A Appendix for Acemoglu and Molina (Not For Publication)

A.1 Proof of Proposition 1

Let $c = \{1, 2, ..., C\}$ designate a country and C be the total number of countries. Suppose that the main outcome of interest, y_{ct}, follows the AR(1) process

$$\Delta y_{ct} = \alpha_c + \rho_c y_{ct-1} + \varepsilon_{ct},$$

so that the joint data generating process for all countries follows equation (1). Let $T_c = \{t | y_{c,t} \in \mathbb{R} \wedge y_{c,t-1} \in \mathbb{R}\}$. Let us first consider the OLS estimate of country-specific persistent parameter $\hat{\rho}_c$ for country c (with country-specific intercepts). This can be written as:

$$\hat{\rho}_c = \frac{\sum_{t \in T_c} (\Delta y_c - \overline{\Delta y}_c)(y_{c,t-1} - \overline{y}_{c,-})}{\sum_{t \in T_c} (y_{c,t-1} - \overline{y}_{c,-})^2}, \tag{5}$$

where $\overline{\Delta y}_c = \frac{\sum_{t \in T_c} \Delta y_{c,t}}{T_c}$ and $\overline{y}_{c,-} = \frac{\sum_{t \in T_c} y_{c,t-1}}{T_c}$. Because the DGP is given by (1), we have $plim \hat{\rho}_c = \rho_c$.

Now consider the OLS estimate of ρ, $\hat{\rho}$, in the cross-country equation of unconditional convergence without country-specific intercepts (country fixed effects) as in equation (2):

$$\hat{\rho} = \frac{\sum_{c \in C} \sum_{t \in T_c} (\Delta y_{c,t} - \overline{\Delta y})(y_{c,t-1} - \overline{y}_{-})}{\sum_{c \in C} \sum_{t \in T_c} (y_{c,t-1} - \overline{y}_{-})^2}, \tag{6}$$

where $\overline{\Delta y} = \frac{\sum_{c \in C} \sum_{t \in T_c} \Delta y_{c,t}}{\sum_{c \in C} T_c}$ and $\overline{y}_{-} = \frac{\sum_{c \in C} \sum_{t \in T_c} y_{c,t-1}}{\sum_{c \in C} T_c}$.

We first prove part 2 of the proposition, thus establishing that

$$\hat{\rho} = \sum_{c=1}^{C} \theta_c \hat{\rho}_c + \theta_0 k. \tag{7}$$

To do so, we expand equation (6), and write the right-hand side, in terms of the $\hat{\rho}_c$'s from equation (5). The numerator of this expression can be written as

$$\begin{aligned}
\sum_{c \in C} \sum_{t \in T_c} (\Delta y_{c,t} - \overline{\Delta y})(y_{c,t-1} - \overline{y}_{-}) &= \sum_{c \in C} \sum_{t \in T_c} \Delta y_{c,t} (y_{c,t-1} - \overline{y}_{-}) \\
&= \sum_{c \in C} \sum_{t \in T_c} \Delta y_{c,t} (y_{c,t-1} - \overline{y}_{c,-}) + \sum_{c \in C} \sum_{t \in T_c} \overline{\Delta y}_c (\overline{y}_{c,-} - \overline{y}_{-}) \\
&= \sum_{c \in C} \sum_{t \in T_c} \Delta y_{c,t} (y_{c,t-1} - \overline{y}_{c,-}) + \sum_{c \in C} T_c \overline{\Delta y}_c (\overline{y}_{c,-} - \overline{y}_{-}),
\end{aligned}$$

where the first equality uses the fact that $\sum_{t \in T_c} (y_{c,t-1} - \overline{y}_{-}) = 0$, while the second equality rearranges terms. The final equality follows from the fact that $\overline{\Delta y}_c$, $\overline{y}_{c,-}$, and $\overline{y}_-$ do not vary with t. Analogously, we can rewrite the denominator as

$$\sum_{c\in C}\sum_{t\in T_c}(y_{c,t-1}-\overline{y}_-)^2 = \sum_{c\in C}\sum_{t\in T_c}(y_{c,t-1}-\overline{y}_{c,-})^2 + \sum_{c\in C} T_c(\overline{y}_{c,t-1}-\overline{y}_-)^2.$$

Now, letting $\sigma_c = \sum_{t\in T_c}(y_{ct-1}-\overline{y}_{c-})^2$, equation (6) can be written as

$$\begin{aligned}\hat{\rho} =& \frac{\sum_{c\in C}\sigma_c\hat{\rho}_c + \sum_{c\in C} T_c\overline{\Delta y}_c(\overline{y}_{c,-}-\overline{y}_-)}{\sum_{c\in T_c}\sigma_c + \sum_{c\in C} T_c(\overline{y}_{c,t-1}-\overline{y}_{t-1})^2}\\ =& \sum_{c=1}^{C}\theta_c\hat{\rho}_c + \theta_0 k,\end{aligned}$$

whereon

$$\begin{aligned}\theta_c &= \frac{\sigma_c}{\sum_{c\in C}\sigma_c + \sum_{c\in C} T_c(\overline{y}_{c,-}-\overline{y}_-)^2}\ \forall c = \{1,...,C\}\\ \theta_0 &= 1-\sum_{c=1}^{C}\theta_c \\ k &= \frac{\sum_{c\in C} T_c\overline{\Delta y}_c(\overline{y}_{c,-}-\overline{y}_-)}{\sum_{c\in C} T_c(\overline{y}_{c,-}-\overline{y}_-)^2}.\end{aligned} \quad (8)$$

Since $plim\hat{\rho}_c = \rho_c$, we can substitute ρ_c for $\hat{\rho}_c$ in the limit, which gives us the desired result. Note also that $\sigma_c \geq 0$ and thus $\theta_c \in [0,1]$.

To prove part 1 of the proposition, assume $\hat{\rho}_c \neq 0\ \forall c$, and then substitute

$$\theta_c = \omega_c - \frac{\theta_0 k}{C\hat{\rho}_c} \quad (9)$$

in equation (7) and simplify terms. It is straightforward to see that ω_c need not be between 0 and 1, and $\sum_{c=1}^{C}\omega_c$ is not in general equal to one. Our empirical results in Section 3.2 provide one instance in which ω_c is not between 0 and 1 for a significant fraction of countries in the sample, and their sum is very different from 1.

Finally, we establish part 3 by proving that, for any c, $\lim_{T_c\to\infty}\sigma_c \in \mathbb{R}$. Let us drop the country subscript c, since we will work with the equation for a single country: $\Delta y_{t-1} = \alpha + \rho y_{t-2} + \varepsilon_{t-1}$, or

$$y_{t-1} = \alpha + \gamma y_{t-2} + \varepsilon_{t-1},$$

where $\gamma = \rho + 1$. Substituting successively, the solution to this equation takes the form

$$y_{t-1} = a + b\gamma^{t-1} + \varepsilon_1\gamma^{t-2} + \varepsilon_2\gamma^{t-3} + ... + \varepsilon_{t-1}, \quad (10)$$

where $b = y_0 - a$ and $a = \frac{\alpha}{1-\gamma}$ (a represents the steady state and b is the distance of the initial value to the steady state). Then

$$\sigma = \sum_{t \in T} (y_{t-1} - \overline{y}_{t-1})^2$$
$$\sigma = \sum_{t \in T} y_{t-1}^2 - \frac{1}{T} \left(\sum_{t=1}^{T} y_{t-1} \right)^2$$
$$\sigma = a^2 T + O_p(1) - [a^2 T + O_p(1)]$$
$$\sigma = O_p(1),$$

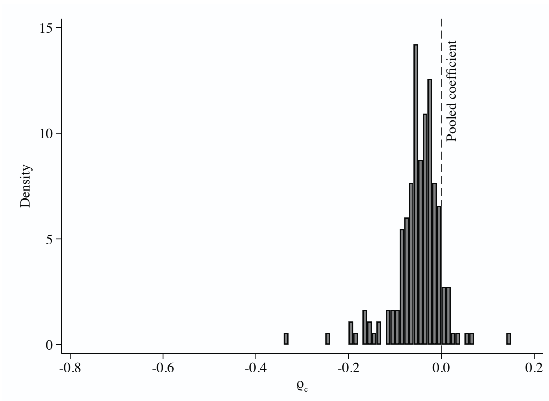
where the second equality follows by rearranging terms. The third equality follows from substituting for equation (10) and using the hypothesis that $-2 < \rho_c < 0$, or equivalently $|\gamma| < 1$, which implies that y_{t-1} converges, and thus both $\sum_{t \in T} y_{t-1}^2$ and $\frac{1}{T} \left(\sum_{t=1}^{T} y_{t-1} \right)^2$ are $O_p(T)$. The fourth equality simply notes that their difference is $O_p(1)$, which also implies that the limit of σ is finite.

The final step is to characterize $plim\theta_c$. The result above implies that if $\underline{T} = \min_c\{T_c\} \to \infty$, then σ_c is finite for any c. Additionally, both $\overline{y}_{c,-}$ and $\overline{y}_{-}$ converge to limit values (since each country converges to its steady state, again my hypothesis). However, because of the α_c terms, these values are country-specific. Consequently, we have that $\lim_{\underline{T} \to \infty} \theta_c = 0 \ \forall c \in \{1, 2, ..., C\}$ and $\theta_0 = 1$. Using an analogous argument, we also have $\lim_{\underline{T} \to \infty} \overline{\Delta y}_c = 0$ and thus $\lim_{\underline{T} \to \infty} k = 0$. Finally, combining these results with equations (9) and (6), we can conclude that $\lim_{\underline{T} \to \infty} \omega_c = 0 \ \forall c \in C$ and $\lim_{\underline{T} \to \infty} \hat{\rho} = 0$.

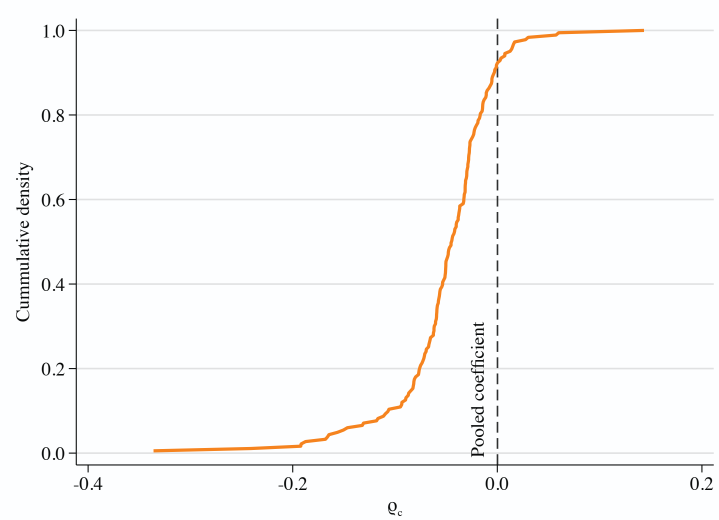
A.2 Replication of Figures and Tables using Penn World Tables

Figure A-1: Empirical distribution of the underlying ρ_c's and ω_c's

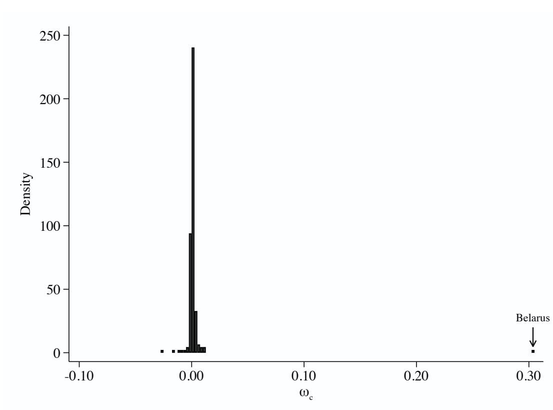
Panel A: Probability density of the ρ_c's



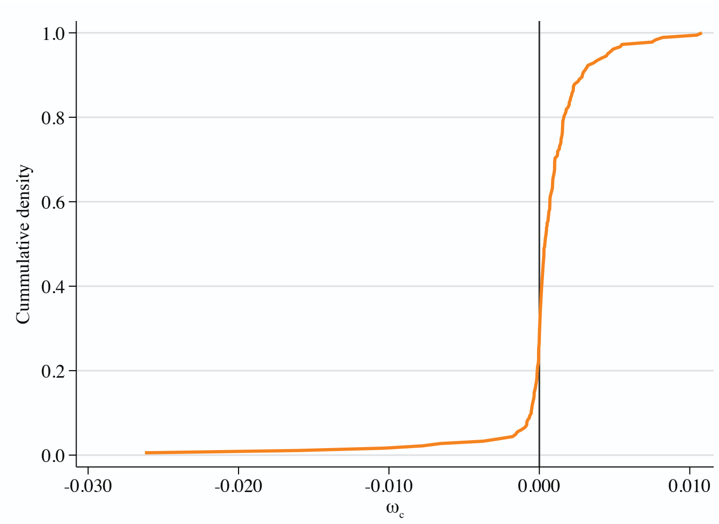
Panel B: Cumulative density of the ρ_c's



Panel C: Probability density of the ω_c's

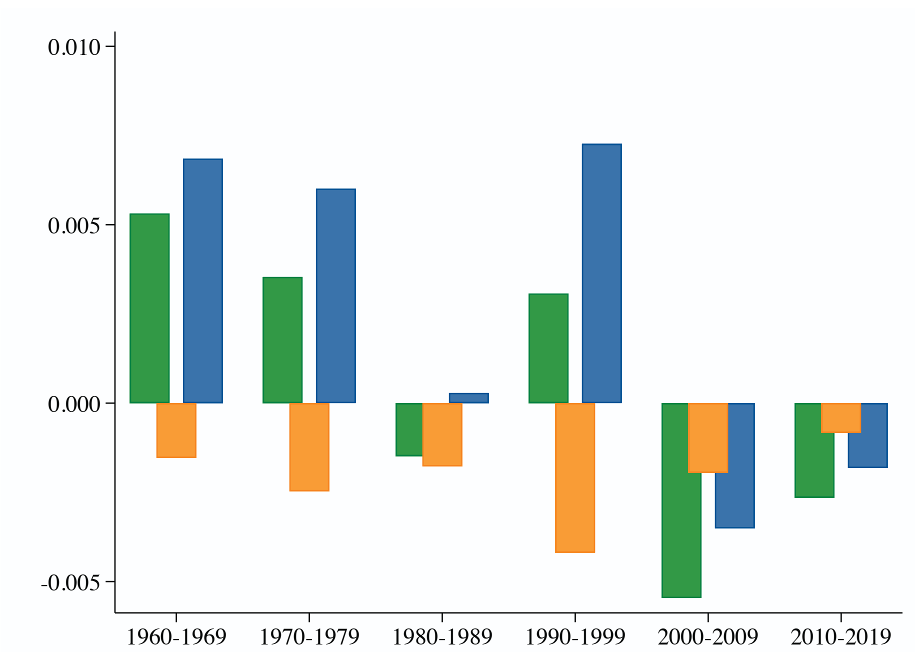


Panel D: Cumulative density of the ω_c's



Notes: The figure presents estimates of the distribution of the country-specific coefficients of convergence (the ρ_c's), and the weights (the ω_c's, defined in Proposition 1). The empirical probability density and the cumulative density are reported on the left and the right-hand side panels respectively. In Panel A, the dashed line indicates the estimate $\hat{\rho}$ from equation (2) (when no country heterogeneity is allowed).

Figure A-2: Decomposition of $\hat{\rho}$ across decades

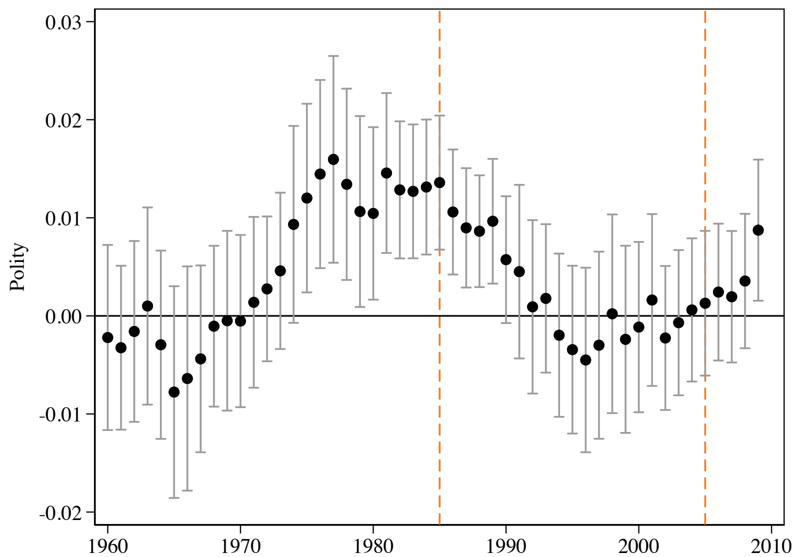


Notes: The figure reports the estimate $\hat{\rho}$ from equation (2) (when no country heterogeneity is allowed) by decade shown by the blue bar. This estimate is decomposed in two terms (see Proposition 1): the underlying distribution of the ρ_c's shown by the orange bar and the bias shown by the green bars.

Figure A-3: Estimates of the relationship of Polity 2 and economic growth over time

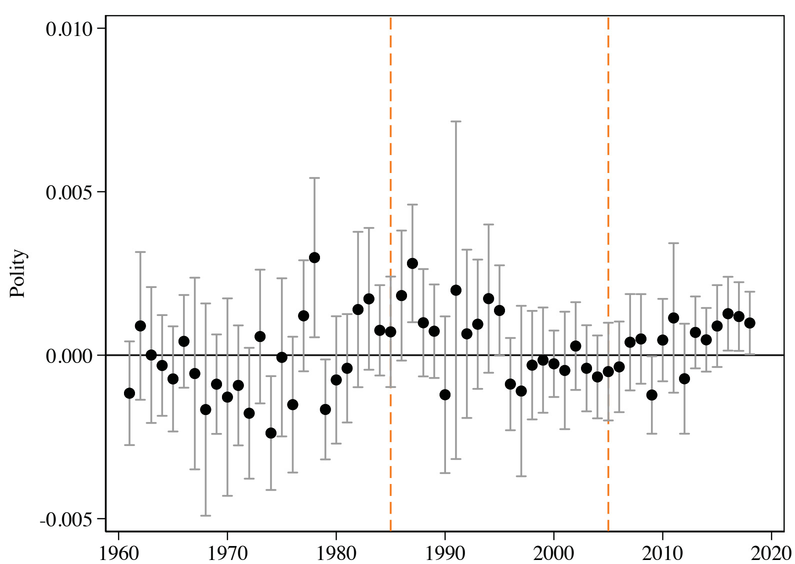
Panel A: Dependent variable is 10-year changes

$$\Delta_{10} y_{ct} = \beta_t x_{ct} + \rho_t y_{ct} + \delta_t + \varepsilon_{ct}$$



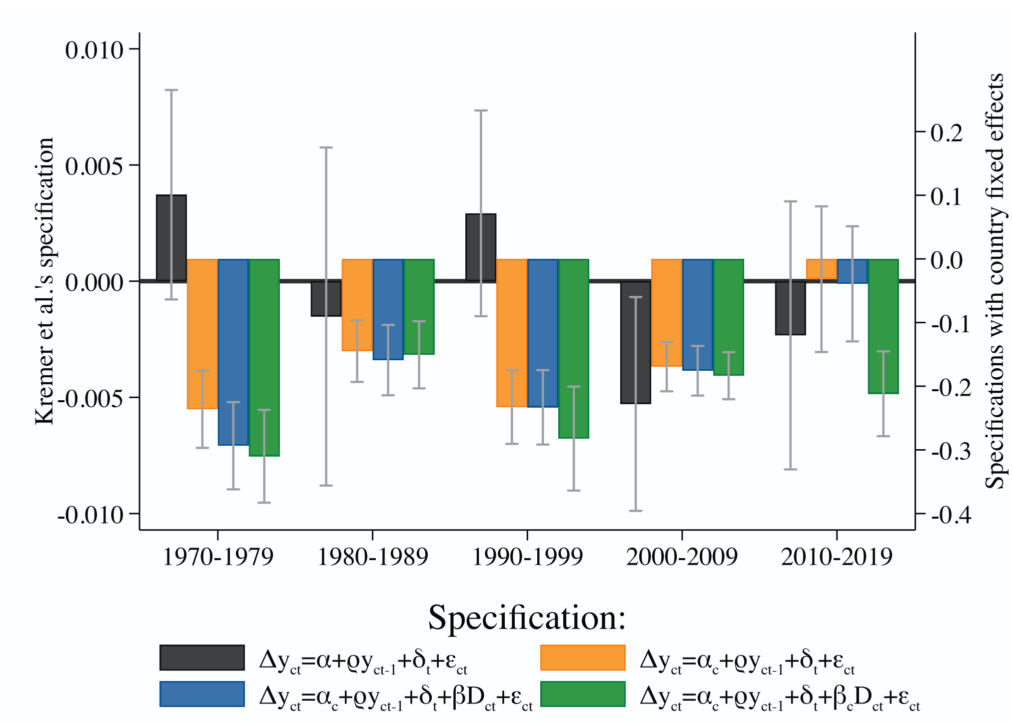
Panel B: Dependent variable is first difference

$$\Delta y_{ct} = \beta_t x_{ct} + \rho_t y_{ct} + \delta_t + \varepsilon_{ct}$$



Notes: The figure reports estimates of the relationship between the Polity 2 score and economic growth and 95% confidence intervals over time, Namely, we plot the coefficient β_t in the equation indicated in the panel labels. Panel A uses as dependent variable the 10-year change in log of GDP per capita, while Panel B uses annual changes (first differences). Standard errors are clustered at the country level.

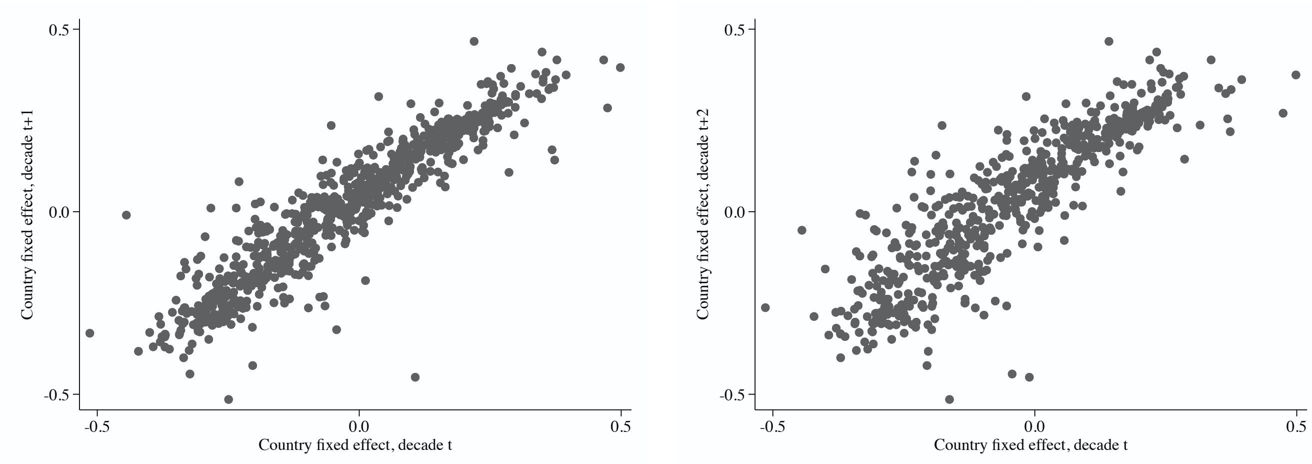
Figure A-4: Estimates of ρ across different specification and decades



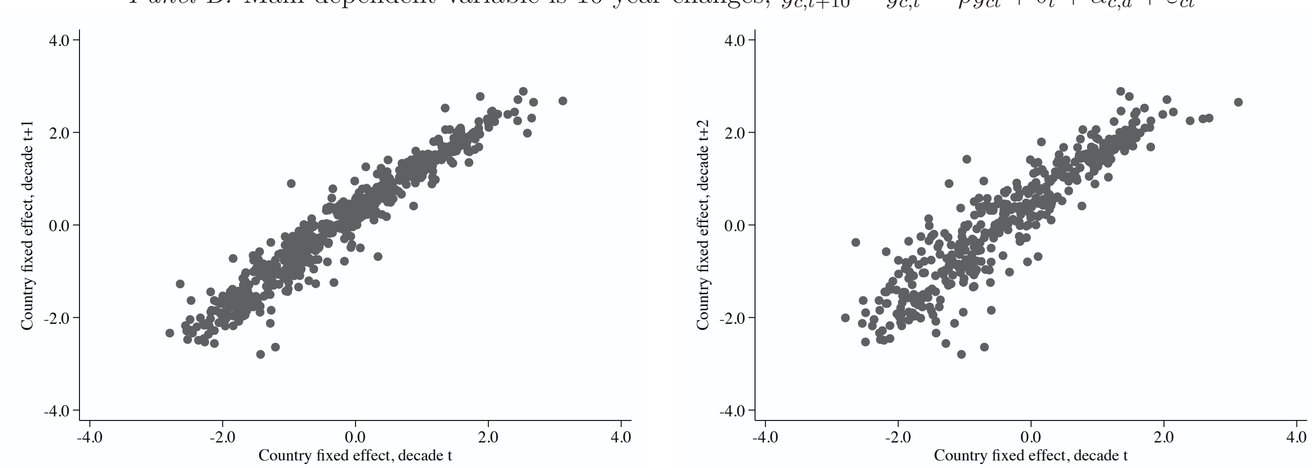
Notes: The figure report estimates of the coefficient of convergence (ρ) as well as 95% confidence intervals across four different specifications. The first is Kremer et al.'s specification of unconditional convergence, which does not include country fixed effects (shown by the black bars). The second is a specification that adds country fixed effects to this baseline (shown by the orange bars). The third adds our dichotomous measure of democracy as a control, focusing on the OLS specification (shown by blue bars). Our final specification allows the effects of democracy to be varying across countries green bars). Standard errors are clustered at the country level.

Figure A-5: Correlation of country fixed effects over decades

Panel A: Main dependent is first difference, $\Delta y_{ct} = \rho y_{ct-1} + \delta_t + \alpha_{c,d} + \varepsilon_{ct}$



Panel B: Main dependent variable is 10-year changes, $y_{c,t+10} - y_{c,t} = \rho y_{ct} + \delta_t + \alpha_{c,d} + \varepsilon_{ct}$



Notes: We estimate the regression specified in the panel label which allows the country fixed effects to vary across decades and then plot these decadal fixed effect estimates against each other. Panels A and B use as dependent variable the first difference (the 10-year change) in log of GDP per capita, while Panels C and D use 10-year changes. The panels to the left (right) compare shows the correlation between the estimated country fixed effect for a decade starting at year t with a decade starting at year t+10 (t+20).

Table A-1: Convergence patterns for the key covariates

	(1)	(2)	(3)	(4)	(5)
	Pooled	Country-level coefficients			
	coefficient	Mean	25th quantile	50th quantile	75th quantile
Log GDP per capita	0.000	-0.050	-0.067	-0.044	-0.023
Polity 2	-0.029	-0.107	-0.165	-0.087	0.000
Rule of law	-0.008	-0.346	-0.487	-0.295	-0.169
Property rights	-0.032	-0.238	-0.360	-0.228	-0.133
Government expenditure	-0.055	-0.215	-0.313	-0.168	-0.069
Credit	0.004	-0.221	-0.403	-0.114	-0.013
Years of schooling	-0.020	-0.506	-0.653	-0.440	-0.321

Notes: The table reports estimates of coefficient of unconditional convergence (ρ, see equation (2)) in column 1. The remaining columns report moments (including the mean and the 25th, 50th and 75th quantiles) of the distribution of the underlying country-specific estimates of convergence (ρ_c's, see equation (1)). We show results for the key variables used by Kremer et al. including (source in parenthesis): Polity 2 (Polity IV Project), Rule of law (Worldwide Governance Indicators), Property rights (Heritage Freedom), Government expenditure (World Development Indicators) and Years of schooling (Barro-Lee).

Table A-2: Estimates of the Effect of democracy on (log) GDP per capita

	(1)	(2)	(3)	(4)
	Dependent variable is Log GDP per capita			
	Estimator...			
	Within	Arellano-Bond	HHK	IV
A. Effect of Democracy on Log GDP				
Democracy	0.009	0.014	0.005	0.005
	(0.004)	(0.006)	(0.007)	(0.010)
Observations	5,961	5,796	5,961	5,937
B. Effect of Democracy on Log GDP and its change over time				
Democracy	0.009	0.014	0.004	0.006
	(0.004)	(0.006)	(0.006)	(0.011)
Democracy × Trend	0.000	0.004	0.001	0.002
	(0.003)	(0.005)	(0.004)	(0.005)
Observations	5,961	5,796	5,961	5,937

Notes: This table presents estimates of the effect of democracy on log GDP per capita following Acemoglu et al. (2019). Democracy is measured as a dichotomous variable to minimize measurement error. Panel A replicates the results in Acemoglu et al. (2019), and Panel B extends the regression by allowing an interaction between the measure of democracy and a linear function of time. Column 1, 2, 3 and 4 present results from the within estimator, the Arellano and Bond's (1991) GMM estimator, the HHK (Hahn et al., 2001) estimator, and an IV (exploiting regional waves of democratization) respectively. All regressions include four lags of log GDP per capita. Standard errors are clustered at the country level.