Economics 270c

Graduate Development Economics

Professor Ted Miguel

Department of Economics

University of California, Berkeley

Economics 270c

Graduate Development Economics

Lecture 9 – March 17, 2009

Macroeconomic growth empirics
Lecture 1: Global patterns of economic growth and development (1/20)
Lecture 2: Inequality and growth (1/27)
The political economy of development
Lecture 3: History and institutions (2/3)
Lecture 4: Corruption (2/10)
Lecture 5: Patronage politics (2/17)
Lecture 6: Democracy and development (2/24)
Lecture 7: War and Economic Development (3/3)
Lecture 8: Economic Theories of Conflict (3/10) – Guest lecture by Gerard Padro
Human resources
Lecture 9: Human capital and income growth (3/17)
Lecture 10: Increasing human capital (3/31)
Lecture 11: Labor markets and migration (4/7)
Lecture 12: Health and nutrition (4/14)
Lecture 13: The demand for health (4/21)
Other topics
Lecture 14: Environment and development (4/28)
Lecture 15: Resource allocation and firm productivity (5/5)
Additional topics for the development economics field exam
-- Ethnic and social divisions
-- The Economics of HIV/AIDS

- Prerequisites: Graduate microeconomics, econometrics
- Grading:
  Four referee reports – 40%
  → Fourth referee report due today, Mar. 17, 2009

  Two problem sets – 20%
  → Problem set 1 to be distributed on Mar. 31, due April 7

  Research proposal – 30%
  Class participation – 10%
  No final exam
- All readings are available online (see syllabus)
- Additional references on syllabus

Lecture 9 outline

(1) Human capital in economic development

(2) Krueger and Lindahl (2001) on education and macroeconomic growth, measurement issues

(3) Duflo (2001) on the returns to schooling in Indonesia

(1) Human capital in economic development

- There have been massive increases in literacy and schooling attainment around the world – Africa, Asia, Latin America – during the past 60 years

- Perhaps unexpectedly, at the regional level increased schooling does not line up well with faster economic growth rates, e.g., sub-Saharan Africa versus South Asia

HDI rank [a]	Human development index (HDI) value 2003	Life expectancy at birth (years) 2003	Adult literacy rate (% ages 15 and above) 2003 [b]	Combined gross enrolment ratio for primary, secondary and tertiary schools (%) 2002/03 [c]	GDP per capita (PPP US$) 2003
Developing countries	0.694	65.0	76.6	63	4,359
Least developed countries	0.518	52.2	54.2	45	1,328
Arab States	0.679	67.0	64.1	62	5,685
East Asia and the Pacific	0.768	70.5	90.4	69	5,100
Latin America and the Caribbean	0.797	71.9	89.6	81	7,404
South Asia	0.628	63.4	58.9	56	2,897
Sub-Saharan Africa	0.515	46.1	61.3	50	1,856
Central and Eastern Europe and the CIS	0.802	68.1	99.2	83	7,939
OECD	0.892	77.7	..	89	25,915
High-income OECD	0.911	78.9	..	95	30,181

HDI rank	GDP US$ billions 2003	GDP PPP US$ billions 2003	GDP per capita US$ 2003	GDP per capita PPP US$ 2003	Annual growth rate (%) 1975–2003	Annual growth rate (%) 1990–2003
Developing countries	6,981.9 T	21,525.4 T	1,414	4,359	2.3	2.9
Least developed countries	221.4 T	895.1 T	329	1,328	0.7	2.0
Arab States	773.4 T	1,683.6 T	2,611	5,685	0.2	1.0
East Asia and the Pacific	2,893.6 T	9,762.2 T	1,512	5,100	6.0	5.6
Latin America and the Caribbean	1,745.9 T	3,947.0 T	3,275	7,404	0.6	1.1
South Asia	902.2 T	4,235.9 T	617	2,897	2.6	3.5
Sub-Saharan Africa	418.5 T	1,227.4 T	633	1,856	–0.7	0.1
Central and Eastern Europe and the CIS	1,189.9 T	3,203.5 T	2,949	7,939	..	0.3
OECD	29,650.5 T	29,840.6 T	25,750	25,915	2.0	1.8
High-income OECD	28,369.5 T	27,601.9 T	31,020	30,181	2.2	1.9

(1) Human capital in economic development

- There have been massive increases in literacy and schooling attainment around the world – Africa, Asia, Latin America – during the past 60 years

- Perhaps unexpectedly, at the regional level increased schooling does not line up well with faster economic growth rates, e.g., sub-Saharan Africa versus South Asia

- This is consistent with the view that institutions and technology (“A”) matter more for growth than physical or human capital investments. But in the short-run boosting capital could still increase income levels

(1) Human capital in economic development

- This week: what is the return to schooling in less developed countries?

(1) Human capital in economic development

- This week: what is the return to schooling in less developed countries?

- Next week: what does the education production function look like? Which inputs lead to more educational production?

- Why focus on education? In many poor countries, education spending is the largest single recurrent discretionary budget expenditure. E.g., in late 1990s Ghana, it was 35% of discretionary expenditures

(1) Different conceptions of education

- Benefits of education could include:

  -- Higher wages ("human capital")

  -- Education as consumption (reading Shakespeare)

  -- Education as a signal of ability

(1) Different conceptions of education

- Benefits of education could include:

  -- Higher wages (“human capital”)

  -- Education as consumption (reading Shakespeare)

  -- Education as a signal of ability

  -- Possible **social** benefits include labor productivity spillovers, a “better” functioning democracy (?), less crime (?), better child health (?), others?

(1) Different conceptions of education

- Benefits of education could include:

  -- Higher wages (“human capital”)

  -- Education as consumption (reading Shakespeare)

  -- Education as a signal of ability

  -- Possible **social** benefits include labor productivity spillovers, a “better” functioning democracy (?), less crime (?), better child health (?), others?

- Costs: Opportunity cost of time studying; tuition costs

- Socially suboptimal investments if spillovers, household agency problems

(1) Estimating Mincerian wage regressions

- The Mincerian wage regression:

  $$ln(w_i) = b_0 + b_1 S_i + b_2 X_i + b_3 X_i^2 + e_i$$

  where w is the individual wage, S is years of schooling, and X is years of experience, for individual i

(1) Estimating Mincerian wage regressions

- The Mincerian wage regression:

  $$ln(w_i) = b_0 + b_1S_i + b_2X_i + b_3X_i^2 + e_i$$

  where w is the individual wage, S is years of schooling, and X is years of experience, for individual i

- This has been run in literally dozens of countries, and estimates of b_1 usually fall in the range 0.05-0.15

- Reliably estimating this equation has been central to labor economics for 40+ years. Possible upward selection / omitted variables bias, and possible downward attenuation bias due to measurement error

(1) Measurement error and attenuation bias

- Imagine the exact (but unmeasured) variable X* is imperfectly captured by the (measured) variable X:

$$X_i = X^*_i + u_i$$

where u_i is an i.i.d. normally distributed random variable. This is classical measurement error

-- X is reported years of schooling

-- X* is real schooling or skills

(1) Measurement error and attenuation bias

- Imagine the exact (but unmeasured) variable X* is imperfectly captured by the (measured) variable X:

  $$X_i = X^*_i + u_i$$

  where u_i is an i.i.d. normally distributed random variable. This is classical measurement error

  -- X is reported years of schooling

  -- X* is real schooling or skills

- We want to run the regression $Y_i = a + bX^*_i + e_i$ but due to data limitations have to run $Y_i = \alpha + \beta X_i + \varepsilon_i$. How does β^{OLS} relate to b?

(1) Measurement error and attenuation bias

- The coefficient of interest is *b*, where OLS delivers: $b^{OLS} = Cov(X^*,Y)/Var(X^*)$

(1) Measurement error and attenuation bias

- The coefficient of interest is b, where OLS delivers:
  $b^{OLS} = Cov(X^*,Y)/Var(X^*)$

- But we end up estimating:
  $\beta^{OLS} = Cov(X,Y)/Var(X)$
  $= [Cov(X^*, Y) + Cov(u, Y)] / [Var(X^*) + Var(u)]$
  $= [Cov(X^*, Y)] / [Var(X^*) + Var(u)]$
  $= [Cov(X^*,Y)*Var(X^*)/Var(X^*)] / [Var(X^*)+Var(u)]$
  $= b^{OLS}*\{Var(X^*)/ [Var(X^*) + Var(u)]\}$

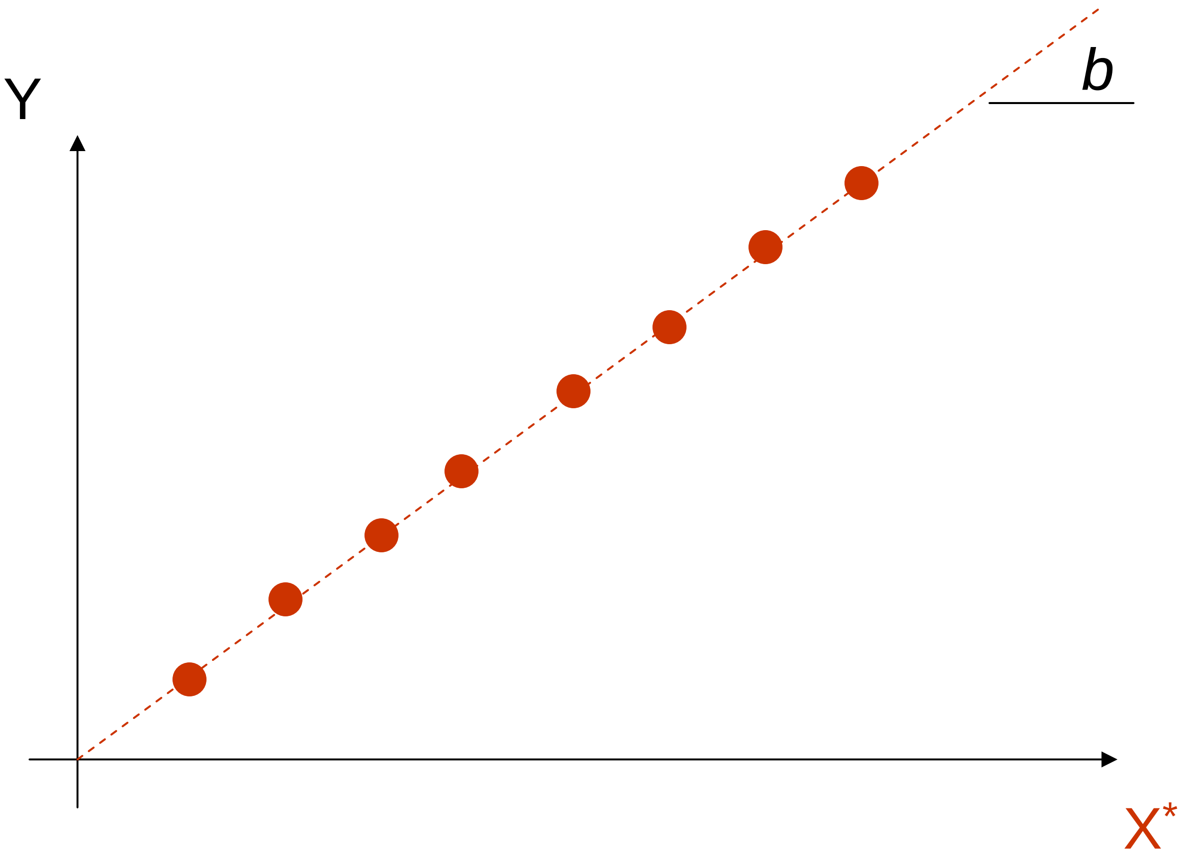
(1) Measurement error and attenuation bias

- The coefficient of interest is *b*, where OLS delivers:
  $b^{OLS} = Cov(X^*,Y)/Var(X^*)$

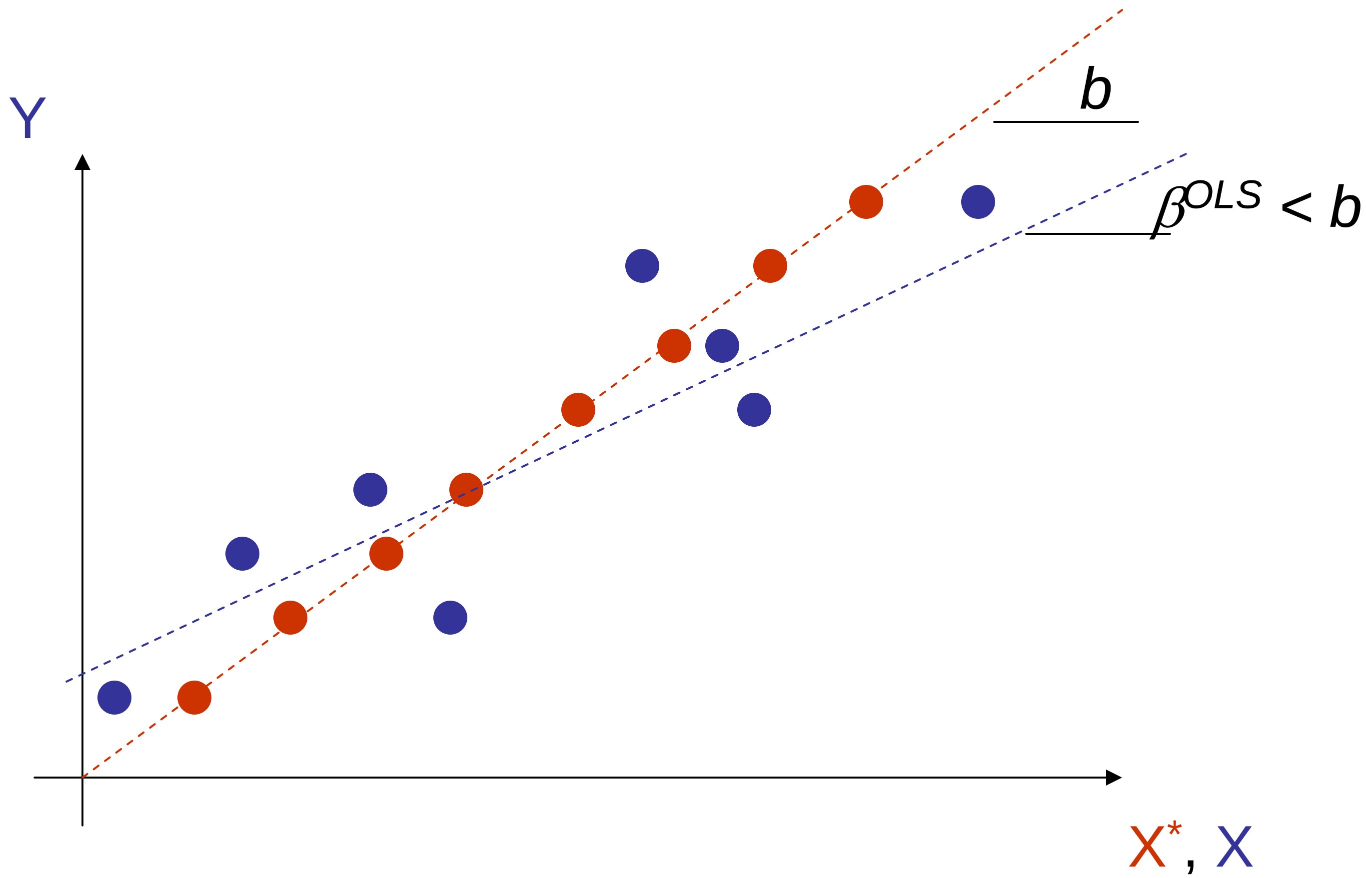
- But we end up estimating:
  $\beta^{OLS} = Cov(X,Y)/Var(X)$
  $= [Cov(X^*, Y) + Cov(u, Y)] / [Var(X^*) + Var(u)]$
  $= [Cov(X^*, Y)] / [Var(X^*) + Var(u)]$
  $= [Cov(X^*,Y)*Var(X^*)/Var(X^*)] / [Var(X^*)+Var(u)]$
  $= b^{OLS}*\{Var(X^*)/ [Var(X^*) + Var(u)]\}$

- Bias towards zero, as a function of the signal-noise ratio, i.e., if half the variance of X is noise, the bias is 50%

(1) Measurement error and attenuation bias



(1) Measurement error and attenuation bias



(1) IV and local average treatment effects

- Another important issue in estimating the returns to schooling arises when using instrumental variables (IV): most IV approaches that rely on exogenous shifts in attained schooling identify effects only for the population affected by the shift in attainment (Angrist, Imbens and Rubin 1996) → local average treatment effect (LATE)

(1) Returns to schooling in poor countries

- Given these concerns over identification, measurement error, and external validity, few studies in developing countries have rigorously estimated returns to schooling in less developed countries. How should we interpret Mincerian regressions?

  -- Duflo (2001) is a notable exception

- Using Mincerian regressions, Paul Schultz has found quite low “returns” to primary schooling across multiple African countries in recent years, although reasonably high returns to secondary schooling

  -- Today’s lecture explores macro and micro estimates

(2) Krueger and Lindahl (2001)

- Some researchers have focused on the macroeconomic evidence using cross-country regression methods

- One possible advantage of the macro approach is the ability to capture social benefits of schooling, e.g., labor productivity spillovers missed using individual data

  -- This would suggest macro estimates should be larger than micro estimates

  -- From a public economics and policy point of view, social benefits are more important to understand than private benefits

(2) Krueger and Lindahl (2001)

- The micro Mincerian regression for person i in country j at time t is: $ln(w_{ijt}) = b_{0jt} + b_{1jt}S_{ijt} + e_{ijt}$

- Now aggregate up to the country level (where Y is now the geometric mean of income rather than the wage):

$$ln(Y_{jt}) = b_{0jt} + b_{1jt}S_{jt} + e_{jt}$$

(2) Krueger and Lindahl (2001)

- The micro Mincerian regression for person i in country j at time t is: $ln(w_{ijt}) = b_{0jt} + b_{1jt}S_{ijt} + e_{ijt}$

- Now aggregate up to the country level (where Y is now the geometric mean of income rather than the wage):

$$ln(Y_{jt}) = b_{0jt} + b_{1jt}S_{jt} + e_{jt}$$

- Now consider changes in log per capita income:

$$\begin{aligned} \Delta ln(Y_{jt}) &= b_0 + b_{1jt}S_{jt} - b_{1jt-1}S_{jt-1} + \Delta e_{it} \\ &= b_0 + b_{1jt}(S_{jt} - S_{it-1}) - (b_{1jt-1} - b_{1jt})S_{jt-1} + \Delta e_{it} \\ &= b_0 + b_{1jt}\,\Delta S_{jt} + \Delta b_{1jt}\,S_{jt-1} + \Delta e_{jt} \end{aligned}$$

(2) Krueger and Lindahl (2001)

- $$\begin{aligned}\Delta ln(Y_{jt}) &= b_0 + b_{1jt}S_{jt} - b_{1jt-1}S_{jt-1} + \Delta e_{it} \\ &= b_0 + b_{1jt}(S_{jt} - S_{it-1}) - (b_{1jt-1} - b_{1jt})S_{jt-1} + \Delta e_{it} \\ &= b_0 + b_{1jt}\,\Delta S_{jt} + \Delta b_{1jt}\,S_{jt-1} + \Delta e_{jt}\end{aligned}$$

- This formulation allows us to distinguish competing growth models, "endogenous growth" vs. "growth accounting". The coefficient on S_{jt-1} reflects changes in returns to schooling over time. Romer (1990) predicts a positive sign, as human capital generates innovations

  -- Lucas (1988) predicts increases in an accumulable factor like human capital (ΔS_{jt}) is associated with higher income, so b_{1jt} > 0, especially considering social returns

(2) Krueger and Lindahl (2001)

- Existing cross-country studies regressing income growth on human capital find positive impacts of lagged schooling stocks on growth, but small and not very large effects of changes in educational attainment, say 4% per year of schooling – not what we would expect

TABLE 1
REPLICATION AND EXTENSION OF BENHABIB AND SPIEGEL (1994)
DEPENDENT VARIABLE: ANNUALIZED CHANGE IN LOG GDP, 1965–85

	Log Schooling			Linear Schooling		
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Δ Log S	–.072 (.058)	.178 (.112)	.614 (.162)	—	—	—
Log S_{65}	—	.010 (.004)	.026 (.005)	—	—	—
ΔS	—	—	—	.012 (.023)	.039 (.024)	.151 (.034)
S_{65}	—	—	—	—	.003 (.001)	.004 (.001)
Log Y_{65}	–.009 (.002)	–.012 (.002)	–.015 (.003)	–.008 (.002)	–.014 (.002)	–.014 (.004)
Δ Log Capital	.523 (.048)	.461 (.052)	—	.521 (.051)	.465 (.052)	—
Δ Log Work Force	.175 (.164)	.232 (.160)	—	.110 (.160)	.335 (.167)	—
R^2	.694	.720	.291	.688	.726	.271

Too large? (0.3)

Notes: All change variables were divided by 20, including the dependent variable. Sample size is 78 countries. Standard errors are in parentheses. All equations also include an intercept. S_{65} is Kyriacou's measure of schooling in 1965; Δ Log S is the change in log schooling between 1965 and 1985, divided by 20; and Y_{65} is GDP per capita in 1965. Mean of the dependent variable is .039; standard deviation of dependent variable is .020.

(2) Krueger and Lindahl (2001)

- Existing cross-country studies regressing income growth on human capital find positive impacts of lagged schooling stocks on growth, but small and not very large effects of changes in educational attainment, say 4% per year of schooling – not what we would expect

- Are the micro estimates just hopeless biased (upwards) by omitted variables / selection?

  -- Or could measurement error in national educational data be (partially) to blame?

(2) Krueger and Lindahl (2001)

- Sources of measurement error in macro education data:

  -- Differences in the quality of schooling across countries (e.g., there are big differences even across U.S. towns)

  -- The widely used UNESCO database, based on Ministry of Education statistics. These may be unreliable due to a lack of trained statistical personnel, resources

  -- UNESCO data use enrollment at start of school year

  -- Children educated abroad not counted

- Measurement error may be exacerbated in first differenced specifications, like growth regressions

(2) Krueger and Lindahl (2001)

- Consider the first differenced regression equivalent to above, ΔY_i on ΔX_i. The estimate of β becomes:

  $\beta^{OLS} = Cov(\Delta X, \Delta Y)/Var(\Delta X)$

  $= b^{OLS} * \{Var(X^*)/ [Var(X^*) + Var(u)*\Omega]\}$

  where $\Omega = (1 - \rho_u) / (1 - \rho_{X^*})$, where ρ captures the extent of serial correlation across time in a variable, i.e., $Corr(u_t, u_{t-1}) = Cov(u_t, u_{t-1})/Var(u)$

(2) Krueger and Lindahl (2001)

- Consider the first differenced regression equivalent to above, ΔY_i on ΔX_i. The estimate of β becomes:

  $$\beta^{OLS} = Cov(\Delta X, \Delta Y)/Var(\Delta X)$$

  $$= b^{OLS} * \{Var(X^*)/ [Var(X^*) + Var(u)^*\Omega]\}$$

  where $\Omega = (1 - \rho_u) / (1 - \rho_{X^*})$, where ρ captures the extent of serial correlation across time in a variable, i.e., $Corr(u_t, u_{t-1}) = Cov(u_t, u_{t-1})/Var(u)$

  -- First differencing exacerbates attenuation when there is more serial correlation in schooling than measurement error. “Differencing out” signal leaves noise. Over short periods, schooling levels are nearly fixed but noise is not

(2) Krueger and Lindahl (2001)

- Recall $\beta^{OLS} = Cov(\Delta X, \Delta Y)/Var(\Delta X)$

  $= b^{OLS}*\{Var(X^*)/ [Var(X^*) + Var(u)*\Omega]\}$

  where $\Omega = (1 – \rho_u) / (1 – \rho_{X^*})$

- An example: if ρ_{X^*}= 0.4 and ρ_u= 0.1, then Ω = (0.9/0.6) = 1.5. If $Var(X^*) = Var(u)$ = 1, then the attenuation bias correction rises from 2 to 2.5.

  -- Using the Barro-Lee and Kyriacou data, they estimate that ρ_{X^*} = 0.97 > 0.61 = ρ_u. This data seems pretty poor, with very non-idiosyncratic errors

(2) Krueger and Lindahl (2001)

- Eliminating “signal” from the key explanatory variable by including additional controls can also exacerbate measurement error

  -- The relative R^2’s of the regressions with and without additional controls determines the extent of attenuation bias towards zero due to these controls. This is particularly important with the investment controls (which are highly correlated with changes in schooling)

  -- Including a highly endogenous (outcome?) variable such as capital stock as a control variable makes it impossible to assess the impact of human capital, in their view (nice discussion on p. 1126)

(2) Krueger and Lindahl (2001)

- The existence of two different cross-country education series (Barro and Lee; Kyriacou) allows them to validate the accuracy of the data. Assume that there is classical measurement error in both series. A higher correlation between the two series → greater reliability

  -- These data series are quite highly correlated in levels, but much less so in first differences. There appears to be substantial measurement error in the first differenced education series, likely leading to major attenuation bias

- The reliability ratio captures the extent of attenuation bias: $R_i = Cov(S_i, S_j) / Var(S_i)$

TABLE 2
RELIABILITY OF VARIOUS MEASURES OF YEARS OF SCHOOLING

A. Estimated Reliability Ratios for Barro-Lee and Kyriacou Data

	Reliability of Barro-Lee Data	Reliability of Kyriacou Data
Average years of schooling, 1965	.851 (.049)	.964 (.055)
Average years of schooling, 1985	.773 (.055)	.966 (.069)
Change in years of schooling, 1965–85	.577 (.199)	.195 (.067)

B. Estimated Reliability Ratios for Barro-Lee and World Values Survey Data

	Reliability of Barro-Lee Data	Reliability of WVS Data
Average years of schooling, 1990	.903 (.115)	.727 (.093)
Average years of secondary and higher schooling, 1990	.719 (.167)	.512 (.119)

Notes: The estimated reliability ratios are the slope coefficients from a bivariate regression of one measure of schooling on the other. For example, the .851 entry in the first row is the slope coefficient from a regression in which the dependent variable is Kyriacou's schooling variable and the independent variable is Barro-Lee's schooling variable. The .964 ratio in the second column is estimated from the reverse regression. In panel B, the reliability ratios are estimated by comparing the Barro-Lee and WVS data. In the WVS data set, secondary and higher schooling is defined as years of schooling attained *after 8 years of schooling.*

Sample size for panel A is 68 countries. Sample size for panel B is 34 countries. Standard errors are reported in parentheses.

(2) Krueger and Lindahl (2001)

- Examine the relationship between economic growth and education growth over different time periods. Since the underlying education stock is slow moving, over shorter intervals Ω is likely to be larger thus exacerbating measurement error

  -- Using the best data, a longer time period (20 years), and correcting for attenuation bias yields a return of 30% to an additional year of education attained (on average)

TABLE 3
THE EFFECT OF SCHOOLING ON GROWTH
DEPENDENT VARIABLE: ANNUALIZED CHANGE IN LOG GDP PER CAPITA

	5-year changes			10-year changes			20-year changes		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
S_{t-1}	.004	—	.004	.003	—	.004	.005	—	.005
	(.001)		(.001)	(.001)		(.001)	(.001)		(.001)
ΔS	—	.031	.039	—	.075	.086	—	.184	.182
		(.015)	(.014)		(.026)	(.024)		(.057)	(.051)
Log Y_{t-1}	–.005	.004	–.006	–.003	.004	–.005	–.010	–.001	–.013
	(.003)	(.002)	(.003)	(.003)	(.001)	(.003)	(.003)	(.002)	(.003)
R^2	.197	.161	.207	.242	.229	.284	.184	.103	.281
N	607	607	607	292	292	292	97	97	97

Notes: First six columns include time dummies. Equations were estimated by OLS. The standard errors in the first six columns allow for correlated errors for the same country in different time periods. Maximum number of countries is 110. Columns 1–3 consist of changes for 1960–65, 1965–70, 1970–75, 1975–80, 1980–85, 1985–90. Columns 4–6 consist of changes for 1960–70, 1970–80, 1980–90. Columns 7–9 consist of changes for 1965–85. Log Y_{t-1} and S_{t-1} are the log GDP per capita and level of schooling in the initial year of each period. ΔS is the change in schooling between $t-1$ and t divided by the number of years in the period. Data are from Summers and Heston and Barro and Lee. Mean (and standard deviation) of annualized per capita GDP growth is .021 (.033) for columns 1–3, .022 (.026) for columns 4–6, and .022 (.020) for columns 7–9.

(2) Krueger and Lindahl (2001)

- Examine the relationship between economic growth and education growth over different time periods. Since the underlying education stock is slow moving, over shorter intervals Ω is likely to be larger thus exacerbating measurement error

  -- Using the best data, a longer time period (20 years), and correcting for attenuation bias yields a return of 30% to an additional year of education attained (on average), where $0.182/0.577 \approx 0.3$

  -- Alternatively one measure IV's for the other (Table 4)

- The social return to education – or endogeneity / OVB?

"Education," as Harbison and Myers (1965) stress, "is both the seed and the flower of economic development." It is difficult to separate the causal effect of education from the positive income demand for education in cross-country data over long time periods. N. G.

(2) Krueger and Lindahl (2001)

- Many other interesting findings on the heterogeneous effects of initial schooling levels on growth; on the linear specification assumption in most cross-country growth models; on IV specifications using multiple schooling measures, …

- Bottom line: there is not much we can say about the causal effect of more schooling on economic growth using cross-country data, due to both measurement and identification issues

(3) Duflo (2001, *AER*)

- The ideal experiment would randomize educational chances (by varying costs or subsidies, perhaps) across individuals, as well as across regions, to estimate externalities

- Duflo (2001) is the most reliable estimate of returns to education in a less developed country

  -- Studies the impact of a massive school building campaign in Indonesia during the oil-rich 1970s. What impact did this expansion have on later schooling attainment? On later wages?

(3) Duflo (2001, *AER*)

- Between 1973-1978 the government built 61,000 additional primary schools, doubling the number of classrooms in the country. The number of teachers also increased by 43% (!) during this period. This could be thought of as a sharp drop in the price of primary education for many households (e.g., travel costs)

- Poor areas were supposed to be targeted, but not exactly following the formula – schools were supposed to be built in proportion to the number of children out of school in 1973 (Table 2)

TABLE 2—THE ALLOCATION OF SCHOOLS

	Log(INPRES schools)[a]
Log of number of children aged 5–14 in the region	0.78
	(0.027)
Log(1 − enrollment rate in primary school in 1973)[b]	0.12
	(0.038)
Number of observations	255
R^2	0.78

Notes: Standard errors are in parentheses.

[a] The dependent variable is the log of the number of INPRES schools built between 1973 and 1978.

[b] The enrollment rate in primary school is the number of children enrolled in primary school in 1973 (obtained from the Ministry of Education and Culture) divided by the number of children aged 5–14 in the region in 1973.

(3) Duflo (2001, *AER*)

- Focuses on the 1995 labor market outcomes of men born between 1950-1972 (using the SUPAS intercensal household survey)

- Difference in differences strategy: compare cohorts too old to benefit to those who benefited from the program, across areas with more versus fewer schools built

- IV-2SLS estimation:

  School construction (instrumental variable)

  → educational attainment (endogenous variable)

  → wages (outcome variable)

(3) Duflo (2001, *AER*)

- Consider the impact of the program on school attainment in the first stage:

$$S_{ijk} = c + \alpha_j + \beta_k + (P_j{*}T_i)\gamma + (Z_j{*}T_i)\delta + \varepsilon_{ijk}$$

where S is the amount of schooling for an individual i, in region j and age cohort k. Let c be a constant, α_j be an indicator for district of individual birth, β_k be cohort indicator variables, P_j denotes program intensity in region j , Z_j are other regional controls, and T is an indicator taking on a value of one if the individual was young enough to benefit from the program

(3) Duflo (2001, *AER*)

- An identification concern is the exclusion restriction: other targeted programs in the same areas

  -- Would there have been convergence across regions even in the absence of the school-building program?

  -- Did quantity and quality of education change?

- The performance of older cohorts in programs districts serves as a sort of internal control to capture local trends

- Bottom line: returns to schooling in Indonesia in 1995 between 5-10% per year

  -- Poor and low density regions appear to benefit most

Table 4—Effect of the Program on Education and Wages: Coefficients of the Interactions Between Cohort Dummies and the Number of Schools Constructed per 1,000 Children in the Region of Birth

First stage

Reduced form

		Dependent variable					
		Years of education			Log(hourly wage)		
	Observations	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Experiment of Interest: Individuals Aged 2 to 6 or 12 to 17 in 1974 (Youngest cohort: Individuals ages 2 to 6 in 1974)							
Whole sample	78,470	0.124	0.15	0.188			
		(0.0250)	(0.0260)	(0.0289)			
Sample of wage earners	31,061	0.196	0.199	0.259	0.0147	0.0172	0.0270
		(0.0424)	(0.0429)	(0.0499)	(0.00729)	(0.00737)	(0.00850)
Panel B: Control Experiment: Individuals Aged 12 to 24 in 1974 (Youngest cohort: Individuals ages 12 to 17 in 1974)							
Whole sample	78,488	0.0093	0.0176	0.0075			
		(0.0260)	(0.0271)	(0.0297)			
Sample of wage earners	30,225	0.012	0.024	0.079	0.0031	0.00399	0.0144
		(0.0474)	(0.0481)	(0.0555)	(0.00798)	(0.00809)	(0.00915)
Control variables:							
Year of birth*enrollment rate in 1971		No	Yes	Yes	No	Yes	Yes
Year of birth*water and sanitation program		No	No	Yes	No	No	Yes

Notes: All specifications include region of birth dummies, year of birth dummies, and interactions between the year of birth dummies and the number of children in the region of birth (in 1971). The number of observations listed applies to the specification in columns (1) and (4). Standard errors are in parentheses.

Wald = 0.027/0.259 ≈ 10%

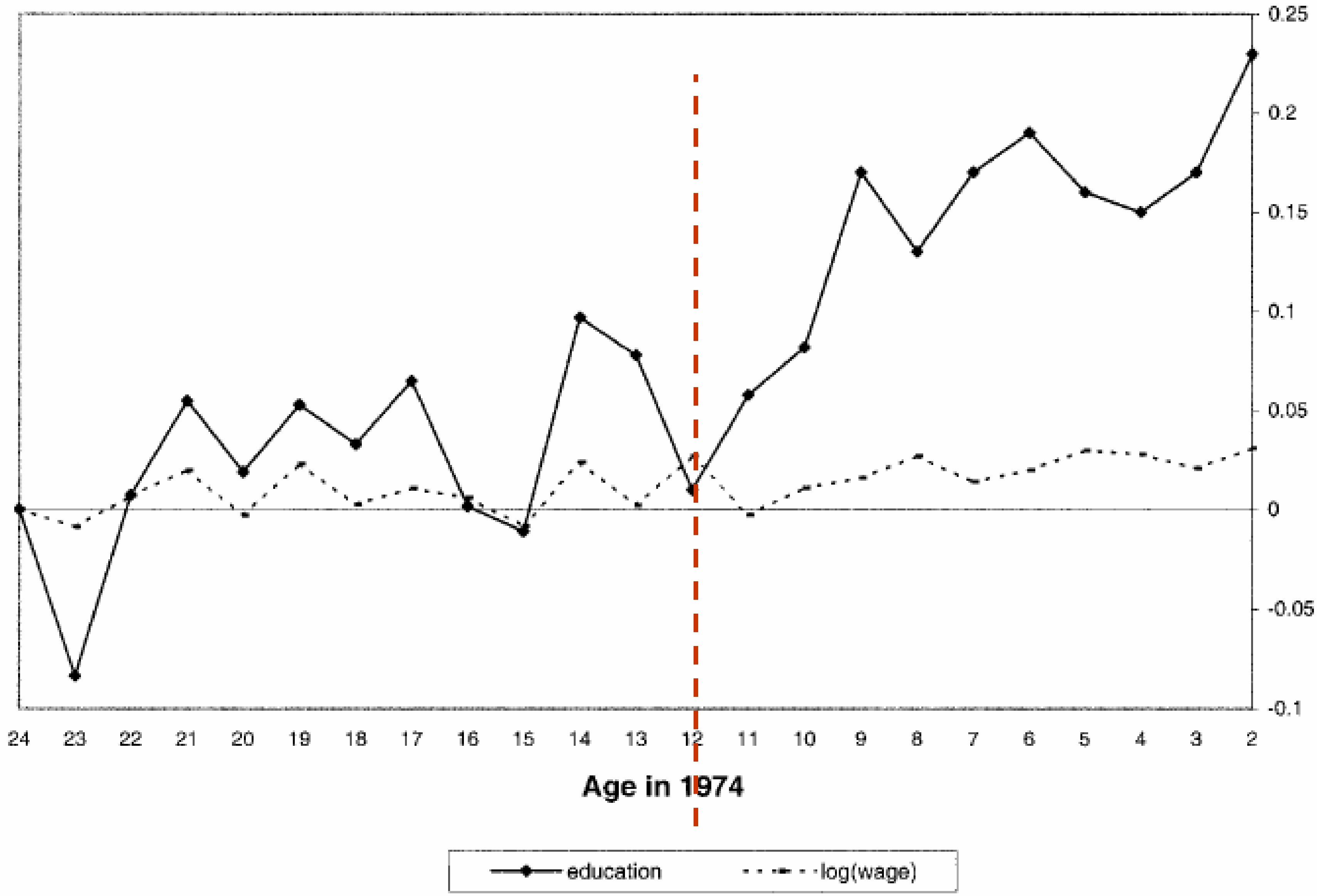


FIGURE 3. COEFFICIENTS OF THE INTERACTIONS AGE IN 1974* PROGRAM INTENSITY IN THE REGION OF BIRTH IN THE WAGE AND EDUCATION EQUATIONS

TABLE 6—PROGRAM EFFECT AND RETURNS TO EDUCATION BY CATEGORIES OF REGION OF BIRTH

		Characteristics of region of birth					
		Density[a]		1976 Poverty[b]		Preprogram education[c]	
	Whole sample (1)	<Median (2)	>Median (3)	High (4)	Low (5)	<Median (6)	>Median (7)
Panel A: Effect of the Program on Education							
Dependent variable: Years of education. Sample: individuals ages 2 to 6 or 12 to 17 in 1974							
Interaction	0.15	0.19	−0.014	0.13	0.083	0.14	0.13
(2–6 in 1974)*program intensity in region of	(0.026)	(0.035)	(0.048)	(0.058)	(0.035)	(0.040)	(0.036)
Panel B: Effect of the Program on Wages							
Dependent variable: log(hourly wage). Sample: individuals ages 2 to 6 or 12 to 17 in 1974 (wage earners)							
Interaction	0.017	0.032	−0.00084	0.051	−0.00083	0.028	0.0046
(2–6 in 1974)*program intensity in region of	(0.0074)	(0.011)	(0.012)	(0.017)	(0.0094)	(0.013)	(0.0095)
Panel C: Returns to Education							
Dependent variable: log(hourly wage). Sample: wage earners							
Years of education	0.078	0.11	No First stage	0.10	No First stage	0.12	0.029
	(0.00062)	(0.026)		(0.028)		(0.032)	(0.052)
	[0.9]	[0.86]		[0.88]		[0.72]	[0.83]

Table 7—Effect of Education on Labor Market Outcomes: OLS and 2SLS Estimates

Control function

Method	Instrument	(1)	(2)	(3)	(4)
Panel A: Sample of Wage Earners					
Panel A1: Dependent variable: log(hourly wage)					
OLS		0.0776	0.0777	0.0767	
		(0.000620)	(0.000621)	(0.000646)	
2SLS	Year of birth dummies*program intensity in region of birth	0.0675	0.0809	0.106	0.0908
		(0.0280)	(0.0272)	(0.0222)	(0.0541)
		[0.96]	[0.9]	[0.93]	[0.9]
2SLS	(Aged 2–6 in 1974)*program intensity in region of birth	0.0752	0.0862	0.104	
		(0.0338)	(0.0336)	(0.0304)	
		(0.0338)	(0.0336)	(0.0304)	
Panel A2: Dependent variable: log(monthly earnings)					
OLS		0.0698	0.0698	0.0689	
		(0.000601)	(0.000602)	(0.000628)	
2SLS	Year of birth dummies*program intensity in region of birth	0.0756	0.0925	0.0913	0.134
		(0.0280)	(0.0278)	(0.0219)	(0.0631)
		[0.73]	[0.63]	[0.58]	[0.7]
Panel B: Whole Sample					
Panel B1: Dependent variable: participation in the wage sector					
OLS		0.0328	0.0327	0.0337	
		(0.00311)	(0.000311)	(0.000319)	
2SLS	Year of birth dummies*program intensity in region of birth	0.101	0.118	0.0892	
		(0.0210)	(0.0197)	(0.0162)	
		[0.66]	[0.93]	[1.12]	
Panel B2: Dependent variable: log(monthly earnings), imputed for self-employed individuals					
OLS		0.0539	0.0539	0.0539	
		(0.000354)	(0.000354)	(0.000355)	
2SLS	Year of birth dummies*program intensity in region of birth	0.0509	0.0745	0.0346	
		(0.0157)	(0.0136)	(0.0138)	
		[0.68]	[0.58]	[1.16]	
Control variables:					
Year of birth*enrollment rate in 1971		No	Yes	Yes	Yes
Year of birth*water and sanitation program		No	No	Yes	No
Propensity score, propensity score squared		No	No	No	Yes

(3) Duflo (2001, *AER*)

- What is the rate of return of the program?

  -- Estimated returns are highly sensitive to post-construction income growth in Indonesia

  -- Under fast growth (like that observed in 1970s-1990s), education investments high rates of return, 8.8 to 12%. Under slow growth, returns to this program probably would have been small or even negative

- Given this finding, forward looking governments' education investments might be endogenous to growth prospects – further complicating cross-country results

TABLE 8—EVALUATION OF THE PROGRAM'S NET RETURN

	Deadweight loss 0.2 (1)	(2)
Panel A: Results		
Control for year of birth*enrollment rate	No	Yes
First year where benefit > costs (discount rate = 5 percent)		
In annual value	1996	1996
In discounted sum	2005	2002
Discounted sum of net benefits in 2050 (growth rate after 1997 = 5 percent, discount rate 5 percent)		
In million 1990 U.S.$	13,025	13,096
As a fraction of Indonesia's GDP in 1973	0.30	0.36
Divided by initial costs	24.1	24.2
Discounted sum of net benefits in 2050 (growth rate after 1997 = 2 percent, discount rate 5 percent)		
In million 1990 U.S.$	6,691	11,589
As a fraction of Indonesia's GDP in 1973	0.18	0.32
Divided by initial costs	12.4	21.4
Discounted sum of net benefits in 2050 (growth rate from 1973 = 2 percent, discount rate 5 percent)		
In million 1990 U.S.$	−631.6	1,200
As a fraction of Indonesia's GDP in 1973	−0.017	0.033
Divided by initial costs	−1.16	2.22
Internal rate of return[a]		
Growth rate after 1997 = 5 percent	0.102	0.118
Growth rate after 1997 = 2 percent	0.088	0.106
Growth rate from 1973 = 2 percent	0.0443	0.059
Panel B: Assumptions and Parameters		
Population growth rate after 1997	0.015	
Yearly teacher's salary in 1973 (1990 U.S. dollars)	363	
Yearly teacher's salary in 1995 (1990 U.S. dollars)	2,467	
Total recurrent costs/teacher salary	1.25	
Total cost of construction (million 1990 U.S. dollars)	522	
Number of schools constructed	61,800	
Lifetime of the schools (years)	20	
Share of labor income in GDP	0.7	

(3) Duflo (2001, *AER*)

- Looking ahead to next week:

  -- If education does have sizeable private (and perhaps even larger social) returns, should public resources be spent on education in less developed countries? If so, what types of investments should be made?

  -- Pupil-teacher ratios, textbooks, the organization of the school system / teacher’s unions, incentives for teachers, students, parents, ….

- Building a sense of national identity and cohesion is a social return to education that may be important but is hard to estimate with microeconometric methods

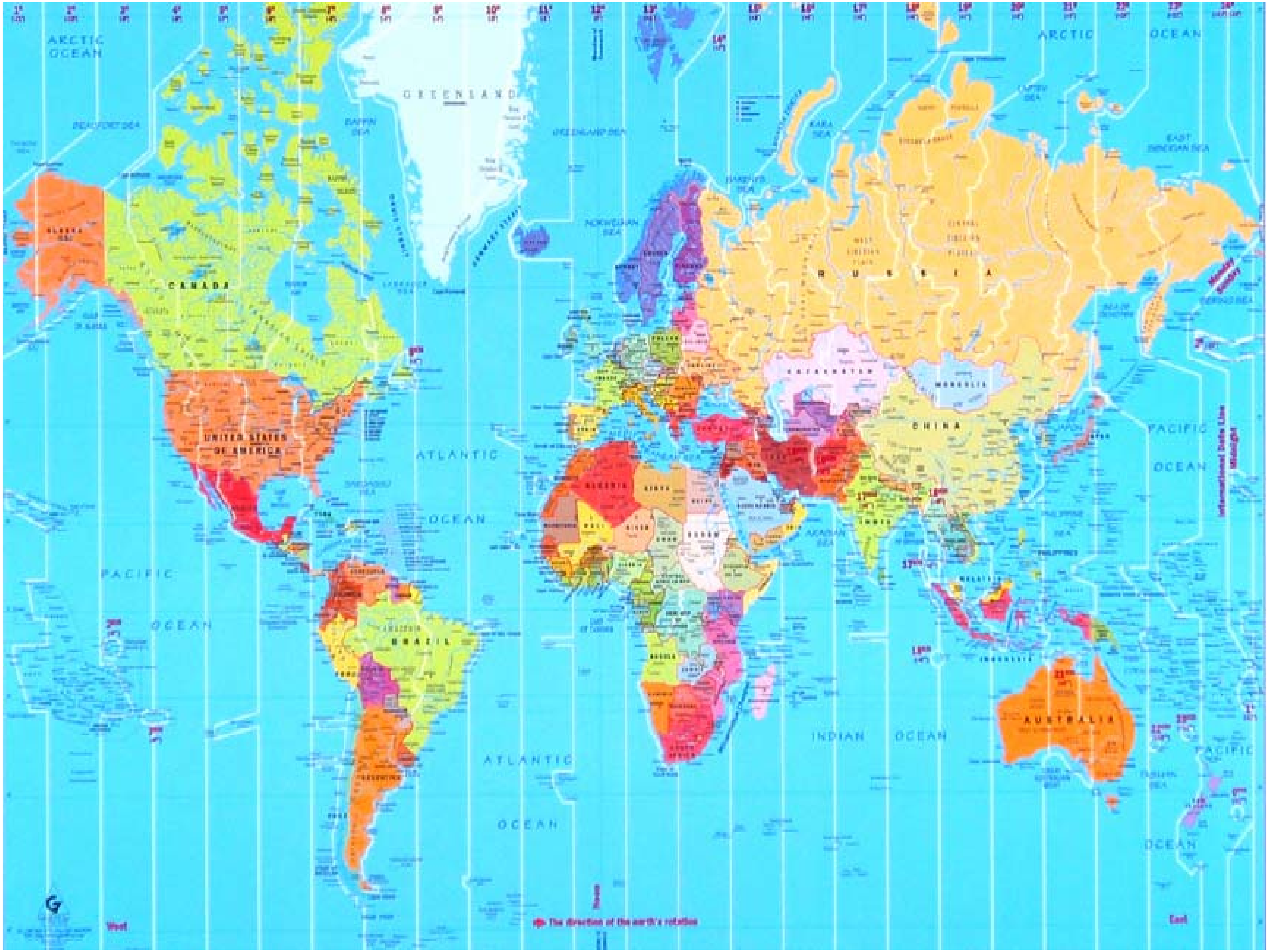
Whiteboard #1

Whiteboard #2

Whiteboard #3

Whiteboard #4

Whiteboard #5


ARCTIC OCEAN
ARCTIC OCEAN
BEAUFORT SEA
BAFFIN SEA
GREENLAND
GREENLAND SEA
NORWEGIAN SEA
BARENTS SEA
KARA SEA
EAST SIBERIAN SEA
ALASKA
CANADA
UNITED STATES OF AMERICA
RUSSIA
KAZAKHSTAN
MONGOLIA
CHINA
INDIA
ALGERIA
LIBYA
NIGER
SUDAN
ARABIAN SEA
ATLANTIC
OCEAN
PACIFIC
OCEAN
BRAZIL
ARGENTINA
ATLANTIC
OCEAN
INDIAN OCEAN
INDONESIA
AUSTRALIA
PACIFIC
OCEAN
PACIFIC
OCEAN
West
East
The direction of the earth's rotation