak exogeneity condition

$$(2) \quad \epsilon_{it} \perp I_{it}, \quad I_{it} := \{(D_{is}, W_{is}, b_s)_{s=1}^{t}, a_i\}.$$

We shall assume that all the variables are stationary over t conditional on the unobserved effects, and the vectors $Z_i := \{(Y_{it}, D_{it}', W_{it}')'\}_{t=1}^{T}$, that collect these variables for the observational unit i, are i.i.d. across i. The main challenge in the estimation of panel data models is how to deal with the unobserved effects. We review two approaches.

B. *Fixed Effects Approach*

This approach treats the unit and time effects as parameters to be estimated by applying OLS in the model:

$$Y_{it} = D_{it}'\alpha + X_{it}'\gamma + \epsilon_{it},$$

where $X_{it} := (W_{it}', Q_i', Q_t')'$, Q_i is an N-dimensional vector of indicators for observational units with a 1 in the i-th position and 0's otherwise, and Q_t is a T-dimensional vector of indicators for observational units with a 1 in the t-th position and 0's otherwise. The elements of Q_i and Q_t are called unit fixed effects and time fixed effects, respectively. The resulting estimator is the fixed effect or least squares dummy variable estimator. It can be seen as an exactly identified GMM estimator with the score function

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

where $M_{it} := (D_{it}', X_{it}')'$.

Under a short panel asymptotic approximation where $N \to \infty$ and T is fixed, the fixed effect estimator of α is inconsistent in general. This is a manifestation of

* Chen: Department of Economics, Boston University, swchen@bu.edu. Chernozhukov: Department of Economics, MIT, vchern@mit.edu. Fernández-Val: Department of Economics, Boston University, ivanf@bu.edu. We gratefully acknowledge research support from the British Academy's visiting fellowships and National Science Foundation. Fernández-Val was visiting UCL while working on this paper, he is grateful for their hospitality.

the so-called *incidental parameter problem* (Neyman and Scott, 1948) that arises in situations where there are many nuisance (non-target) parameters such as the fixed effects. In this case, the number of nuisance parameters $p := \dim(\gamma) = \dim(\beta) + T + N$ is non-negligible compared to the number of observations $n = NT$. A way to tackle this problem is to adopt an alternative large panel asymptotic approximation where both $N \to \infty$ and $T \to \infty$. Under these sequences, the fixed effect estimator is consistent, but has a small-sample bias that affects inference. We show in Section I.D how to reduce this bias.

C. GMM Approach

This approach eliminates the unit effects a_i by taking differences across time and uses moment conditions for the variables in differences. Specifically, define the differencing operator Δ acting on doubly indexed random variables V_{it} by creating the difference $\Delta V_{it} = V_{it} - V_{it-1}$. Apply this operator to both sides of (1) to obtain:

$$\Delta Y_{it} = \Delta D_{it}'\alpha + \Delta X_{it}'\gamma + \Delta\epsilon_{it}, \tag{3}$$

where $X_{it} = (W_{it}', Q_t')'$. Note that by (2),

$$\Delta\epsilon_{it} \perp (D_{is}, W_{is})_{s=1}^{t-1}, \quad t = 2, \ldots, T.$$

This means that estimation and inference can be done using an overidentified GMM with score function

$$g(Z_i, \alpha, \gamma) = \{(\Delta Y_{it} - \Delta D_{it}'\alpha - \Delta X_{it}'\gamma) M_{it}\}_{t=2}^{T},$$

where $M_{it} = [(D_{is}', W_{is}')_{s=1}^{t-1}, Q_t']$. This is the Arellano and Bond (1991) estimator.

The Arellano-Bond estimator is consistent under short-panel asymptotics, but can be biased in large panels due to the *many instrument problem.* In this case the number of moment conditions $m = \dim(g(Z_i, \alpha, \gamma)) = T + T(T-1)(\dim(\alpha) + \dim(\beta))/2$ can be a non-negligible fraction of the number of observations $n = NT$. We show how to reduce the bias in the next section.

D. Bias Corrections

In the fixed effect approach, the dimension of α is low, but the dimension of γ might be high. We can approximate this situation as $p = \dim(\gamma) \to \infty$ when $n \to \infty$, while $\dim(\alpha)$ is fixed. In the GMM approach, the number of moment conditions, $m = \dim(g(Z_i, \alpha, \gamma))$, could be high, so we can approximate this situation as $m \to \infty$ when $n \to \infty$. Under this high dimensional asymptotic approximation, there exist regularity conditions such that if the square of the dimension of the nuisance parameter or the number of moments is small compared to the sample size, namely that:

$$(p \vee m)^2/n \to 0 \text{ as } n \to \infty, \tag{4}$$

then the approximate normality and consistency results of the GMM estimator continue to hold:

$$\sqrt{n}(\hat{\alpha} - \alpha) \overset{\text{a}}{\sim} N(0, V_{11}), \tag{5}$$

where V_{11} is the $d_\alpha \times d_\alpha$ upper-left block of the asymptotic variance of the GMM estimator corresponding to $\hat{\alpha}$.[1]

The rate condition (4) has a simple practical message: p^2 and m^2 should be small compared to n. This is not the case in the fixed effect approach where $p^2 = O(N^2 + T^2)$ and $n = NT$, and might not provide a good approximation to the GMM approach when T is large because $m^2 = O(T^4)$ and $n = NT$. To understand where (4) comes from, let us focus on the exactly identified case where $p = m$. An asymptotic second order expansion of $\hat{\alpha}$ around α gives

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

where $Z_n \overset{\text{a}}{\sim} N(0, V_{11})$, $b = O(p)$ is a first order bias term coming from the quadratic term of the expansion, and r_n is the higher order remainder such as $r_n = O_p((p/n)^{3/2} +$

[1] Sufficient conditions are given, for example, by Newey and Windmeijer (2009) for GMM problems with $m \to \infty$ and p fixed; and by Hahn and Newey (2004), Hahn and Kuersteiner (2011) and Fernández-Val and Weidner (2018) for nonlinear panel data models where $m \propto p \to \infty$.

$p^{1/2}/n$). Then, (5) holds if both

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

and

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

The previous derivation shows that the bias is the bottleneck. If we remove the bias somehow, then we can improve the requirement from $p^2/n \to 0$ to a weaker condition. There are several ways of removing the bias:

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

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

b) *Split-sample bias correction*, where we split the sample into two parts, compute the estimator on the two parts $\hat{\alpha}_{(1)}$ and $\hat{\alpha}_{(2)}$ to obtain $\bar{\alpha} = (\hat{\alpha}_{(1)} + \hat{\alpha}_{(2)})/2$, and then set

$$\check{\alpha} = \hat{\alpha} - (\bar{\alpha} - \hat{\alpha}) = 2\hat{\alpha} - \bar{\alpha}.$$

In some cases we can average over many splits to reduce variability, and it is also possible to use the bootstrap and leave-one-out methods for bias correction.

Why does the sample-splitting method work? Assuming that we estimate the same number of nuisance parameters and use the same number of moment conditions in all the parts of the sample, and that these parts are homogenous, then the first order biases of $\hat{\alpha}$, $\hat{\alpha}_{(1)}$, and $\hat{\alpha}_{(2)}$ are

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

so that the first order bias of $\check{\alpha}$ is

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

With the bias correction, the resulting rate conditions are weaker. In particular, there exist regularity conditions such that if

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

then the approximate normality and consistency results for the bias-corrected GMM estimator continue to hold:[2]

$$\sqrt{n}(\check{\alpha} - \alpha) \overset{a}{\sim} N(0, V_{11}).$$

To implement the analytical bias correction, we need to characterize the first order bias. For the fixed effect approach, Nickell (1981) showed that

$$Hb = -\frac{1}{T}\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{s=1}^{t-1}\mathrm{E}[D_{it}\epsilon_{is}],$$

where

$$H = \frac{1}{NT}\sum_{k=1}^{N}\sum_{t=1}^{T}\mathrm{E}[\widetilde{D}_{it}\widetilde{D}_{it}'],$$

and $\widetilde{D}_{it}$ is the residual of the linear projection of D_{it} on X_{it}. Note that $b = O(N)$ because the source of the bias is the estimation of the N unit fixed effects and the order of the bias is $b/n = O(T^{-1})$ because there are only T observations that are informative about each unit fixed effect.[3] An estimator of the bias can be formed as

$$\hat{H}\hat{b} = -\sum_{t=1}^{T-1}\sum_{s=t+1}^{(t+M)\wedge T}\frac{D_{is}\hat{\epsilon}_{it}}{T-s+t},$$

where $\hat{\epsilon}_{it}$ is the fixed effect residual,

$$\hat{H} = \frac{1}{NT}\sum_{k=1}^{N}\sum_{t=1}^{T}\widetilde{D}_{it}\widetilde{D}_{it}',$$

and M is a trimming parameter such that $M/T \to 0$ and $M \to \infty$ as $T \to \infty$ (Hahn and Kuersteiner, 2011).

To implement the split-sample bias correction, we need to determine the par-

[2]Sufficient conditions are given in Kiviet (1995), Hahn and Kuersteiner (2002) and Chudik, Pesaran and Yang (2018) for dynamic linear panel data models and Fernández-Val and Weidner (2016) and Fernández-Val and Weidner (2018) for nonlinear panel data models.

[3]There is no bias coming from the estimation of the time fixed effects because the model is linear and we assume independence across i.

tition of the data. In the fixed effect approach, we split the panel along the time series dimension because the source of the bias is the estimation of the unit fixed effects. Thus, following Dhaene and Jochmans (2015), the parts contain the observations $\{i = 1, \dots, N; t = 1, \dots, \lceil T/2 \rceil\}$ and $\{i = 1, \dots, N; t = \lfloor T/2 \rfloor, \dots, T\}$, where $\lceil \cdot \rceil$ and $\lfloor \cdot \rfloor$ are the ceiling and floor functions. This partition preserves the time series structure and delivers two panels with the same number of unit fixed effects, where there are $T/2$ observations that are informative about each unit fixed effect. In the GMM approach, we split the panel along the cross section dimension because the source of the bias is the number of moment conditions relative to the sample size. Thus, the parts contain the observations $\{i = 1, \dots, \lceil N/2 \rceil; t = 1, \dots, T\}$ and $\{i = \lfloor N/2 \rfloor, \dots, N; t = 1, \dots, T\}$. This partition delivers two panels where the number of observations relative to the number of moment conditions is half of the original panel. Note that there are multiple possible partitions because the ordering of the observations along the cross section dimension is arbitrary. We can therefore average across multiple splits to reduce variability.

II. Democracy and Growth

We revisit the application to the causal effect of democracy on economic growth of Acemoglu et al. (forthcoming) using the econometric methods described in Section I. We use a balanced panel of 147 countries over the period from 1987 through 2009 extracted from the data set used in Acemoglu et al. (forthcoming). The outcome variable Y_{it} is the logarithm of GDP per capita in 2000 USD as measured by the World Bank for country i at year t. The treatment variable of interest D_{it} is a democracy indicator constructed in Acemoglu et al. (forthcoming), which combines information from several sources including Freedom House and Polity IV. This indicator captures a bundle of institutions that characterize electoral democracies such as free and competitive elections, checks on executive power, an inclusive political process that permits various groups of society to be represented politically, and expansion of civil rights. Table 1 reports some descriptive statistics of the variables used in the analysis. The unconditional effect of democracy on GDP is 134% in this period.

We control for unobserved country effects, time effects and rich dynamics of GDP using the linear panel model (1), where W_{it} includes four lags of Y_{it}. The weak exogeneity condition (2) implies that democracy and past GDP are orthogonal to contemporaneous and future GDP shocks, and that the error ϵ_{it} is serially uncorrelated once we include four lags of GDP. In addition to the instantaneous or short-run effect of a transition to democracy to economic growth measured by the coefficient α, we are interested in a permanent or long-run dynamic effect. This effect in the dynamic linear panel model (1) is

$$\alpha/(1 - \sum_{j=1}^{4} \beta_j), \qquad (6)$$

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

Table 2 presents the results.[4] We consider estimators based on the fixed effect and GMM approaches. For example, FE is the fixed effect estimator and AB is a one-step Arellano-Bond estimator. For each estimator, we report analytical standard errors clustered at the country level and bootstrap standard errors based on resampling countries with replacement. The estimates of the long-run effect are obtained by plugging-in estimates of the coefficients in the expression (6). We use the delta method to construct analytical standard errors clustered at the country level, and resample countries with replacement to construct bootstrap standard errors. FE finds that a transition to democracy increases economic growth by almost 2% in the first year and 16% in the long run, while AB finds larger impacts of 4% and 21% but less precisely estimated. In results not reported in the table, we do not find evidence to re-

[4] We obtained the estimates with the commands `plm` and `pgmm` of the package `plm` in `R`.

TABLE 1—DESCRIPTIVE 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

ject the J-test of overidentifying restrictions based on AB.[5]

AB relies on $m = 632$ moment conditions to estimate $p = 169$ parameters with $n = 147 \times 18 = 2,646$ observations, after using the first five periods as initial conditions. Thus, $(m \vee p)^2/n \approx 150$ and therefore the rate condition (4) is unlikely to provide a good approximation. We consider two versions of the split-sample bias correction. SBC1 uses one random split, whereas SBC5 uses the average of five random splits. A similar situation arises for FE because it estimates $p = 170$ parameters with $n = 147 \times 19 = 2,793$ observations, after using the first four periods as initial conditions, such that $(m \vee p)^2/n \approx 10$. We consider both analytical and split sample bias corrections. ABC4 implements the analytical correction with $M = 4$, whereas SBC implements the split-sample bias correction for the fixed effect approach. We report bootstrap standard errors for the corrected estimators. There is no need to recompute the analytical standard errors, because the ones obtained for the uncorrected estimators remain valid for the bias corrections.

We find that the bias corrections change the estimates by an amount that can be significant relative to the corresponding standard error. The analytical standard errors are smaller than the bootstrap standard errors for the split-sample bias corrections. These differences might indicate that the analytical standard errors miss the additional sampling error introduced by the estimation in smaller panels. The analytical correction produces more precise estimates than the split-sample correction. Focusing on the effect of democracy, the corrected estimators find that a permanent transition to democracy increases economic growth by about 2-5% in the first year, and about 25-26% in the long run. Interestingly, the fixed effect and GMM approaches produce very similar estimates of the long run effect after the correction, which provides further support to the robustness of the results of I to the econometric methodology and data selection.

[5]The J-test statistic based on the twos-step Arellano-Bond estimator is 130.23, which corresponds to a p-value of 1.00 under a chi-square distribution with 463 degrees of freedom.

REFERENCES

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo, and James A. Robinson. forthcoming. "Democracy Does Cause Growth." *Journal of Political Economy.*

Arellano, Manuel, and Stephen Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *The Review of Economic Studies*, 58(2): 277–297.

Chudik, Alexander, M. Hashem Pesaran, and Jui-Chung Yang. 2018. "Half-panel jackknife fixed-effects estimation of linear panels with weakly exogenous regressors." *Journal of Applied Econometrics*, 33(6): 816–836.

Dhaene, Geert, and Koen Jochmans. 2015. "Split-panel Jackknife Estimation of Fixed-effect Models." *The Review of Economic Studies*, 82(3): 991–1030.

Fernández-Val, Iván, and Martin Weidner. 2016. "Individual and time effects in nonlinear panel models with large N, T." *Journal of Econometrics*, 192(1): 291–312.

Table 2—Effect of Democracy on Economic Growth

	Fixed Effects			GMM		
	FE	ABC-M4	SBC	AB	SBC1	SBC5
Democracy	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]
L1.log(gdp)	1.15	1.23	1.30	1.00	0.98	1.03
	(0.05)			(0.06)		
	[0.05]	[0.05]	[0.08]	[0.06]	[0.07]	[0.08]
L2.log(gdp)	-0.12	-0.14	-0.13	-0.06	-0.05	-0.07
	(0.06)			(0.06)		
	[0.05]	[0.05]	[0.08]	[0.06]	[0.07]	[0.07]
L3.log(gdp)	-0.07	-0.09	-0.13	-0.04	-0.04	-0.06
	(0.04)			(0.04)		
	[0.04]	[0.04]	[0.06]	[0.04]	[0.04]	[0.04]
L4.log(gdp)	-0.08	-0.08	-0.08	-0.08	-0.08	-0.08
	(0.02)			(0.03)		
	[0.02]	[0.03]	[0.04]	[0.03]	[0.03]	[0.03]
Long-run	16.05	25.91	25.69	20.97	26.46	25.24
(democracy)	(6.67)			(9.51)		
(×100)	[6.63]	[9.31]	[12.12]	[9.38]	[10.72]	[11.29]

Note: All the specifications include country and year effects. Clustered standard errors at the country level in parentheses. Bootstrap standard errors in brackets based on 500 replications.

Fernández-Val, Iván, and Martin Weidner. 2018. “Fixed Effects Estimation of Large-TPanel Data Models.” *Annual Review of Economics*, 10(1): 109–138.

Hahn, Jinyong, and Guido Kuersteiner. 2011. “Bias reduction for dynamic nonlinear panel models with fixed effects.” *Econometric Theory*, 27(06): 1152–1191.

Hahn, Jinyong, and Guido M. Kuersteiner. 2002. “Asymptotically Unbiased Inference for a Dynamic Panel Model with Fixed Effects When Both N and T Are Large.” *Econometrica*, 70(4): 1639–1657.

Hahn, Jinyong, and Whitney Newey. 2004. “Jackknife and Analytical Bias Reduction for Nonlinear Panel Models.” *Econometrica*, 72(4): 1295–1319.

Kiviet, Jan F. 1995. “On bias, inconsistency, and efficiency of various estimators in dynamic panel data models.” *J. Econometrics*, 68(1): 53–78.

Newey, Whitney K., and Frank Windmeijer. 2009. “Generalized Method of Moments With Many Weak Moment Conditions.” *Econometrica*, 77(3): 687–719.

Neyman, J., and E. L. Scott. 1948. “Consistent Estimates Based on Partially Consistent Observations.” *Econometrica*, 16(1): 1–32.

Nickell, Stephen. 1981. “Biases in Dynamic Models with Fixed Effects.” *Econometrica*, 49(6): 1417–1426.

Appendix

The online supplemental Appendix contains the data and code in `R` and `Stata` for the empirical application.