"Eppur si muove": Convergence Within and Between Countries
Online Appendix

Giovanni Lombardo
BIS

October 22, 2026

A The correlated-random-effects machinery

A.1 The Mundlak device

KWY, AM and PS denote Kremer et al. (2022), Acemoglu and Molina (2022) and Pesaran and Smith (2026), as in the main text. Consider the static panel $y_{it} = x'_{it}\beta + a_i + \varepsilon_{it}$ with $E[\varepsilon_{it} \mid x_i, a_i] = 0$. Random-effects estimation requires a_i independent of the regressors; fixed-effects estimation drops the assumption by demeaning, at the cost of discarding all between variation and estimating N intercepts. Mundlak (1978) replaces the independence assumption with a linear projection of the country effect on the within-country means,

$$a_i \;=\; \bar{x}'_i\gamma + u_i, \qquad E[u_i \mid x_i] = 0, \tag{A.1}$$

so that the augmented model $y_{it} = x'_{it}\beta + \bar{x}'_i\gamma + u_i + \varepsilon_{it}$ satisfies the random-effects orthogonality by construction. Generalized least squares on the augmented model returns the within estimator for β exactly in the balanced case, whatever the variance-ratio weighting; the between regression of $\bar{y}_i$ on $\bar{x}_i$ has slope $\beta + \gamma$; and, with strictly exogenous regressors, the test of $\gamma = 0$ is the regression form of the specification test of Hausman (1978). That equivalence does not carry to the dynamic specification used here. The country mean contains the future income that defines growth, so the corresponding likelihood-ratio test rejects whether or not steady states differ; Section 4.2 reports what it can and cannot establish. Three properties make the device the natural vehicle here: it estimates two parameters (γ, σ_u) instead of N intercepts; it keeps the between margin inside the model rather than absorbing it, so within and between slopes come from one likelihood with a joint covariance; and it transfers unchanged to maximum-likelihood and Bayesian settings where a dummy-variable fixed-effects estimator is awkward or inconsistent.

The specification is a reparameterization of the within–between (or "hybrid") form used in the multilevel literature (Bell and Jones, 2015; McNeish and Kelley, 2019), in which the level-1 predictor enters group-mean centered and the group mean enters as a level-2 predictor: since $\beta x_{it} + \gamma \bar{x}_i = \beta(x_{it} - \bar{x}_i) + (\beta + \gamma)\bar{x}_i$, the two fixed parts span the same column space, the within–between coefficients are $(\beta_W, \beta_B) = (\beta, \beta + \gamma)$, and the test of $\gamma = 0$ is the test of $\beta_W = \beta_B$. Estimates, fit, and tests are identical; the forms differ only in which quantity is a raw parameter — the between slope in the within–between form, the within–between contrast in the Mundlak form — in numerical conditioning, where the centered form's orthogonal regressors are friendlier, and in models with random slopes, where a random coefficient on x_{it} and one on $x_{it} - \bar{x}_i$ imply different covariance structures. The models we estimate have no random slopes in the baseline, so the choice is one of parameterization only; the random-slope extension is reported in Section A.6.

A.2 The Wooldridge initial condition

The convergence equation is dynamic: the regressor $\log y_{i,t}$ is (a transform of) the lagged dependent variable, so the country effect a_i, which enters every realization of income, is correlated with the entire observed income path — including its first observation — by construction. Conditioning the likelihood on the initial observation without modeling that correlation misspecifies the random-effects distribution. The device of Wooldridge (2005) specifies the country effect conditionally on the initial observation and on the exogenous covariates. The model here has no exogenous covariates; it projects instead on the within-country means of its own time-varying regressors, which are not exogenous (Section 3.2):

$$a_i \;=\; \mathbf{m}_i'\boldsymbol{\omega} + \delta \log y_{i,0_i} + u_i, \qquad u_i \mid (\mathbf{m}_i, \log y_{i,0_i}) \sim N(0, \sigma_u^2), \tag{A.2}$$

where $\mathbf{m}_i$ collects the means, over country i's observed windows, of the time-varying regressors of Equation 4: the six products $\mathbf{1}\{t \in d\} \log y_{i,t}$, the six products $\mathbf{1}\{t \in d\} \overline{\log y}_i$, and the start-year dummies. The initial-condition term conditions the likelihood on the first observed income, as in Wooldridge's original form, which assumes strictly exogenous covariates; the correlation of the mean terms with the disturbance is the subject of Sections 3.2 and A.7. In our unbalanced panel $y_{i,0_i}$ is the first income observed for country i. The panel's four entry cohorts, entering in 1963–73, 1973–83, 1993–2003 and 2013–23, have distinct patterns of observed windows, so the means of the start-year dummies span the cohort indicators exactly and absorb any level shift common to a cohort. The projection is jointly significant ($\chi^2(10) = 323$, Section 4.2). The test rejects even under a single common steady state, so the statistic measures the mechanical correlation of the means with the disturbance together with any economic one.

One result from the subsequent literature concerns the device. Wooldridge (2019) extends the Mundlak–fixed-effects equivalence to unbalanced panels with means over observed periods — the construction used here, under which the within slopes of the correlated-random-effects model equal the fixed-effects slopes exactly (to 10^{-12} in our estimates) — while cautioning that the projection coefficients may vary with panel length and that the device presumes entry

into the panel ignorable given the heterogeneity; both cautions are taken up in Section A.6.

A.3 The pooled estimator: attenuation by steady-state heterogeneity

The direction of the pooled estimator's bias follows from the omitted-variable formula once the regressor's dependence on the omitted effect is made explicit. With the level dynamics $\log y_{i,t+1} = \rho \log y_{i,t} + a_i + \zeta_{i,t}$ and $\rho < 1$, income decomposes at stationarity into a steady-state component and a transitory deviation, $\log y_{i,t} = m_i + w_{i,t}$ with $m_i = a_i/(1-\rho)$, so $\operatorname{cov}(a_i, \log y_{i,t}) = (1-\rho)\,\sigma_m^2 > 0$ mechanically: countries with a high permanent component are permanently rich. The omitted-variable term $\operatorname{cov}(a_i, \log y_{i,t})/\sigma^2_{\log y}$ is therefore positive while the convergence parameter is negative, and substituting it out gives the pooled probability limit in attenuation form,

$$\hat{\beta}_{\text{pooled}} \;=\; \beta \cdot \frac{\sigma_w^2}{\sigma_m^2 + \sigma_w^2}, \tag{A.3}$$

because at stationarity growth depends only on the deviation $w_{i,t}$: the steady-state component is regressor variation orthogonal to the outcome, exactly as in classical measurement-error attenuation. The shrink factor lies strictly in $(0, 1)$, so in this benchmark the pooled slope is compressed toward zero but never sign-flipped; the limiting case of Proposition 1(3) in AM is the between-dominated limit of the same expression.

The benchmark is quantitatively tight in our panel. The between-country variance of log income is 1.23 against a within-country variance of 0.24, an attenuation factor of 0.16; applied to the mean within slope of -2.6 it predicts a pooled slope of -0.41, which the estimated pooled decade slopes (-0.72 to 0.50) straddle. The deviations from the benchmark are informative in themselves: out of stationarity $\operatorname{cov}(a_i, w_{i,t}) \neq 0$ — countries sitting further below high steady states grow faster, pushing the pooled slope positive, plausibly part of the 1960s value — and the between margin need not be flat, in which case the pooled coefficient is a variance-weighted average of the within slope and a nonzero, moving between slope rather than a pure attenuation. Both departures are estimated, not assumed, in the correlated-random-effects model: the projection coefficients measure the net projection of the country effect on the income history, and β_d^B traces the between slope directly.

A.4 The Nickell bias in this setting

Because the Mundlak-augmented estimator reproduces the within estimator, it inherits the finite-T dynamic-panel bias of Nickell (1981): the Mundlak projection conditions on the same within-country mean that the within transformation subtracts, so the equivalence of Section A.1 transfers the bias unchanged.

Write the decade-frequency autoregression around country effects as

$$y_{i,t} \;=\; \rho\, y_{i,t-1} + a_i + \zeta_{i,t}, \qquad t = 1, \ldots, T, \tag{A.4}$$

with $y_{i,t}$ log income, $0 < \rho < 1$, $\zeta_{i,t}$ serially uncorrelated with variance σ_ζ^2, a stationary initial condition, and T observed transitions per country. The within estimator is least squares on demeaned data: with $\tilde{y}_{i,t} = y_{i,t} - \bar{y}_i$, $\tilde{y}_{i,t-1} = y_{i,t-1} - \bar{y}_{i,-1}$ where $\bar{y}_{i,-1} = T^{-1}\sum_{s=1}^{T} y_{i,s-1}$, and $\tilde{\zeta}_{i,t} = \zeta_{i,t} - \bar{\zeta}_i$, the country effect drops out and the demeaned model is $\tilde{y}_{i,t} = \rho\,\tilde{y}_{i,t-1} + \tilde{\zeta}_{i,t}$. Substituting it into the least-squares slope gives

$$\hat{\rho}_{\mathrm{FE}} \;=\; \frac{\sum_{i,t}\tilde{y}_{i,t-1}\,\tilde{y}_{i,t}}{\sum_{i,t}\tilde{y}_{i,t-1}^2} \;=\; \rho \;+\; \frac{\sum_{i,t}\tilde{y}_{i,t-1}\,\tilde{\zeta}_{i,t}}{\sum_{i,t}\tilde{y}_{i,t-1}^2}\,: \tag{A.5}$$

the estimation error is the sample covariance between the demeaned lag and the demeaned error, over the demeaned lag's variance. Countries are independent draws, so with T fixed and $N \to \infty$ both averages converge to their expectations,

$$\operatorname*{plim}_{N\to\infty}\left(\hat{\rho}_{\mathrm{FE}} - \rho\right) \;=\; \frac{T^{-1}\sum_t E\left[\tilde{y}_{i,t-1}\,\tilde{\zeta}_{i,t}\right]}{T^{-1}\sum_t E\left[\tilde{y}_{i,t-1}^2\right]}, \tag{A.6}$$

and the estimator is inconsistent exactly if the numerator is nonzero. It is. In levels the lag is predetermined, $E[y_{i,t-1}\,\zeta_{i,t}] = 0$, but demeaning entangles the two through the country means: $\bar{\zeta}_i$ contains the shocks dated up to $t-1$ that are embedded in $y_{i,t-1}$, and $\bar{y}_{i,-1}$ contains the incomes dated t and later that embed $\zeta_{i,t}$; both cross terms are positive and enter with minus signs. Carrying out the sums under stationarity leaves the numerator in closed form:

$$\frac{1}{T}\sum_{t=1}^{T} E\left[(y_{i,t-1} - \bar{y}_{i,-1})(\zeta_{i,t} - \bar{\zeta}_i)\right] \;=\; -\frac{\sigma_\zeta^2}{T(1-\rho)}\,\phi_T(\rho), \qquad \phi_T(\rho) \;=\; 1 - \frac{1-\rho^T}{T(1-\rho)} \;\in\; (0,1), \tag{A.7}$$

a strictly negative covariance of order $1/T$. Evaluating the denominator under the same stationary moments and dividing gives the exact inconsistency of Nickell (1981),

$$\operatorname*{plim}_{N\to\infty}\left(\hat{\rho}_{\mathrm{FE}} - \rho\right) \;=\; -\frac{1+\rho}{T-1}\,\phi_T(\rho) \,\Big/\, \left[1 - \frac{2\rho}{(1-\rho)(T-1)}\,\phi_T(\rho)\right], \tag{A.8}$$

whose leading term as T grows ($\phi_T \to 1$, bracket $\to 1$) is the familiar $-(1+\rho)/(T-1)$.

The sign and the size have an accounting interpretation. Demeaned deviations sum to zero within each country by construction: the country mean is estimated from the same handful of observations it is then subtracted from, so a spell of good shocks drags the estimated mean up and mechanically pushes the remaining observations below it. Time above the own-sample mean must be repaid with time below it — which is what mean reversion looks like. Any series demeaned over a short window acquires this appearance, a pure random walk included; the estimator books the artifact as convergence. The polar case $T = 2$ shows the size: there the exact bias reduces to $-(1+\rho)/2$, so a true persistence of 0.65 is estimated as oscillation ($\operatorname{plim}\hat{\rho}_{\mathrm{FE}} = -0.18$); six transitions are better only by degree.

Three features of the formula govern how the within estimates are used. First, the bias is

negative throughout the stationary region: estimated persistence is too low. In the growth parameterization the mapping is exact — $g_{i,t} = \beta_d \log y_{i,t} + \dots$ is the decade-frequency autoregression $\log y_{i,t+10} = (1 + \beta_d/10) \log y_{i,t} + \dots$ under an affine transformation of the dependent variable, so $\hat{\beta}_d = 10(\hat{\rho}_{10} - 1)$ and the slope bias is ten times the persistence bias — a within slope too negative, that is, spurious evidence of convergence. Second, the bias does not diminish as countries are added: with T fixed, every additional country replicates the same short-panel covariance rather than averaging it away. What does shrink it is a longer history — but, as Equation A.9 below shows, longer in calendar time rather than in number of observations. Third, it presumes no misspecification: the formula is a property of the within estimator under the model itself, and (as Section A.10 notes) of the posterior built on the same likelihood.

A benchmark for our decade panel: the 2003–13 within slope of -3.32 implies $\hat{\rho}_{10} = 0.67$, and at that persistence the exact formula gives a bias of -3.2 to -6.1 in the units of β_d for histories of six down to three transitions, the same order as the estimate itself. The benchmark is indicative, since the panel is unbalanced and the model includes time effects, but it settles how the within results are read. The level of β_d is not a convergence speed, and its profile is descriptive, because the bias varies with T and with composition as countries enter. The cross-country conclusions rest on the between margin β_d^B, whose profile the correlation of the country mean with the disturbance moves by far less than the observed decline (Section A.7).

Bias corrections exist: analytical (Kiviet, 1995), the split-panel jackknife (Dhaene and Jochmans, 2015), and difference and system GMM (Arellano and Bond, 1991; Blundell and Bond, 1998). The democracy literature AM draw on uses difference GMM and a bias-corrected instrumental-variables estimator at annual frequency (Acemoglu et al., 2019). Two likelihood-based estimators are also consistent at fixed T: the transformed likelihood of Hsiao et al. (2002), which needs distributional assumptions on the initial conditions, and the orthogonalized-effects posterior of Lancaster (2002), constructed below. Section 4.3 takes them to the decade grid.

The size of this benchmark is governed by the span of the sample rather than by the number of observations in it, and refining the grid reduces it only toward a limit the span fixes. Let the annual autoregression have decay factor ρ_1, so that a regression run on observations h years apart estimates $\rho_h = \rho_1^h$ from $T = S/h$ transitions over a span of S years. The leading term of the bias is then

$$\hat{\rho}_h - \rho_h \;\simeq\; -\frac{1 + \rho_h}{T - 1} \;\simeq\; -\left(1 + \rho_1^h\right) \frac{h}{S}, \tag{A.9}$$

which vanishes as the grid is refined. But ρ_h is a property of the sampling interval, not of the process: the annual object is recovered by inversion, $\hat{\rho}_1 = \hat{\rho}_h^{1/h}$. Writing $\kappa = (1 + \rho_1^h)\, h/S$ for the bias and factoring ρ_1^h out,

$$\hat{\rho}_1 \;=\; \left(\rho_1^h - \kappa\right)^{1/h} \;=\; \rho_1 \left(1 - \kappa\, \rho_1^{-h}\right)^{1/h}, \qquad \log \hat{\rho}_1 \;=\; \log \rho_1 \;+\; \frac{1}{h} \log\left(1 - \kappa\, \rho_1^{-h}\right). \tag{A.10}$$

The argument of the logarithm departs from one by a term of order h, so $\log(1-x) = -x+O(x^2)$ applies and the factor $1/h$ cancels the h carried by κ:

$$\log \hat{\rho}_1 \;=\; \log \rho_1 \;-\; \frac{1+\rho_1^h}{S\,\rho_1^h} \;+\; O(h) \;\longrightarrow\; \log \rho_1 \;-\; \frac{2}{S}. \tag{A.11}$$

In terms of the annual convergence rate $r = -\log \rho_1$ — the object the classical literature places near two percent — this reads $\hat{r} = r + 2/S$: the estimated speed is too fast by $2/S$ a year at any sampling interval. Both are statements about r, not about β_d: the two are measured in the same units and are not the same quantity. A finer grid lowers the bias in $\hat{\rho}_h$ in proportion to h and moves ρ_h toward unity in the same proportion; annualizing divides by h, and the two cancel. Consistency in the annualized rate therefore requires the span to grow; adding observations within a fixed span does not deliver it.

Carrying the same substitution through the exact expression rather than its leading term replaces $2/S$ by $2\,\phi_\infty(c)/\{S\,[\,1-2\phi_\infty(c)/c\,]\}$, where $c = rS$ is the span measured in units of the decay time and $\phi_\infty(c) = 1-(1-e^{-c})/c$ is the limit of $\phi_T(\rho)$ along $\rho_h = \rho_1^h$, $T = S/h$. This is the expression of Barro (2015, eq. 1, p. 914), written for the bias in r rather than for its proportion of r. At the benchmark — $r = 4.0$ percent a year, a half-life of 17 years, so $c = 2.42$ over the sample's 60 years — that factor is 1.29 and the limit is 4.29 percentage points a year in the rate r, against 3.33 for the leading term. Finite intervals approach it from above, all in rate units: 6.56 at $h = 10$, the decade grid with $T = 6$; 5.10 at $h = 5$; 4.42 at annual frequency. Refining the grid over the same sixty years does reduce the bias in the annualized rate, but only toward a limit the span fixes, 4.29 in continuous time, so most of the bias remains. Two units are used. The slope β_d is linear in the persistence by construction, $\beta_d = 10(\rho_{10}-1)$, and biases expressed in slope units follow that convention — so does the Maddison contrast of Section 4.3. The annualized *rate* $r = -\log \rho_1$ is a log object, and every figure in this appendix compared with the $2/S$ limit is annualized that way. Applying the linear form to the rate understates the decade-grid bias by about half, which is the error the comparison invites. Only the span moves it materially, at rate $1/S$, again in rate units: 2.0 percent a year at $S = 120$, 0.9 at $S = 240$. The five-year grid of Section 4.3 is this statement at $h = 5$: it doubles T and leaves S unchanged. On the panel itself the five-year grid behaves as the formula predicts. Evaluated at the decade-grid fixed-effects slope of -2.94, that is at $\hat{\rho}_{10} = 0.71$ rather than at the benchmark above, the exact formula gives a bias of 5.15 on the five-year grid against 6.37 on the decade grid, both in rate units: doubling T lowers the bias by 19 percent. Meanwhile the estimated within slope steepens from -2.94 to -4.76 (s.e. 0.51), the signature of measurement-error mean reversion at shorter horizons and the consideration that leads KWY to ten-year windows.

One corollary of the exact expression is used in Section 4.3. Because the bias falls as the true rate rises while the reported rate $\hat{r} = r + \text{bias}(r, S)$ rises throughout, what the estimator reports is smallest when the truth is smallest. Taking $c = rS \to 0$ in the exact expression, $\phi_\infty(c) \to c/2 - c^2/6$ and the bias tends to $3/S - r/4$, so the reported rate $3/S + 3r/4$ tends to $3/S$ — not to the leading term's $2/S$, which describes the bias at a given r rather than

the floor across r.[1] This is a *minimum reportable speed* of $3/S$ a year in rate units: on a span of S years, a demeaned estimator applied to data generated by this model cannot report an annualized speed below that figure, whatever the true speed. At $S = 60$ that is 5.00 and at $S = 70$ it is 4.29.

The Lancaster construction, in the notation of this appendix: on the balanced decade grid the model is the autoregression above, $y_{i,t} = \rho_{10}\, y_{i,t-1} + a_i + \zeta_{i,t}$, with y log income at the seven grid years, a_i a constant country effect, and the time effects removed by demeaning each cross section. Fitting the N effects jointly with ρ_{10} from six transitions each is what generates the bias; the repair is the reparameterization $a_i = \eta_i\, e^{-\psi(\rho_{10})}$ with

$$\psi(\rho) \;=\; \frac{1}{T}\sum_{t=1}^{T-1}\frac{T-t}{t}\,\rho^t, \tag{A.12}$$

chosen so that the transformed effects η_i are information-orthogonal to $(\rho_{10}, \sigma_\zeta^2)$. Flat priors on the η_i and $p(\rho_{10}, \sigma^2) \propto 1/\sigma^2$ then integrate out in closed form, leaving the one-dimensional marginal posterior

$$\log p(\rho_{10} \mid \text{data}) \;=\; N\,\psi(\rho_{10}) \;-\; \frac{N(T-1)}{2}\log \mathrm{RSS}(\rho_{10}) \;+\; \text{const}, \tag{A.13}$$

where $\mathrm{RSS}(\rho_{10})$ is the within residual sum of squares. The second term alone is the fixed-effects objective, so the entire difference between the biased estimator and the bias-free posterior is the increasing factor $e^{N\psi(\rho_{10})}$, which shifts mass toward higher persistence by exactly the amount the demeaning manufactures: on simulated constant-target panels with $\rho_{10} = 0.65$ and $T = 6$ — where the bias formula above puts the within expectation near 0.33 — the posterior mode is 0.65. Because $e^{N\psi(\rho)}$ keeps growing while the RSS term grows only logarithmically, the posterior is not integrable beyond the stationary region: its support ends at $\rho_{10} = 1$, and mass at that corner is the estimator's answer. The correction is exact only under constant a_i; under heterogeneously trending steady states the likelihood is misspecified.

The minimum reportable speed rests on the stationary initial condition of Equation A.4, and two departures from that model bear on it differently. Heterogeneous persistence leaves it in place. With ρ_i and $\sigma_{\zeta,i}$ varying across countries, the numerator and denominator of Equation A.6 become cross-country averages, and conditional on $(\rho_i, \sigma_{\zeta,i})$ each country's expected covariance is its homogeneous limit times its expected variance. The fixed-effects estimate therefore converges to a variance-weighted average of the homogeneous limits, and no such average exceeds their value at the unit root, $1 - 3/(T+1)$, or 0.57 at $T = 6$. Four heterogeneous designs simulated at $T = 6$ give 0.31 to 0.55. A start further from the targets than the stationary dispersion removes the floor. A deterministic path $y_{i,t} = \rho^t y_{i,0}$ satisfies the autoregression exactly, so a large initial deviation pulls the estimate toward the true

[1] That the leading bias term is of order $3/T$ at $\rho = 1$, rather than the familiar $(1+\rho)/T$, is due to Phillips and Sul (2007); see Smith (2024). The statement here is the same correction carried into the annualized rate, where the governing quantity is the span S rather than the number of observations.

persistence. At $\rho_{10} = 0.82$, initial deviations three and five times the stationary standard deviation raise the fixed-effects limit to 0.66 and 0.74. Two results hold whatever the start. At $\rho = 1$ the initial level is absorbed by a_i, so the distribution of the fixed-effects estimate does not depend on it. On the unbalanced structure of the full panel of 164 countries it averages 0.55 (s.d. 0.04), and none of 2,000 replications reaches the observed 0.71. The Lancaster posterior conditions on $y_{i,0}$, and on simulated constant-target panels of 111 countries it recovers the true persistence under any start. At $\rho_{10} = 0.82$ with initial deviations five times the stationary dispersion, its mode averages 0.82 (s.d. 0.02) across 200 replications, where fixed effects give 0.75. The posterior of Figure A.1 is computed on the balanced panel of 111 countries.

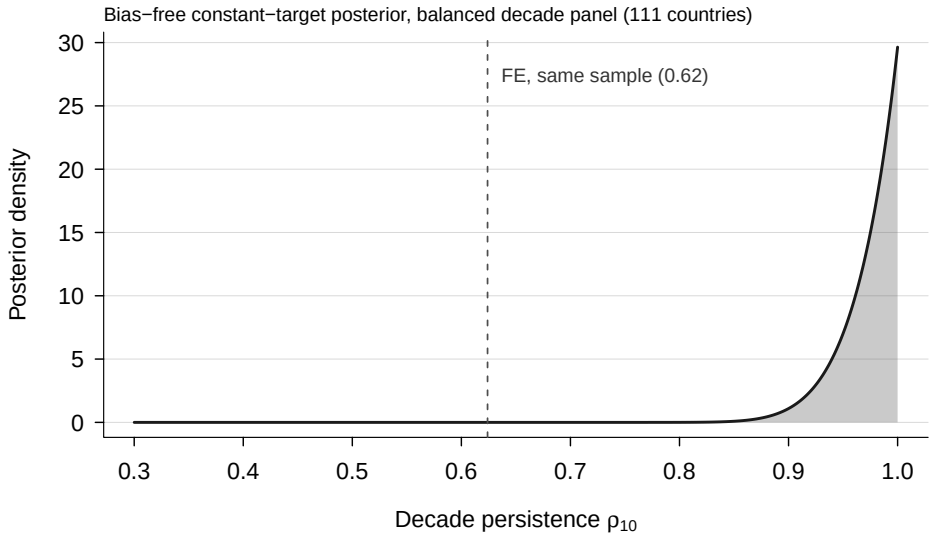


Figure A.1: The bias-free constant-target posterior for the decade persistence ρ_{10} (Lancaster orthogonalized effects, balanced decade panel of 111 countries). The log-posterior rises monotonically over the entire stationary region, so the mass concentrates at the unit root; the dashed line marks the fixed-effects estimate on the identical sample. Beyond the stationary region the correction factor grows without bound and the posterior is not integrable.

A.5 The drift in the fitted targets

The drifting-target nulls of Section A.7 are calibrated from the decade changes in the fitted targets $\hat{y}^*_{i,t}$ of Section 4.4, over the 110 countries observed in at least three windows. Their mean is $\bar{g}^* = 42$ percent per decade, the common drift averaged across countries rather than across country–windows. Their dispersion splits: a single country–decade change has standard deviation 0.55, of which $\tau_b = 0.20$ is the spread of country means and $\tau_w = 0.51$ is movement within a country across decades, the second accounting for 86 percent of the variance. The country means are themselves estimated from four changes apiece, so their spread is not separable from its own sampling noise — the average sampling variance of a country mean exceeds the observed variance across countries — which is why the nulls are run at $\tau_b = 0$ alongside the measured value.

A finer decomposition identifies the persistent component those nulls use. A country's decade change in $\hat{y}^*$ is written as a persistent country component, a permanent innovation to the level, and a transitory disturbance that reverses. The variance and the first two autocovariances identify the three, since at two lags the permanent and transitory parts drop out and only the persistent component survives. Computed on the deviations of the decade changes from their window means — 333 adjacent pairs for the variance and the first autocovariance, 223 for the second — those moments are 0.221, −0.058 and 0.011, giving standard deviations of 0.10, 0.27 and 0.26. Allowing the transitory component to decay gradually rather than reverse within one decade changes little: the persistence implied by the third autocovariance is −0.07, and requiring the three variances to be non-negative admits no value above 0.16.

A.6 Anatomy and robustness of the profiles

The columns of Table 1 are defined as follows. The within column is $\beta^W_{d,\bar{\gamma}}$, the within slope under the constant-wedge restriction $\gamma_d \equiv \bar{\gamma}$ of Equation 4; it is the debate's own estimator and equals the fixed-effects column exactly. The between column, $\beta^B_d = \beta_d + \gamma_d + \bar{\omega}$, is the income gradient holding the initial position fixed: the response of growth in window d to a unit rise of log income in every window of a country observed throughout (Section 3.2). The unconditional column, $\beta^U_d = \beta^B_d + \delta$, is the gradient with the initial condition moving with the level. The between and unconditional columns come from the decade-varying variant under a common δ, with CR2 cluster-robust standard errors of these linear combinations; the freed-δ_d profiles are reported later in this section. The within column comes from a different fit, so the within column and the other two do not difference into γ_d. Split levels are treated as projections throughout. The fixed-effects column drops one singleton country.

Once a country's income history and initial income are accounted for, no permanent growth difference is left: the residual country effect has estimated standard deviation $\hat{\sigma}_u = 0.00$, at the boundary, against the decade-level shock $\hat{\sigma}_\varepsilon = 2.12$ and the cross-country spread of the *level* term, 3.4 percentage points. The Bayesian fit of Section A.10 puts the posterior median of σ_u at 0.16, with a 95 percent interval ending at 0.46. That is what a steady-state interpretation predicts: the role of the country effect a_i in Equation 2 is to set the steady state, and mean income is, asymptotically, that steady state. It also bounds what the linearity of the projection can cost, since whatever a linear projection of the country effect on the income history misses would appear in σ_u.

Three extensions probe the heterogeneity the baseline restricts. First, heterogeneous convergence rates: the data-generating process of AM allows country-specific slopes β_c, which Equation 4 restricts to a common β^W_d profile. Adding a random country slope on the demeaned regressor — the parameterization that keeps the slope deviation out of the between margin (Section A.1) — the dispersion is substantial and unambiguous: the standard deviation of country convergence rates around the common profile is 1.6 against mean within slopes of −2.4 to −3.6 (likelihood ratio $\chi^2(1) = 23$, $p = 2.0 \times 10^{-6}$ before the boundary correction, which would lower it further). The intercept–slope correlation is estimated at its boundary by REML (a singular fit at +1); the Bayesian fit with an LKJ prior regularizes it smoothly

— posterior mean 0.78, 95 percent interval $(0.45, 0.97)$ — leaving the dispersion essentially unchanged, the boundary case of Section A.10 in practice. The decade profile keeps its shape — steepest in 2003–13, retreating in 2013–23 — with levels shifted down by 0.3 to 0.4 as the heteroskedastic weighting reweights countries, so the time-variation conclusions are not an artifact of imposing rate homogeneity, while the level caveats of Section 4.3 apply with extra force under slope heterogeneity (Pesaran and Smith, 1995).

Second, the unbalanced-panel caution of Wooldridge (2019): with panel lengths of one to six observations, the projection coefficients may vary with panel length. Interacting the projection with panel-length groups does not improve the fit ($\chi^2(3) = 4.2$, $p = 0.24$). The panel-length indicator and its interactions with the window-dummy means are already spanned by those means, and the three interactions that remain, with the income means of the last three windows, are jointly insignificant. The within profile cannot move, since the within slopes equal the fixed-effects slopes under either projection.

Third, the stability of the initial-condition coefficient. A likelihood-ratio test of $\delta_d \equiv \delta$ rejects the common value ($\chi^2(4) = 13.7$, $p = 0.008$). Freeing it by decade gives $\hat{\delta}_d = -0.8, -1.1, -1.6, -2.0, 0.6$ from 1973–83 onward (s.e. 0.7–0.9; the 1963–73 window is structurally excluded), against the common -1.0 — within about two standard errors of one another through 2003–13, and changing sign in the pandemic window. The sums barely notice: the unconditional profile moves by at most 0.04 in the pre-pandemic windows, and in 2013–23 from -0.76 to -0.63 — a fifth of its size, within one standard error, and part of the pandemic-window caution of Section 4.1. The addends notice everything: the freed between column swings to -1.24 in the final window while $\hat{\delta}_d$ changes sign — an allocation the ridge cannot decide, so the division between gradient and initial-condition channel is not robust in any window. With δ freed, the between column rises from 0.71 in 1973–83 to 1.11 in 2003–13, and falls by 0.33 from 1993–2003 to 2003–13, against 0.76 under the common coefficient. The movement is carried by the unconditional profile, which the restriction leaves in place: under both, it steps into convergence at 2003–13. Refitting the constant-steady-state null of Section A.7 with δ freed, the step is 0.74 against a manufactured 0.09 (s.d. 0.16), and the fall from 1963–73 is 1.31 against 0.13; no replication reaches either.

A.7 Alternatives to a change in the world

The decline of the between gradient (Section 4.2.1) is a movement in an estimated coefficient, and this section asks what else could produce it. The candidates are taken in turn: the mechanical endogeneity of the country mean; the changing composition of the panel; the cyclical position of the single-year income anchor; steady states that themselves move; more than one regime, together with the influence of single countries and country groups; and a common factor with country-specific loadings, including the design PS recommend. Each is calibrated to this panel. The simulations, the jackknife and the regime splits estimate the between slopes by least squares on the regressors of Equation 4, which on the data gives a fall of 1.30, identical to the REML fit of Table 1 because $\hat{\sigma}_u = 0$, and each is set against that value.

The fall is neither mechanical nor compositional. Because $\overline{\log y_i}$ is the mean of a lagged dependent variable, decade shocks enter the later incomes it averages. This endogeneity is larger for early decades, whose future holds more of the sample, so it declines over time by construction. The *flat-profile null* measures it. A Monte Carlo of 500 replications generates data with a flat profile around constant steady states, on the panel's exact country-year structure, with decade deviations following an AR(1) of persistence 0.90 matched to the within-country variance of log income. It manufactures a decline of 0.12 (simulation s.d. 0.22), against the 1.30 observed. That persistence is not identified (Section 4.3), so the null is run across its range. The manufactured decline peaks at 0.21 near $\rho_{10} = 0.5$–0.65 and falls away on either side, and at no value does it reach the observed decline in more than 1 percent of replications. On the balanced subpanel of the 111 countries observed in all six windows, which holds composition fixed, the fall is 1.31, measured on β_d^U, whose differences equal those of β_d^B under a common δ and which remains identified when every country enters in 1963–73.

Nor is it the anchor. Income enters at a single year while growth is measured over the following decade, so the anchor carries that year's cyclical position; replacing it by a centered three- or five-year average — symmetric, so that no trend drift is traded for the noise removed — leaves the profile where it was, with declines of 1.32 and 1.32 against 1.30 on the single-year anchor. Every window's between slope rises, by 0.11 to 0.13, so the profile shifts up bodily and the decline is untouched. Classical error in the anchor attenuates that slope toward zero, and removing it raises each level.[2] The cyclical component is a negligible share of the anchor's variation: its standard deviation falls from 1.215 to 1.212 under five-year smoothing, cross-country income dispersion being an order of magnitude larger than the cycle.

The result also survives when the null's steady states are allowed to move. Repeating it with country-specific target drift — at the persistent component of about 11 percent a decade identified in Section A.5, and at the raw 0.204, with the drift uncorrelated with income or correlated with it at ± 0.3 — manufactures declines between -0.34 and 0.23 against the 1.30 observed. The negative values mean the simulated profile rises rather than falls.

Another alternative is that the sample is not one regime at all. Under multiple steady states the cross-sectional slope is governed by a covariance between initial and steady-state income within the sample (Bernard and Durlauf, 1996), so it moves with composition while no structural parameter changes (Durlauf et al., 2005, §4.2.2); a placebo that holds a single linear projection fixed cannot speak to it. We test for more than one regime as follows. Countries are split into two regimes by a characteristic fixed in time, every coefficient is free to differ between the regimes, and the split point is chosen to maximize the likelihood ratio against a single regime. On initial income the statistic is 118.7, at a threshold of 7.52 in logs or about $1,800 a head. On initial schooling it is 98.4, at 1.76 standard deviations. A country-cluster bootstrap puts these at $p = 0.000$ and $p = 0.000$. The sample has more than one regime.

A shift confined to one decade invites a question a trend would not: whether one country

[2] A trailing average, by contrast, sits below the current level on a trending series and moves the level down by 0.29, which is why the symmetric window is the right one.

or one group produces it. It does not. Leaving out one country at a time, the shift ranges from 1.16 to 1.39 across the 164 countries, against 1.30 on the full sample; China is the most influential single country and removing it leaves 1.16. Dropping whole groups does no more: 1.24 without resource exporters, 1.11 without the transition economies, and 0.95 without East Asia, which is the most influential group.

We split at the median initial income, which fixes the two regimes in advance rather than choosing them by search, and estimate the profile within each. The between gradient falls by 1.94 in the poorer half and 1.03 in the richer, against 1.30 pooled. At the searched threshold the falls are 2.67 in the poorer regime and 0.74 in the richer.

Splitting into regimes costs most of the precision. A country-cluster bootstrap around the whole split-and-refit gives standard errors of 0.91 and 0.73 on the two within-regime declines, against 0.29 for the pooled decline, close to its cluster-robust standard error of 0.29 (model-based 0.26, Section A.9). The decline in the richer regime is not separated from zero; in the poorer regime the 95 percent interval, $[0.18, 3.82]$, narrowly excludes it. Neither differs significantly from the pooled decline ($p = 0.33$ and 0.58). The split point is itself an estimate, and holding it fixed across replicates understates the uncertainty. Re-estimating it inside every bootstrap replicate widens the intervals further. Were the decline driven by that covariance — the sample's mix of countries shifting across regimes, with no parameter changing — it would weaken once the regime is held fixed. Within the regimes the decline is larger than the pooled value in the poorer one and smaller in the richer, and neither difference is significant. At this precision the exercise cannot exclude it.

A common factor is the harder case, since Section 3.2 invokes one and the flat-profile null has none. It is also the null that speaks to a one-decade event: the factor is drawn afresh in every window, so each cross-section is tilted by a shock whose incidence varies with income, which is what a single episode would look like. Adding a factor drawn afresh each replication, with country loadings and an innovation dispersion scaled window by window to the panel's own, leaves the average manufactured decline small — between 0.07 and 0.24 across nine configurations — but widens its dispersion sharply when the loadings correlate with the income level. With uncorrelated loadings one replication in four hundred reaches the observed decline, at the largest loading dispersion, and none at the two smaller ones. Correlating the loadings with income at 0.5, the share that does depends on how large the factor is made: 0.5 percent at a loading dispersion of 0.05 log points, 3.5 percent at 0.10 and 12.5 percent at 0.20. Those dispersions are not free parameters to set aside: they contribute 0.71, 1.41 and 2.83 percentage points a year to the cross-country dispersion of decade growth, against the 2.70 observed — so the configuration that reaches double figures posits a single common shock larger than the entire spread of growth rates it would be explaining, and the one before it a shock carrying a quarter of the cross-country variance of growth. At a factor of ordinary size, a loading dispersion of 0.05, the decline is manufactured twice in four hundred replications with loadings correlated with income at 0.5, and never at -0.5. A common shock therefore produces the decline only at a size that would displace the phenomenon it accounts for. With the nulls above, this rules out the endogeneity of the country mean, the changing composition

of the panel and a common shock of ordinary size. More than one regime is the alternative the data cannot settle.[3]

The sharpest version of that alternative is the design PS themselves recommend. Their mechanism needs three ingredients the placebo above lacks: a factor that trends rather than one drawn afresh each window, loadings that differ across countries and correlate with income, and dynamics that differ across countries. Building all three into the flat-profile null — country-specific decade persistence at dispersions up to 0.25, and no parameter of the simulation varying by decade — the manufactured decline lies between -0.54 and -0.34. The simulated profile rises rather than falls, and in none of those replications does it reach the observed value. Across every configuration examined the largest share of replications reaching the observed decline is 3 percent. Only two configurations have any replication reaching it, and both generate samples whose cross-country income dispersion is flat or narrowing, against the widening these data show. Dynamic heterogeneity on its own moves the manufactured decline down rather than up, from 0.15 at homogeneous dynamics to 0.13 at the widest dispersion tried: Heterogeneous dynamics make the null less able to produce the result.

The same simulation bounds what the levels can mean. Under the flat-profile null, where nothing in the data-generating process ties income to growth, the projection on the country-level regressors alone delivers $\hat{\beta}_d^B \approx 2.2$ and $\hat{\delta} \approx -2.2$, larger in absolute value than every observed between level (at most 1.42) and than the observed $\hat{\delta} = -0.98$. Levels of that size therefore say nothing about a gradient, and the between and initial-condition coefficients are read throughout as projections rather than as structural quantities.

Collinearity is a second and smaller issue, and it bears on the split rather than on the level. For a country observed in all six windows the initial income is six times the element of $\mathbf{m}_i$ that averages first-window income. On the 111 countries of the balanced 1963–73 cohort the likelihood is therefore flat along the direction that raises every β_d^B by the same amount and lowers δ by it, which leaves β_d^U unchanged, and the level of β_d^B is identified by the 53 countries that enter later. That margin is what a structural reading of β_d^B would have to rest on, and it is why the division between β_d^B and δ depends on which restriction is imposed (Section A.6).

The movement is what the null cannot produce. Its β_d^B decline is the 0.12 above (simulation s.d. 0.22), against the 1.30 observed. Because δ is common across decades it cancels in any difference of $\beta_d^U = \beta_d^B + \delta$, so that same comparison is a statement about the unconditional gradient itself: the null moves it from 0.04 to -0.07 where the data move it from $+0.45$ to -0.85. That reading does not depend on the split: freeing δ_d leaves the decline in β_d^U essentially unchanged while the between and initial-condition terms trade places (Section A.6). The within estimator moves the same way: $\beta_{d,\bar{\gamma}}^W$ steepens from -1.94 to -3.32, flatness rejected ($\chi^2(5) = 47.3$, $p = 5.0 \times 10^{-9}$).[4] The between movement is a cross-sectional fact.

[3] An income-proportional loading cannot be estimated apart from a moving between slope on this grid, and that is the direction the placebo bounds. Loading heterogeneity orthogonal to income is estimable and substantial, and it leaves the profile intact (Section A.9).

[4] The likelihood-based Wald tests use the model-based covariance. Cluster-robust versions (CR2 sandwich with Satterthwaite small-sample F) reject flatness equally: $F = 8.5$ for the within profile and $F = 8.7$ for the

The within movement must be interpreted through Section 4.3: the level of these slopes is not a convergence speed, and their drift mixes any change in adjustment with the changing composition and panel length of the sample.

A.8 Correlates and growth: profiles and decomposition

This section reports the decade profiles of the correlate slopes of Equation 7 and the omitted-variable decomposition summarized in Section 4.5.

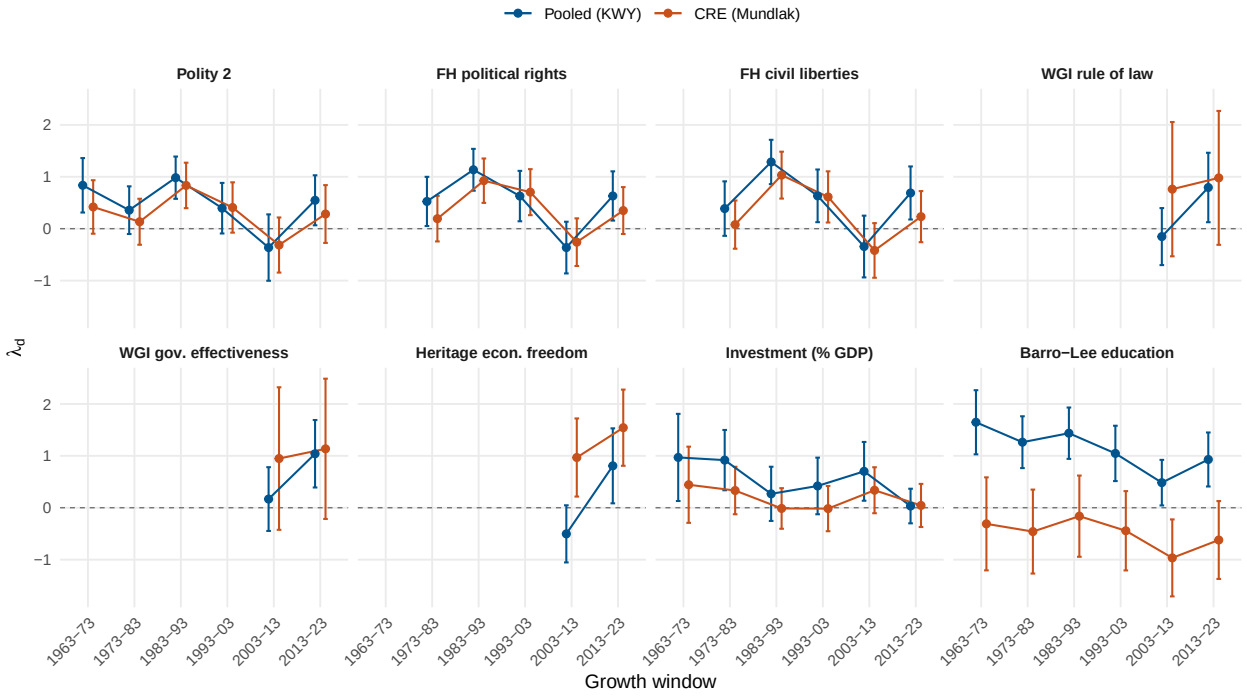


Figure A.2: Decade profiles of the correlate–growth slope λ_d (correlates in the Kremer–Willis–You normalization, growth over the subsequent decade, income-controlled), estimated pooled and with the Mundlak correction of Equation 7. Bars: 95% confidence intervals. Growth and income from PWT 11.0; correlates from the Kremer–Willis–You replication package. The Worldwide Governance Indicators and the Heritage score begin in 1996, so only the last two windows are available for them.

KWY report that the growth payoff to institutions and policies has flattened since the 1990s, and read that flattening as the reason absolute convergence caught up with conditional convergence. The pooled profiles of Figure A.2 reproduce it and qualify it. For the political-institutions block, the pooled slope falls from around 1 in the 1980s windows to zero or below in 2003–13, and a Wald test rejects a flat profile for every one of the eight correlates (Table 3,

between profile, both $p < 0.001$.

pooled p column).[5] In the 2013–23 window the institutional slopes rebound: on the extended data the flattening is a trough centered on the 2000s rather than a terminal state.

The split matters here for the reason it did for convergence. The correlate regression carries the same income-correlated country effects, so its pooled coefficient mixes a within-country effect with whatever co-varies with the correlate across countries, and a movement in that coefficient can come from either margin. KWY's flattening is a statement about the pooled coefficient and the stability AM report is a statement about the within one, so the split is what makes the two distinguishable: a decline in λ_d says the within-country effect itself changed. A decline in the pooled coefficient while λ_d stays flat says something else: the pooled coefficient is an average of the within-country effect and the between-country association, weighted by how much of the correlate's variation is of each kind, so if neither effect moves and the pooled average still falls, what moved is the weighting — the correlate's variation shifted from between countries toward within them, or the reverse. Neither margin carries a causal reading, and only λ_d carries the decade index, so no decline is attributed to the between component here.

The Mundlak correction then separates two cases. For the Solow correlates the pooled association does not survive the split as a within-country effect: education's within slope is negative in every window and its country-mean term, averaged over the windows, strongly positive ($\xi = 1.01$, s.e. 0.40), so its apparent growth effect is between-country heterogeneity, and investment is weak on both margins ($\xi = -0.04$, s.e. 0.23). Flatness cannot be rejected for either in the CRE specification ($p = 0.59$ for investment, 0.08 for education). Education's within slope is not evidence that schooling fails to raise growth: only 24 percent of its within variation survives the year effects, and a country-specific linear path accounts for 96 percent of what remains, so the coefficient is barely identified by this panel. Investment's null is informative, education's is not.

For the institutional block the pattern reverses. The Mundlak term is small and insignificant for the political-institutions measures, so their within-country slope and between-country association cannot be told apart, and the within slope survives: Polity 2's λ_d is 0.83 in 1983–93, falls to -0.31 in 2003–13, returns to 0.28 in 2013–23, and flatness is rejected ($p = 0.006$), a rejection that survives Bonferroni (footnote above), as do those for Freedom House political rights and civil liberties, both below 0.001 in this column. This does not settle the democracy result AM report, whose variable is the dichotomous indicator of Acemoglu et al. (2019) estimated on annual data, where a country contributes close to forty transitions against six here — by the criteria of Section 4.3 the better identified design. What the decade grid establishes is narrower: KWY's own measure, on KWY's own windows, moves over time

[5] Eight correlates are tested, so the rejections are adjusted for multiplicity within each specification. Bonferroni holds the chance of any false rejection across the eight tests below five percent, and Benjamini–Hochberg the expected share of false rejections among those made; the first is the stricter. In the pooled column Benjamini–Hochberg leaves all eight rejections standing and Bonferroni seven, investment being the exception at 0.27. In the Mundlak column four of the eight reject unadjusted; Benjamini–Hochberg leaves all four, Bonferroni three, with the Heritage score the one that turns, at 0.10 against 0.024 under Benjamini–Hochberg.

and does not flatten terminally. The movement in the pooled coefficients is genuine; their level mixes the within-country effect with between-country heterogeneity, which for education accounts for the entire association.

For a single correlate, in each growth window d, estimate three cross-sectional regressions on the countries with the correlate observed:

$$g_i = \alpha_d + b_d \log y_i + e_i, \qquad g_i = \alpha_d^* + b_d^* \log y_i + \theta_d x_i + e_i^*, \qquad x_i = \alpha_d^x + \pi_d \log y_i + v_i, \tag{A.14}$$

so that b_d is the absolute convergence slope, b_d^* and θ_d the income and correlate slopes of the conditional regression, and π_d the correlate–income gradient. Lowercase b's distinguish these cross-sectional slopes from the panel coefficients of Section 3, and θ_d is the single-window analogue of the panel slope λ_d of Equation 7. These are the regressions linked by the omitted-variable identity of Equation 8.



Figure A.3: Absolute (b_d) and conditional (b_d^*) convergence coefficients by growth window, per correlate; the shaded band is the conditioning gap $\pi_d\theta_d$ of Equation 8. Cross-sectional regressions on the sample with the correlate observed; growth and income from PWT 11.0, correlates from the Kremer–Willis–You replication package.

Three facts emerge (Figure A.3). First, the mechanism of KWY is confirmed on the extended data: in the 1983–93 window the conditioning gap runs from 0.44 to 0.81 across the three institutional correlates available then — absolute convergence was substantially weaker than conditional — and by 2003–13 the gap is zero or negative: absolute convergence converged

to conditional convergence. Decomposing the change in the gap, from each correlate's first window to 2003–13, into $\theta\Delta\pi + \pi\Delta\theta + \Delta\pi\,\Delta\theta$ shows the flattening of θ_d doing most of the work, as they argue. It does not do all of it. For Polity 2 the flattening contributes $\pi\Delta\theta = -0.27$, and the convergence of the correlate itself — π_d falls from 0.62 to 0.29 as poor countries catch up on institutions — contributes $\theta\Delta\pi = -0.11$, some two fifths as much, with an offsetting interaction of 0.15. The correlates did converge: β-convergence in the correlate is significant in 21 of 26 correlate-window pairs, echoing the 17-of-20 count in KWY.

Second, the same accounting reframes the retracement after 2006. In the 2013–23 window the conditioning gap reopens to 0.10–0.96 for the institutional block while conditional convergence stays strong (b_d^* between -0.31 and -1.23 across correlates). The reopening gap carries the whole weakening of absolute convergence in the pandemic window for the Heritage score and the two Worldwide Governance Indicators, 67 and 70 percent of it for the two Freedom House scores, and 26 percent for Polity 2, whose conditional slope weakens from -0.70 to -0.31.

Third, the conditioning that matters happens across countries. Repeat the exercise on within variation: estimate the convergence slope with country and year fixed effects, once without and once with the correlate, on the same sample (the omitted-variable identity links the two exactly there as well). Adding the correlate then barely moves the slope: across the eight correlates the change is at most 0.34, against within slopes of -2.5 to -5.6 in this sample (the range across correlates is wide because samples differ — the WGI and Heritage panels cover only the last two windows — and shorter panels amplify the finite-T bias). Whether one controls for a country's institutions is thus nearly irrelevant for how it converges to its own path. The correlates channel of Equation 8 — the gap that closed by the 2000s and reopened in 2013–23 — operates through *which countries* have better institutions and how that lines up with income.

A.9 Cross-sectional dependence and factor structure

A common factor with country-specific loadings (Section 3.2) would also leave the residuals correlated across countries, which bears on the inference. The standard errors of the main text are clustered by country: they allow $\mathrm{corr}(\varepsilon_{i,t}, \varepsilon_{i,s}) \neq 0$ within a country across its windows, and impose $\mathrm{corr}(\varepsilon_{i,t}, \varepsilon_{j,t}) = 0$ between countries in the same window. Positive dependence across countries would make them too small, so that is the condition to test, net of the window effects μ_t which already absorb whatever moves every country together. The signed CD statistic of Pesaran (2004) cannot do it: with μ_t removed the residuals sum to approximately zero across countries within a window, which pushes each pairwise correlation toward $-1/(N-1)$ and the statistic toward -1.55 on this grid and -5.07 on one-year growth, so neither its sign nor its size is read against a standard normal. The mean absolute pairwise correlation can be, against the value independent series with the same overlaps would give, simulated for the panel's own distribution of overlap lengths: 0.465 against 0.458, a ratio of 1.02. The dependence surviving μ_t is therefore small relative to what short overlaps produce by chance, though at $T \leq 6$ the two cannot be separated with precision. Clustering by country is not exact under any such dependence, and the cluster-robust versions of the main text do not repair it — they relax

the model-based covariance within a country and impose the same restriction across countries — so it is the factor placebo of Section A.7 that prices the alternative.

Estimation and simulation cover different directions here. A loading *proportional to income* cannot be estimated apart from a moving between slope on this grid: the fixed part already contains $\overline{\log y}_i$ interacted with the six window dummies, and those columns span $\overline{\log y}_i$ times any window-indexed vector, so an income-proportional $\ell_i f_d$ lies exactly in the design (the regression of one on the other returns $R^2 = 1$). That direction is the one the placebo bounds. Loading heterogeneity *orthogonal* to income is estimable, and substantial: giving each country a random loading on the cross-sectional income average — the mixed-model analogue of Pesaran (2006), feasible at $T = 6$ where a country-by-country regression is not — yields a loading standard deviation of 2.45, worth about one percentage point a year of growth, and a likelihood-ratio statistic of 30.2. It leaves the profile intact. The decline moves from 1.30 (s.e. 0.26) to 1.23 (s.e. 0.29), and every window's between slope falls, by 0.05 to 0.13. A loading on the cross-sectional growth average has zero estimated variance.

A.10 Bayesian estimation as regularization

The mixed-effects models of the main text are estimated both by REML and by full Bayesian inference (Stan/brms; four chains, weakly informative $N(0, 25^2)$ priors on regression coefficients, half-Student-t defaults on variance components; diagnostics in Section A.13). The posterior medians and intervals track the REML estimates closely: the medians of β_d^B, β_d^U, their decade differences and δ lie within 0.07 of the REML values, as expected with 881 observations. The exception is σ_u, discussed below. The value of the Bayesian machinery here is regularization of the dimensions the likelihood identifies weakly, and finite-sample uncertainty for derived quantities such as β_d^B, obtained directly from posterior draws rather than by the delta method.

Two dimensions benefit. First, variance components: restricted maximum likelihood can return boundary estimates ($\hat{\sigma}_u = 0$, a "singular fit"), as it does for both correlated-random-effects fits here, and standard errors at the boundary are unreliable. A continuous prior puts posterior mass on small positive values and propagates the uncertainty about σ_u instead of pinning it: the posterior median is 0.16 in the constant-wedge fit and 0.17 in the decade-varying one, with 95 percent intervals ending at 0.46 and 0.47. Second, the decomposition of the profile. The level of β_d^B is separated from δ only by the countries that enter after 1963–73 (Section A.7), so the likelihood is nearly flat along the direction that raises every β_d^B and lowers δ by the same amount. The frequentist β_d^B carry standard errors of about 0.6, against 0.16 to 0.20 for β_d^U and 0.22 to 0.26 for the decade differences of β_d^B whose falls the main text reports. Gaussian priors act as ridge-type regularization along that direction without informing β_d^U or the differences; at the $N(0, 25^2)$ priors used here the effect is negligible, and the posterior standard deviations of β_d^B, 0.61 to 0.63, match the REML standard errors. Hierarchical priors shrinking the decade slopes toward a common mean would regularize the profile itself and are not used, so that the trend tests are not prior-driven.

What Bayesian estimation does not deliver is a cure for the Nickell bias: the bias is a property

of the sampling model, and a weakly informative prior placed on the same likelihood shifts nothing systematically. Regularization is variance reduction on weakly identified dimensions; the bias discussion of Section A.4 applies to the posterior exactly as to the REML estimates.

A.11 Income series for levels and growth

Growth is measured from `rgdpna` and levels from `rgdpo`, as in KWY. The split follows their footnote 5 and PWT's own guidance — `rgdpna` for growth over time, the chained-PPP series for level comparisons across countries (Feenstra et al., 2015) — and it matters for the dynamic within margin. The within slope on the correlates sample of Section 4.4 is −2.20 under this design; measuring both level and growth from `rgdpna` gives −2.32, and both from `rgdpo` −3.92. Starting the growth window one year after the level is observed — so that transitory error in the anchor-year level cannot enter both sides — leaves the same-series estimates at −2.48 and −4.24: the component specific to `rgdpo` (a one-standard-deviation within-country discrepancy of 22 percent on the decade grid) is persistent rather than year-specific and mean-reverts at decade scale, which reusing `rgdpo` on both sides registers as convergence. Pesaran and Smith (2026) raise the same design point and formalize it: entering the level discrepancy as a separate regressor, KWY's specification is the restriction that its coefficient equals minus the coefficient on initial income, which is testable.

A.12 Reproduction details

On KWY's replication data, re-estimating their specifications[6] reproduces every coefficient, standard error, and the sample size of their Table 1 exactly (pooled slope −0.2702, s.e. 0.1182; trend interaction −0.0251, s.e. 0.0056; $N = 863$). Re-estimated on PWT 11.0, the profile matches throughout — every movement under one standard error of the published value, and all but the 1970s window under half of one, with significant convergence confined to the 2000s and the 2007–17 window — so the published result is robust to the data revision and the slightly larger sample; the table below reports it in full.

Re-estimating Equation 1 on PWT 11.0 then isolates the effect of the data revision (table above): the pooled slope is −0.261 against −0.270, the trend interaction −0.028 against −0.025, and the decade profile matches, the only sign flips being the two near-zero, insignificant 1970s–1980s coefficients,[7] with significant convergence confined to the 2000s (−0.644, s.e. 0.144) and the 2007–17 window. The published result is robust to the PWT 10.0–11.0 revision and

[6] All estimation is carried out in R 4.3.3 (R Core Team, 2024): the pooled and fixed-effects regressions with `fixest` (Bergé et al., 2026); the GMM check with `plm` (Croissant and Millo, 2008); the mixed-effects models with `lme4` (Bates et al., 2015) and `lmerTest` (Kuznetsova et al., 2017); the Bayesian models with `brms` (Bürkner, 2017) and Stan (Carpenter et al., 2017) via `cmdstanr` (Gabry et al., 2025); figures with `ggplot2` (Wickham, 2016).

[7] Re-estimating on PWT 11.0 restricted to the authors' exact country-years separates the two sources: the data revision alone moves the 1970s coefficient from −0.075 to 0.02 and the 1980s one from 0.106 to 0.05, and the country-years added in the new vintage carry them the rest of the way. All movements are under half a standard error; the significant decades move by at most 0.02 under the revision.

Table A.1: Absolute convergence 1960–2017: Table 1 of KWY re-estimated on PWT 11.0. Dependent variable: average annualized GDP per capita growth (%) over the subsequent decade; sample years 1960, 1970, 1980, 1990, 2000, 2007. Year fixed effects throughout; standard errors clustered by country in parentheses. Published column transcribed from KWY, Table 1.

	PWT 10.0 (published)	PWT 11.0 (this paper)
Pooled: log(GDPpc)	-0.270 (0.118)	-0.261 (0.106)
Trend: log(GDPpc)	0.449 (0.224)	0.543 (0.201)
Trend: log(GDPpc)× (year−1960)	-0.025 (0.006)	-0.028 (0.005)
By decade: 1960s	0.532 (0.191)	0.546 (0.190)
By decade: 1970s	-0.075 (0.293)	0.142 (0.281)
By decade: 1980s	0.106 (0.246)	-0.005 (0.243)
By decade: 1990s	-0.127 (0.221)	-0.082 (0.201)
By decade: 2000s	-0.651 (0.168)	-0.644 (0.144)
By decade: 2007–17	-0.764 (0.146)	-0.821 (0.146)
Observations	863	878

the slightly larger sample (878 versus 863 country–years).

A.13 Estimation diagnostics

All REML fits we report converged without warnings. The correlated-random-effects fits of Table 1, their panel-length extension and the baseline of Section A.9 are singular, with $\hat{\sigma}_u = 0$; there REML coincides with least squares on the same regressors, to 10^{-9}, and the Bayesian fits of Section A.10 carry the uncertainty about σ_u. The correlated random-intercept–random-slope fit is singular as well, with its intercept–slope correlation at the +1 boundary, and is therefore reported in its uncorrelated frequentist and LKJ-regularized Bayesian versions (Section A.6). The table below collects the sampler diagnostics of the Bayesian runs: four chains each, 2,500 post-warmup draws per chain.

Table A.2: Sampler diagnostics of the Bayesian estimations. $\hat{R}$ and bulk effective sample size are the worst values across all parameters of each model; divergent transitions are counted over all post-warmup draws. Per-run values are in the replication outputs.

Model	max $\hat{R}$	min bulk ESS	divergences
CRE, constant wedge	1.002	2579	0
CRE, decade-varying wedge	1.002	2088	0
Random slopes (LKJ)	1.002	1593	0

References

Acemoglu, D. and Molina, C. (2022). Comment on "Converging to Convergence". *NBER Macroeconomics Annual*, 36:425–442.

Acemoglu, D., Naidu, S., Restrepo, P., and Robinson, J. A. (2019). Democracy does cause growth. *Journal of Political Economy*, 127(1):47–100.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, 58(2):277–297.

Barro, R. J. (2015). Convergence and modernisation. *The Economic Journal*, 125(585):911–942.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Bell, A. and Jones, K. (2015). Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods*, 3(1):133–153.

Bergé, L. R., Butts, K., and McDermott, G. (2026). Fast and user-friendly econometrics estimations: The R package fixest. arXiv preprint arXiv:2601.21749.

Bernard, A. B. and Durlauf, S. N. (1996). Interpreting tests of the convergence hypothesis. *Journal of Econometrics*, 71(1–2):161–173.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115–143.

Bürkner, P.-C. (2017). brms: An R package for bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1):1–28.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1):1–32.

Croissant, Y. and Millo, G. (2008). Panel data econometrics in R: The plm package. *Journal of Statistical Software*, 27(2):1–43.

Dhaene, G. and Jochmans, K. (2015). Split-panel jackknife estimation of fixed-effect models. *Review of Economic Studies*, 82(3):991–1030.

Durlauf, S. N., Johnson, P. A., and Temple, J. R. W. (2005). Growth econometrics. In Aghion, P. and Durlauf, S. N., editors, *Handbook of Economic Growth*, volume 1A, pages 555–677. Elsevier, Amsterdam.

Feenstra, R. C., Inklaar, R., and Timmer, M. P. (2015). The next generation of the Penn World Table. *American Economic Review*, 105(10):3150–3182.

Gabry, J., Češnovar, R., Johnson, A., and Bronder, S. (2025). *cmdstanr: R Interface to CmdStan*. R package version 0.9.0.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46(6):1251–1271.

Hsiao, C., Pesaran, M. H., and Tahmiscioglu, A. K. (2002). Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods. *Journal of Econometrics*, 109(1):107–150.

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

Kremer, M., Willis, J., and You, Y. (2022). Converging to convergence. *NBER Macroeconomics Annual*, 36:337–412.

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13):1–26.

Lancaster, T. (2002). Orthogonal parameters and panel data. *Review of Economic Studies*, 69(3):647–666.

McNeish, D. and Kelley, K. (2019). Fixed effects models versus mixed effects models for clustered data: Reviewing the approaches, disentangling the differences, and making recommendations. *Psychological Methods*, 24(1):20–35.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica*, 46(1):69–85.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49(6):1417–1426.

Pesaran, M. H. (2004). General diagnostic tests for cross section dependence in panels. Cambridge Working Papers in Economics 0435, University of Cambridge.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4):967–1012.

Pesaran, M. H. and Smith, R. P. (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics*, 68(1):79–113.

Pesaran, M. H. and Smith, R. P. (2026). The output convergence debate revisited: Lessons from recent developments in the analysis of panel data models. arXiv preprint 2602.04060.

Phillips, P. C. B. and Sul, D. (2007). Bias in dynamic panel estimation with fixed effects, incidental trends and cross section dependence. *Journal of Econometrics*, 137:162–188.

R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Smith, R. P. (2024). Econometric aspects of convergence: A survey. *Open Economies Review*, 35(4):701–721.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York.

Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics*, 20(1):39–54.

Wooldridge, J. M. (2019). Correlated random effects models with unbalanced panels. *Journal of Econometrics*, 211(1):137–150.