Development From the Viewpoint of Convergence

1. Stylized Facts of Economic Growth

1.1 A Long-run View

- A central fact of aggregate economic activity across countries:
  - massive divergence in living standards occurring over the last several centuries.
- A snapshot of the world in 1700 would show all countries to be poor assessed by today's living standards.
- Over 18th and 19th centuries, growth rates increased slightly in the UK and other countries in Western Europe.
  - Annual growth rates remained low by modern standards even in the midst of the Industrial Revolution.
  - But since this growth was sustained over time, GDP per capita rose steadily.
- Outcome: the UK, countries in Western Europe and their offshoots – Australia, New Zealand, Canada, USA – gradually advanced further ahead of the rest of the world.
- What was happening elsewhere?

- Pritchett (1997) argues:

  Even in the absence of National accounts data, we can be almost certain that rapid productivity growth was never sustained in the poorer regions of the world.

  – The argument proceeds by extrapolating backwards from their current levels of GDP per capita, using a fast growth rate.

  – This quickly implies earlier levels of income that would be too low to support human life.

- Today's overall inequality across countries is partly the legacy of rapid growth in a small group of Western economies, and its absence elsewhere.

- There have been important deviations from this general pattern.

  – Some developing countries have grown at rates that are unprecedented (based on the experiences of the advanced economies of Europe and North America).

    ○ The tiger economies of East Asia have seen GDP per worker grow at around 5% a year for the best part of forty years.

  $\Rightarrow$ GDP per worker has risen more than sevenfold as in Hong Kong, Singapore, South Korea and Taiwan.

1.2 Differences in Levels of GDP per Worker

- Table 1 shows GDP per worker, relative to the USA, for 1960 and 2000.
  - The major economies of Western Europe have either maintained their positions relative to the USA (the UK) or substantially improved it (France, Italy, Spain).
  - Among the poorer nations (in 1960),
    - some countries have improved their relative positions dramatically (Japan, South Korea, Thailand);
    - others have performed badly (Argentina, Nigeria).
  - The mean and median of relative GDP per worker has shown a moderate increase.
- Figure 1 shows the kernel density plot of the distribution of GDP per worker in 1960 and 2000, relative to the USA's 1960 value as the benchmark.
  - The rightwards movement reflects the growth that took place over this period.
  - Also noticeable is a thinning in the middle of the distribution:

    the "Twin Peaks" phenomenon identified in Quah (1993a, 1993b, 1996a, 1996b, 1996c, 1997).

Table 1
International disparities in GDP per worker

Country	Population (m, 2000)	R1960	R2000
USA	275	1	1
United Kingdom	60	0.69	0.69
Argentina	37	0.62	0.40
France	60	0.60	0.76
Italy	58	0.55	0.84
South Africa	43	0.47	0.34
Mexico	97	0.44	0.38
Spain	40	0.40	0.68
Iran	64	0.30	0.30
Colombia	42	0.27	0.18
Japan	127	0.25	0.60
Brazil	170	0.24	0.30
Turkey	67	0.17	0.24
Philippines	76	0.17	0.13
Egypt	64	0.17	0.21
Korea, Republic of	47	0.15	0.57
Bangladesh	131	0.10	0.10
Nigeria	127	0.08	0.02
Indonesia	210	0.08	0.14
Thailand	61	0.07	0.20
Pakistan	138	0.07	0.11
India	1016	0.06	0.10
China	1259	0.04	0.10
Ethiopia	64	0.04	0.02
Mean		0.29	0.35
Median		0.21	0.27

Note: R is GDP per worker as a fraction of that in the USA.

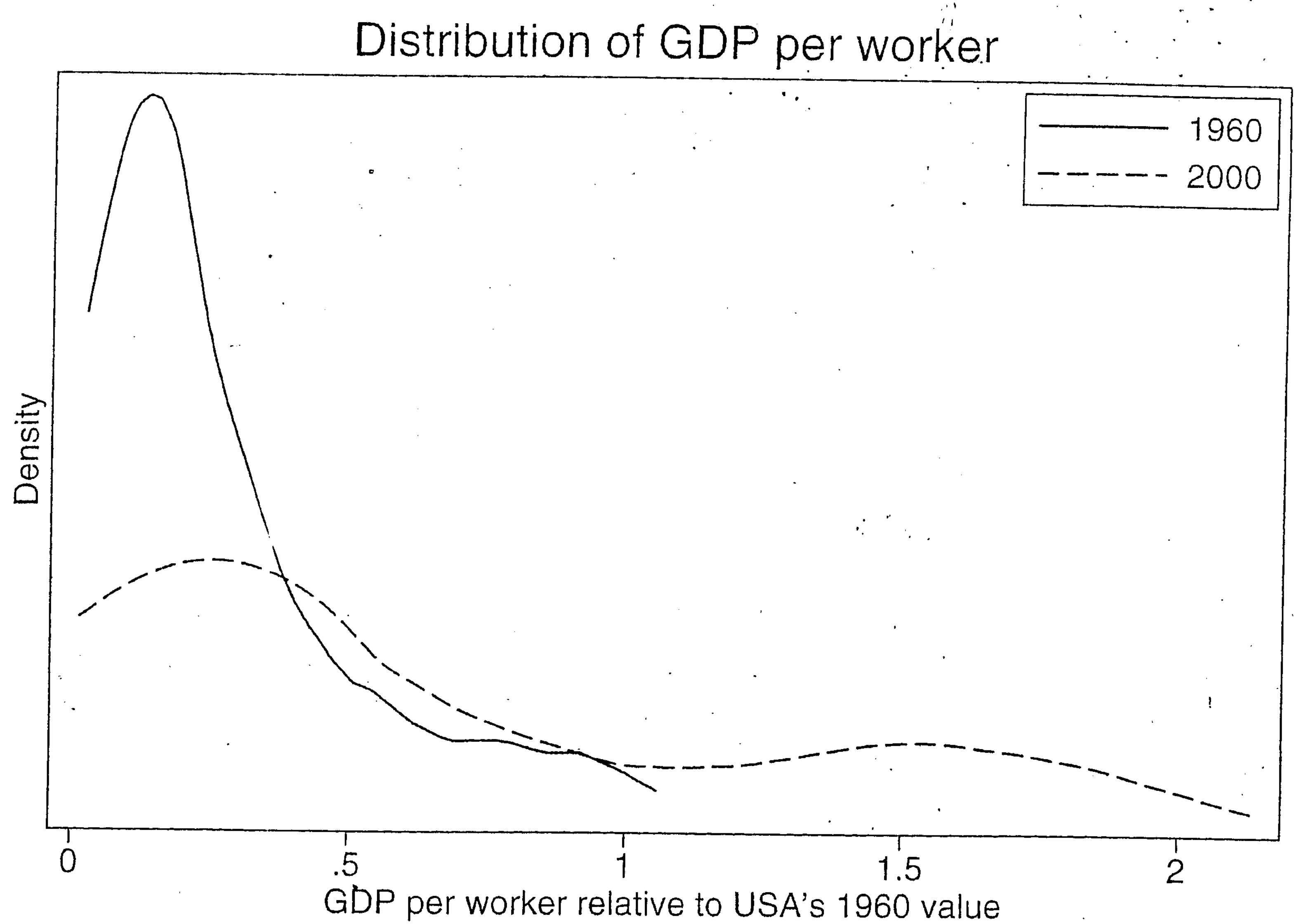
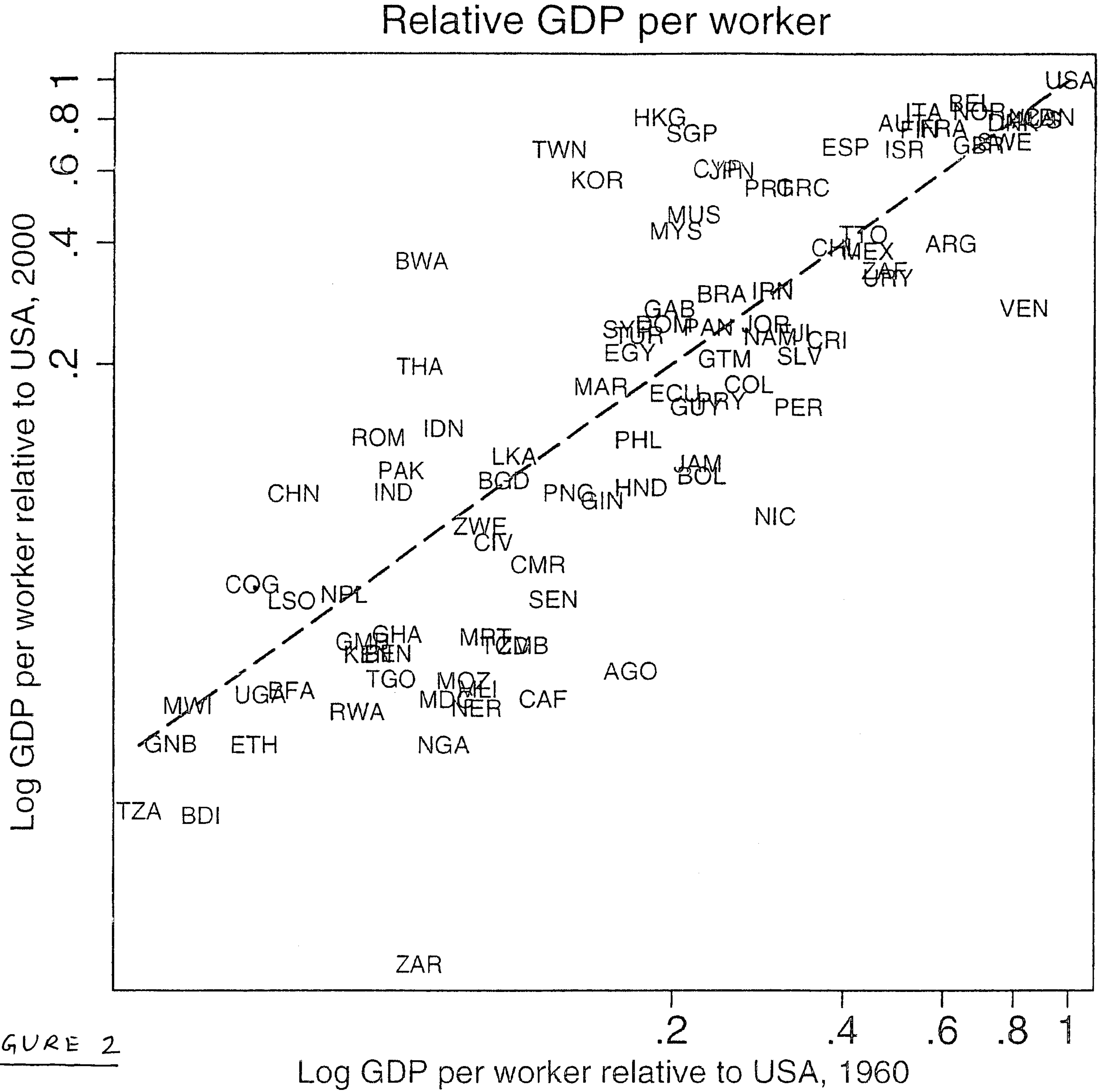


Figure 1. Cross-country density of output per worker.

- Is the position in the league table of GDP per worker in 1960 a good predictor of that in 2000?
  - The Spearman rank correlation is 0.84.
  - This pattern is shown in more detail in Figure 2.



Relative GDP per worker
Log GDP per worker relative to USA, 2000
Log GDP per worker relative to USA, 1960
.2
.4
.6
.8
1
USA
HKG
SGP
TWN
KOR
ESP
ISR
ARG
VEN
BWA
THA
ROM
IDN
PAK
IND
CHN
LKA
BGD
PHL
NIC
CMR
SEN
AGO
CAF
RWA
NGA
ETH
GNB
TZA
BDI
ZAR

FIGURE 2

1.3 Growth Miracles and Disasters

- Despite some stability in relative positions, it is easy to pick out countries that have done exceptionally well and others that have done badly.
- There is an enormous range in observed growth rates between 1960 and 2000.
  - Tables 2 and 3 present a list of 15 best and worst performers, respectively.
- Growth miracles and disasters show a regional pattern:
  - The best performing countries are mainly located in East Asia and Southeast Asia.
  - Growth disasters – many instances of “negative growth” – occur predominantly in countries in sub-Saharan Africa.

Table 2
Fifteen growth miracles, 1960–2000

Country	Growth 1960–2000	Factor increase
Taiwan	6.25	11.3
Botswana	6.07	10.6
Hong Kong	5.67	9.09
Korea, Republic of	5.41	8.24
Singapore	5.09	7.29
Thailand	4.50	5.83
Cyprus	4.30	5.39
Japan	4.13	5.04
Ireland	4.10	5.00
China	3.99	4.77
Romania	3.91	4.63
Mauritius	3.88	4.58
Malaysia	3.82	4.48
Portugal	3.48	3.93
Indonesia	3.34	3.72

Table 3
Fifteen growth disasters, 1960–2000

Country	Growth 1960–2000	Ratio
Peru	0.00	1.00
Mauritania	−0.11	0.96
Senegal	−0.26	0.90
Chad	−0.43	0.84
Mozambique	−0.50	0.82
Madagascar	−0.60	0.79
Zambia	−0.61	0.78
Mali	−0.77	0.74
Venezuela	−0.88	0.70
Niger	−1.03	0.66
Nigeria	−1.21	0.62
Nicaragua	−1.30	0.59
Central African Republic	−1.56	0.53
Angola	−2.04	0.44
Congo, Democratic Rep.	−4.00	0.20

1.4 Convergence?

- Figure 3 plots the growth rates over 1960-2000 against the 1960 level of real GDP per worker, relative to the USA.
  - The most obvious lesson is the diversity of growth rates, especially at low levels of development.
  - Does not provide much support for the idea that countries are converging to a common level of income.
    - That would require evidence of a downward sloping relationship between growth and initial income.

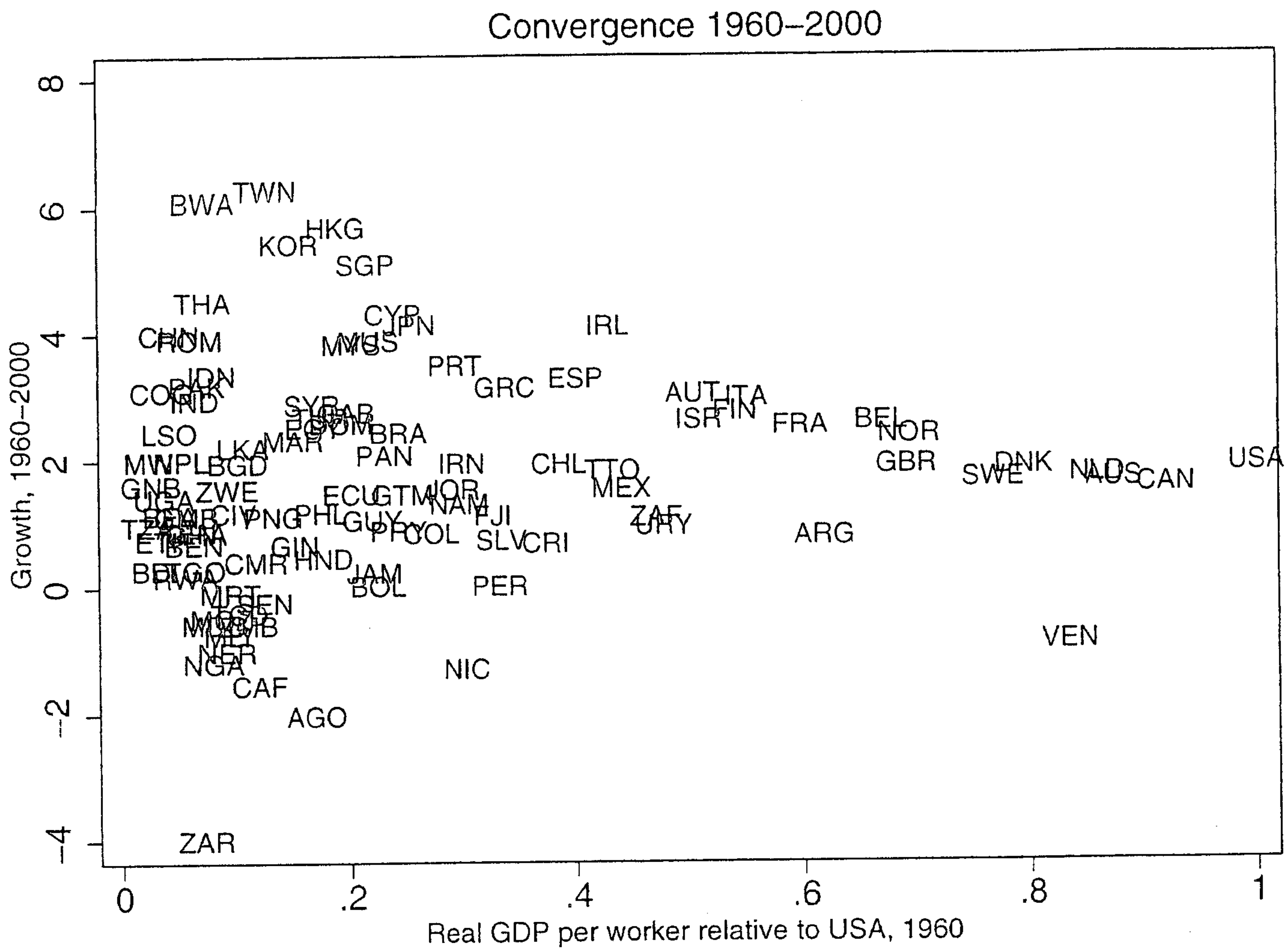


Figure 3. Growth versus initial income: 1960–2000.

1.5 The Growth Slowdown

- Figure 6 shows kernel density estimates of international distribution of growth rates for the two subperiods: 1960-1980 and 1980-2000.
  - The mass of the distribution has shifted leftwards $\Rightarrow$ slower growth.
  - The variance has increased.
- Figure 7 highlights the growth slowdown by plotting the growth rates in 1980-2000 against that in 1960-1980.
  - Countries above the 45 degree line have seen growth increase, below $\rightarrow$ growth decline.
  - Clearly more countries where growth has declined over time.
    - Crucial exceptions: China and India, seen dramatic improvements.

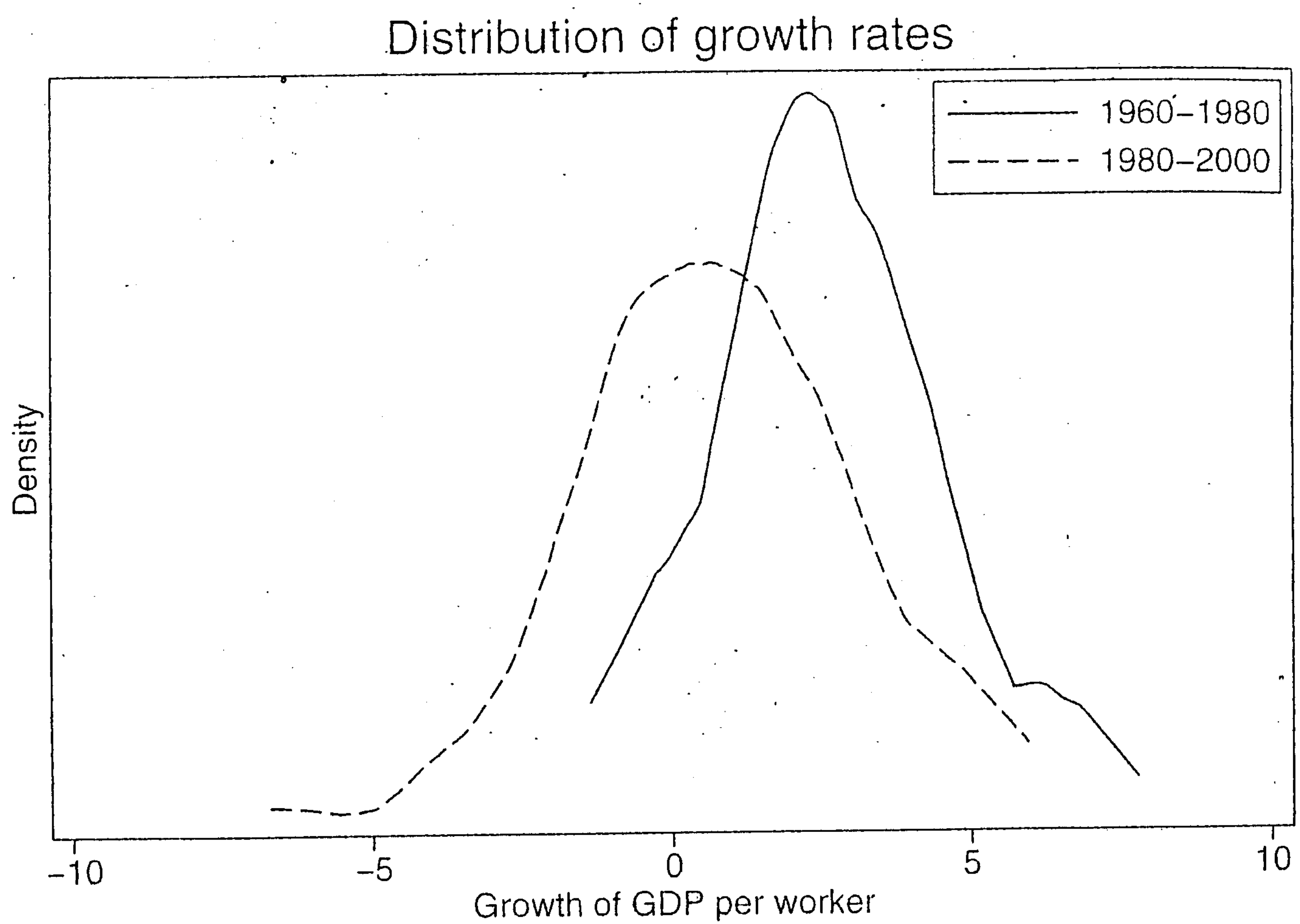


Figure 6. Density of growth rates across countries.

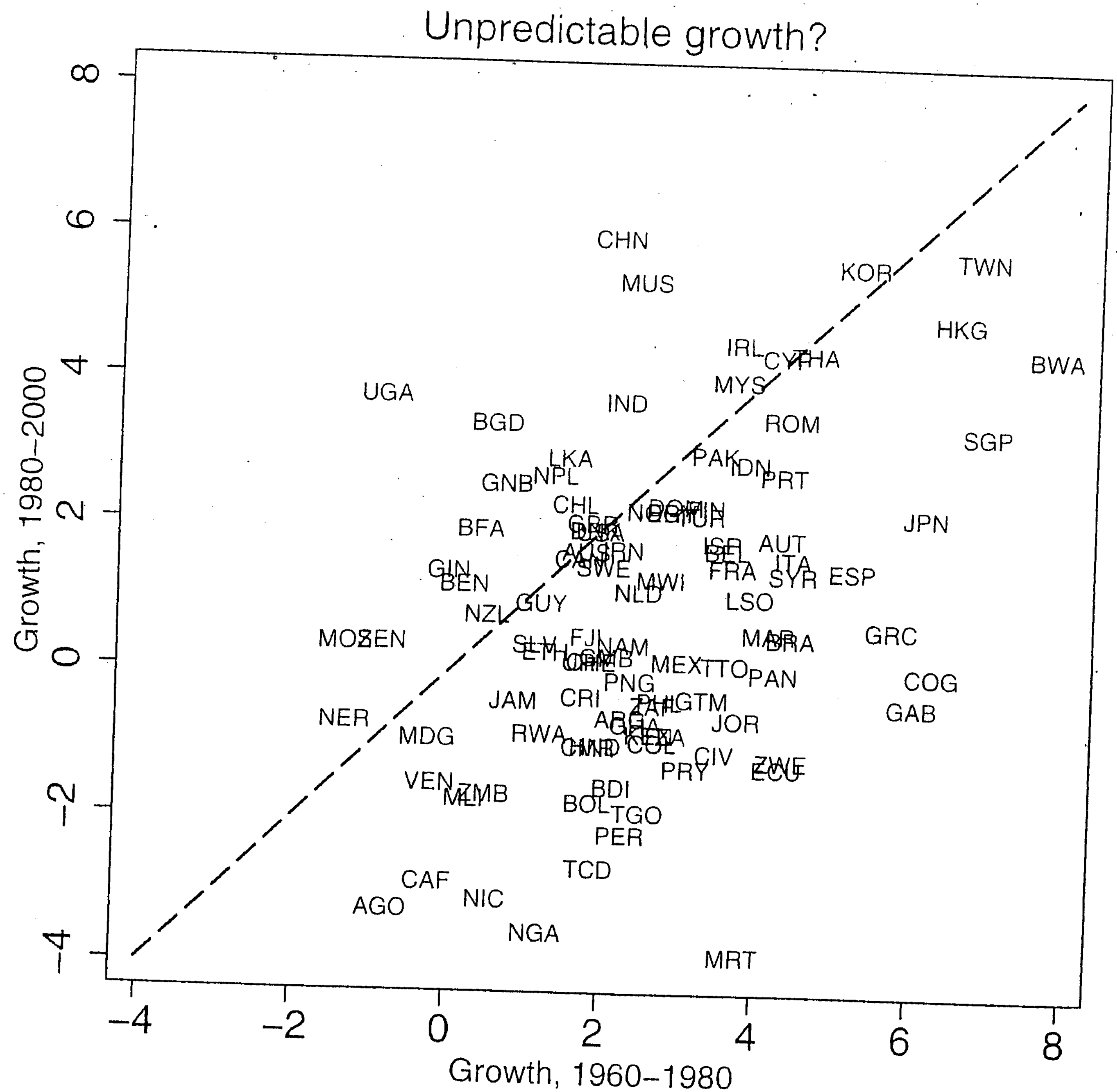


Figure 7. Growth rates in 1960–1980 versus 1980–2000.

1.6 Growth Differences by Development Level and Geographic Region

- Table 6 shows the 25th, 50th and 75th percentiles of the distribution of growth rates for countries at various levels of income relative to the USA in 1960.
  - Take the 22 countries which began between 5% and 10% of GDP per worker.
    - Growth rate at the 25th percentile is negative, but is 2.9% at the 75th percentile.
    - This diversity of experience shows up throughout the distribution; less pronounced for the richest group.
- Table 7 shows the quartiles of growth rates for countries in different regions.
  - Sub-Saharan Africa is the weakest performer.
    - Within sub-Saharan Africa, even the country at the 75th percentile shows growth of just 1.3%.
  - Performance is slightly better for South and Central America, but still quite weak.
  - Against this background, the record of East and Southeast Asia looks all the more remarkable.

Table 6
Growth, 1960–2000, by initial relative income

Percentile	N	25th	Median	75th
All	102	0.7	1.6	2.7
Relative income:				
$R \leqslant 0.05$	10	1.0	1.5	2.4
$R > 0.05$ & $R \leqslant 0.10$	22	−0.5	0.9	2.9
$R > 0.10$ & $R \leqslant 0.25$	33	0.4	1.9	2.7
$R > 0.25$ & $R \leqslant 0.50$	19	0.8	1.5	3.1
$R > 0.50$	18	1.6	1.9	2.6

Notes: This table shows the 25th, 50th and 75th percentiles of the distribution of growth rates for countries at various levels of development in 1960.
R is GDP per worker in 1960 relative to the US level.

Table 7
Growth, 1960–2000, by country groups

Group	N	25th	Median	75th
Sub-Saharan Africa	36	−0.5	0.7	1.3
South and Central America	21	0.4	0.9	1.5
East and Southeast Asia	10	3.8	4.3	5.4
South Asia	7	1.9	2.2	2.9
Industrialized countries	19	1.7	2.4	3.0

Note: This table shows the 25th, 50th and 75th percentiles of the distribution of growth rates for various groups of countries.

1.7 Stagnation and Output Volatility

- Some countries did not record fast growth even in the boom of the 1960s; some have simply stagnated or declined.
  - Out of the 102 countries in the Penn World Table (PWT 6.1)
    - 9 have never exceeded their 1960 level of GDP per worker by more than 30%;
    - 26 never exceeded their 1960 level by more than 60%.
  - [Note: 2% annual growth rate over 40 years $\Rightarrow$ GDP per worker increase by 120%.]
- Major collapse in output is quite common in less developed countries.
  - Calculate the largest percentage drop in output over 3 years recorded for each country using data from 1960 to 2000:

$$100 \cdot \left[1 - \min \left(\frac{Y_{1963}}{Y_{1960}}, \frac{Y_{1964}}{Y_{1961}}, \ ..., \ \frac{Y_{2000}}{Y_{1997}}\right)\right].$$

  - Table 8 shows the largest ten output falls.
    - Several collapses associated with intense civil war (Rwanda, Angola, Congo).

Table 8
Output collapses

Country	Largest 3-year drop	Dates
Chad	50%	1980–83
Rwanda	47%	1991–94
Angola	46%	1973–76
Romania	37%	1977–80
Dem. Rep. Congo	36%	1992–95
Mauritania	34%	1985–88
Tanzania	34%	1987–90
Mali	34%	1985–88
Cameroon	33%	1987–90
Nigeria	32%	1997–00

Note: This table shows the ten countries with the largest output collapses over a three-year period, using data on GDP per worker between 1960 and the latest available year.

- Output collapses are more widespread than explained by extreme events.
  - 50 LDC's showed at least one 3-year output collapse of 15% or more;
  - 65 LDC's experienced at least one collapse of 10% or more.

- Contrast with the largest 3-year output collapse (between 1960 and 2000) in developed economies:
  - USA: 5.4% (1979-82)
  - UK: 3.6% (1979-82)

- Finally consider some evidence of long-run output volatility.

- Table 9 reports standard deviation of annual growth rates between 1960 and 2000.
  - Industrialized countries are relatively stable.
  - Sub-Saharan Africa is by far the most volatile region, followed by South and Central America.

Table 9
Volatility, 1960–2000, by regions

Group	*N*	25th	Median	75th
Sub-Saharan Africa	36	5.5	7.4	9.3
South and Central America	21	3.9	4.8	5.4
East and Southeast Asia	10	3.8	4.1	4.7
South Asia	7	3.0	3.3	5.2
Industrialized countries	19	2.3	2.9	3.5

Note: This table shows the 25th, 50th and 75th percentiles of the distribution of the standard deviation of annual growth rates, using data from the earliest available year until the latest available, between 1960 and 2000.

2. Cross-country Growth Regressions: From Theory to Empirics

- The stylized facts of economic growth have led to two major themes in the development of formal econometric analyses of growth.

1. Convergence: Are contemporary differences in aggregate economies transient over sufficiently long time horizons?
2. Identification of growth determinants: Which factors seem to explain observed differences in growth?

- These two questions are closely related.
  - Each requires the specification of a statistical model of cross-country growth differences from which the effects on growth of various factors may be identified.

2.1 Growth Dynamics: Basic Ideas

- Develop a general theoretical framework for understanding growth dynamics.
  - The framework is explicitly neoclassical and represents the basis for most empirical works on growth.
- Consider economy i at time t.
  - $Y_{i,t}$: output
  - $L_{i,t}$: labour force
    - $L_{i,t} = L_{i,0} \cdot e^{n_i t}$, n_i is constant population growth rate
  - $A_{i,t}$: efficiency level of each worker
    - $A_{i,t} = A_{i,0} \cdot e^{g_i t}$, g_i is constant rate of labour augmenting technological progress
  - Two main per capita notions:
    - $y^E_{i,t} = \dfrac{Y_{i,t}}{A_{i,t} \cdot L_{i,t}}$: output per efficiency unit of labour input
    - $y_{i,t} = \dfrac{Y_{i,t}}{L_{i,t}}$: output per labour unit

- The generic one-sector growth model, in either its Solow-Swan or Ramsey-Cass-Koopmans variant, implies, to a first-order approximation, that

$$\log\ y_{i,t}^{E} = \left(1 - e^{-\lambda_i t}\right) \log\ y_{i,\infty}^{E} + e^{-\lambda_i t} \log\ y_{i,0}^{E}. \tag{1}$$

  - $y_{i,\infty}^{E}$: steady-state value of $y_{i,t}^{E}$ and $\lim_{t\to\infty} y_{i,t}^{E} = y_{i,\infty}^{E}$
  - The parameter λ_i (must be positive) measures the *rate of convergence* of $y_{i,t}^{E}$ to its steady-state value.
    - Depends on the other parameters of the model.
  - Given $\lambda_i > 0$, the value of $y_{i,\infty}^{E}$ is independent of $y_{i,0}^{E}$ so that *initial conditions do not matter in the long-run*.
- [For the derivation of (1) refer to Barro and Sala-i-Martin (2004):
  - footnote (26), page 58, for the Solow-Swan model, and
  - equation (2.40), page 111 and Appendix 2A, for the Ramsey-Cass-Koopmans model.]

- Growth dynamics in terms of the observable variable $y_{i,t}$:

$$\log\ y_{i,t} = g_i t + \left(1 - e^{-\lambda_i t}\right) \log\ y_{i,\infty}^E + \left(1 - e^{-\lambda_i t}\right) \log\ A_{i,0} + e^{-\lambda_i t} \log\ y_{i,0}. \tag{2}$$

- This description of the dynamics of output provides the basis for describing the dynamics of growth.
  - $\gamma_i \equiv \dfrac{\log\ y_{i,t} - \log\ y_{i,0}}{t}$: growth rate of output per worker between 0 and t.
  - (2) implies

$$\gamma_i = g_i + \beta_i \left(\log\ y_{i,0} - \log\ y_{i,\infty}^E - \log\ A_{i,0}\right), \tag{3}$$

  where $\beta_i = -\dfrac{1}{t}\left(1 - e^{-\lambda_i t}\right) < 0$.

- (3) decomposes the growth rate in country i into two distinct components:

1. g_i: measures growth due to technological progress,
2. $\beta_i \left(\log\ y_{i,0} - \log\ y_{i,\infty}^E - \log\ A_{i,0}\right)$: measures growth due to the gap between initial output per worker and the steady-state value, both measured in terms of efficiency units of labour.

- ○ This second source of growth is what is meant by “catching up” in the literature.
- ○ As $t \to \infty$ the importance of the catch-up term, which reflects the role of initial conditions, diminishes to zero.

- If the rates of technological progress and the λ_i parameters are assumed to be the same across countries, $g_i = g$ and $\lambda_i = \lambda, \forall i,$ (3) may be rewritten as

$$\gamma_i = g + \beta \log\ y_{i,0} - \beta \log\ y^E_{i,\infty} - \beta \log\ A_{i,0}. \tag{4}$$

– In a cross-section of countries, we should observe a negative relationship between average rates of growth and initial level of output over any time period.

– Countries that start out below their balanced growth path must grow relatively quickly if they are to catch up with other countries that have the same levels of steady-state output per effective worker and initial efficiency.

– This is closely related to the *conditional convergence* hypothesis:

Countries converge to parallel growth paths the levels of which are assumed to be a function of a small set of variables.

2.2 Cross-country Growth Regressions

- Equation (4) provides the motivation for the standard cross-country growth regression.
  - Typically, these regression specifications start with (4) and append a random error term υ_i:

$$\gamma_i = g + \beta \log\ y_{i,0} - \beta \log\ y^E_{i,\infty} - \beta \log\ A_{i,0} + \upsilon_i. \tag{5}$$

- Implementation of (5) requires the development of empirical analogs for $\log y^E_{i,\infty}$ and $\log A_{i,0}$.
- **Mankiw, Romer and Weil (1992)**, a pioneering analysis, shows how to do this in a way that produces a growth regression model that is linear in observable variables.
- Three-factor Cobb-Douglas production function:

$$Y_{i,t} = K^{\alpha}_{i,t} H^{\phi}_{i,t} \left(A_{i,t} L_{i,t}\right)^{1-\alpha-\phi},$$

  - $K_{i,t}$: physical capital
  - $H_{i,t}$: human capital

- Physical and human capital accumulation equations:

$$\dot{K}_{i,t} = s_{K,i} Y_{i,t} - \delta K_{i,t},$$
$$\dot{H}_{i,t} = s_{H,i} Y_{i,t} - \delta H_{i,t},$$

  – δ: depreciation rate

  – $s_{K,i}$: saving rate for physical capital

  – $s_{H,i}$: saving rate for human capital

$\Rightarrow$ The steady-state value of output per effective worker is

$$y_{i,\infty}^{E} = \left[\frac{s_{K,i}^{\alpha} s_{H,i}^{\phi}}{(n_i + g + \delta)^{\alpha+\phi}} \right]^{\frac{1}{1-\alpha-\phi}} .$$

- Using this $y_{i,\infty}^{E}$ in (5) produces a cross-country growth regression of the form

$$\gamma_i = g + \beta \log \ y_{i,0} + \beta \frac{\alpha+\phi}{1-\alpha-\phi} \log \ (n_i + g + \delta) - \beta \frac{\alpha}{1-\alpha-\phi} \log \ s_{K,i} \\ -\beta \frac{\phi}{1-\alpha-\phi} \log \ s_{H,i} - \beta \log \ A_{i,0} + v_i. \tag{6}$$

- Mankiw, Romer and Weil (1992) assume that
  - $(g+\delta)$is known;
  - $A_{i,0}$ is unobservable, and should be interpreted as reflecting not just technology, but country-specific influences on growth such as resource endowments, climate and institutions.
    - These differences vary randomly:

$$\log A_{i,0} = \log A + e_i,$$

e_i: country-specific shock distributed independently of n_i, $s_{K,i}$, $s_{H,i}$.

- Substituting into (6) and defining $\varepsilon_i = \upsilon_i - \beta e_i$, we have the regression relationship

$$\begin{aligned} \gamma_i = g - \beta \log\ A + \beta \log\ y_{i,0} + \beta \frac{\alpha+\phi}{1-\alpha-\phi} \log\ (n_i + g + \delta) \\ - \beta \frac{\alpha}{1-\alpha-\phi} \log\ s_{K,i} - \beta \frac{\phi}{1-\alpha-\phi} \log\ s_{H,i} + \varepsilon_i. \end{aligned} \quad (7)$$

- Using data from 98 countries over 1960-1985, Mankiw, Romer and Weil estimates

$$\hat{\beta} = -0.299,\ \hat{\alpha} = 0.48,\ \text{and}\ \hat{\phi} = 0.23.$$

- Many cross-country regression studies have attempted to extend Mankiw, Romer and Weil by adding additional control variables Z_i to equation (7).
  - Such studies may be understood as allowing for predictable heterogeneity in the steady-state growth term g_i and initial technology term $A_{i,0}$ that are assumed constant across i in (7):
    - The $(g_i - \beta \log A_{i,0})$ term in (3) are replaced with $(g - \beta \log \ A + \pi Z_i - \beta e_i)$ rather than with $(g - \beta \log \ A - \beta e_i)$.
- This produces the cross-country growth regression

$$\gamma_i = g - \beta \log \ A + \beta \log \ y_{i,0} + \beta \frac{\alpha + \phi}{1 - \alpha - \phi} \log \ (n_i + g + \delta) \\ -\beta \frac{\alpha}{1 - \alpha - \phi} \log \ s_{K,i} - \beta \frac{\phi}{1 - \alpha - \phi} \log \ s_{H,i} + \pi Z_i + \varepsilon_i. \quad (8)$$

- The canonical cross-country growth regression may be understood as a version of (8) when the cross-coefficient restrictions embedded in (8) are ignored.
  - Usually the case in most empirical works.

- A generic representation of the **canonical cross-country growth regression**:

$$\gamma_i = \beta \log \ y_{i,0} + \psi X_i + \pi Z_i + \varepsilon_i, \tag{9}$$

  - X_i contains a constant, $\log \left(n_i + g + \delta\right)$, $\log s_{K,i}$ and $\log s_{H,i}$.
    - X_i and $\log y_{i,0}$ represent those growth determinants suggested by Solow model.
  - Z_i represents growth determinants outside Solow's original theory.
- Distinction between Solow variables and Z_i is important in understanding the empirical literature:
  - Solow variables appear in different empirical studies.
    - Reflects treatment of the Solow model as a baseline for growth analysis.
  - Choices concerning which Z_i variables to include vary widely.
- Regressions represented by (9) are sometimes referred to as *Barro regressions*.
  - Barro has extensively used such regressions to study alternative growth determinants starting with Barro (1991).
- This regression model has been the workhorse of empirical growth research.

3. The Convergence Hypothesis

- The effect of initial conditions on long-run outcomes represents one of the primary empirical questions explored by growth economists.
  - The claim that effects of initial conditions eventually disappears is the heuristic basis for what is known as the convergence hypothesis.
- This literature aims to answer two questions:
  1. Are cross-country differences in per-capita incomes *temporary* or *permanent*?
  2. If they are permanent, does that permanence reflect *structural heterogeneity* or the *role of initial conditions* in determining long-run outcomes?
- If per-capita income differences are
  - temporary, *unconditional convergence* (to a common long-run level) is occurring;
  - permanent solely because of cross-country structural heterogeneity, *conditional convergence* is occurring;
  - permanent at least in part determined by initial conditions, and countries with similar initial conditions exhibit similar long-run outcomes → *convergence clubs*.

- Statistical analyses of convergence have largely focused on the properties of β in regressions of the form (9).
- **$\boldsymbol{\beta}$-convergence** is defined as $\beta < 0$:

  If two countries have common steady-state determinants and are converging to a common balanced growth path, the country that begins with a relatively low level of initial income per capita will tend to grow faster.

  - The poorer country has a lower capital-labour ratio, and, by the law of diminishing returns, has a higher marginal product of capital.

  $\Rightarrow$ A given addition to the capital stock will have a larger impact on per-capita income translating into relatively faster growth.

- There now exists a large body of studies of β-convergence, differentiated by country set, time period and choice of control variables.
- Typically, the **unconditional $\boldsymbol{\beta}$-convergence** hypothesis is supported when applied to data from relatively homogeneous group of economic units such as

- the states of the US,
- the OECD,
- the regions of Europe.

- In contrast there is generally no correlation between initial income and growth for data taken from more heterogeneous groups such as a broad sample of countries of the world.

- In moving from unconditional to **conditional β-convergence**, complexities arise in terms of the specification of steady-state income.
  - Reason is the dependence of the steady-state on Z.
  - Theory is not always a good guide in the choice of elements of Z.

- Differences in formulations of (9) have led to a “growth regression industry” as researchers have added plausibly relevant variables to the baseline Solow specification.
  - variants of (9) exist where convergence occurs as $\hat{\beta} < 0$;
  - variants where divergence appears to occur, $\hat{\beta} > 0$.

- One class of efforts to address choice of control variables has led to confirmatory evidence of conditional β-convergence.
  - This approach assigns probabilities to alternative formulations of (9) and uses these probabilities to construct statements about β that average across different models.
  - Doppelhofer, Miller and Sala-i-Martin (2004) conclude that the posterior probability that initial income is part of the linear growth model is 1.
  - Fernandez, Ley and Steel (2001a) also find this posterior probability to be 1 despite using a different set of potential models and different priors on model parameters.
- Conclusion: Evidence for conditional β-convergenc appears to be robust with respect to choice of controls.

4. Determinants of Growth

- One primary focus of empirical growth papers may be thought of as a general exploration of potential growth determinants.

$\Rightarrow$ Identification of appropriate variables to include in linear growth regressions.

$\rightarrow$ Specification of Z in equation (9).

- Appendix B of Durlauf, Johnson and Temple (2005) provides a survey table of different regressors that have been proposed in the growth literature.
  - The table contains 145 different regressors, the vast majority of which have been found to be statistically significant using conventional standards.
  - The number of growth regressors approaches the number of countries available in even the broadest samples!
    - This regressor list does not consider cases where interaction between variables or nonlinear transformations of variables have been included as regressors.

- This plethora of potential growth regressors illustrates one of the fundamental problems with empirical growth research:

  the absence of any consensus on which growth determinants ought to be included in a growth model.

- Brock and Durlauf (2001a) identifies the root cause:

  the growth theories are ‘open-ended’.

  – Theory open-endedness means that the growth theories are typically compatible with one another.

    ◦ For example, a theory that institutions matter for growth is not logically inconsistent with a theory that emphasizes the role of geography in growth.

  $\Rightarrow$ If one has a set of K potential growth theories, all of which are logically compatible with one another, then there exist $2^K - 1$ potential theoretical specifications of the form (9).

- For different approaches to resolving this problem of model uncertainty refer to Durlauf, Johnson and Temple (2005), section 5.1.

5. Sources of Convergence or Divergence

- Two sources have been scrutinized in the literature:
  1. Factor accumulation,
  2. Total factor productivity (TFP).
- **Factor Accumulation:**

  The "neoclassical revival" (Barro, 1991; Mankiw, Romer and Weil, 1992) adopts the view that convergence is driven by diminishing returns to factors of production.

  – If each country has access to the same aggregate production function, long-run differences in output reflect differences in the determinants of accumulation, not differences in technology.

  – Even if the assumption of same production function is relaxed, convergence in growth rates can still occur so long as
    - each country's production function is concave in capital per efficiency unit of labour, and
    - each country experiences the same labour-augmenting technical change.

- **Total Factor Productivity (TFP):**

  Klenow and Rodriguez-Clare (1997a) challenge this “neoclassical revival” with results suggesting that

  - differences in factor accumulation are, at best, no more important than differences in productivity in explaining cross-country distribution of output per capita.
  - They find:
    - only about half of the cross-country variation in 1985 level of output per worker is due to variation in human and physical capital inputs,
    - just about 10% of the variation in growth rates from 1960 to 1985 reflects differences in the growth of these inputs.
  - Prescott (1998), Hall and Jones (1999) and Caselli (2005) confirm the view that differences in inputs are unable to explain observed differences in output.
  - Easterly and Levine (2001): “the ‘residual’ (total factor productivity, TFP) rather than factor accumulation accounts for most of the income and growth differences across countries”.

- For a critical evaluation of sources of convergence or divergence refer to section 4.5 of Durlauf, Johnson and Temple (2005) which concludes as follows:
  - Despite the concerns and differences in the precise estimates found by different researchers, it is clear that cross-country variation in inputs falls short of explaining the observed cross-country variation in output.
  - The result that the TFP residual, a "measure of our ignorance", is an important (perhaps the dominant) source of cross-country differences in long-run economic performance is useful but hardly satisfying.
    - Need for a theory of TFP expressed by Prescott (1998) is well founded.

6. Criticism of the Conventional Growth Theory

- Banerjee and Duflo (2005) is of the view that the failure of the conventional growth theory to explain the lack of convergence is intimately tied to the failure of the assumptions that underlie the construction of aggregate production function.
  - Two implications of the neoclassical model are at the root of the convergence result:
    1. Rates of returns should be higher in poor countries;
    2. Investment rates should be higher in poor countries.
    - Citing a number of studies on physical capital, human capital and health, Banerjee and Duflo show that neither rates of return nor investment are, on average, much higher in poor countries.
  - Convergence relies on diminishing returns to capital $\rightarrow$ share of capital in national income does give us rough estimates of the concavity of production in capital.
    - Problem: The resulting concavity understates observed variation in cross-country income by orders of magnitude.

- · Parente and Prescott (2000) calibrate a basic Cobb-Douglas production function by using reasonable estimates of the share of capital in national income (0.25).
- · Huge variations in savings rate do not change world income by much:

  Doubling the savings rate leads to a change in steady-state income by a factor of 1.25, inadequate to explain an observed range of around 20:1 (PPP).

– Lucas (1990) observes that the discrepancy actually appears in a more primitive way, at the level of the production function.

○ For the same simple production function to fit the data on per-capita income differences, a poor country would have to have enormously higher rates of return to capital:

- · Around 60 times higher if it is one-fifteenth as rich.
- · **This is implausible!**

● Begins the hunt for other factors that might explain the difference.

– What did we not control for, but should have?

- This describes the methodological approach:
  - The convergence benchmark must be pitted against the empirical evidence on world income distributions, savings rates, or rates of return to capital.
  - The two will usually fail to agree.
  - Then look for the parametric differences that will bridge the model to the data.
- "Human capital" is often used as a first port of call:

  Can differences here account for observed cross-country variation?
  - The easiest way to slip differences in human capital in Solow equations is to renormalize labour.
  - This exercise does not take us very far.
    - Lucas exercise still predicts that the rate of return to capital is 6 times higher in India than in the US (see Banerjee and Duflo (2005), page 492).
    - Parente and Prescott (2000)) exercise predicts that per-capita income differences are only around half as much (or less) as they truly are (pages 61-62).

- The rest must be attributed to that familiar black box: “technological differences” (TFP).
  - That slot can be filled in a variety of ways:
    - externalities arising from human capital,
    - incomplete diffusion of technology,
    - excessive government intervention,
    - within-country misallocation of resources, ...
- All of these are interesting candidates, but by now we have wandered far from the original convergence model.
  - If that model still continues to illuminate at all, it is by way of occasional return to the recalibration exercise after choosing plausible specifications for each of these potential explanations.
- This model serves as a quick and ready fix on the world and organizes a search for possible explanations.
  - Viewed as a first pass, such an exercise can be very useful.

- Playing the game too seriously reveals a particular world-view:
  - It suggests a fundamental belief that the world economy is ultimately a great leveller.
  - If the levelling is not taking place, we must search for that explanation in "parameters" that are somehow structurally rooted in a society.
- Suggests that societies are somehow basically different – with regard to underlying attitudes, preferences or culture, and these account for their differences in economic development.
  - Some may believe that Confucianism or the Protestant ethic breeds economic success.
  - It may be argued that the feudal *zamindari* system in West Bengal was inimical to the development of entrepreneurship.
- These views may or may not be correct in some sense.
  - The way they are phrased suggests a comfortable distinction between our notions of what is endogenous and what is exogenous, leading to the *endogeneity problem*.

- **The Endogeneity Problem:**

  The parameters identified in the growth regression or calibration exercises do go hand in hand with underdevelopment, yet there is no ultimate causal chain.

  – Two variables may be correlated but jointly determined by a third variable.

    ○ For example, variables such as growth and political stability could be seen as jointly determined equilibrium outcomes associated with a particular set of institutions.

    ○ A correlation between growth and political stability, even if robust in statistical terms, is not informative about the structural determinants of growth.

  – There are many instances in growth research when explanatory variables are clearly endogenously determined in the economic sense.

    ○ Consider a regression that relates growth to the ratio of investment to GDP.

      · This may tell us that the investment share and growth are associated, but stops short of identifying a causal effect.

      · Even if we are confident that a change in investment would affect growth, this just pushes the relevant question further back: what determines investment?

- · In short, the conventional approach truncates our willingness to go further.

– When variables are endogenously determined in economic sense, there is also a strong chance that they will be correlated with the disturbances in the structural equation for growth.

  ○ Example: Suppose political instability lowers growth, but slower growth feeds back into political instability.

    · The estimated regression coefficient will tend to conflate these two effects and will be an inconsistent estimate of the causal effect of instability.

● The most common response to the endogeneity of growth determinants has been the application of instrumental variable procedures.

– For a critical evaluation of the use of instrumental variables refer to section 6.4 of Durlauf, Johnson and Temple (2005) which is of the following opinion:

  ○ Many instrumental variables used in the empirical growth literature are invalid as there are many ways in which they could be correlated with the disturbances in a growth model.

- Referring to the endogeneity problem, Ray (2008) concludes the discussion of development from the viewpoint of convergence as follows:
  - The convergence predicted by technologically diminishing returns to inputs should not blind us to the possibility of nonconvergent behaviour when all variables are treated as they should be
    - as variables that potentially make for underdevelopment, but also as variables that are profoundly affected by the development process.
  - This leads to a different way of asking the “big questions”
    - one that is not grounded in any presumption of convergence.

- Banerjee and Duflo (2005) sum up their criticism of the convergence approach as follows:
  - Lucas’ (1990) question about why capital does not flow from the US to India was, in some sense, where it all started.

- But, from the vantage point of what we know today, this is in some ways the lesser problem.
    - There exist differences in the marginal product of capital within the same economy that dwarf the gap that Lucas calculated comparing India and the US.
- The harder question is why capital flows do not eliminate these differences.
- Lucas' resolution of the puzzle was to give up the key neoclassical postulate of equal TFP across countries.
    - Focus on reasons for technological backwardness in poor countries.
- But this is not consistent with the empirical evidence (see section 3.1 of Banerjee and Duflo, 2005):
    - Many firms in poor countries do use the latest technologies, while others in the same country use obsolete models of production.
- What we need to explain is less the overall technological backwardness and more why some firms do not adopt profitable technologies that are available to them.
    - Why the market allows this to be the case?

– The premise of the conventional approach to growth:

  Markets function well enough within countries that we can largely ignore the inefficiency and unequal access to resources within an economy when we are interested in dynamics at the country level.

– The evidence suggests that is not true:

  ○ The cross-country differences in marginal products or technology that we want to explain are of the same order of magnitude as the differences we observe within each economy.

– A theory of cross-country differences has to be based on an understanding and an acknowledgement of the reasons why rates of returns vary so much within each country.

7. Criticism: Distribution Dynamics and Polarization

- Quah (1993a, 1993b, 1996a, 1996b, 1996c, 1997) has persuasively criticized standard regression approaches to studying convergence issues for being unable to shed light on important issues of
  - mobility, stratification and polarization in the world income distribution.
- The following description is from Quah (*Journal of Economic Growth*, 1997).
- The patterns of growth across countries that Quah attempts documents are given in the stylized features of Figure 1.
- Figure 1 records, for different time points, (the densities corresponding to) cross-country per capita income distributions.
  - The distribution at time t shows most countries having a medium level per capita income.
    - Few are very rich, and few very poor.

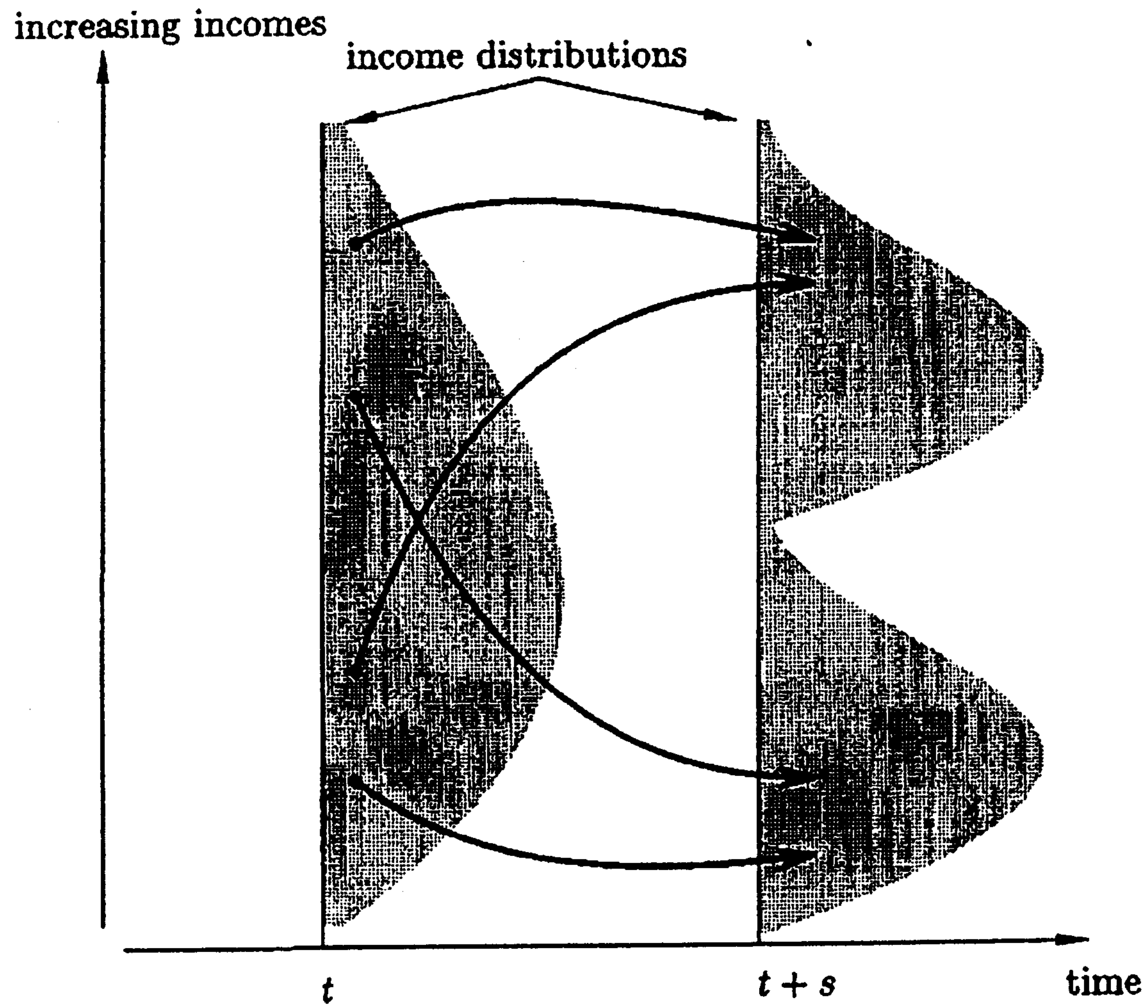


Figure 1. Emerging twin peaks.

- Over time, cross-country income distributions fluctuate:
  - Figure 1 makes explicit the distribution at time $t + s$.
  - In general, there is one such distribution for each time period.
  - Figure 1 is like a time-series plot, except that instead of recording the trajectory of a scalar or vector quantity, it comprises the trajectory of an entire distribution.
- Figure 1 can address the question: whether poor countries are catching up with rich ones.
  - That would happen if, for example, the sequence of distributions collapses over time to a degenerate point limit.
  - But in general that need not occur, and there are other ways whereby the poor can catch up with the rich
    - as illustrated, for instance, by the criss-crossing arrows.
- Figure 1 shows the distribution at time $t + s$ to have a *twin-peaks* property:
  - there is a clustering together of the very rich, a clustering together of the very poor, and a vanishing of the middle income class.

⇒ *Polarization*: those portions of the underlying population of economies collecting in the different peaks may be said to be polarized, one group versus another.

– If more than two peaks emerged, it might be natural to call the situation *stratification* or *convergence clubs* of countries catching up with one another but only within particular subgroups.

- Arrows drawn in Figure 1 indicate a rich spectrum of intradistribution dynamics:

– Persistence:

Some countries rich at time $t + s$ had already been rich at time t; similarly, others poor at $t + s$ had already been poor at t.

– Mobility:

Some countries rich at $t+s$ had begun poor; some of those poor at $t+s$ had begun rich.

– Separation:

Some groups of economies originally close together in the middle class have subsequently separated, with some becoming much richer than others.

- Sum up: there are both *shape* and *mobility* dynamics in the distributions in Figure 1.
- Data: Per capita incomes across 105 countries drawn from Summers-Heston (1991) dataset.
- **Shapes:**
  - Figure 3.1 shows a sequence of kernel-smoothed densities of relative per capita incomes taken at roughly decade-long intervals.
    - Data definition: 1/2 on the horizontal axis indicates one-half the world average per capita income; 2 indicates twice the world average; and so on.
    - In 1961 a nascent twin-peakedness – the first mode at a little less than 1, the second at slightly greater than 2.5 – was beginning to be visible.
    - By 1988 that second peak had become pronounced;
      - The relative income distance between the peaks doubled from about 1.5 in 1961 to more than 3 in 1988.

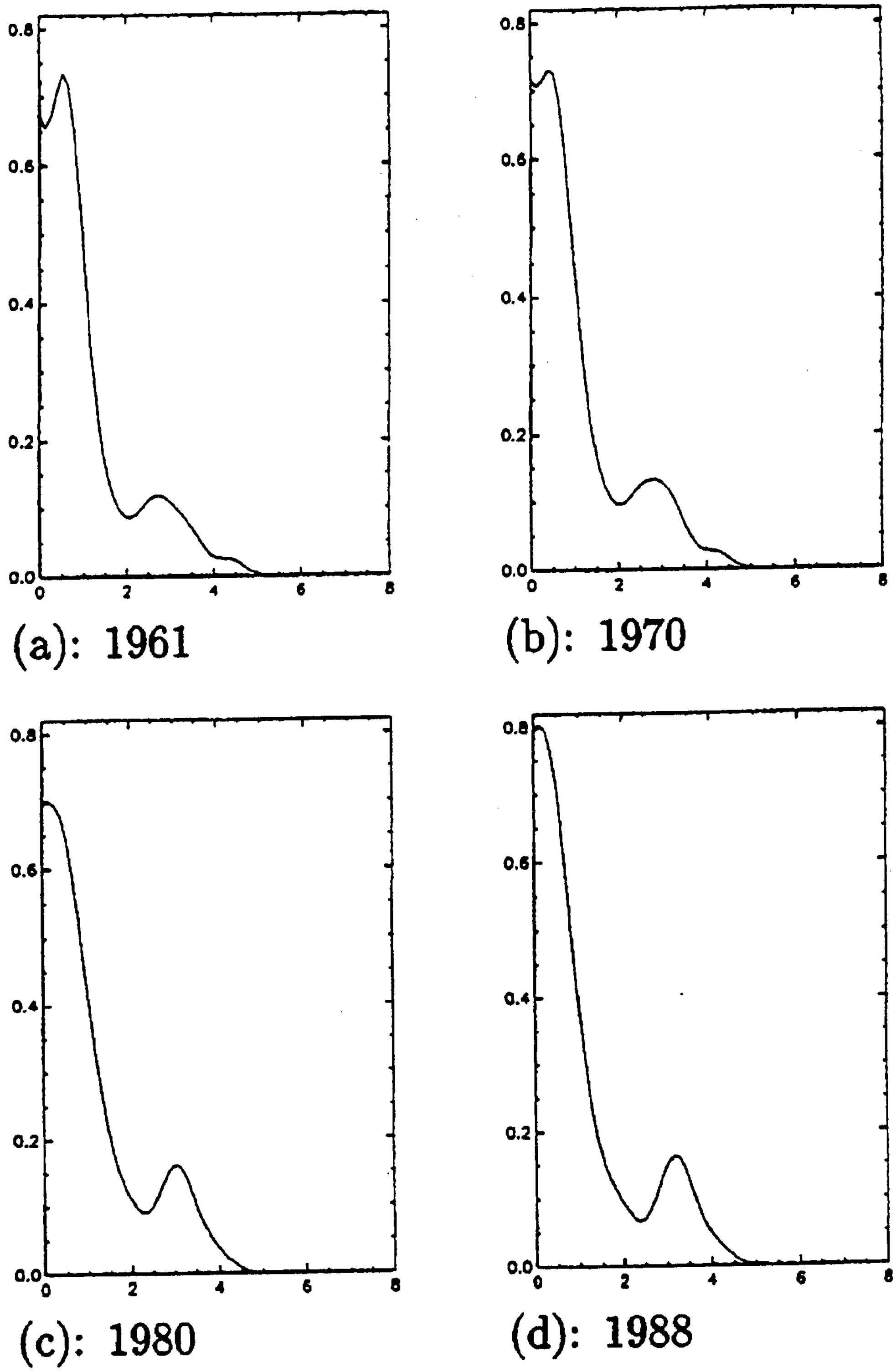


Figure 3.1. Densities of relative (per capita) incomes across 105 countries.

- **Mobility:**
  - An easy way to quantify intradistribution dynamics in a sequence of distributions is:
    - discretize the space of income values, and then
    - count the observed transitions out of and into distinct discrete cells.
  - In Figure 4 – which reproduces the essential features of Figure 1 – one might
    - add up the number of transitions out of cell II into cells I and III respectively (and everywhere else), and then
    - normalize those counts by the total number of observations.
  - Using discrete cells that span the space of all possible realizations, one can then construct a *transition probability matrix* as in Quah (*European Economic Review*, 1993):

$$F_{t+1} = M \cdot F_t,$$

    - F_t: cross-country income distribution at time t
    - F_{t+1}: cross-country income distribution at time $t + 1$
    - M: transition matrix that maps F_t onto F_{t+1}.

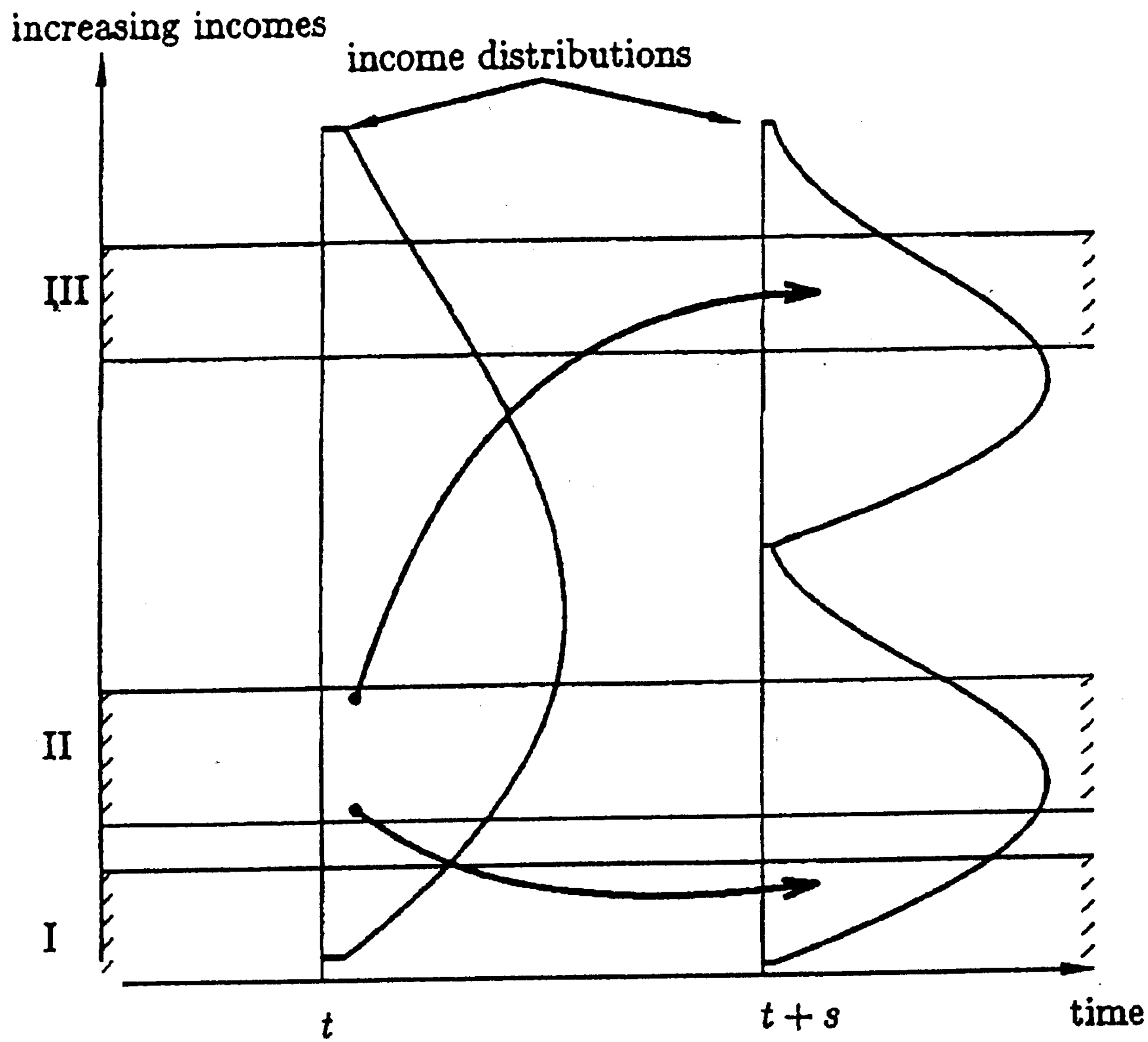


Figure 4. Discretization for intradistribution dynamics.

  - Each row of M is a probability mass function describing the distribution over states of the system after one transition given that the system is currently in the state corresponding to that row.
- The goal is to know M, which determines the evolution of the distribution.
- Assuming M to be time invariant, iteration yields (a predictor for) future cross-section distributions:

$$F_{t+s} = (M \cdot M \cdot \ ...M) \cdot F_t = M^s \cdot F_t,$$

  where s is any particular length of time (number of years, say).
- Taking this limit as $s \to \infty$, one can characterize the likely long-run (ergodic) distribution of cross-country incomes (if it exists).
  - Convergence might manifest in $\{F_{t+s}\}$ tending towards a point mass.
  - The world partitioning into haves and have-nots might manifest in $\{F_{t+s}\}$ tending towards a two-point or bimodal distribution.

- Quah (1993) calibrates M using per capita income of 118 countries from 1962 to 1985.

- Takes each country's per capita income relative to the world average as basic data.
- Operator M is approximated by discretizing the set of possible values of relative incomes into intervals at $\frac{1}{4}, \frac{1}{2}, 1, 2$ and ∞.
  - State 1 ($\frac{1}{4}$): per capita incomes no greater than one-fourth the world average,
  - State 2 ($\frac{1}{2}$): per capita incomes between one-fourth and one-half the world average, ...
  - State 5 (∞): per capita incomes greater than twice the world average.
- All relevant properties of M are described by the 5×5 transition matrix in Table 1:
  - (j, k)-th entry is the probability that an economy in state j transits to state k.

- Table 1, First Panel:
  - Contains the one-step annual transition matrix estimated by averaging the observed one-year transitions over every year from 1962-63 through 1984-85.
  - The first column gives the total number of transitions with starting points in that income state.

Table 1

Real GDP per capita (relative to world average).
First-order, time-stationary (1962 to 1984);
grid: (0, 1/4, 1/2, 1, 2, ∞); states: 5.

	Upper endpoint				
(Number)	1/4	1/2	1	2	∞
(456)	0.97	0.03	0.00	0.00	0.00
(643)	0.05	0.92	0.04	0.00	0.00
(639)	0.00	0.04	0.92	0.04	0.00
(468)	0.00	0.00	0.04	0.94	0.02
(508)	0.00	0.00	0.00	0.01	0.99
Ergodic	0.24	0.18	0.16	0.16	0.27
	From second-order specification				
Ergodic	0.28	0.15	0.12	0.15	0.30
23-year transition (1962–1985)					
(17)	0.76	0.12	0.12	0.00	0.00
(29)	0.52	0.31	0.10	0.07	0.00
(35)	0.09	0.20	0.46	0.26	0.00
(17)	0.00	0.00	0.24	0.53	0.24
(20)	0.00	0.00	0.00	0.05	0.95
Ergodic	0.16	0.05	0.10	0.12	0.57
	From second-order specification (22-year transition)				
Ergodic	0.20	0.09	0.13	0.12	0.47
Stationary estimate, iterated 23 times					
1/4	0.61	0.27	0.09	0.03	0.00
1/2	0.37	0.32	0.20	0.09	0.02
1	0.14	0.23	0.31	0.25	0.07
2	0.04	0.11	0.25	0.39	0.22
5	0.00	0.01	0.04	0.12	0.82

- For example, the second row shows that over the entire sample – across 118 countries and 23 years – 643 observations fell in state 2, i.e., had incomes between one-fourth and one-half the world average.
  - Of these, 92% remained in that same state in the following year, 4% moved up to incomes between one-half and the world average, while 5% moved down to incomes no greater than one-fourth the world average.
- Over this one-year horizon, the predominant feature is high *persistence*:
  - All diagonal entries exceed 90%.
  - Other entries are non-zero only for the first state off the main diagonal.

• Table 1, Second Panel:

- Describes the one 23-year transition from 1962 through 1985.
- Again we see *persistence*, although less pronounced.
  - 7% of the economies originally at incomes between one-fourth and one-half of the world average (state 2) transited to incomes at world average or higher (state 4) over this longer horizon.

- This mobility is not all favourable:
  - Of those same economies originally in state 2, over one-half *dropped* to even lower incomes.
- Looking down the neighbourhood of the main diagonal suggests
  - At low incomes (states 1 and 2) the greater tendency is to become even poorer,
    - although some possibility for upward mobility always remains.
  - At middle to higher incomes (states 3 and 4) upward and downward mobilities just about balance.
  - The highest income state appears persistent:
    - 95% probability of the richest remaining richest.

- This informal description suggests cross-country incomes tending towards extremes at both high and low endpoints.
  - Table 1 makes this precise by examining the long-run (ergodic) distributions implied by these transition mappings.
    - The ergodic distributions are given below the corresponding transition matrices.

◦ Quah (1993) emphasizes:

· The steady-state distributions should not be read as forecasts of what will happen in the future.

· Rather, these distributions should be interpreted simply as characterizations of tendencies in the post-War history that actually realized.

- For both of the first two panels – M estimated by averaging annual transitions across time and over the long horizon – the implied ergodic distributions show:

1. a thinning in the middle, and
2. an accumulation in *both* the low and high tails.

– A higher probability at the upper tail – especially so for M estimated over the long horizon.

- Robustness of Tendency of Incomes Towards the Two Extremes:

  Quah (1993) experimented with Markov chain transition models having dynamics beyond the first order.

- The implied ergodic distributions from richer structures almost invariably carry the same message:
  - thinning out of the middle-income economies in favour of the very rich and very poor.
- The second ergodic distribution given in the first two panels of Table 1 is generated from a second-order Markov chain.
  - The bimodal property of the ergodic distribution remains.
  - The tendency for a poverty pile-up is now even more pronounced.

- Consistency of the Short- and Long-run First-order Models:

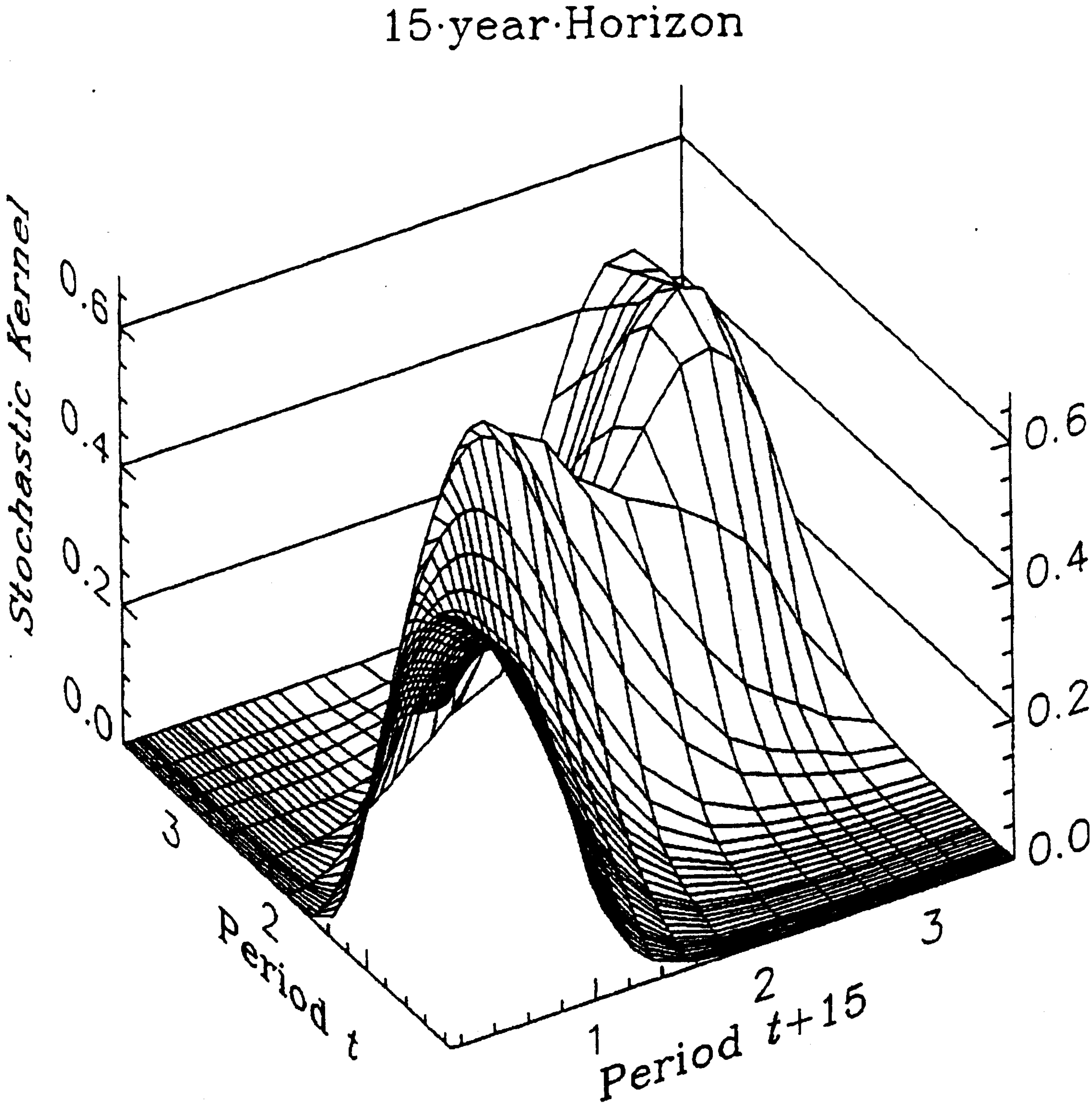
  Quah (1993) iterated the one-period transition function of the short-run model to cover the same time span (23 years) as that of the long-run one.

  - Third panel of Table 1 reports this iterated one-step transition.
  - Comparing diagonal entries in the second and third panels shows consistent under-prediction:
    - The data display greater long-run persistence than predicted by the best-fitting first-order model.

- Every entry in the iterated transition function, except the two extreme endpoints, is strictly positive.
  - By contrast, the long-run model shows no transitions between very low- and very high-income states.
    - 9 entries in the opposing off-diagonal corners are zero in the second panel.
- Again, this indicates higher persistence – lower cross-sectional mobility – in the data over long horizons.

- Kremer, Onatski and Stock (2001) update Quah's analysis using more recent data.
  - Find evidence of twin-peaks in the long-run distribution of per capita incomes.
  - Find the rich peak to be much higher than the poor peak.
  - Kremer, Onatski and Stock's point estimates imply that most countries will ultimately move to the rich state.
    - But, during the transition period, which could last hundreds of years, polarization in the income distribution may worsen.

- Quah (1997) recognizes that discretization of the relative income intervals as in Quah (1993) can distort dynamics in important ways when the underlying observations are continuous variables.
- Quah (1997) considers the following solution:
  - Not to use a discretization at all, but to retain the original set of continuous income observations in quantifying intradistribution dynamics.
  - Doing that is like allowing the number of distinct cells {I, II, III, ...} in Figure 4 to tend to infinity and then to the continuum.
  - The corresponding transition probability matrix tends to a matrix with a continuum of rows and columns.
    - It becomes a stochastic kernel, as graphed in Figure 5.1.
- Figure 5.1 shows the stochastic kernel for 15-year transitions in relative-income data, averaging over 1961 through 1988.
  - From any point on the axis marked Period t extending parallel to the axis marked Period $t + 15$ the stochastic kernel is a probability density function:
    - the projection traced out is nonnegative and integrates to unity.



Relative income dynamics across 105 countries.

- That projection is similar to a row of a transition probability matrix:
  - such a row has all entries nonnegative and summing to 1.
- Roughly speaking, this probability density describes transitions over 15 years from a given income value in period t.

• Figure 5.1 shows how the cross-sectional distribution at time t evolves into that at $t + 15$.

- If most of the graph were concentrated along the 45-degree diagonal, then elements in the distribution remain where they began.
- If, by contrast, most of the mass in the graph were rotated 90 degrees counterclockwise from that 45-degree diagonal,
  - then substantial overtaking occurs – the rich become poor, and the poor rich, periodically over 15-year horizons.
- If most of the graph were concentrated around the 1-value of the Period $t + 15$ axis – extending parallel to the Period t axis,
  - then over a 15-year horizon, the cross-section distribution converges towards equality.

- Dynamics over longer horizons can be studied by recursively applying a given stochastic kernel.

- In Figure 5.1 the *twin-peaks property* manifests again.
  - Over the 15-year horizon, a large portion of the probability mass remains clustered around the main diagonal.
  - Along that principal ridge, a dip appears in the middle-income portion while the kernel itself rises towards local maxima in both poor and rich parts of the income range.
  - The two peaks in the stochastic kernel correspond to what Durlauf and Johnson (1995) call “basins of attraction”.
  - At the same time, while the middle-income class is vanishing, portions of the cross section do transit from high to low and from low to high:
    - the stochastic kernel is positive almost everywhere, and communicates across the entire range of income values.

- The estimated ergodic densities presented by Bulli (2001) and Johnson (2004) support Quah’s conclusions.

- The findings of ‘persistence’ and ‘bi-modality’ are valuable additions to the known stylized facts about cross-country growth regularities.
  - Growth theories now need to explain these facts.
  - The Markov analysis itself does not help in the explanation.
    - It is another type of reduced form analysis.
    - The transition matrix M is memory-less, and no growth theory is required for its estimation.
      - No structure is imposed on the data.
    - Quah makes it explicit that it is his intention not to be restricted by assumptions of long-run growth.
  - Quah (*Economic Journal*, 1996) has made suggestions about the directions in which explanations for these stylized facts have to be sought.
    - One explanation is along models of multiple equilibrium that yield ‘club convergence’.

- Quah (1997) suggests that the dynamics towards bi-modality can be explained by 'spatial spillovers'
  - which in turn are determined not so much by 'openness' of an economy as by 'who trades with whom', that is, the trading partners.
- More research will be necessary to resolve these issues.

8. Criticism: Structural Differences Vs. Initial Conditions: Convergence Clubs

- Recall the conceptual distinction:

  If cross-country per-capita income differences are

  - permanent solely because of cross-country structural heterogeneity, *conditional convergence* is occurring;
  - permanent at least in part determined by initial conditions, and countries with similar initial conditions exhibit similar long-run outcomes $\rightarrow$ *convergence clubs*.

- Several papers have attempted to disentagle the roles of heterogenous structural characteristics and initial conditions in determining growth performance.
  - Despite a wide variety of statistical methods being employed, each of these studies provide evidence of the existence of convergence clubs even after accounting for variation in structural characteristics.

- Durlauf and Johnson (1995):

  Allude to theoretical models of multiple equilibrium and observe:

  - Convergence in large samples (*global* convergence) does not hold (or proves weak) because countries belonging to different equilibrium (or ‘regimes’) are lumped together.
  - The proper thing is to identify country groups whose members share the same equilibrium,
    - and then to check whether convergence holds within these groups (*local* convergence).
  - Durlauf and Johnson use initial levels of income and literacy rate (as a proxy for human capital) to group the countries.
  - Using the same cross-country data set as Mankiw, Romer and Weil (1992), Durlauf and Johnson tests:
    - Null hypothesis: a common growth regime;
    - Alternative hypothesis: a growth process with multiple regimes in which economies with similar initial conditions tend to converge to one another.

- Result: Reject the single regime model required for global convergence.
- Interpretation: Even after controlling for the structural heterogeneity, there is a role for initial conditions in explaining variation in cross-country growth behaviour.

- Durlauf and Johnson's findings of multiple convergence clubs appear to be reinforced by subsequent research:
  - The sample splits estimated by Papageorgiou and Masanjala (2004) divide the data in four distinct growth regimes and are broadly consistent with those found by Durlauf and Johnson.
  - Tan (2004) finds strong evidence that measures of institutional quality and ethnic fractionalization define convergence clubs across a wide range of countries.
- Desdoigts (1999):

  Uses projection pursuit methods in an attempt to separate the roles of microeconomic heterogeneity and initial conditions in the growth experiences of a group of countries.
  - Identifies groups of countries with relatively homogeneous growth experiences based on data about the characteristics and initial conditions of each country.

- The idea of projection pursuit is to find the orthogonal projections of the data into lower-dimensional spaces that best display some interesting feature of the data.
  - One should consider as many dimensions as needed to account for “most” of the clustering in the data.
- Desdoigts finds several interesting clusters.
  - The first is the OECD countries.
  - The two projections identifying this cluster put most of their weight on
    - the primary and secondary school enrollment rates,
    - the 1960 income gap, and
    - the rate of growth in the labour force.
  - The prominence of variables that Desdoigts argues are proxies for initial conditions among those defining the projections leads him to conclude:
    - Initial conditions are more important in defining this cluster than are other country characteristics.

- Reapplication of the clustering method to the remaining (non-OECD) countries yield three sub-clusters:
  - Africa, Southeast Asia, and Latin America.
  - Here the projections put most weight on
    - government consumption,
    - the secondary school enrollment rate, and
    - investment in electrical machinery and transportation equipment.
  - Most of these variables are argued to proxy for structural characteristics.
    - Suggests that structural characteristics, rather than initial conditions, are responsible for the differences in growth experiences across these three geographic sub-clusters.
- Note that this approach relies on the judgement of the researcher in determining which variables proxy for initial conditions and which proxy for structural characteristics.

- For the details of the empirical literature on multiple regimes refer to section 5.3 of Durlauf, Johnson and Temple (2005).

9. References

- Sections 1, 2, 3, 4, 5, 8 and the endogeneity problem discussed in section 6 of this note are based on

1. Durlauf, Steven, Paul Johnson and Jonathan Temple (2005), "Growth Econometrics", *Handbook of Economic Growth*, Vol. 1A, eds. Philippe Aghion and Steven Durlauf, Amserdam: Elsevier, 555-677.

- Section 6 is based on

2. Ray, Debraj (2008), "Development Economics", in *The New Palgrave Dictionary of Economics* (section 2), edited by L. Blume and S. Durlauf.
3. Mookherjee, Dilip and Debraj Ray (2000), Introduction to *Readings in the Theory of Economic Development*, London: Blackwell: section 2, and
4. Banerjee, Abhijit and Esther Duflo (2005), "Growth Theory through the Lens of Development Economics", *Handbook of Economic Growth*, Vol. 1A, eds. Philippe Aghion and Steven Durlauf, Amserdam: Elsevier, 473-552: sections 2 and 3.

- Section 7 is based on

5. Quah, Danny (1993), "Empirical Cross-section Dynamics in Economic Growth", *European Economic Review*, 37, 426-434.
6. Quah, Danny (1997), "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs", *Journal of Economic Growth*, 2, 27-59, and
1. Durlauf, Steven, Paul Johnson and Jonathan Temple (2005), "Growth Econometrics", *Handbook of Economic Growth*, Vol. 1A, eds. Philippe Aghion and Steven Durlauf, Amserdam: Elsevier, 555-677: section 4.3.3.