Why Don't We See Poverty Convergence?

Martin Ravallion

Development Research Group,

World Bank

Two “stylized facts” about development

1. The advantage of backwardness—the mean convergence property whereby diminishing returns to aggregate capital imply higher growth rates in countries starting out with a low mean.

2. The advantage of growth—economic growth is distribution-neutral on average => economic growth reduces poverty. Strong form: “economic growth is the main driving force for poverty reduction.”

Qualified empirical support for both stylized facts

- Considerable empirical support for both views in the literature, though various qualifications.
  - Conditional vs unconditional convergence
  - Variance in poverty impacts of growth and some exceptions.

- That support has typically come from regressions that assume that the parameters of the dynamic processes for growth and poverty reduction are independent of the initial level of poverty.

But an implication has largely gone unnoticed.

And that implication is <u>not</u> supported by the data.

We should see "poverty convergence"

- It has largely gone unnoticed in the literature that, when taken together, these stylized facts imply "poverty convergence:"
  - a catching up process whereby the poorest countries should experience a higher rate of progress against poverty.
- And poverty rates should converge at the same speed as means!
- Simple expository log-linear model:

Growth model for the mean: $\Delta \ln \mu_{it} = \alpha_i + \beta_i \ln \mu_{it-1} + \varepsilon_{it}$

Poverty model: $\ln H_{it} = \delta_i + \eta_i \ln \mu_{it} + \nu_{it}$

Growth model for poverty: $\Delta \ln H_{it} = \alpha_i^* + \beta_i \ln H_{it-1} + \varepsilon_{it}^*$

$$\begin{cases} \alpha_i^* = \alpha_i \eta_i - \beta_i \delta_i \\ \varepsilon_{it}^* = \varepsilon_{it} \eta_i + \nu_{it} - (1+\beta_i)\nu_{it-1} \end{cases}$$

$\Rightarrow$The parameter determining the speed of convergence should be the same for the mean as the poverty measure

However, we do not find poverty convergence in the data

- As this paper will show, there is no correlation across countries between the initial levels of poverty and subsequent proportionate rates of poverty reduction.
- The overall incidence of poverty is falling in the developing world but no faster in the poorest countries.

So why don't we see poverty convergence?

In terms of the model above, it will be shown that α_i is a decreasing function of the initial poverty rate while $-\eta_i$, is a decreasing function of the initial level of poverty.

Outline

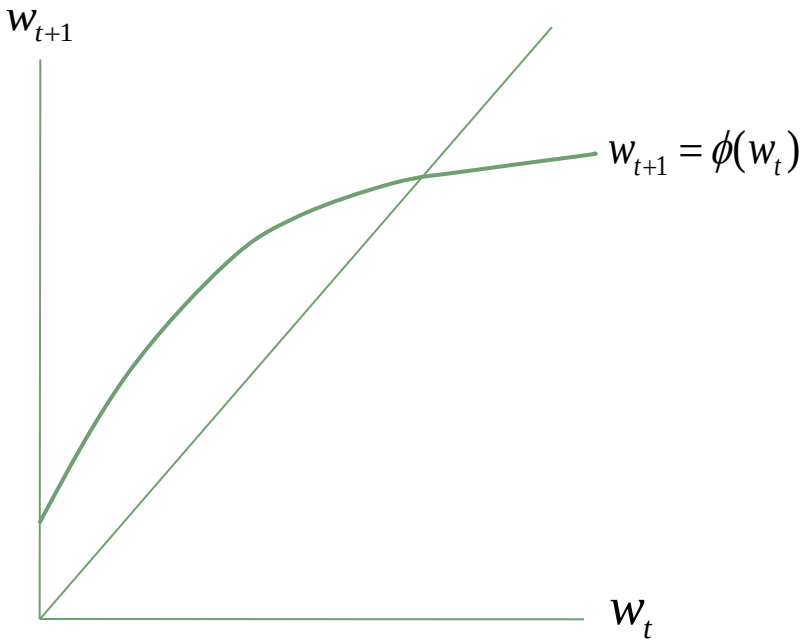
- Theories of distribution-dependent growth
- Past evidence on growth and initial distribution
- Data and descriptive statistics
- Testing the relevance of initial distribution to growth
- Initial distribution and the effect of growth on poverty
- Implications for poverty convergence

Theories of distribution-dependent growth based on credit-market failures

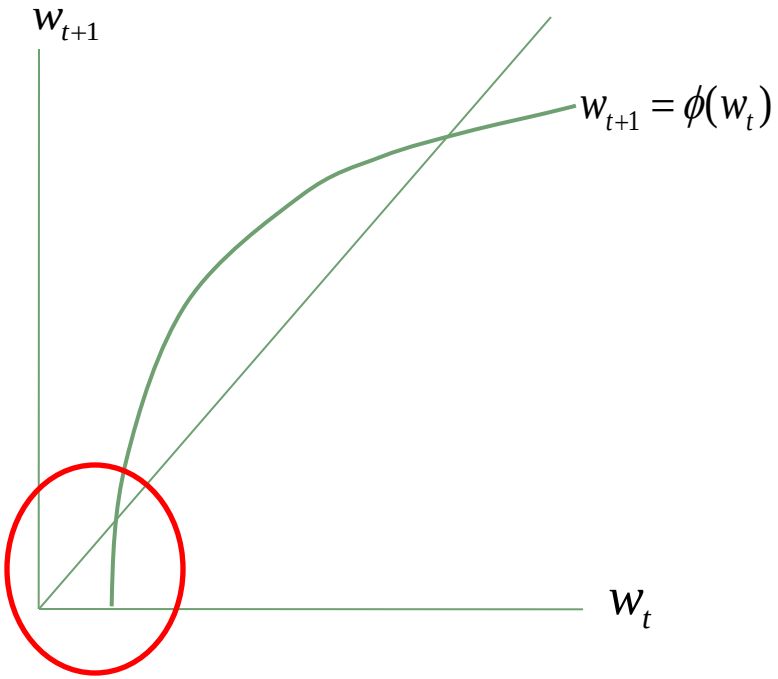
- Market failure attributed to information asymmetries, notably that lenders are imperfectly informed about borrowers.

- Key analytic feature: a suitably nonlinear relationship between an individual's initial wealth and her future wealth (the "recursion diagram").
  - With diminishing marginal products of capital, the mean future wealth will be a quasi-concave function of the distribution of current wealth
  - Thus higher current inequality implies lower future mean wealth at a given value of current mean wealth, i.e., lower growth.
  - Examples: Galor and Zeira (1993), Benabou (1996), Aghion and Bolton (1997) and Banerjee and Duflo (2003).

Nonlinear dynamics: High inequality handicaps growth



Unique equilibrium



Multiple equilibria

Dynamic implication of borrowing constraints: poverty also impedes growth

- Example: Banerjee and Duflo (2003) provide a simple but insightful growth model with a borrowing constraint.
  - Those with sufficient wealth will reach their unconstrained optimum, equating the marginal product of capital with the interest rate.
  - But the "wealth poor," for whom the borrowing constraint is binding, will not be able to do so.
  - Banerjee and Duflo show that higher inequality in such an economy implies lower growth.
- However, they do not observe that their model also implies that higher current wealth poverty for a given mean also implies lower growth.
- *In the Banerjee-Duflo model an unambiguously higher initial headcount index of poverty holding the initial mean constant implies a lower growth rate.*

Other theories of distribution-dependent growth

- Inequality restricts efficiency-enhancing cooperation amongst people, such that public goods needed for growth are underprovided or efficiency-enhancing policy reforms are blocked (Bardhan et al., 2000).

- Political-economy models of redistribution argue that high inequality leads democratic governments to implement distortionary redistributive policies, e.g., Alesina and Rodrik (1994).

Other theories cont.,

Other theoretical models => Poverty itself can impede growth and (hence) poverty reduction

1. Lasting (adverse) productivity effects of poor nutrition, esp., in childhood (Dasgupta and Ray, 1986; Cunha and Heckman, 2007).

2. Lopez and Servén (2009) introduce a subsistence consumption requirement into the utility function in the model of Aghion et al. (1999) and show that higher poverty incidence (failure to meet the subsistence requirement) implies lower growth.

Past evidence on growth and inequality

- Empirical support for the view that a higher Gini index of inequality impedes growth; Alesina and Rodrik (1994), Persson and Tabellini (1994), Birdsall et al., (1995), Clarke (1995), Perotti (1996), Deininger and Squire (1998) and Knowles (2005)
- However, not all the evidence has been supportive; also see Li and Zou (1999), Barro (2000) and Forbes (2000).
- The main reason why these studies have been less supportive appears to be that they have allowed for country-level fixed effects.

Data issues

Growth of what?

- Past growth empirics has almost only used GDP growth.
- Given the focus on poverty this study uses instead household expenditure on consumption at purchasing power parity, rather than income or GDP.
  - Tests for sensititivity to using private consumption from NAS
- Real consumption is a money metric of welfare under standard assumptions, but those assumptions can be questioned (non-market goods; inter-household inequality).

Inequality of what?

- Most past work has used growth and inequality in the space of incomes rather than wealth, which is closer to the concept implied by theory.
- An exception: Ravallion (1998) studied wealth inequality as a determinant of growth in China.)

Data issues cont.,

What inequality measure?

- The Gini index—half the mean absolute difference between all pairs of incomes normalized by the overall mean—has been (by far) the most popular inequality measure,
- This owes more to its availability in secondary data compilations than any intrinsic relevance to the economic arguments.

Is inequality acting as a proxy for poverty (at given mean)?

- As Lopez and Servén observe, the significance of the Gini index in past studies may well reflect an omitted variable bias,
- given that one expects that inequality will be highly correlated with poverty at a given mean.

Data issues cont.,

What control variables?

- The specification choices in past empirical work have lacked clear theoretical justification.
- Consider three of popular predictors of growth, namely human development, the investment share, and financial development.
  1. Basic schooling and health attainments (often significant in growth regressions) are arguably one of the channels linking initial distribution to growth; see original Galor and Zeria (1993) model.
  2. Share of investment in GDP (robust predictors of growth rates) is one of the main channels through which distribution affects growth from the theoretical literature.
  3. Private credit (as a share of GDP) has been used as a measure of "financial sector development" in explaining growth and poverty reduction (Beck et al., 2000, 2007). But theories based on borrowing constraints imply that the aggregate flow of credit in the economy depends on the initial distribution.

Data for this study

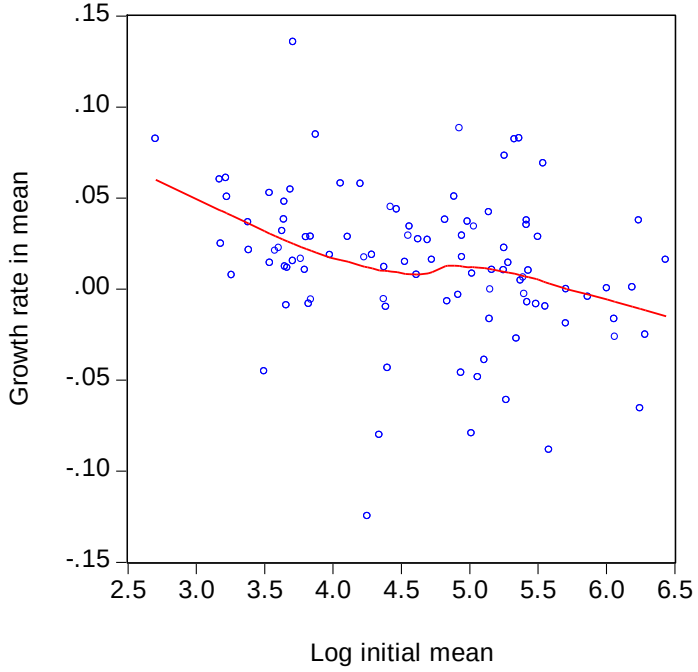
- 99 developing countries with at least two nationally representative surveys since 1980; 92 in which the earliest sample survey finds that at least some households lived below the poverty line.
- The longest spell between two surveys was used for each country.
- Both surveys used the same welfare indicator, either consumption or income per person, following standard measurement practices.
- When both were available, consumption was generally preferred. Three-quarters of the spells use consumption.
- Comparability problems between surveys remain, such as differences in recall periods and imputation/valuation methods.
- Median year of the first survey is 1991; median for the second is 2004. Median interval between surveys is 13 years and the interval varies from three to 27 years.
- All changes between the surveys are annualized.
- National accounts data were mapped as closely as possible to the survey dates, interpolating as need be.
- All monetary measures are in constant 2005 prices (using country-specific Consumer Price Indices) and all international comparisons are at PPP.

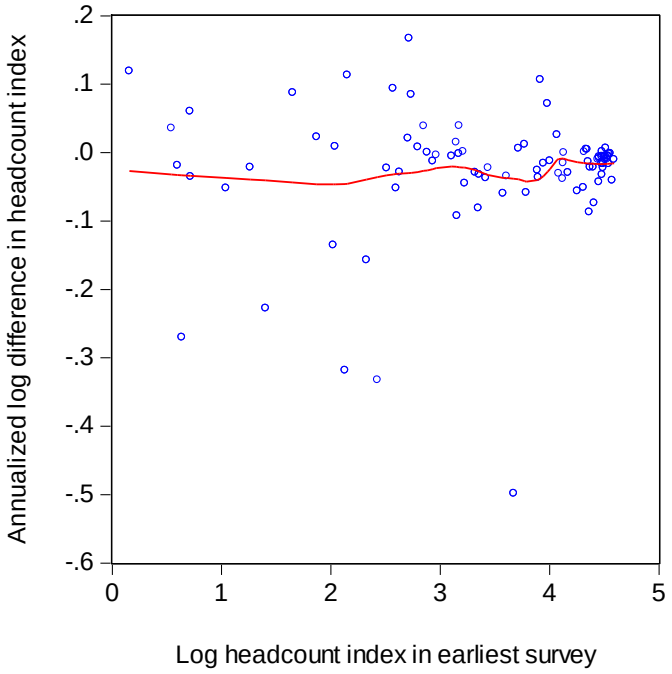
Measures of distribution

- Poverty is mainly measured by the headcount index for $2.00 per day at 2005 PPP, which is the median poverty line amongst developing countries.
  - Also lower line of $1.25 a day (mean of poorest 15 countries) and a much higher line of $13 a day in 2005 (US line)
- The size of the middle class is measured by the proportion of the population living in the interval $2 to $13 a day at 2005 purchasing power parity (PPP) (Ravallion, 2010).
- Those living above $13 a day can be thought of as the "middle class" by Western standards.
- Inequality is measured by the usual Gini index.

Convergence?

- The survey means exhibit (unconditional) convergence.
- But the poverty measures do not.





Estimated convergence parameters

	(1) Surveys means (full sample)	(2) Surveys means (consumption surveys only)	(3) Consumption per capita from NAS	(4) Headcount index ($2.00 a day)	(5) Headcount index ($1.25 a day)
Un-conditional	-0.013**	-0.010	-0.007	0.005	-0.005
	(-3.413)	(-1.882)	(-1.743)	(0.542)	(-0.393)
Conditional	-0.042**	-0.040**	-0.026**	-0.015	-0.028
	(-7.435)	(-4.928)	(-4.431)	(-1.035)	(-1.734)

Note: Controls included initial consumption per capita from the NAS, primary school enrollment rate, life expectancy at birth, and the price index of investment goods from Penn World Tables (6.2), which is a widely-used measure of market distortions; all three variables are matched as closely as possible to the date of the earliest survey.

Poverty and growth

- Benchmark regression:

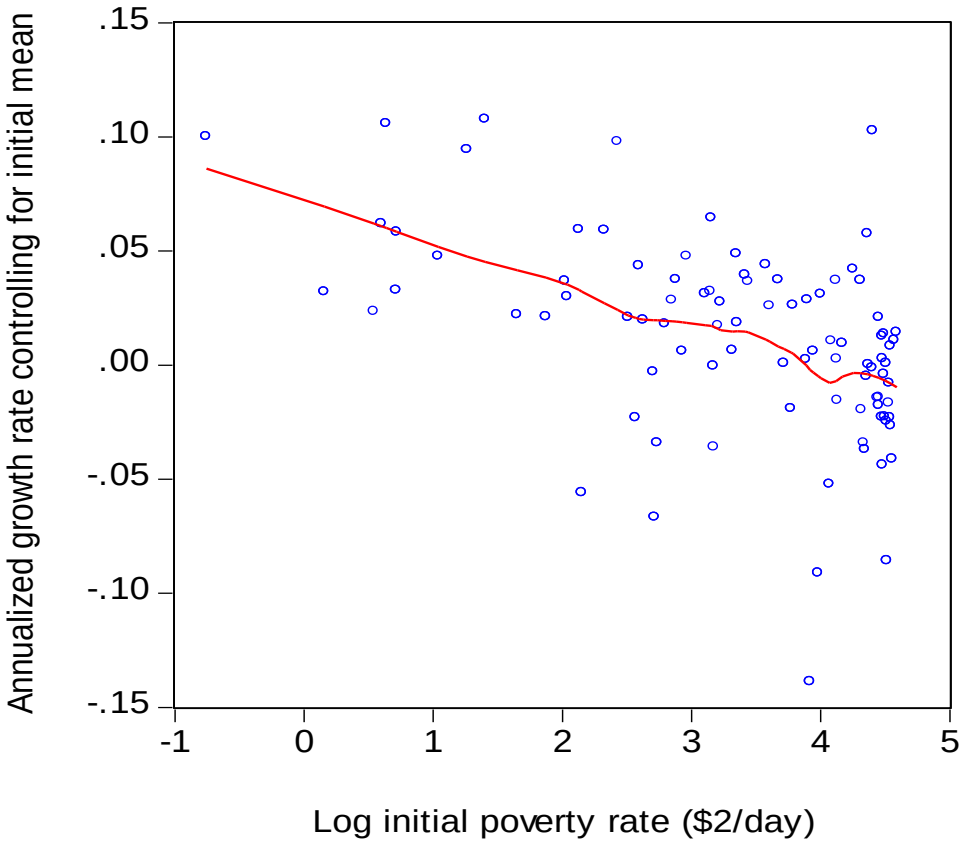
$$g_i(\mu_{it}) = \underset{(5.183)}{0.234} - \underset{(-5.131)}{0.035} \ln \mu_{it-\tau} - \underset{(-3.626)}{0.017} \ln H_{it-\tau} + \hat{\varepsilon}_{it}$$

where $g_i(\mu_{it}) \equiv \ln(\mu_{it} / \mu_{it-\tau}) / \tau$

- The regression is consistent with a derivative of current mean with respect to lagged mean that is less than unity, but fades toward zero at sufficiently long gaps between survey rounds.

Functional form test

A simple graphical test for misspecification is to plot $g_i(\mu_{it}) + 0.035 \ln \mu_{it-\tau}$ against $\ln H_{it-\tau}$ =>



The relationship is close to linear in the log poverty rate.

Alternative specifications

$$g_i(\mu_{it}) = \alpha + \beta \ln \mu_{it-\tau} + \gamma \ln H_{it-\tau} + \varepsilon_{it}$$

	(1)	(2)	(3)
		Sample with two surveys	
	Full sample	Consumption surveys only	NAS consumption per capita
Intercept	0.024**	0.300**	0.151**
	(5.183)	(5.850)	(3.705)
Log initial mean	-0.035**	-0.044**	-0.020**
	(-5.131)	(-5.318)	(-3.037)
Log initial headcount index	-0.017**	-0.025**	-0.011**
	(-3.626)	(-4.845)	(-2.711)
R^2	0.147	0.201	0.128
N	92	70	81

Using countries with 3+ surveys

	(4) Means from first two surveys used as initial conditions	(5) GMM estimator with IVs from earliest survey rounds	(6) As for (5) but using NAS consumption instead of survey means
Intercept	0.235**	0.180**	0.169**
	(4.569)	(2.772)	(3.517)
Log initial mean	-0.029**	-0.020*	-0.014*
	(-3.264)	(-1.994)	(-2.017)
Log initial headcount index	-0.022**	-0.020**	-0.022**
	(-6.305)	(-3.381)	(-4.749)
N	77	64	58

The subsample 70+ countries with 3+ surveys cont.

Inverted U in past changes in inequality?

- Banerjee and Duflo: it is not the level of initial inequality that matters to growth but past changes in inequality and that this has an inverted-U effect, whereby changes in inequality in either direction tend to reduce the growth rate.

- Test: annualized growth rates between the most recent and the middle survey and replacing the Gini index for the earliest survey by a quadratic function of the change in the Gini index between the earliest survey and the middle survey.
  - Coefficients for the quadratic function of the change in the lagged Gini index were individually and jointly insignificant in the regressions for both growth rates.
  - Nor was there any sign of an inverted U relationship with the lagged changes in the poverty rate.

The subsample 70+ countries with 3+ surveys cont.

Non-robustness to fixed effects in growth rates

- FE addresses the problem of time-invariant latent heterogeneity but it is unlikely to have much power for detecting the true relationships given that the changes in growth rates have a low signal-to-noise ratio.
- Simulation studies: coefficients on growth determinants are heavily biased toward zero in fixed-effects growth regressions (Hauk and Wacziarg, 2009).
- For eaxmple, using FE in growth rates the elasticity drops to an implausibly low figure, undoubtedly reflecting an attenuation bias

		$\hat{\beta}$	n	R^2
Levels for latest survey	$\ln \mu_{it} = \alpha + \beta \ln C_{it} + \varepsilon_{it}$	0.747 (21.463)	97	0.823
Fixed effects in levels	$g_i(\mu_{it}) = \beta g_i(C_{it}) + \varepsilon_{it}$	0.508 (4.936)	92	0.208
Fixed effects in growth rates	$\Delta g_i(\mu_{it}) = \beta \Delta g_i(C_{it}) + \varepsilon_{it}$	0.094 (7.389)	65	0.069

Encompassing regressions

	(1) Survey Means	(2) Consumption from NAS	(3) Survey Means	(4) Consumption from NAS
Intercept	0.442	0.617	0.26	-0.275
	(0.795)	(1.234)	(1.279)	(-1.914)
Initial mean	-0.058**	-0.035**	-0.060**	-0.030**
	(-5.961)	(-3.657)	(-6.912)	(-3.764)
Initial poverty rate	-0.027**	-0.017**	-0.027**	-0.014**
	(-5.482)	(-3.033)	(-5.750)	(-3.024)
Initial Gini index	-0.020	-0.081	0	0
	(-0.400)	(-1.784)		
Initial income share of middle	-0.117	-0.167*	-0.091**	0
three quintiles	(-1.477)	(-2.167)	(-3.985)	
Initial share of population in	-0.102*	-0.128**	-0.110**	-0.133**
Western middle class	(-2.284)	(-2.815)	(-2.432)	(-3.691)
Initial primary school	0.007	0.003	0	0
enrolment rate	(0.700)	(0.271)		
Initial life expectancy	0.117**	0.154**	0.129**	0.139**
	(2.768)	(3.653)	(3.068)	(3.665)
Initial price of investment	-0.014**	-0.016**	-0.014**	-0.017**
	(-2.650)	(-3.140)	(-2.698)	(-3.434)
N	0.434	0.470	0.430	0.453
R^2	88	84	88	87

Initial distribution and growth elasticity of poverty reduction

- In general, the growth elasticity of poverty reduction will depend on the initial distribution.
- This can be thought of as the direct effect of the initial distribution on the rate of poverty reduction, as distinct from the indirect effect via the rate of growth.
- Past work has focused on inequality as the relevant aspect of initial distribution.

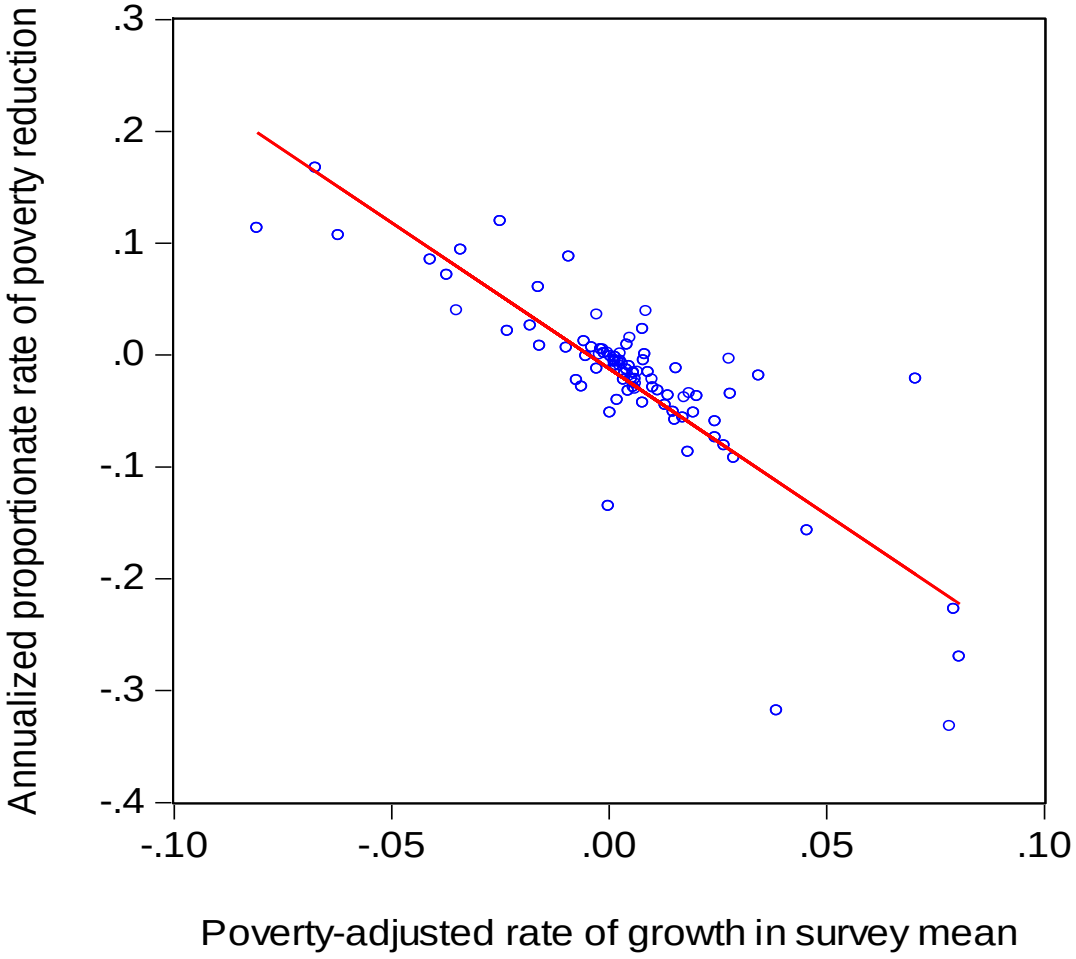
Regressions for rate of poverty reduction

	(1) OLS	(2) IVE	(3) OLS	(4) IVE	(5) OLS	(6) IVE
Intercept	0.002	0.008	-0.012	-0.005	-0.012**	-0.008
	(0.078)	(0.267)	(-1.908)	(0.607)	(-2.175)	(-1.365)
Initial poverty rate	-0.004	0.008	0	0	0	0
	(-0.792)	(0.267)				
Growth rate (annualized	-2.674**	-3.564**	-2.615**	-3.323**	0	0
change in log mean)	(-6.660)	(-4.325)	(-6.608)	(-4.560)		
Growth rate interacted with	2.780**	3.492**	2.621**	3.101**	0	0
initial poverty rate	(5.206)	(3.650)	(4.915)	(3.746)		
(1-Poverty rate) *times*	0	0	0	0	-2.613**	-3.294**
growth rate					(-7.273)	(-4.585)
N	91	86	91	86	91	86
R^2	0.537	0.439	0.535	0.458	0.535	0.466
Homogeneity test	0.673	-0.215	0.037	-0.620	n.a.	n.a.

Homogeneity tests passes => the relevant growth rate is the poverty-adjusted rate, as given by the growth rate *times* one minus the poverty rate.

$$g_i(H_{it}) = \eta(1 - H_{it-\tau})g_i(\mu_{it}) + \varepsilon_{it}$$

Rate of poverty reduction is proportional to the distribution-corrected rate of growth



Poverty makes growth less pro-poor

- The (absolute) growth elasticity of poverty reduction tends to be lower in countries with a higher initial poverty rate.
- Poorer countries tend to experience lower proportionate effects on their poverty measures from any given rate of growth.
- At an initial poverty rate of 10% (about one standard deviation below the mean) the elasticity is about -3 (using the IVE) while it falls to about -0.7 at a poverty rate of 80% (about one standard deviation above the mean).
- The interaction effect with the poverty rate is stronger than that with the partial elasticity of poverty reduction derived analytically.

So why don't we see poverty convergence?

Preferred model:

$$g_i(H_{it}) = \eta(1 - H_{it-\tau})g_i(\mu_{it}) + \varepsilon_{it}$$

$$g_i(\mu_{it}) = \alpha + \beta \ln \mu_{it-\tau} + \gamma \ln H_{it-\tau} + \upsilon_{it}$$

Poverty convergence elasticity:

$$\frac{\partial g_i(H_{it})}{\partial \ln H_{it-\tau}} = \eta\beta(1 - H_{it-\tau})\left(\frac{\partial \ln H_{it-\tau}}{\partial \ln \mu_{it-\tau}}\right)^{-1} + \eta\gamma(1 - H_{it-\tau}) - \eta g_i(\mu_{it})H_{it-\tau}$$

Total effect =0.006	Mean convergence Effect=-0.038	Direct effect of poverty=0.024	Poverty elasticity effect=0.020

Conclusions

- Consistently with theoretical models of economic growth incorporating borrowing constraints, there is an adverse effect on consumption growth of high initial poverty at a given mean.
- A high initial incidence of poverty also entails a lower subsequent rate of progress against poverty at a given growth rate (and poor countries tend to experience less steep increases in poverty during recessions).
- For many poor countries, the growth advantage of starting out with a low mean ("conditional convergence") is lost due to their high poverty rates.
- High current inequality is only a handicap if it entails a high incidence of poverty relative to mean consumption.