Development Economics

Slides 5

Debraj Ray, NYU

Convergence

- **A central prediction of the Solow growth model: unconditional convergence:**
  - The incomes of countries move ever closer to one another
  - Based on the deep legacy of diminishing returns …
  - …and the equality of s, n, π across countries

Convergence

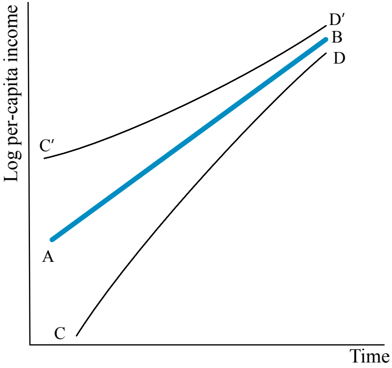
- **A central prediction of the Solow growth model: unconditional convergence:**
  - The incomes of countries move ever closer to one another
  - Based on the deep legacy of diminishing returns …
  - …and the equality of s, n, π across countries
- Does this sound trivial to you, or totally wild?
  - "Convergence versus divergence" pervades this entire course.

Unconditional and Conditional Convergence

- Unconditional convergence presumes the steady states are the same
- Conditional convergence allows them to vary across countries.

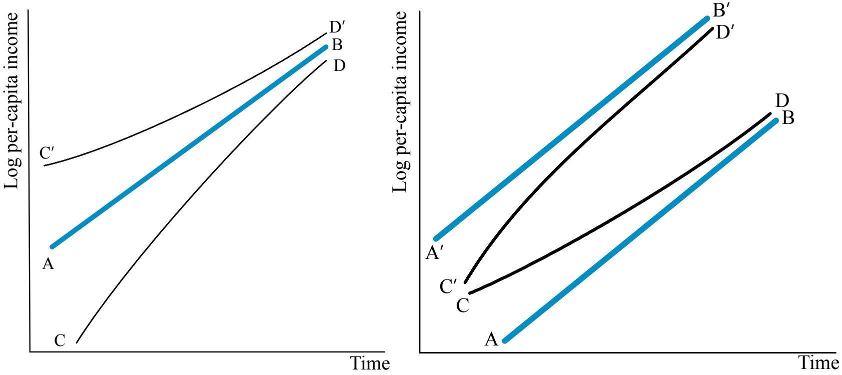
Unconditional and Conditional Convergence

- Unconditional convergence presumes the steady states are the same
- Conditional convergence allows them to vary across countries.



Unconditional and Conditional Convergence

- Unconditional convergence presumes the steady states are the same
- Conditional convergence allows them to vary across countries.



Testing for Convergence

- **Act I**: Convergence? 1870–1979
- **Interlude**: Conditional Convergence
- **Act II**: Modern Convergence?

Act I: Convergence? 1870–1979

- Baumol (AER 1986) studied 16 countries:
  - among the richest in the world today.
  - In order of poorest to richest in 1870: Japan, Finland, Sweden, Norway, Germany, Italy, Austria, France, Canada, Denmark, USA, Netherlands, Switzerland, Belgium, UK, and Australia.

Act I: Convergence? 1870–1979

- Baumol (AER 1986) studied 16 countries:
  - among the richest in the world today.
  - In order of poorest to richest in 1870: Japan, Finland, Sweden, Norway, Germany, Italy, Austria, France, Canada, Denmark, USA, Netherlands, Switzerland, Belgium, UK, and Australia.
- Why just 16?
  - The Maddison project (Angus Maddison 1982, 1991, 2007)
  - As of 2020: 169 countries up to 2018, with over 60 going back to 1870.
  - But not when Baumol wrote this paper.

Act I: Convergence? 1870–1979

- Idea: regress 1870–1979 growth rate on 1870 incomes.

$$\ln y_i^{1979} - \ln y_i^{1870} = A + b \ln y_i^{1870} + \epsilon_i$$

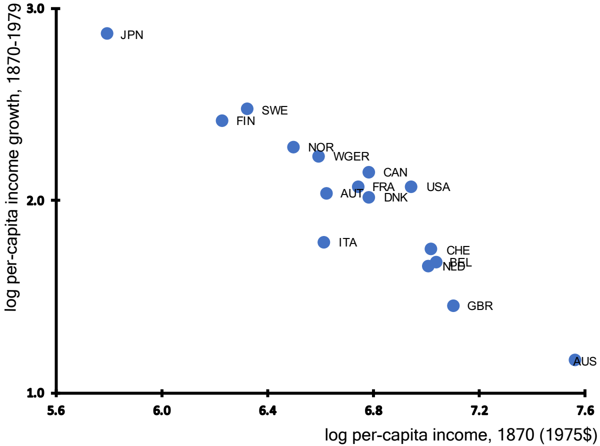
- Unconditional convergence $\Rightarrow b \simeq -1$. Get $b = -0.995$, $R^2 = 0.88$.

Act I: Convergence? 1870–1979

- Idea: regress 1870–1979 growth rate on 1870 incomes.

$$\ln y_i^{1979} - \ln y_i^{1870} = A + b \ln y_i^{1870} + \epsilon_i$$

- Unconditional convergence $\Rightarrow b \simeq -1$. Get $b = -0.995$, $R^2 = 0.88$.



Act I: Convergence? 1870–1979

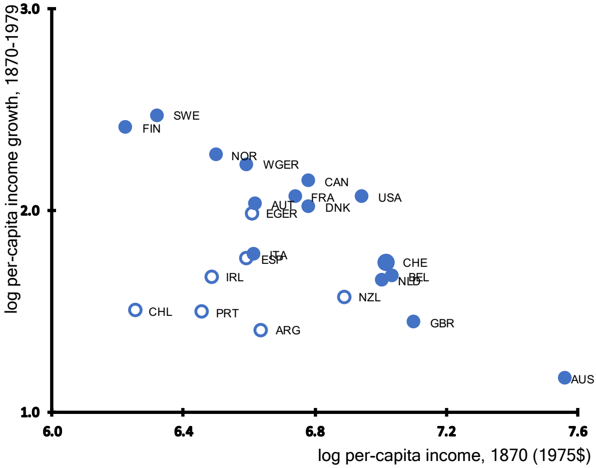
- **De Long critique (AER 1988):**

- Add seven more countries to Maddison's 16.
- In 1870, they had as much claim to membership in the "convergence club" as any included in the 16: Argentina, Chile, East Germany, Ireland, New Zealand, Portugal, and Spain.
- New Zealand, Argentina, and Chile were in the top-10 list for British and French overseas investment (in per capita terms) as late as 1913.

Act I: Convergence? 1870–1979

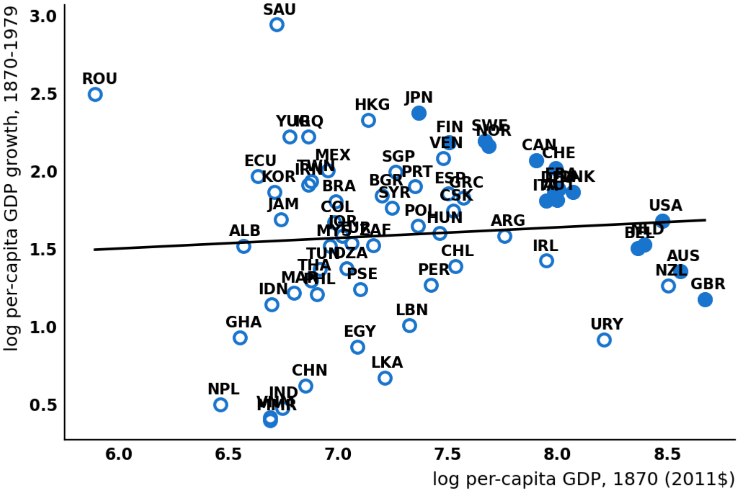
- **De Long critique (AER 1988):**
  - Add seven more countries to Maddison's 16.
  - In 1870, they had as much claim to membership in the "convergence club" as any included in the 16: Argentina, Chile, East Germany, Ireland, New Zealand, Portugal, and Spain.
  - New Zealand, Argentina, and Chile were in the top-10 list for British and French overseas investment (in per capita terms) as late as 1913.
  - All had per capita GDP higher than Finland in 1870.
  - Strategy: drop Japan (why?), add the 7.

Act I: Convergence? 1870–1979



- Slope still negative, though loses significance.

- Updated Maddison dataset 2020, 66 countries:



Act I: Convergence? 1870–1979

- Lant Pritchett (1997) called it "**divergence, big time**."
  - **Assumption**: no country can fall below $250 per capita (1985 PPP dollars)

Act I: Convergence? 1870–1979

- Lant Pritchett (1997) called it "**divergence, big time**."
  - **Assumption**: no country can fall below $250 per capita (1985 PPP dollars)
  - Defense 1: lowest 5-year average ever is Ethiopia $275 (1961–65).

Act I: Convergence? 1870–1979

- Lant Pritchett (1997) called it "**divergence, big time**."
  - **Assumption**: no country can fall below $250 per capita (1985 PPP dollars)
  - Defense 1: lowest 5-year average ever is Ethiopia $275 (1961–65).
  - Defense 2: $250 per capita is below extreme nutrition-based poverty lines actually used in poor countries (say, pegged at 2000Kcal, see Ravallion, Datt and van de Walle 1991).

Act I: Convergence? 1870–1979

- Lant Pritchett (1997) called it "**divergence, big time**."
  - **Assumption**: no country can fall below $250 per capita (1985 PPP dollars)
  - Defense 1: lowest 5-year average ever is Ethiopia $275 (1961–65).
  - Defense 2: $250 per capita is below extreme nutrition-based poverty lines actually used in poor countries (say, pegged at 2000Kcal, see Ravallion, Datt and van de Walle 1991).
  - Defense 3: at any lower income, population too unhealthy to grow. Child mortality rate estimated to climb well above barrier of 600 per 1000.

Act I: Convergence? 1870–1979

- Claim: the $250 bound "proves" divergence over long-run.

Act I: Convergence? 1870–1979

- Claim: the \$250 bound “proves” divergence over long-run.
  - The US grew @1.7% p.a., so by 4 times from 1870 to 1960.
  - Thus, any country whose income was not fourfold higher in 1960 than it was in 1870 grew more slowly than the United States.
  - 42 out of 125 countries in the PWT have pcy below \$1,000 in 1960.

Act I: Convergence? 1870–1979

- Claim: the $250 bound "proves" divergence over long-run.
  - The US grew @1.7% p.a., so by 4 times from 1870 to 1960.
  - Thus, any country whose income was not fourfold higher in 1960 than it was in 1870 grew more slowly than the United States.
  - 42 out of 125 countries in the PWT have pcy below $1,000 in 1960.
- Or try this:
  - extrapolate back so poorest country in 1960 hits exactly $250 in 1870.
  - US: use actual figures.
  - preserve the relative rankings of all other countries (see his footnote 11)

Act I: Convergence? 1870–1979

	1870	1960	1990
USA (P$)	2063	9895	18054
Poorest (P$)	250 (assumption)	257 (Ethiopia)	399 (Chad)
Ratio of GDP per capita of richest to poorest country	8.7	38.5	45.2
Average of seventeen "advanced capitalist" countries from Maddison (1995)	1757	6689	14845
Average LDCs from PWT5.6 for 1960, 1990 (imputed for 1870)	740	1579	3296
Average "advanced capitalist" to average of all other countries	2.4	4.2	4.5
Standard deviation of natural log of per capita incomes	.51	.88	1.06
Standard deviation of per capita incomes	P$459	P$2,112	P$3,988
Average absolute income deficit from the leader	P$1286	P$7650	P$12,662

Conditional Convergence

- Unconditional convergence assumes all parameters are the same.
  - Way too strong

Conditional Convergence

- Unconditional convergence assumes all parameters are the same.
  - Way too strong
- **Conditional convergence:**
  - Control for parameters such as s and n
  - Or any parameter that systematically varies across countries

Conditional Convergence

- Unconditional convergence assumes all parameters are the same.
  - Way too strong
- **Conditional convergence:**
  - Control for parameters such as s and n
  - Or any parameter that systematically varies across countries
- **Calibration:**
  - recall our steady state equation

$$\frac{f(\hat{k}^*)}{\hat{k}^*} \simeq \frac{n+\delta+\pi}{s},$$

  - and then move to the Cobb-Douglas production function.

Calibration

- Been there, done that, but let's review:

$$Y = AK^a(eL)^{1-a}, \text{ where } e(t) = (1+\pi)^t.$$

Calibration

- Been there, done that, but let's review:

$$Y = AK^a(eL)^{1-a}, \text{ where } e(t) = (1+\pi)^t.$$

- In effective labor units:

$$\hat{y} = \frac{Y}{eL} = \frac{AK^a(eL)^{1-a}}{eL} = A\left(\frac{K}{eL}\right)^a = A\hat{k}^a (= f(\hat{k})).$$

Calibration

- Been there, done that, but let's review:

$$Y = AK^a(eL)^{1-a},\ \text{where } e(t) = (1+\pi)^t.$$

- In effective labor units:

$$\hat{y} = \frac{Y}{eL} = \frac{AK^a(eL)^{1-a}}{eL} = A\left(\frac{K}{eL}\right)^a = A\hat{k}^a (= f(\hat{k})).$$

- Combine with steady state equation:

$$\frac{n+\delta+\pi}{s} \simeq \frac{f(\hat{k}^*)}{\hat{k}^*} = A\hat{k}^{*a-1}$$

$\Rightarrow$

$$\hat{k}^* = \left(\frac{sA}{n+\delta+\pi}\right)^{1/(1-a)} \text{ and } \hat{y}^* = A\left(\frac{sA}{n+\delta+\pi}\right)^{a/(1-a)}.$$

Calibration

$$\hat{y}^* = A\left(\frac{sA}{n+\delta+\pi}\right)^{a/(1-a)} = A^{1/(1-a)}\left(\frac{s}{n+\delta+\pi}\right)^{a/(1-a)}.$$

- It follows that if two countries have similar π, n and δ,

$$\frac{y_1}{y_2} = \left(\frac{A_1}{A_2}\right)^{1/(1-a)}\left(\frac{s_1}{s_2}\right)^{a/(1-a)}.$$

Calibration

$$\hat{y}^* = A\left(\frac{sA}{n+\delta+\pi}\right)^{a/(1-a)} = A^{1/(1-a)}\left(\frac{s}{n+\delta+\pi}\right)^{a/(1-a)}.$$

- It follows that if two countries have similar π, n and δ,

$$\frac{y_1}{y_2} = \left(\frac{A_1}{A_2}\right)^{1/(1-a)} \left(\frac{s_1}{s_2}\right)^{a/(1-a)}.$$

- a is share of capital (why?).
- 0.25 (Parente-Prescott) – 0.40 (Lucas), so $a/(1-a) \leq 2/3$.
- Doubling $s \Rightarrow$ income ratio approx $2^{2/3}$, around **60%**.
- 1990–2025, average per capita income (PPP) of richest 10% about **40+ times** corresponding figure for the poorest 10%.

Calibration

- Technology differentials give us a better chance; recall:

$$\frac{y_1}{y_2} = \left(\frac{A_1}{A_2}\right)^{1/(1-a)} \left(\frac{s_1}{s_2}\right)^{a/(1-a)} .$$

- That is, A-differences are more amplified than s-differences:

$$\frac{y_1}{y_2} = \left(\frac{A_1}{A_2}\right)^{1/(1-a)} .$$

  - When $a = 1/3$, $(A_1/A_2)^{3/2}$ as opposed to $(s_1/s_2)^{1/2}$.
  - Better, but still not close.
- **Variation in incomes just too high relative to the basic theory.**

Regression Approach

- **Deeper dive** Mankiw, Romer and Weil (QJE 1992). Steady state again:

$$\hat{y}^* = A^{1/(1-a)} \left(\frac{s}{n + \delta + \pi} \right)^{a/(1-a)} .$$

- so that

$$y(t) = A^{1/(1-a)} (1+\pi)^t \left(\frac{s}{n + \delta + \pi} \right)^{a/(1-a)} .$$

Regression Approach

- **Deeper dive** Mankiw, Romer and Weil (QJE 1992). Steady state again:

$$\hat{y}^* = A^{1/(1-a)} \left(\frac{s}{n+\delta+\pi} \right)^{a/(1-a)} .$$

- so that

$$y(t) = A^{1/(1-a)} (1+\pi)^t \left(\frac{s}{n+\delta+\pi} \right)^{a/(1-a)} .$$

- Take logarithms:

$$\ln y(t) = \left[\frac{1}{1-a} \ln A + t \ln(1+\pi) \right] + \frac{a}{1-a} \ln s - \frac{a}{1-a} \ln(n+\delta+\pi).$$

Regression Approach

$$\ln y(t) = \left[\frac{1}{1-a}\ln A + t\ln(1+\pi)\right] + \frac{a}{1-a}\ln s - \frac{a}{1-a}\ln(n+\delta+\pi).$$

- Motivates the regression that we need to run:

$$\ln y_i(t) = [C + Dt] + b_1 \ln s_i + b_2 \ln(n+\delta+\pi)_i + \epsilon_i.$$

Regression Approach

$$\ln y(t) = \left[\frac{1}{1-a}\ln A + t\ln(1+\pi)\right] + \frac{a}{1-a}\ln s - \frac{a}{1-a}\ln(n+\delta+\pi).$$

- Motivates the regression that we need to run:

$$\ln y_i(t) = [C + Dt] + b_1 \ln s_i + b_2 \ln(n+\delta+\pi)_i + \epsilon_i.$$

- And also pins down what we should expect to find:
- $b_1 > 0$, $b_2 < 0$, and $b_1 = -b_2 = a/(1-a) \simeq 0.6$.

Regression Approach

$$\ln y(t) = \left[\frac{1}{1-a}\ln A + t\ln(1+\pi)\right] + \frac{a}{1-a}\ln s - \frac{a}{1-a}\ln(n+\delta+\pi).$$

- Motivates the regression that we need to run:

$$\ln y_i(t) = [C + Dt] + b_1 \ln s_i + b_2 \ln(n+\delta+\pi)_i + \epsilon_i.$$

  - And also pins down what we should expect to find:
  - $b_1 > 0$, $b_2 < 0$, and $b_1 = -b_2 = a/(1-a) \simeq 0.6$.
- **Implementation:** take $\delta + \pi = 0.05$ (exact numbers don't matter much).
  - Regress y^{1985} on parameter averages over 1960–1985.
  - Get $b_1 = 1.42$ and $b_2 = -1.97$. Signs ok, but way too big!

But What Does a Really Mean?

- $a =$ **share of capital in national income**
  - The larger it is, the greater the spread we can calibrate or predict.
  - But a measures the share of physical capital, which is not close to 1.
  - That's the **heart of the difficulty** with the Solow model.

But What Does a Really Mean?

- $a =$ **share of capital in national income**
  - The larger it is, the greater the spread we can calibrate or predict.
  - But a measures the share of physical capital, which is not close to 1.
  - That's the **heart of the difficulty** with the Solow model.
- But other inputs, such as **human capital**, can also be accumulated.
  - Their income shares need to be considered as well.

An Example With Multiple Inputs

- **A three-input model**:

$$Y = AK^a U^b H^c$$

- where U is unskilled labor and H is educated labor.

An Example With Multiple Inputs

- **A three-input model**:

$$Y = AK^aU^bH^c$$

- where U is unskilled labor and H is educated labor.
- Divide through by U; then

$$y = Ak^ah^c$$

An Example With Multiple Inputs

- **A three-input model**:

$$Y = AK^a U^b H^c$$

  - where U is unskilled labor and H is educated labor.

- Divide through by U; then

$$y = Ak^a h^c$$

- Now there are two accumulation equations:
  - Savings: $K(t+1) = (1-\delta_k)K(t) + s_k Y(t)$
  - Education: $H(t+1) = (1-\delta_h)H(t) + s_h Y(t)$

- No technical progress for simplicity. Just divide by U; then

$$(1+n)k(t+1) = (1-\delta_k)k(t) + s_k y(t),$$

$$(1+n)h(t+1) = (1-\delta_h)h(t) + s_h y(t),$$

- No technical progress for simplicity. Just divide by U; then

$$(1+n)k(t+1) = (1-\delta_k)k(t) + s_k y(t),$$

$$(1+n)h(t+1) = (1-\delta_h)h(t) + s_h y(t),$$

- In steady state $k(t) = k(t+1) = k^*$, $h(t) = h(t+1) = h^*$, $y(t) = y^*$:

$$k^* = \frac{s_k y^*}{n+\delta_k}$$

$$h^* = \frac{s_h y^*}{n+\delta_h}$$

- No technical progress for simplicity. Just divide by U; then

$$(1+n)k(t+1) = (1-\delta_k)k(t) + s_k y(t),$$

$$(1+n)h(t+1) = (1-\delta_h)h(t) + s_h y(t),$$

- In steady state $k(t) = k(t+1) = k^*$, $h(t) = h(t+1) = h^*$, $y(t) = y^*$:

$$k^* = \frac{s_k y^*}{n+\delta_k}$$

$$h^* = \frac{s_h y^*}{n+\delta_h}$$

- Recall $y = Ak^a h^c$; combining:

$$y^* = A{k^*}^a {h^*}^c = A\left(\frac{s_k y^*}{n+\delta_k}\right)^a \left(\frac{s_h y^*}{n+\delta_h}\right)^c, \text{ or}$$

$$y^* = A^{1/(1-a-c)}\left(\frac{s_k}{n+\delta_k}\right)^{a/(1-a-c)}\left(\frac{s_h}{n+\delta_h}\right)^{c/(1-a-c)}.$$

- Take logarithms:

$$\ln y^* = \frac{\ln A}{1-a-c} + \frac{a \ln s_k}{1-a-c} + \frac{c \ln s_h}{1-a-c} - \frac{a \ln(n+\delta_k)}{1-a-c} - \frac{c \ln(n+\delta_h)}{1-a-c}.$$

- As before, motivates the regression we need to run:

$$\ln y_i = C + b_1 \ln s_{ki} + b_2 \ln s_{hi} + b_3 \ln(n+\delta_k)_i + b_4 \ln(n+\delta_h)_i + \epsilon_i.$$

- Take logarithms:

$$\ln y^* = \frac{\ln A}{1-a-c} + \frac{a \ln s_k}{1-a-c} + \frac{c \ln s_h}{1-a-c} - \frac{a \ln(n+\delta_k)}{1-a-c} - \frac{c \ln(n+\delta_h)}{1-a-c}.$$

- As before, motivates the regression we need to run:

$$\ln y_i = C + b_1 \ln s_{ki} + b_2 \ln s_{hi} + b_3 \ln(n+\delta_k)_i + b_4 \ln(n+\delta_h)_i + \epsilon_i.$$

- **Predictions**: $b_1 = \frac{a}{1-a-c}$, $b_2 = \frac{c}{1-a-c}$, and coefficient on $\ln n$ is $b_3 + b_4 = -\frac{a+c}{1-a-c}$.

- Now income differences higher than that predicted by a alone.

- Take logarithms:

$$\ln y^* = \frac{\ln A}{1-a-c} + \frac{a \ln s_k}{1-a-c} + \frac{c \ln s_h}{1-a-c} - \frac{a \ln(n+\delta_k)}{1-a-c} - \frac{c \ln(n+\delta_h)}{1-a-c}.$$

- As before, motivates the regression we need to run:

$$\ln y_i = C + b_1 \ln s_{ki} + b_2 \ln s_{hi} + b_3 \ln(n+\delta_k)_i + b_4 \ln(n+\delta_h)_i + \epsilon_i.$$

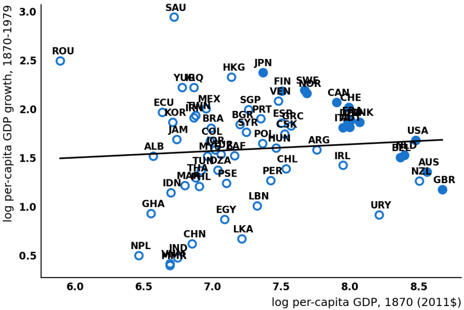
- **Predictions**: $b_1 = \frac{a}{1-a-c}$, $b_2 = \frac{c}{1-a-c}$, and coefficient on $\ln n$ is $b_3 + b_4 = -\frac{a+c}{1-a-c}$.
- Now income differences higher than that predicted by a alone.
- The coefficients on s_k and n will be larger than before.
- If $a = c = 1/3$, $b_1 = b_2 = 1$, and $b_3 + b_4 = -2$.

TABLE II
ESTIMATION OF THE AUGMENTED SOLOW MODEL

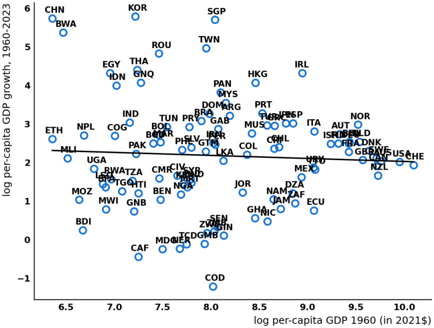
Dependent variable: log GDP per working-age person in 1985			
Sample:	Non-oil	Intermediate	OECD
Observations:	98	75	22
CONSTANT	6.89	7.81	8.63
	(1.17)	(1.19)	(2.19)
ln(I/GDP) (i.e., ln s_k)	0.69	0.70	0.28
	(0.13)	(0.15)	(0.39)
$\ln(n + g + \delta)$	−1.73	−1.50	−1.07
	(0.41)	(0.40)	(0.75)
ln(SCHOOL) (i.e., ln s_h)	0.66	0.73	0.76
	(0.07)	(0.10)	(0.29)
$\overline{R}^2$	0.78	0.77	0.24

Source: Mankiw, Romer and Weil (1992).

Act II: Modern Convergence?



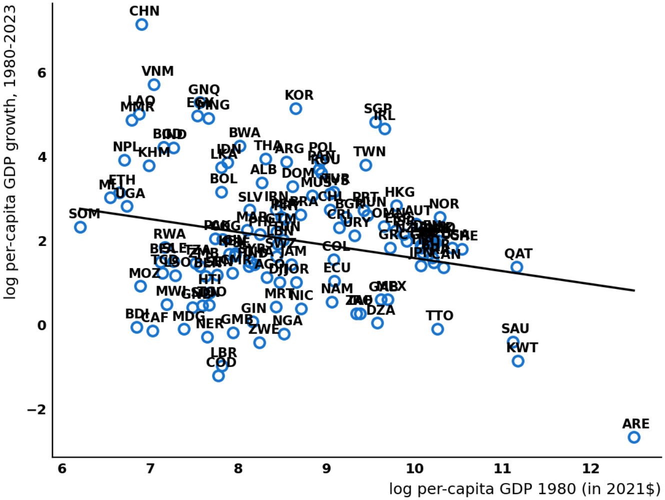
1870–1979



1960–2023

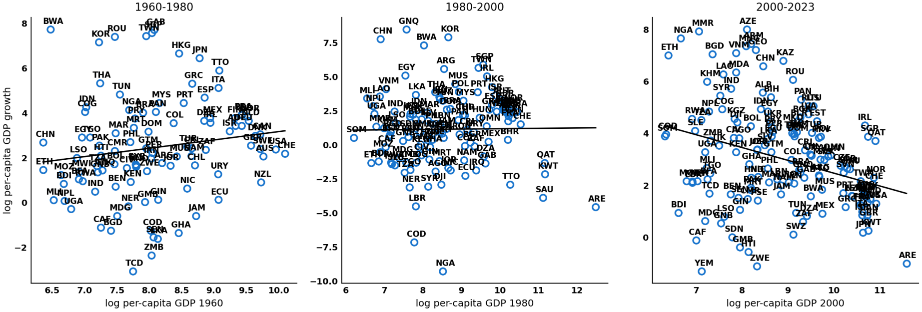
- Looks about the same, right?
- But there is a difference

Act II: Modern Convergence?

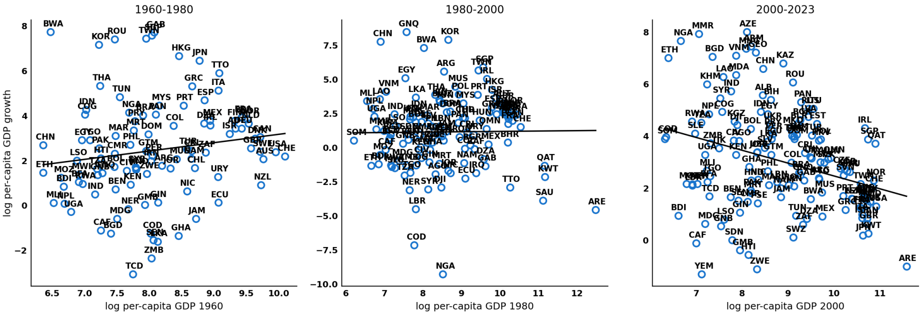


- Between 1980–2023, some evidence of **unconditional convergence**.

Act II: Modern Convergence?



Act II: Modern Convergence?



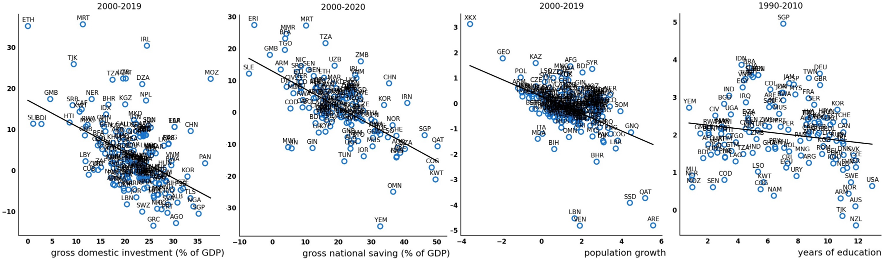
Growth Rate of GDP per capita

	1870–1979	1960–2023	1960–1980	1980–2000	2000–2023
Coefficient β	-0.068	-0.078	0.372	0.027	-0.493
	(0.110)	(0.150)	(0.241)	(0.204)	(0.123)

Regression coefficient β of per-capita GDP growth on baseline log GDP per-capita $y(s)$. Standard errors in parentheses. Sources: Maddison database, Penn World Table 11.0 (2021 US$).

Act II: Modern Convergence?

- Some recent evidence for **parametric convergence**:



- These are hopeful signs, but it is way too early to be sure:
  - The above parameters are only the basic Solow parameters
  - Institutions move far more slowly
  - We’ve had divergence for far too long. 20 years does not fully reverse that.

What Does This Exercise Achieve?

- No evidence for unconditional convergence:
- Until very recently
- Otherwise, convergence contradicts the facts

What Does This Exercise Achieve?

- No evidence for unconditional convergence:
  - Until very recently
  - Otherwise, convergence contradicts the facts
- Conditional convergence does a lot better:
  - Conditioning on s_k, s_h and n.

What Does This Exercise Achieve?

- No evidence for unconditional convergence:
  - Until very recently
  - Otherwise, convergence contradicts the facts
- Conditional convergence does a lot better:
  - Conditioning on s_k, s_h and n.
- But we should be trying to understand variations in those parameters to begin with.
  - Turtles (?!)
  - Later: theories of divergence.