

Why Don't We See Poverty Convergence?

Martin Ravallion

The World Bank
Development Research Group
Director's Office
June 2009



Abstract

We are not seeing faster progress against poverty amongst the poorest developing countries. Yet this is implied by widely accepted “stylized facts” about the development process. The paper tries to explain what is missing from those stylized facts. Consistently with models of economic growth incorporating borrowing constraints, the analysis of a new data set for 100 developing countries reveals an adverse effect on consumption growth of high initial poverty incidence at a given initial mean. A high incidence of poverty also entails a lower subsequent rate of progress against poverty at any given

growth rate (and poor countries tend to experience less steep increases in poverty during recessions). Thus, for many poor countries, the growth advantage of starting out with a low mean (“conditional convergence”) is lost due to their high poverty rates. The size of the middle class—measured by developing-country, not Western, standards—appears to be an important channel linking current poverty to subsequent growth and poverty reduction. However, high current inequality is only a handicap if it entails a high incidence of poverty relative to mean consumption.

This paper—a product of the Director's Office, Development Research Group—is part of a larger effort in the department to understand why some countries are more successful in the fight against poverty than others. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The author may be contacted at mravallion@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Why Don't We See Poverty Convergence?

Martin Ravallion¹

*Development Research Group, World Bank
1818 H Street NW, Washington DC, 20433, USA*

¹ The author is grateful to Shaohua Chen and Prem Sangraula for help in setting up the data set used here. Helpful comments were received from Francois Bourguignon, Sebastian Galiani, Karla Hoff, Aart Kraay, Kalle Moene, Luis Servén, Dominique van de Walle, participants at presentations at the World Bank, the Courant Research Center, Göttingen, the 2009 meeting in Buenos Aires of the Network on Inequality and Poverty, Cornell University, the Center for Global Development, the University of California Berkeley, the University of Oslo and the Paris School of Economics. These are the views of the author and should not be attributed to the World Bank or any affiliated organization.

1. Introduction

Two prominent stylized facts about economic development are that there is an advantage of backwardness, such that on comparing two otherwise similar countries the one with the lower initial mean income will tend to see the higher rate of economic growth, and that there is an advantage of growth, whereby a higher mean income tends to come with a lower incidence of absolute poverty. Past empirical support for both stylized facts has almost invariably assumed that the dynamic processes for growth and poverty reduction do not depend directly on the initial level of poverty. Under that assumption, the two stylized facts imply that we should see poverty convergence: countries starting out with a high incidence of absolute poverty should enjoy a higher subsequent growth rate in mean consumption and (hence) a higher proportionate rate of poverty reduction.

That poses a puzzle. The data on poverty measures over time for 90 developing countries assembled for this paper reveal little or no sign of poverty convergence. Figure 1 plots the proportionate rate of change in poverty—specifically the annualized log difference between household surveys in the percentage of each country’s population living below \$2 a day at 2005 purchasing power parity—against its initial level. (The data are described later.) There is no sign of convergence; the regression line has a slope of 0.006, with a t-ratio of 0.590, based on a robust (White) standard error. The overall poverty rate of the developing world has been falling since at least 1980 (Chen and Ravallion, 2010), but the proportionate rate of decline has been no higher in its poorest countries.

Clearly something important is missing from the story. Intuitively, one hypothesis is that either the growth rate in the mean, or the impact of growth on poverty, depends directly on the initial poverty rate, in a way that nullifies the “advantage of backwardness.” To test this hypothesis, the paper estimates a model in which the proportionate rate of progress against poverty depends on the rate of growth in mean consumption and the poverty rate, while the rate of growth in the mean depends in turn on the initial poverty rate as well as the initial mean.

The results suggest that mean-convergence is counteracted by two distinct “poverty effects.” First, there is an adverse direct effect of high initial poverty on growth—working against convergence in mean consumption. Second, high initial poverty dulls the impact of growth on poverty. On balance there is little or no systematic effect of starting out poor on the proportionate rate of poverty reduction.

In the process of documenting these findings the paper also explores the role played by other aspects of the initial distribution discussed in the literature, including inequality. These are found to play no more than a subsidiary role. For example, high initial inequality only matters to growth and poverty reduction in so far as it entails a high initial incidence of poverty relative to the mean. And the paper confirms that countries starting out with a small middle class—judged by developing country rather than Western standards—face a handicap in promoting growth and poverty reduction, but this too is largely accountable to differences in the incidence of poverty.

After reviewing the literature in the next section, section 3 provides a simple theoretical model of how higher initial poverty can impede economic growth, by building on an existing model in the literature. The data are then described in section 4 while section 5 tests for convergence in both the mean and the poverty rate. The main results, including various tests of their robustness, are then presented in sections 6 (on how poverty affects growth) and 7 (on how it affects the elasticity of poverty to growth). Section 8 brings these elements together to calibrate a decomposition of the speed of convergence in poverty, which answers the question in the paper's title. Section 9 concludes.

2. Past theories and evidence

A number of papers have demonstrated that an economy's growth path can depend on parameters of the initial distribution of income. The parameter that has received most attention is inequality. One way that high inequality can reduce an economy's aggregate output is when borrowing constraints stemming from credit market failures leave unexploited opportunities for investment in physical and human capital (Galor and Zeira, 1993, Bénabou, 1996; Aghion and Bolton, 1997).² With diminishing marginal products of capital, 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 current mean wealth. A similar result can be obtained when high inequality prompts distortionary policy responses (as in Alesina and Rodrik, 1994) or restricts efficiency-enhancing cooperation, such that key public goods are underprovided or efficiency-enhancing reforms are blocked (as in the models reviewed in Bardhan et al., 2000).

² Also see the discussions in Perotti (1996), Hoff (1996), Aghion et al. (1999), Bardhan et al. (2000), Banerjee and Duflo (2003), Azariadis (2006) and World Bank (2006, Chapter 5).

Motivated by these theoretical arguments, a subset of the (large) empirical literature on the determinants of economic growth has included explicit measures of inequality as regressors for growth, although inferences are clouded by the fact that the regressions often also include variables that are implicitly functions of inequality, such as human development attainments, aggregate investment shares and measures of financial-sector development.³ A number of empirical papers have reported adverse effects of inequality on growth.⁴ Typically a single aggregate inequality index is used and the measure that has received most attention is the Gini index.⁵ Voitchovsky (2005) argues that a single aggregate measure of inequality might well miss the impacts; with some support from data for developed countries, she argues that inequality amongst low incomes is bad for growth, but that the opposite holds for “high-end” inequality.

Another strand of the literature has argued that the size of a country’s middle class matters to economic growth, by fostering entrepreneurship, or shifting the composition of consumer demand, or making it more politically feasible to attain policy reforms and institutional changes conducive to growth.⁶

While this literature has focused on inequality or the middle class, arguments can also be made suggesting that poverty may well be the more relevant parameter of the initial distribution.

- Models of economic growth with borrowing constraints, such as the model in Abhijit Banerjee and Esther Duflo (2003), also imply that a higher current incidence of poverty—defined by any poverty line up to the minimum level of initial wealth needed to

³ Basic schooling and health attainments (often significant in growth regressions) are one of the potential channels linking initial distribution to growth, as in Galor and Zeria (1993). Similarly, while the share of investment in GDP has often been used as a predictor of growth rates (Levine and Renelt, 1992), this is one of the channels identified in the theoretical literature linking inequality to growth. The same argument can be made about private credit (as a share of GDP) as a measure of “financial sector development” (Beck et al., 2007); growth theories based on borrowing constraints suggest that the aggregate flow of credit depends on the initial distribution.

⁴ Support for the view that higher initial inequality impedes growth has been reported by Alesina and Rodrik (1994), Persson and Tabellini (1994), Birdsall et al., (1995), Clarke (1995), Perotti (1996), Deininger and Squire (1998), Ravallion (1998), Knowles (2005) and Voitchovsky (2005) (amongst others). Not all the evidence has been supportive; also see Li and Zou (1999), Barro (2000) and Forbes (2000). The main reason why the latter studies have been less supportive appears to be that they have allowed for additive country-level fixed effects in growth rates; I will return to this point.

⁵ Wealth inequality is arguably more relevant though this has rarely been used due to data limitations. An exception is Ravallion (1998), who studies the effect of geographic differences in the distribution of wealth on growth in China and finds evidence that high wealth-inequality impedes growth.

⁶ Analyses of the role of the middle class in promoting entrepreneurship and growth include Acemoglu and Zilibotti (1997) and Doepke and Zilibotti (2005). Middle-class demand for higher quality goods plays a role in the model of Murphy et al. (1989). Birdsall et al. (2000) conjecture that support from the middle class is crucial to reform. Sridharan (2004) describes the role of the Indian middle class in promoting reform. Easterly (2001) finds evidence that a larger income share controlled by the middle three quintiles promotes economic growth.

not be liquidity constrained in investment choices—yields lower growth at a given level of mean current wealth. This is demonstrated in the following section.

- 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.
- This is also suggested by models of poverty traps based on impatience for consumption—high time preference rates associated with low life expectancy—leading to low savings and investment rates by the poor.⁷ Here too, while the theoretical literature has focused on initial inequality, it can also be argued that a higher initial incidence of poverty implies a higher proportion of impatient consumers and hence lower growth.
- Yet another example is found in how work productivity is affected by past nutritional and health status. Only when past nutritional intakes have been high enough (above basal metabolic rate) will it be possible to do any work, but diminishing returns will set in later; see the model in Dasgupta and Ray (1986). Following Cunha and Heckman (2007), this type of argument can be broadened to include other aspects of child development that have lasting impacts on learning ability and earnings as an adult. By implication, having a larger share of the population who grew up in poverty will have a lasting negative impact on an economy's aggregate output.

These arguments point to the importance of poverty as a constraint on growth, which is our first clue as to why we do not find poverty convergence. However, while all these arguments suggest that the growth rate may depend on parameters of the initial distribution, it is unclear whether inequality, poverty or the size of the middle class is the most relevant parameter. The fact that very few encompassing tests are found in the literature,⁸ and that these different measures of distribution are clearly not independent, leaves one in doubt about what aspect of distribution really matters. For example, when the initial value of mean income is included in a growth regression alongside initial inequality, but initial poverty is an excluded but relevant variable, the inequality measure may pick up the effect of poverty rather than inequality *per se*.

⁷ See, for example, Azariadis (2006), though Kraay and Raddatz (2007) argue that poverty traps arising from low savings (high time preference rates) in poor countries are hard to reconcile with the data.

⁸ By an encompassing test I mean that a nested test of the competing hypotheses is employed. In this instance, the encompassing test entails putting all the parameters of the initial distribution in the growth regression.

A second clue to the puzzle of why we do not see poverty convergence can be found in the literature on the effects of growth on poverty in developing countries. The consensus in that literature is that higher growth rates tend to yield more rapid rates of absolute poverty reduction.⁹ There is also evidence that inequality matters to how much impact a given growth rate in the mean has on poverty.¹⁰ Intuitively, in high inequality countries the poor will tend to have a lower share of the gains from growth in the mean. Ravallion (1997, 2007) examines this issue empirically using household surveys for multiple countries over time and finds evidence of a strong interaction effect between initial inequality and the growth rate in the mean when explaining the proportionate rate of poverty reduction. In the most parsimonious specification, which also fits the data for developing countries well, the expected value of the log difference in the poverty rate over time is directly proportional to the “distribution-corrected” growth rate, given by the ordinary growth rate in the mean times one minus an index of inequality.

However, here too, one can question whether inequality is the only relevant parameter of the initial levels distribution. If consumption is log-normally distributed and inequality (the variance of consumption) is unchanged with growth in the mean then the (absolute) elasticity of the poverty rate to mean consumption will be strictly decreasing in the poverty rate as well as inequality.¹¹ This is of course as special case, but it at least suggests that poverty may be a relevant predictor of the elasticity, though this has never been tested to my knowledge.¹²

3. A model of aggregate growth with micro borrowing constraints

Banerjee and Duflo (2003) provide a simple but insightful growth model with borrowing constraints. Someone who starts her productive life with sufficient wealth is able to invest her unconstrained optimal amount, equating the (declining) marginal product of her capital with the interest rate. But the “wealth poor,” for whom the borrowing constraint is binding, are unable 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

⁹ See World Bank (1990, 2000), Ravallion (1995, 2001, 2007), Fields (2001) and Kraay (2006). Also see the review of the arguments and evidence on this point in Ferreira and Ravallion (2009). (Relative poverty measures are less responsive to growth since the poverty lines rises with the mean.)

¹⁰ See Ravallion (1997, 2007); World Bank (2000, 2006); Bourguignon (2003) and Lopez and Servén (2006).

¹¹ This claim can be readily verified from the results for the log-normal distribution in Bourguignon (2003).

¹² This point is anticipated by Easterly (2009) who argues that countries with high poverty rates (with specific reference to Sub-Saharan Africa) will require higher growth rates to attain the same proportionate rate of poverty reduction (as required by the Millennium Development Goal of halving the 1990 poverty rate by 2015). Easterly does not present any evidence to support this view.

a given mean wealth also implies lower growth. The following discussion uses the Banerjee-Duflo model to illustrate this hypothesis, which will be tested later in the paper.

The basic set up of the Banerjee-Duflo model is as follows. Current wealth, w_t , is distributed across individuals according to the cumulative distribution function, $p = F_t(w)$, giving the population proportion p with wealth lower than w at date t . It will be analytically easier to work with the quantile function, $w_t(p)$ (the inverse of $F_t(w)$). The credit market is imperfect, such that individuals can only borrow up to λ times their wealth. Each person has a strictly concave production function yielding output $h(k)$ from a capital stock k . Given the rate of interest r (taken to be fixed) the desired capital stock is k^* , such that $h'(k^*) = r$. Those with initial wealth less than $k^*/(\lambda + 1)$ are credit constrained in that, after investing all they can, they still find that $h'(k_t) > r$, while the rest are free to implement k^* . A share $1 - \beta \in (0,1)$ of current wealth is consumed, leaving β for the next period.

Under these assumptions, the recursion diagram takes the form:

$$w_{t+1} = \phi(w_t) = \beta[h((\lambda + 1)w_t) - \lambda r w_t] \text{ for } w_t \leq k^*/(\lambda + 1) \quad (1.1)$$

$$= \beta[h(k^*) + (w_t - k^*)r] \text{ for } w_t > k^*/(\lambda + 1) \quad (1.2)$$

Plainly, $\phi(w_t)$ is strictly concave up to $k^*/(\lambda + 1)$ and linear above that. Mean future wealth is:

$$\mu_{t+1} = \int_0^{\infty} \phi[w_t(p)] dp \quad (2)$$

By standard properties of concave functions, we have:

Proposition 1: (Banerjee and Duflo, 2003, p.277): “An exogenous mean-preserving spread in the wealth distribution in this economy will reduce future wealth and by implication the growth rate.”

However, the Banerjee-Duflo model has a further implication concerning poverty, as another aspect of the initial distribution. Let $H_t = F_t(z)$ denote the headcount index of poverty (“poverty rate”) in this economy when the poverty line is z . I assume that $z \leq k^*/(\lambda + 1)$ and let $H_t^* \equiv F_t[k^*/(\lambda + 1)]$. Using (1.1) and (1.2) we can re-write (2) as:

$$\mu_{t+1} = \beta \int_0^{H_t^*} [h((\lambda + 1)w_t(p)) - \lambda r w_t(p)] dp + \beta \int_{H_t^*}^1 [h(k^*) + (w_t(p) - k^*)r] dp \quad (3)$$

This holds for any given economy. Now consider the implications of a difference in the poverty rate between two economies. H_t^* is higher in one economy and I assume that no individual with wealth less than $k^*/(\lambda + 1)$ is any better off in that economy. Thus $\partial w_t(p)/\partial H_t^* \leq 0$ for all $p \leq H_t^*$. If this holds then we can say that poverty is unambiguously higher. It is readily verified that:¹³

$$\frac{\partial \mu_{t+1}}{\partial H_t^*} = \beta \int_0^{H_t^*} [h'((\lambda + 1)w_t(p))(\lambda + 1) - \lambda r] \frac{\partial w_t(p)}{\partial H_t^*} dp + \beta r \int_{H_t^*}^1 \frac{\partial w_t(p)}{\partial H_t^*} dp \quad (4)$$

The sign of (4) cannot be determined under the assumptions so far.¹⁴ However, suppose further that the two economies have the same initial wealth, $\mu_t = \bar{\mu}$, implying that:

$$\int_0^{H_t^*} \frac{\partial w_t(p)}{\partial H_t^*} dp + \int_{H_t^*}^1 \frac{\partial w_t(p)}{\partial H_t^*} dp = 0$$

(We can think of this as a difference in relative poverty.) Then equation (4) simplifies to:

$$\left[\frac{\partial \mu_{t+1}}{\partial H_t^*} \right]_{\mu_t = \bar{\mu}} = \beta \int_0^{H_t^*} [h'((\lambda + 1)w_t(p)) - r](\lambda + 1) \frac{\partial w_t(p)}{\partial H_t^*} dp < 0 \quad (5)$$

Thus we also have:

Proposition 2: *In the Banerjee-Duflo model an unambiguously higher initial headcount index of poverty holding the initial mean constant implies a lower growth rate.*

This model implies an aggregate efficiency cost of a high incidence of poverty. But a number of points should be noted. An inequality effect is still present—separately to the poverty effect. And the less poverty there is, the less important overall inequality is to subsequent growth prospects. Also note that the theoretical prediction concerns the level of poverty at a given initial value of mean wealth. Without controlling for the initial mean, the sign of the effect of higher poverty on growth is ambiguous. Two opposing effects can be identified. The first is the usual conditional convergence property, whereby countries with a lower initial mean (and hence higher initial poverty) tend to have higher subsequent growth. Against this, there is an adverse

¹³ Note that the function ϕ defined by equations (1.1) and (1.2) is continuous at $k^*/(\lambda + 1)$.

¹⁴ If there is (unrestricted) first-order dominance, whereby $\partial w_t(p)/\partial H_t^* \leq 0$ for all $p \in [0, 1]$, then $\partial \mu_{t+1}/\partial H_t^* \leq 0$. However, first-order dominance is ruled out by the fact that the mean is held constant in this “though experiment;” there is a redistribution from the “wealth poor” to the “wealth nonpoor”.

distributional effect of higher poverty (Proposition 2). Which effect dominates is an empirical question.

4. Data and descriptive statistics

In keeping with the bulk of the literature, the country is the unit of observation.¹⁵ However, unlike past data sets in the literature on growth empirics, this one is firmly anchored to the household surveys, in keeping with the focus on the role played by poverty and inequality, which is measured from surveys. By calculating the poverty and inequality statistics directly from the primary data, at least some of the comparability problems found in existing data compilations from secondary sources can be eliminated. However, there is no choice but to use household consumption or income, rather than the theoretically preferable concept of wealth.

I found almost 100 developing and transition countries with at least two suitable household surveys since about 1980. Virtually all of the surveys are nationally representative.¹⁶ For the bulk of the analysis I restrict the sample to the 90 countries in which the earliest available survey finds that at least some households lived below the average poverty line for developing countries (described below).¹⁷ This happens mechanically given that log transformations are used. However, it also has the defensible effect of dropping a number of the countries of Eastern Europe and Central Asia (EECA) (including the former Soviet Union); indeed, all of the countries with an initial poverty rate (by developing country standards) of zero are in EECA. As is well known, these countries started their transitions from socialist command economies to market economies with very low poverty rates, but poverty measures then rose sharply in the transition.¹⁸ The earliest available surveys pick up these low poverty rates, with a number of countries having no sampled household living below the poverty lines typical of developing countries. With the subsequent rise in poverty incidence, this looks like “convergence,” but it has little or nothing to do with neoclassical growth processes—rather it is a “policy convergence”

¹⁵ It is known that aggregation can hide the true relationships between the initial distribution and growth, given the nonlinearities involved at the micro level (Ravallion, 1998); identifying the deeper structural relationships would require micro data, and even then the identification problems can be formidable.

¹⁶ The only exception was that urban surveys were also used (for both the first and last survey) for Uruguay where over 90% of the population lives in urban areas. Results were robust to dropping these urban surveys.

¹⁷ The data set was constructed from [PovcalNet](#) in December 2008.

¹⁸ Prior to the global financial crisis there were signs that poverty measures were finally falling in the region, since the later 1990s; see Chen and Ravallion (2010).

effect associated with the transition. The experience of these countries is clearly not typical of the developing world.¹⁹

The longest available spell between two surveys is used for each country. Both surveys use the same welfare indicator, either consumption or income per person, following standard measurement practices. When both are available, consumption is preferred, in the expectation that it is both a better measure of current economic welfare and that it is likely to be measured with less error than incomes.²⁰ Three-quarters of the spells use consumption.

For about two-thirds of the countries there are three or more surveys. This sub-sample will be used in testing robustness to relaxing various specification assumptions.

Naturally the time periods between surveys are not uniform. The median year of the first survey is 1991 while the median for the second is 2004. The median interval between surveys is 13 years and it varies from three to 27 years. All changes between the surveys are annualized. Given the most recent survey for date t_i in country i and the earliest available survey for date $t_i - \tau_i$, the growth rate for the variable x is $g_i(x_{it}) \equiv \ln(x_{it} / x_{it-\tau}) / \tau$ (dropping the i subscript on t and τ for brevity). National accounts and social indicators are also used, matched as closely as possible to survey dates. All monetary measures are in constant 2005 prices (using country-specific Consumer Price Indices) and are at Purchasing Power Parity (PPP) using the individual consumption PPPs from the 2005 International Comparison Program (World Bank, 2008).

Poverty is mainly measured by the headcount index (H_{it}), given by the proportion of the population living in households with consumption per capita (or income when consumption is not available) below the poverty line. For the bulk of the analysis the poverty line is set at \$2.00 per person per day at 2005 PPP, which is the median poverty line amongst developing countries based on the compilation of national poverty lines in Ravallion et al. (2009). \$2 a day is also very close to the median consumption per person in the developing world for 2005; see Chen and Ravallion (2010) which also describes the methods used here in measuring poverty and inequality. The (unweighted) mean poverty rate for the \$2 line fell from 43.6% in the earliest round of surveys to 38.3% in the latest rounds. This line is clearly somewhat arbitrary; for example, there is no good reason to suppose that \$2 a day corresponds to the point where credit

¹⁹ Nor are they typical of developed countries; for this reason, Voitchovsky (2005) chose to drop these countries from her sample for industrialized countries.

²⁰ The only exception was Peru, for which incomes allowed a much longer time period.

constraints cease to bite, but nor is there any obviously better basis for setting a threshold. I will also consider a lower line of \$1.25 a day and a much higher line of \$13 a day in 2005, corresponding to the US poverty line. The \$1.25 line is the expected value of the poverty line in the poorest countries in terms of consumption per person;²¹ the \$13 per person per day is the official poverty line in the US for a family of four in 2005.²²

Inequality is measured by the usual Gini index (G_{it}). The initial index ranged from 19.4% (Czech Republic) to 62.9% (Sierra Leone), both around 1990, and about one quarter of the sample had a Gini index below 30% while one quarter had an index above 50%. Between the earliest and latest surveys, the mean Gini index stayed roughly unchanged at about 42%.

Four measures of the middle class are used. The first is the population share living between \$2 and \$13 a day, denoted $MC_{it} \equiv F_{it}(13) - F_{it}(2)$ where $F_{it}(z)$ is the distribution function for country i at date t (so $H_{it} = F_{it}(2)$). This is interpreted as the middle class by developing-country standards; while the bounds are somewhat arbitrary, this definition appears to accord roughly with the idea of what it means to be “middle class” in China and India (Ravallion, 2010). By contrast, those living above \$13 a day can be thought of as the “middle class and rich” by Western standards. These are absolute measures. The third measure uses a relative definition of the middle class, namely the consumption or income share controlled by the middle three quintiles, denoted MQ_{it} , as used by Easterly (2001). Finally, I will also consider the “miser index” proposed by Lind and Moene (2010), given by $H_{it}(\mu_{it} - \mu_{it}^H)$ where μ_{it} is the overall mean and μ_{it}^H is the mean below the poverty line. Lind and Moene propose this as a measure of polarization between the rich and poor, but it can equally well be thought of as an inverse measure of the middle class.

As one would expect, there are some strong correlations amongst these parameters of the initial distribution.²³ The Gini index is highly correlated with MQ ($r=-0.969$ for the earliest surveys). The poverty measures are also strongly correlated with the survey means; $\ln H_{it-\tau}$ and $\ln \mu_{it-\tau}$ have a correlation of -0.743. The least-squares elasticity of $H_{it-\tau}$ with respect to the

²¹ This is based on a threshold model calibrated to national poverty lines, as documented in Ravallion et al., (2009). \$1.25 is also the mean poverty line for the poorest 15 countries.

²² See [Department of Health and Human Services](#).

²³ The following descriptive statistics apply to the sample of 90 countries used in the main econometric analysis, for which a non-zero proportion of the population live below \$2 a day.

initial survey mean (i.e., the regression coefficient of $\ln H_{it-\tau}$ on $\ln \mu_{it-\tau}$) is -1.062 (t=4.836). (All t-ratios in this paper are based on White standard errors.) There is a positive correlation between the poverty measure and the initial Gini index, though not strong ($r=0.248$ between $\ln H_{it-\tau}$ and $\ln G_{it-\tau}$). However, there is a strong multiple correlation between the poverty measures (on the one hand) and the log mean and log inequality (on the other); for example, regressing $\ln H_{it-\tau}$ on $\ln \mu_{it-\tau}$ and $\ln G_{it-\tau}$ one obtains $R^2=0.607$. The log Gini index also has a strong partial correlation with the log of the poverty rates holding the log mean constant (t=2.764).

The size of the middle class is also highly correlated with the poverty rate; the correlation coefficient between $MC_{it-\tau}$ and $H_{it-\tau}$ is -0.979 (-0.717 in the logs); 96% of the variance in the initial size of the middle class is accountable to differences in the initial poverty rate. Across countries, 82% of the variance in the changes over time in MC_{it} can be attributed to the changes in H_{it} .²⁴ The absolute and relative measures of the size of the middle class are positively correlated but not strongly so.

Since the time period between surveys (τ) figures in the calculation of the growth rates it might be conjectured that poorer countries have longer periods between surveys, biasing this paper's results. However, the correlation coefficients between τ and the various measures of initial distribution are all small (most well under 0.2 in absolute value).

5. Testing for convergence in mean consumption and poverty

Intuitively, the twin stylized facts that there is convergence in mean consumption and that growth in the mean consumption reduces the incidence of absolute poverty imply that we should see poverty convergence, as discussed in the introduction. Indeed, an even stronger result is implied by the standard log-linear models for growth and poverty reduction found in the literature, with parameters independent of the initial level of poverty. Then the speed of convergence will be the same for the mean as the poverty measure. To see this, consider the most common empirical specification for the growth process in the mean:

$$\Delta \ln \mu_{it} = \alpha_i + \beta_i \ln \mu_{it-1} + \varepsilon_{it} \quad (6)$$

²⁴ $R^2=0.818$ for the regression of $MC_{it} - MC_{it-\tau}$ on $F_{it}(2) - F_{it-\tau}(2)$; the regression coefficient is -0.888 (t=-24.412; n=90), which is significantly different from -1 (t=3.077).

where α_i is a country-specific effect, β_i is a country-specific convergence parameter and ε_{it} is a zero-mean error term. (To simplify notation I assume evenly spaced data for now.) Next let the headcount index of poverty be a log-linear function of the mean:

$$\ln H_{it} = \delta_i + \eta_i \ln \mu_{it} + \nu_{it} \quad (7)$$

Where δ_i is a country-specific effect, η_i is interpretable as the (country-specific) elasticity of poverty to the mean (with the expectation that $\eta_i < 0$) and ν_{it} is a zero-mean error term. The implied model of the growth rate in poverty is then:

$$\Delta \ln H_{it} = \alpha_i^* + \beta_i^* \ln H_{it-1} + \varepsilon_{it}^* \quad (8)$$

for which it is readily verified that $\alpha_i^* = \alpha_i \eta_i - \beta_i \delta_i$, $\beta_i^* = \beta_i$ and $\varepsilon_{it}^* = \varepsilon_{it} \eta_i + \nu_{it} - (1 + \beta_i) \nu_{it-1}$.

The parameters of (6) and (7) ($\alpha_i, \beta_i, \delta_i, \eta_i$) can vary across counties but (for the sake of this argument) suppose they do so independently of $H_{it-\tau}$. Comparing (6) and (8) it can be seen that the “speed of convergence” for the poverty measures, $\partial \Delta \ln H_{it} / \partial \ln H_{it-1} = \beta_i$, is the same as that for the mean, $\partial \Delta \ln \mu_{it} / \partial \ln \mu_{it-1} = \beta_i$. Thus we have:

Proposition 3: *In standard log-linear models for growth and poverty reduction, with parameters independent of the initial level of poverty, the speed of convergence will be the same for the mean as the poverty measure.*

Turning to the data, we do see signs of mean convergence. Table 1 gives standard convergence tests for mean consumption based on the regression coefficient of $g_i(\mu_{it})$ on $\ln \mu_{it-\tau}$, with and without controls.²⁵ The controls included initial consumption per capita from the national accounts, 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. It can be seen from Table 1 that the survey means exhibit convergence; the β coefficient is -0.013 (t=-3.412) without the controls and -0.042 (t=-7.435) with them. Unconditional

²⁵ Alternatively one can estimate the convergence parameter using a nonlinear regression $g(\mu) = \alpha - [(1 - e^{\beta\tau}) / \tau] \ln \mu_{-\tau} + \varepsilon$ (as in Barro and Sala-i-Martin, 1992). This gave a very similar result to (1) in Table 1, namely $\hat{\beta} = -0.012$. (t=-2.865). Clearly, the approximation that $e^{\beta\tau} = 1 + \beta\tau$ (linearizing the nonlinear regression specification) works well.

convergence is weaker using means from consumption surveys only (Column 2) or national accounts (Column 3), though conditional convergence is still evident.

The data also suggest that higher growth rates are associated with higher (proportionate) rates of poverty reduction. The regression coefficient of $g_i(H_{it})$ on $g_i(\mu_{it})$ is -1.085 (t=-4.128) with $R^2=0.377$. The (absolute) elasticity is lower using the growth rate of consumption per capita from national accounts (an elasticity of -0.797, t=-3.089) while it is higher using a lower poverty line; for the \$1.25 line the elasticity with respect to growth in the survey mean is -1.385 (t=-3.735).

However, as we saw in Figure 1, there is no sign of convergence for the \$2 a day poverty rate, with proportionate rates of poverty reduction roughly orthogonal to initial levels.²⁶ This appears to be robust to the choice of poverty line.²⁷ It is also robust to including the same set of controls, although then there are slightly stronger signs of conditional convergence at lower poverty lines; for the \$1.25 a day line the conditional convergence parameter for the poverty rate is -0.025, but this is still not significantly different from zero at the 10% level (t=-1.575).

In summary, while there is mean convergence and growth tends to reduce poverty, there are no signs of poverty convergence. The rest of this paper will try to explain why. It will be argued that the initial poverty rate matters to the subsequent rate of poverty reduction through two distinct channels, namely the growth rate in mean consumption and the elasticity of poverty to the mean. First, it will be shown in the following section that the parameter α_i (in equation 6) is a decreasing function of the initial poverty rate. Second, section 7 will be shown that the elasticity of poverty to the mean, $-\eta_i$ (equation 7), is a decreasing function of the initial level of poverty. Section 8 will bring these two elements together to answer the question posed in the title to this paper.

6. The relevance of initial poverty to the growth rate in the mean

The section begins with “benchmark regressions” for growth, and then tests its robustness to various changes in specification.

²⁶ Recall that poverty convergence is defined in proportionate rather than absolute terms. The absence of poverty convergence by this definition implies that poorer countries tend to see larger absolute reductions in their poverty rate.

²⁷ This was also true for the \$13 line, for which the convergence parameter was -0.009 (t=-0.480). Similarly, for a lower line of \$1.25 a day the parameter was -0.005 (t=-0.393).

6.1 Growth regressions with poverty as an initial condition

Table 2 gives estimates of the following regression:

$$g_i(\mu_{it}) = \alpha + \beta \ln \mu_{it-\tau} + \gamma \ln H_{it-\tau} + \varepsilon_{it} \quad (9)$$

(The regressors are assumed to be exogenous, though this will be relaxed later.) The result using the full sample is given in column (1) of Table 2.²⁸ This suggests that differences in the initial poverty rate have sizeable negative impacts on the growth rate at any given initial mean. A one standard deviation increase in $\ln H_{it-\tau}$ comes with a decline of 0.025 (2.5% points) in the growth rate for the survey mean.

As can be seen in column (2) of Table 2, this poverty effect on growth is even stronger if one confines attention to consumption surveys (dropping the 21 surveys for which only income was available). The result is also robust to using consumption from the national accounts (column 3), although the headcount index based on the \$1.25 line is a slightly stronger predictor of the national accounts consumption growth rate,²⁹ and the coefficient on the poverty rate is lower, as is the convergence parameter. (I will return to discuss the estimates in columns (4)-(6).)

It might be conjectured that the poverty measure (at given initial mean) is picking up some other aspect of the initial distribution, such as inequality (the variable identified in much of the empirical literature referred to in section 2). Indeed, if we imagine $\ln H_{it-\tau}$ to be a linear function of $\ln \mu_{it-\tau}$ and $\ln G_{it-\tau}$ (which fits the data quite well as noted in section 3) then one can re-write (9) in a reduced form similar to the past papers in the literature which found that inequality impedes growth at a given mean (section 2). However, an encompassing test—adding the log of the initial Gini index to equation (9)—does not change the result; then the coefficient on the Gini index is not significantly different from zero for growth rates in either the survey means or national accounts consumption and the coefficient on $\ln H_{it-\tau}$ remains (highly) significant in the augmented version of (9). It is poverty not inequality that is doing the work.

To investigate this further, I added inequality ($\ln G_{it-\tau}$), the income share of the middle three quintiles ($\ln MQ_{it-\tau}$), the share of the Western middle class and rich ($1 - F_{it-\tau}(13)$), the

²⁸ The regressions are consistent with a derivative of $\ln \mu_{it}$ with respect to $\ln \mu_{it-\tau}$ that is less than unity, but fades toward zero at sufficiently long gaps between survey rounds; for example, column (1) in Table 2 implies a derivative that is positive but less than unity for $\tau < 26$ years; the largest value of τ in the data is 27 years.

²⁹ Using the \$2 line, the coefficient on the poverty measure in column (3) becomes -0.012 (t=-2.245).

“miser index”, primary school enrollment rate, life expectancy at birth, and the relative price index of investment goods.³⁰ Table 3 gives the augmented models using both survey means and consumption from national accounts. The table also gives restricted forms that pass comfortably.

We see that the initial poverty rate remains a strong and significant predictor of growth. The Gini index also has a negative effect on the growth rate, but it is only significant (and only at the 10% level) when using consumption from the national accounts. Significant predictors of the growth rate at a given initial mean and poverty rate are the size of the Western middle class (negatively), life expectancy (positively) and the price of investment (negatively). The relative share of the middle quintiles is significant for the growth rates in national accounts consumption (but not the survey means), though with a negative sign; it appears that Easterly’s (2001) findings on the positive effect on growth of a larger income share held by the middle quintiles vanishes once one introduces the other distributional parameters considered here. The miser index has no significant effect on growth (as also found by Lind and Moene, 2010).

The negative coefficients on both the poverty rate and the share of the Western middle class imply that a higher population share in the developing-world middle class—those living between \$2 and \$13 per day—is growth enhancing. Thus the data can also be well described by a model relating growth to the population share of the developing world’s middle class. As one would expect, replacing $\ln H_{it-\tau}$ and $1 - F_{it-\tau}(13)$ by $\ln[F_{it-\tau}(13)/H_{it-\tau}]$ gave a very similar overall fit, though not quite as good as Table 3. The negative (conditional) effect of the poverty rate may well be transmitted through differences in the size of the middle class.

While the above results appear to be convincing in suggesting that it is high poverty not inequality that retards growth, it is important to recall that the poverty effect only emerges when one controls for the initial mean. The between-country differences in the incidence of poverty at a given mean reflect differences in relative distribution. To the extent that higher overall inequality comes with higher poverty at a given mean it yields lower growth rates.

6.2 *Further tests of robustness*

There are a number of alternative specifications. One might allow for nonlinearities and interaction effects. The log-linear form in equation (9) appears to give the best fit amongst

³⁰ This is a common measure of policy distortions, derived from Penn World Tables (following Lopez and Servén, 2009). As noted in section 2, schooling and health attainments can also be interpreted as channels linking initial distribution to growth rather than as independent effects, so the interpretation of the poverty coefficient in these augmented regressions is not strictly the same as for the benchmark regression.

obvious (parametric) options. The negative effect is still evident, but not as strong statistically, if one uses the linear poverty measure rather than its log; for example, replacing $\ln H_{it-\tau}$ by $H_{it-\tau}$ in (9) the regression coefficient on the latter is -0.071 with a t-ratio of -4.770.³¹ It might be conjectured that the effect of $\ln H_{it-\tau}$ in (9) reflects a misspecification of the functional form for the convergence effect, noting that the poverty measure is a nonlinear function of mean income. To test for this, I re-estimated (9) using cubic functions of $\ln \mu_{it-\tau}$ to control for the initial mean. However, this adds little to the explanatory power (the adjusted R^2 falls) and makes little difference to the main results on the significant negative effect of initial poverty (with or without controls). Nor did I find evidence of any significant interaction effects between initial inequality and the initial mean or between initial inequality and the initial poverty measure.

There is also the choice of poverty line and poverty measure. The \$2 line appears to fit best. On replacing $\ln H_{it-\tau}$ by $\ln F_{it-\tau}(1.25)$ in (9), the poverty rate still has a negative coefficient (-0.010) but is only significant at the 2% level ($t=-2.413$). I also estimated an encompassing specification, including both poverty measures; clearly these are highly correlated, with $r=0.974$, but it is still possible to disentangle their effects. The coefficient for the \$1.25 line was insignificant ($t=0.918$) while that the \$2 line remained negative and significant though only at the 7% level. The results are also robust to using the poverty gap index instead of the headcount index; the corresponding version of (9) is similar, with a coefficient on the log of the poverty gap index of -0.014, with t-ratio of -3.821. However, the fit is better using the headcount index.

The subsample of 70 countries with at least three surveys can also be used to test for robustness. One can use this sample to form inter-temporal averages, to help reduce the effects of measurement error. Equation (9) can be re-estimated in the form:

$$g_i(\mu_{it}) = \alpha + \beta \ln M_i(\mu_{it-\tau_2}) + \gamma \ln M_i(H_{it-\tau_2}) + \varepsilon_{it} \quad (10)$$

where $M_i(x_{it-\tau_2}) \equiv (x_{it-\tau_2} + x_{it-\tau_1-\tau_2})/2$. Column (4) of Table 3 gives the results, which are robust to this change in specification.

Another way of using the extra survey rounds is as a source of instrumental variables (IVs). Growth rates between the middle and last rounds are regressed on the mean and

³¹ This is for the same sample used in Table 2. If one also includes the extra seven countries for which $H_{it-\tau} = 0$ then the coefficient is -0.063 ($t=-3.314$).

distributional variables for the middle round but treating them as endogenous and using earliest round as a source of IVs. Letting τ_i now denote the length of spell i ($=1,2$), the model becomes:

$$g_i(\mu_{it}) = \alpha + \beta \ln \mu_{it-\tau_2} + \gamma \ln H_{it-\tau_2} + \varepsilon_{it} \quad (11)$$

Column (5) of Table 2 reports the Generalized Methods of Moments (GMM) estimates of this specification, using $\ln \mu_{it-\tau_1-\tau_2}$, $\ln C_{it-\tau_1-\tau_2}$, $\ln C_{it-\tau_2}$, $\ln G_{it-\tau_1-\tau_2}$, $\ln F_{it-\tau_1-\tau_2}(z)$ ($z=1.25, 2.00$) and τ_1 as IVs (where C denotes consumption per capita from national accounts).³² Column (6) gives the corresponding result using growth rates of national accounts consumption (for which I dropped $\ln C_{it-\tau_2}$ from the set of IVs). It can be seen that the GMM results for the sub-sample with at least three surveys are more similar between the survey means and national accounts consumption than is the case using the full sample.

Overall, the finding of a direct effect of initial poverty on growth rates is robust to allowing for the endogeneity of the initial mean and initial poverty rate, subject to the usual assumption that the above instrumental variables are excludable from the main regression.

The above regressions do not include geographic effects. The two regional effects that have been identified in the literature on growth empirics are for Sub-Saharan Africa (negatively) and East Asia (positively). In testing augmented versions of the regressions in Tables 2 and 3, with dummy variables for these two regions, I find no sign of significant SSA or East Asia effects. Of course (as noted), there are unconditional effects on growth in both regions. But these are largely captured within the model.

One can also use the subsample with three surveys to allow for country-fixed effects, which sweep up any confounding latent heterogeneity in growth rates at country level. The main results are not robust to this change. Regressing $g_i(\mu_{it}) - g_i(\mu_{it-\tau_2})$ on $\ln(\mu_{it-\tau_2} / \mu_{it-\tau_1-\tau_2})$ and $\ln(H_{it-\tau_2} / H_{it-\tau_1-\tau_2})$, the coefficient on the former remains significant but the poverty rate ceases to be so. However, it is hard to take fixed-effects growth regressions seriously with these data. While this specification addresses the problem of time-invariant latent heterogeneity it is unlikely to have much power for detecting the true relationships given that the changes over time in growth rates will almost certainly have a low signal-to-noise ratio. The simulations reported

³² The significant effect of initial poverty was still evident if one dropped $\ln C_{it-\tau_2}$ as an IV in the regression in column (5); the coefficient becomes -0.020, with $t=-3.315$.

by Hauk and Wacziarg (2009) indicate that the coefficients on growth determinants are heavily biased toward zero in fixed-effects growth regressions. I suspect that the problem of time-varying measurement errors in both growth rates and initial distribution is even greater in the present data set, possibly reflecting survey comparability problems over time.

The problem of noise in the changes in growth rates can be illustrated if we consider the relationship between the two measures of the mean used in this study, namely that from the surveys (μ_{it}) and that from the private consumption component of domestic absorption in the national accounts (C_{it}). Using a log-log regression in the levels gives an elasticity of μ_{it} to C_{it} of 0.75 ($R^2=0.82$) for the latest survey rounds. Using a country-fixed effects specification in the levels, the elasticity drops to 0.51 ($R^2=0.21$). However, when one also includes fixed-effects in the growth rates in the mean (using the subsample with at least three surveys) the elasticity drops to 0.09 ($R^2=0.07$), which must be considered an implausibly low figure, undoubtedly reflecting substantial attenuation bias due to measurement error in the changes in growth rates.

7. Initial poverty and the growth elasticity of poverty reduction

We have seen that countries starting with a higher poverty rate tend to see slower growth at a given initial mean. Now I turn to the second channel—how the growth elasticity of poverty reduction depends on initial distribution. This can be thought of as the direct effect of the initial distribution on the pace of poverty reduction, as distinct from the indirect effect via the rate of growth in the mean. Again I focus on the \$2 line, although the \$1.25 line gives similar results.

Table 4 gives regressions of the annualized change in the log of the \$2 a day poverty rate against both the annualized growth rate in the mean and its interaction with the initial poverty rate. Columns (1) and (2) give unrestricted estimates of an encompassing test:

$$g_i(H_{it}) = \delta_0 + \delta_1 \ln H_{it-\tau_2} + (\eta_0 + \eta_1 H_{it-\tau_2})g(\mu_{it}) + \nu_{it} \quad (12)$$

Results are given for both OLS and IVE; the IVE method uses the growth rate in private consumption per capita from the national accounts as the instrument for the growth rate in the survey mean; following Ravallion (2001), this allows for the possibility that a spurious negative correlation exists due to common measurement errors (given that the poverty measure and the mean are calculated from the same surveys).

The results in Table 4 indicate that the (absolute) growth elasticity of poverty reduction tends to be lower in countries with a higher initial poverty rate. There is no sign of conditional

convergence in poverty; the null that $\delta_1 = 0$ is easily accepted. Table 4 also gives homogeneity tests for the null $\eta_0 + \eta_1 = 0$; the tests pass comfortably, indicating that the relevant growth rate is the “poverty-adjusted rate,” as given by the growth rate in the mean *times* one minus the poverty rate.³³ Columns (5) and (6) give this more parsimonious specification.

There is a sizeable attenuating effect of a higher initial poverty rate on the elasticity of poverty reduction to growth. At an initial poverty rate of 10% (about one standard deviation below the mean) the elasticity is about -2.2 (almost 3 using the IVE estimate) while it falls to about -0.5 at a poverty rate of 80% (about one standard deviation above the mean).

There is also a strong interaction effect with the size of the middle-class:

$$g_i(H_{it}) = -0.008 + (0.177 - 0.029 MC_{it-\tau})g_i(\mu_{it}) + \hat{v}_{it} \quad R^2=0.674, n=89 \quad (13)$$

(-1.738) (0.951) (-5.243)

However, this interaction effect is largely attributable to $H_{it-\tau}$. Letting $H_{it-\tau}$ and $F_{it-\tau}$ (13) enter separately (recalling that $MC_{it} = F_{it}(13) - H_{it}$) only $H_{it-\tau}$ is significant:

$$g_i(H_{it}) = -0.009 + (-0.137 - 0.015 F_{it-\tau}(13) + 0.028 H_{it-\tau})g_i(\mu_{it}) + \hat{v}_{it} \quad R^2=0.675, n=89 \quad (14)$$

(-2.035) (0.295) (-0.352) (3.856)

One cannot reject the null hypotheses that the interaction effect with $F_{it-\tau}$ (13) has no impact, though nor can one reject the null that the coefficients on the two interaction effects add up to zero (their sum of 0.013 has a t-ratio of 0.347)—implying again that it is the population share of the middle class (by developing country standards) that matters.

So it is a dead heat statistically between a model in which it is a larger middle class that determines how much impact a given rate of growth has on poverty and a model in which it is the initial poverty rate that matters. However, given that the main way people in developing countries enter the middle class is by escaping poverty—recall that over 80% of the variance in changes in the size of the middle class is accountable to changes in the poverty rate—it seems more reasonable to think of poverty as the relevant primary factor.

I also tried adding extra interaction effects with the initial Gini index, the partial elasticity of poverty reduction holding the Lorenz curve constant, the primary school enrollment rate, life expectancy, the price of investment goods and regional dummy variables for SSA and East Asia. (Growth elasticities of poverty reduction are significantly lower in SSA, but this is entirely due

³³ I also used the subsample with three survey rounds to implement an IVE using the same instruments as before. Again, the homogeneity restriction is easily accepted ($t=-0.447$). The IVE of the regression coefficient of $g_i(H_{it})$ on $(1 - H_{it-\tau_2})g_i(\mu_{it})$ is -2.929 ($t=-3.227$; $n=63$).

to the region's above-average poverty incidence.) These are individually and jointly insignificant.

Nor does the relationship differ according to whether growth is positive or negative. On stratifying the parameters according to whether the mean is increasing or not, and re-estimating the regressions in Table 4, I found a positive interaction effect during spells of contraction in the mean as well as expansions; the homogeneity restriction passes in both cases and one cannot reject the null that the coefficients are the same for expansions versus contractions ($F=2.978$, $\text{prob.}=0.062$).

So the key proximate determinant of the rate of poverty reduction is the “poverty adjusted growth rate” $((1 - H_{it-\tau})g_i(\mu_{it}))$ rather than the ordinary growth rate $(g_i(\mu_{it}))$. The regression coefficient of the rate of poverty reduction $(g_i(H_{it}))$ against the poverty-adjusted growth rate in the survey mean $((1 - H_{it-\tau})g_i(\mu_{it}))$ is -2.468 (Table 4, column 5), which more than twice as high as that for the ordinary growth rate, namely -1.085. And allowing for initial poverty rate adds almost 30 percentage points to the share of the variance in the rate of poverty reduction that can be explained by the rate of growth.

8. So why don't we see poverty convergence?

We can now combine the main results from the last two sections to explain why the speed of poverty convergence, $\partial g_i(H_{it})/\partial \ln H_{it-\tau}$, is close to zero, despite the fact that there is mean convergence and that growth tends to reduce poverty. Based on the various encompassing tests above my empirically-preferred model takes the form:

$$g_i(H_{it}) = \eta(1 - H_{it-\tau})g_i(\mu_{it}) + \nu_{it} \quad (15.1)$$

$$g_i(\mu_{it}) = \alpha + \beta \ln \mu_{it-\tau} + \gamma \ln H_{it-\tau} + \varepsilon_{it} \quad (15.2)$$

The regressors in (15.2) are not, of course, independent; as we also saw in Section 3, countries with a higher initial mean tend to have a lower poverty rate.³⁴ I shall allow for this by assuming that $\ln H_{it-\tau}$ varies linearly as a function of $\ln \mu_{it-\tau}$ consistently with the data. We can then derive the following decomposition of the poverty convergence elasticity:

³⁴ Nonetheless, as we have also seen, the differences across countries in initial distributions entail that $\ln \mu_{it-\tau}$ and $\ln H_{it-\tau}$ are not so highly correlated as to prevent disentangling their effects.

$$\frac{\partial g_i(H_{it})}{\partial \ln H_{it-\tau}} = \underbrace{\eta\beta(1-H_{it-\tau})\left(\frac{\partial \ln H_{it-\tau}}{\partial \ln \mu_{it-\tau}}\right)^{-1}}_{\text{(Mean convergence effect)}} + \underbrace{\eta\gamma(1-H_{it-\tau})}_{\text{(Direct effect of poverty)}} - \underbrace{\eta g_i(\mu_{it})H_{it-\tau}}_{\text{(Poverty elasticity effect)}} \quad (16)$$

On evaluating all variables at their sample means and using the estimates in column (1) of Table 3 and column (5) from Table 4, and using the OLS elasticity of elasticity of the initial headcount index with respect to the initial survey mean of -1.085, one finds that the mean convergence effect is -0.047, while the direct effect of poverty is 0.026 and the poverty elasticity effect is 0.016. The mean convergence effect is almost exactly cancelled by the combination of the two “poverty effects,” which are roughly equal in size.

Naturally, different data points and parameter estimates give different magnitudes for this decomposition, though all share the feature that the two poverty effects work in opposition to the (conditional) mean convergence effect. Evaluating the decomposition at a higher initial headcount index increases the poverty elasticity effect while reducing the other two components. The estimates using only the consumption surveys give a higher direct effect of poverty, as do the estimates from the subsample with three surveys; in the latter case the poverty convergence elasticity is larger due to both a lower mean convergence component and the higher direct effect.

9. Conclusions

The most interesting thing about the fact that we do not see poverty convergence in the developing world is what it tells us about the underlying process of economic growth and its impact on poverty. The lack of poverty convergence—despite mean convergence and that growth generally reduces poverty—suggests that something about the initial distribution is offsetting both the “advantage of backwardness” and the “advantage of growth.”

That something turns out to be poverty itself. The paper’s findings point to three distinct consequences of being a poor country for subsequent progress against poverty. The usual neoclassical convergence effect entails that countries starting with a lower mean, and so (typically) a higher poverty rate, grow faster and (hence) enjoy faster poverty reduction than otherwise similar countries. Against this, there is an adverse direct effect of poverty on growth, such that countries with a higher initial incidence of poverty tend to experience a lower rate of growth, controlling for the initial mean. Additionally, a high poverty rate makes it harder to

achieve any given proportionate impact on poverty through growth in the mean. (And, by the same token, the proportionate impact of economic contraction on poverty tends to be smaller in countries with a higher poverty rate.)

The two “poverty effects” work against the mean convergence effect, leaving little or no correlation between the initial incidence of poverty and the subsequent rate of progress against poverty. In terms of the pace of poverty reduction, the growth advantage for countries starting with a low capital endowment (given diminishing returns to aggregate capital) is largely wiped out by the high level of poverty that tends to accompany a low initial mean. This dynamic “disadvantage of poverty” appears to exist independently of other factors impeding growth and poverty reduction, including human underdevelopment and policy distortions.

The evidence is mixed on the role played by other aspects of distribution. Initial inequality has at most a weak effect once one controls for initial poverty. Of course, initial inequality can still matter via its bearing on the initial incidence of poverty. A larger middle class by developing country standards—but not Western standards—promotes economic growth in developing countries, but this is largely attributable to lower poverty, which is what drives the expansion in the middle class. A larger middle class by developing-country standards also makes growth more poverty-reducing, but this too can be interpreted as the effect of a lower poverty rate.

Knowing more about the “reduced form” empirical relationship between growth, poverty reduction and the parameters of the initial distribution will not, of course, resolve the policy issues at stake. The policy implications of distribution-dependent poverty reduction depend on why countries starting out with a higher incidence of poverty tend to face worse growth prospects and enjoy less poverty reduction from a given rate of growth. The initial level of poverty may well be picking up other factors, such as the distribution of human and physical capital; indeed, some of the underlying theories point more to “wealth poverty” than consumption or income poverty. The control variables used here for schooling, life-expectancy and the price of investment goods do not explain the direct adverse effects of high initial poverty on the scope for reducing poverty. However, the cross-country empirical relationships reported here do point to the importance in future work of a deeper understanding of the specific handicaps faced by poor countries in their efforts to become less poor.

References

- Acemoglu, Daron and Fabrizio Zilibotti, 1997, "Was Prometheus Unbound by Chance?" *Journal of Political Economy* 105(4): 709-751.
- Aghion, Philippe and Patrick Bolton, 1997, "A Theory of Trickle-Down Growth and Development," *Review of Economic Studies* 64: 151-172.
- Aghion, Philippe, Eve Caroli and Cecilia Garcia-Penalosa, 1999, "Inequality and Economic Growth: The Perspectives of the New Growth Theories", *Journal of Economic Literature* 37(4): 1615-1660.
- Aghion, Philippe, Abhijit Banerjee and Thomas Piketty, 1999, "Dualism and Macroeconomic Volatility," *Quarterly Journal of Economics* 114(4): 1359-1397.
- Alesina, Alberto and Dani Rodrik, 1994, "Distributive Politics and Economic Growth", *Quarterly Journal of Economics* 108: 465-90.
- Azariadis, Costas, 2006, "The Theory of Poverty Traps: What Have we Learned?," in Samuel Bowles, Steven Durlauf and Karla Hoff (eds) *Poverty Traps*, Princeton: Princeton University Press.
- Banerjee, Abhijit and Esther Duflo, 2003, "Inequality and Growth: What Can the Data Say?" *Journal of Economic Growth* 8(3): 267-99.
- Bardhan, Pranab, Samuel Bowles and Herbert Ginitis, 2000, "Wealth Inequality, Wealth Constraints and Economic Performance," in A.B. Atkinson and F. Bourguignon (eds) *Handbook of Income Distribution Volume 1*, Amsterdam: North-Holland.
- Barro, Robert, 2000, "Inequality and Growth in a Panel of Countries," *Journal of Economic Growth* 5(1): 5-32.
- Barro, Robert and Xavier Sala-i-Martin, 1992, "Convergence," *Journal of Political Economy* 100(2): 223-251.
- Beck, Thorsten, Asli Demirguc-Kunt and Ross Levine, 2007, "Finance, Inequality and the Poor," *Journal of Economic Growth* 12: 27-49.
- Bénabou, Roland, 1996, "Inequality and Growth", in Ben Bernanke and Julio Rotemberg (eds) *National Bureau of Economic Research Macroeconomics Annual*, Cambridge: MIT Press.
- Birdsall, Nancy, Carol Graham and Stefano Pettinato, 2000, "Stuck in the Tunnel: Is Globalization Muddling the Middle Class?" Center on Social and Economic Dynamics,

- Working Paper 14, Brookings Institution, Washington DC.
- Birdsall, Nancy, D. Ross and R. Sabot, 1995, "Inequality and Growth Reconsidered: Lessons from East Asia", *World Bank Economic Review* 9(3): 477-508.
- Bourguignon, François, 2003, "The Growth Elasticity of Poverty Reduction: Explaining Heterogeneity across Countries and Time Periods," in T. Eicher and S. Turnovsky (eds) *Inequality and Growth: Theory and Policy Implications*, Cambridge: MIT Press.
- Chen, Shaohua and Martin Ravallion, 2010, "The Developing World is Poorer than we Thought, but no less Successful in the Fight against Poverty," *Quarterly Journal of Economics* 125(4): 1577-1625.
- Clarke, George R.G., 1995, "More Evidence on Income Distribution and Growth", *Journal of Development Economics* 47: 403-428.
- Cunha, Flavio and James Heckman, 2007, "The Technology of Skill Formation," *American Economic Review, Papers and Proceedings*, 97(2): 31-47.
- Dasgupta, Partha and Debraj Ray, 1986, "Inequality as a Determinant of Malnutrition and Unemployment", *Economic Journal* 96: 1011-34.
- Deininger, Klaus and Lyn Squire, 1998, "New Ways of Looking at Old Issues: Inequality and Growth," *Journal of Development Economics* 57(2): 259-87.
- Doepke, Matthias and Fabrizio Zilibotti, 2005, "Social Class and the Spirit of Capitalism," *Journal of the European Economic Association* 3(2-3): 516-524.
- Easterly, William, 2001, "The Middle Class Consensus and Economic Development," *Journal of Economic Growth* 6(4): 317-335.
- _____, 2009, "How the Millennium Development Goals are Unfair to Africa," *World Development* 37(1): 26-35.
- Fields, Gary, 2001, *Distribution and Development*. New York: Russell Sage Foundation.
- Ferreira, Francisco and Martin Ravallion, 2009, "Poverty and Inequality: The Global Context", in Wiemer Salverda, Brian Nolan and Tim Smeeding (eds), *The Oxford Handbook of Economic Inequality*, Oxford: Oxford University Press.
- Forbes, Kristin J., 2000. "A Reassessment of the Relationship Between Inequality and Growth," *American Economic Review* 90(4): 869-87.
- Galor, Oded, and Joseph Zeira, 1993, "Income Distribution and Macroeconomics." *Review of Economic Studies* 60(1): 35-52.

- Hauk, William R. and Romain Wacziarg, 2009, "A Monte Carlo Study of Growth Regressions," *Journal of Economic Growth* 14(2): 103-147.
- Hoff, Karla, 1996, "Market Failures and the Distribution of Wealth: A Perspective from the Economics of Information," *Politics and Society* 24(4): 411-432.
- Kraay, Aart, 2006, "When is Growth Pro-Poor? Evidence from a Panel of Countries," *Journal of Development Economics* 80: 198-227.
- Kraay, Aart and Claudio Raddatz, 2007, "Poverty Traps, Aid and Growth," *Journal of Development Economics* 82(2): 315-347.
- Knowles, Stephen, 2005, "Inequality and Economic Growth: The Empirical Relationship Reconsidered in the Light of Comparable Data," *Journal of Development Studies* 41(1): 135-159.
- Levine, Ross and D. Renelt, 1992, "A Sensitivity Analysis of Cross-Country Growth Regressions," *American Economic Review* 82: 942-963.
- Li, Hongyi and Heng-fu Zou, 1998, "Income Inequality is not Harmful to Growth: Theory and Evidence," *Review of Development Economics* 2(3): 318-334.
- Lind, Jo Thori and Karl Moene, 2010, "Miserly Developments," mimeo, Department of Economics, University of Oslo.
- Lopez, Humberto and Luis Servén, 2006, "A Normal Relationship? Poverty, Growth and Inequality," Policy Research Working Paper 3814, World Bank.
- _____ and _____, 2009, "Too Poor to Grow," Policy Research Working Paper 5012, World Bank, Washington DC.
- Murphy, Kevin, Andrei Schleifer and Robert Vishny, 1989, "Industrialization and the Big Push," *Journal of Political Economy* 97(5): 1003-1026.
- Perotti, R., 1996, "Growth, Income Distribution and Democracy: What the Data Say," *Journal of Economic Growth*, 1(2): 149-187.
- Persson, Torsten and Guido Tabellini, 1994, "Is Inequality Harmful for Growth?" *American Economic Review* 84: 600-621.
- Ravallion, Martin, 1995, "Growth and Poverty: Evidence for Developing Countries in the 1980s," *Economics Letters*, 48, 411-417.
- _____, 1997, "Can High Inequality Developing Countries Escape Absolute Poverty?" *Economics Letters*, 56, 51-57.

- _____, 1998, “Does Aggregation Hide the Harmful Effects of Inequality on Growth?” *Economics Letters*, 61(1): 73-77.
- _____, 2001, “Growth, Inequality and Poverty: Looking Beyond Averages,” *World Development*, 29(11): 1803-1815.
- _____, 2007, “Inequality is Bad for the Poor,” in J. Micklewright and S. Jenkins (eds.), *Inequality and Poverty Re-Examined*. Oxford: Oxford University Press.
- _____, 2009, “Why Don’t we see Poverty Convergence?” Policy Research Working Paper 4974, World Bank, Washington DC (<http://econ.worldbank.org/docsearch>).
- _____, 2010, “The Developing World’s Bulging (but Vulnerable) ‘Middle Class’”, *World Development*, 38(4): 445-454.
- Ravallion, Martin, Shaohua Chen and Prem Sangraula, 2009, “Dollar a Day Revisited,” *World Bank Economic Review* 23(2): 163-184.
- Sridharan, E., 2004, “The Growth and Sectoral Composition of India’s Middle Class: Its Impact on the Politics of Economic Liberalization,” *India Review* 3(4): 405-428.
- Voitchovsky, Sarah, 2005, “Does the Profile of Income Inequality Matter for Economic Growth?” *Journal of Economic Growth* 10: 273-296.
- World Bank, 1990, *World Development Report: Poverty*, New York: Oxford University Press.
- _____, 2000, *World Development Report: Attacking Poverty*, New York: Oxford University Press.
- _____, 2006, *World Development Report: Equity and Development*, New York: Oxford University Press.
- _____, 2008, *Global Purchasing Power Parities and Real Expenditures. 2005 International Comparison Program*, Washington DC: World Bank.

Figure 1: The lack of poverty convergence amongst developing countries: Growth in the poverty rate plotted against its initial value

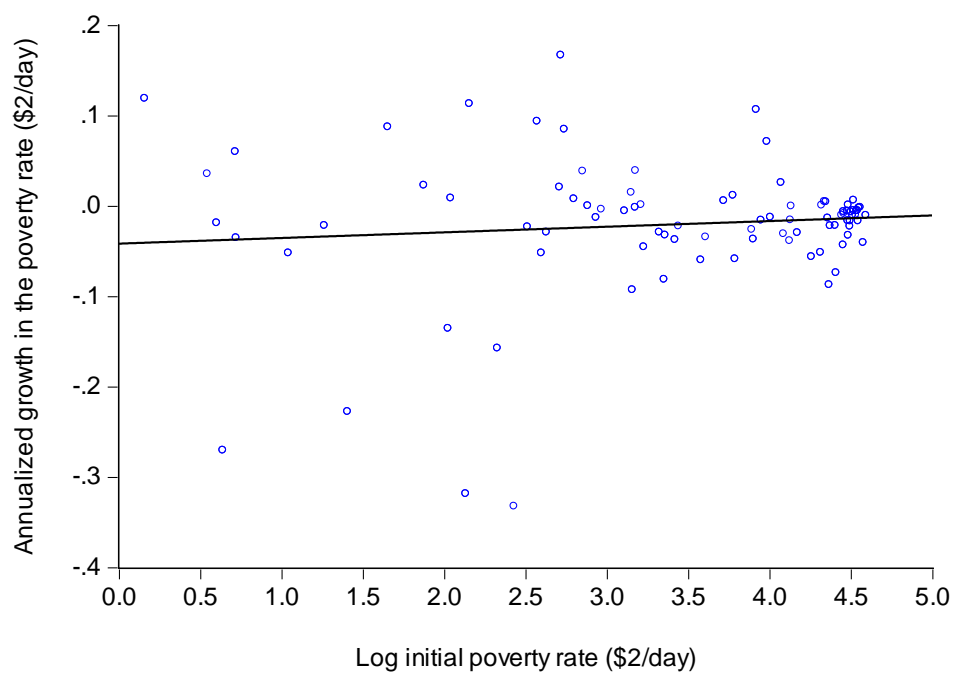


Table 1: Convergence tests for mean consumption

	(1) Surveys means (full sample)	(2) Surveys means (consumption surveys only)	(3) Consumption per capita from national accounts
Unconditional	-0.017*** (-3.202; n=97)	-0.016* (-2.167; n=73)	-0.007* (-1.782; n=90)
Conditional	-0.047*** (-10.602; n=90)	-0.046*** (-8.848; n=67)	-0.026*** (-4.437; n=88)

Notes: The table gives $\hat{\beta}$ in the regression $g_i(\mu_{it}) = \alpha + \beta \ln \mu_{it-\tau} + \gamma X_{it-\tau} + \varepsilon_{it}$ where μ_{it} denotes the surveys mean (Columns 1 and 2) or consumption from the national accounts (column 3). T-ratios based on White standard errors (corrected for heteroskedasticity). The conditional estimates include controls (all for earliest survey date) comprising log mean consumption per capita from national accounts (for the survey means), log primary school enrollment rate; log life expectancy; log relative price index of investment goods. * denotes significant at the 10% level; ** denotes significant at the 5% level; *** denotes significant at the 1% level.

Table 2: Regressions of growth rates on initial mean and initial headcount index of poverty

	(1) Full sample	(2) Full sample: Consumption surveys only	(3) Consumption per capita from national accounts	(4) Means from first two surveys used as initial conditions	(5) Sub-sample with three surveys: GMM estimator of equation (6) using IVs from earliest surveys	(6) As for col. (5) but using national accounts consumption
Intercept	0.262*** (8.470)	0.299*** (13.309)	0.154*** (3.771)	0.235*** (4.573)	0.199*** (3.028)	0.184*** (4.202)
Log initial mean	-0.039*** (-8.410)	-0.044*** (-12.892)	-0.020*** (-3.130)	-0.029*** (-3.271)	-0.022** (-2.341)	-0.015** (-2.234)
Log initial headcount index	-0.020*** (-5.513)	-0.025*** (-7.818)	-0.012*** (-2.761)	-0.022*** (-6.325)	-0.022*** (-3.991)	-0.026*** (-5.254)
R ²	0.275	0.384	0.136	0.134	n.a.	n.a.
N	90	69	80	76	63	58

Notes: These are estimates of equation (9). The dependent variable is the annualized difference in log of the survey mean ($g_i(\mu_{it})$) for columns (1), (2), (4) and (5); annualized difference in log private consumption per capita from national accounts ($g_i(C_{it})$) for (3) and (6). The initial mean corresponds to the same measure used for the growth rate in each regression. The poverty rate is \$2.00 for survey means and \$1.25 for national accounts consumption in (3). The t-ratios in parentheses are based on robust standard errors; ** denotes significant at the 5% level; *** denotes significant at the 1% level. GMM estimates used time series weighting matrix (to allow for serial correlation of the error term) and a quadratic kernel with fixed bandwidth in calculating the weighting matrix. The first-stage regressions for Column (5) had $R^2=0.884$ ($F=59.97$) and $R^2=0.797$ ($F=30.76$) for $\ln \mu_{it-\tau_2}$ and $\ln H_{it-\tau_2}$ respectively. The first-stage regressions for Column (6) had $R^2=0.852$ ($F=53.86$) and $R^2=0.753$ ($F=28.50$) for $\ln C_{it-\tau_2}$ and $\ln H_{it-\tau_2}$ respectively.

Table 3: Regressions for consumption growth rates on the initial poverty rate augmented with extra control variables

	(1)	(2)	(3)	(4)
	Survey Means	Growth rates based on: Consumption from national accounts	Survey Means	Consumption from national accounts
	Complete specification:		Dropping weak predictors:	
Intercept	0.446 (0.831)	0.623 (1.255)	0.259 (1.318)	0.635 (1.298)
Initial mean ($\ln \mu_{it-\tau}$ for (1) and (3) and $\ln C_{it-\tau}$ for (2) and (4))	-0.057*** (-6.139)	-0.034*** (-4.307)	-0.050** (-12.107)	-0.034*** (-4.374)
Poverty rate ($\ln H_{it-\tau}$)	-0.027*** (-5.536)	-0.018*** (-3.239)	-0.025** (-5.991)	-0.017*** (-3.310)
Gini index ($\ln G_{it-\tau}$)	-0.018 (-0.372)	-0.076* (-1.727)	-	-0.078* (-1.800)
Income share of middle three quintiles ($\ln MQ_{it-\tau}$)	-0.117 (-1.505)	-0.167** (-2.183)	-0.084** (-4.245)	-0.169** (-2.244)
Share of population in Western middle class ($1 - F_{it-\tau}(13)$)	-0.127 (-2.781)	-0.154*** (-3.545)	-0.144** (-3.005)	-0.152*** (-3.572)
The miser index (x100)	0.066 (1.108)	0.018 (1.038)	-	-
Primary school enrolment rate (log)	0.007 (0.729)	0.002 (0.253)	-	-
Life expectancy (log)	0.112*** (2.795)	0.147*** (3.894)	0.110** (3.122)	0.151*** (4.164)
Price of investment (log)	-0.013*** (-2.482)	-0.015*** (-3.099)	-0.014** (-2.827)	-0.017*** (-3.246)
N	0.539	0.503	0.531	0.502
R ²	89	85	89	85

Notes: The dependent variable is the annualized change in log mean ($g_i(\mu_{it})$ for (1) and (3) and $g_i(C_{it})$ for (2) and (4)). The initial mean corresponds to the same measure used for the growth rate in each regression. The share of the Western middle class is not logged given that 11 observations are lost because of zeros. The t-ratios in parentheses are based on robust standard errors; * denotes significant at the 10% level; ** denotes significant at the 5% level; *** denotes significant at the 1% level. The hyphen indicates that the variable is dropped.

Table 4: Regressions for proportionate change in poverty rate as a function of the growth rate and initial poverty rate

	(1) OLS Complete specification:	(2) IVE Complete specification:	(3) OLS Dropping initial poverty rate:	(4) IVE Dropping initial poverty rate:	(5) OLS Imposing homogeneity:	(6) IVE Imposing homogeneity:
Intercept	0.009 (0.426)	0.012 (0.411)	-0.009* (-1.935)	0.002 (0.202)	-0.008** (-1.986)	-0.004 (-0.870)
Initial poverty rate ($\ln H_{it-\tau}$)	-0.005 (-1.062)	-0.004 (-0.504)	-	-	-	-
Growth rate (annualized change in log survey mean, $g_i(\mu_{it})$)	-2.587*** (-7.070)	-3.273*** (-4.645)	-2.519*** (-6.780)	-3.198*** (-4.779)	-	-
Growth rate interacted with initial poverty rate ($g_i(\mu_{it}).H_{it-\tau}$)	2.812*** (5.875)	3.067*** (2.953)	2.669*** (5.298)	2.594*** (2.525)	-	-
(1-Poverty rate) <i>times</i> growth rate ($g_i(\mu_{it}).(1-H_{it-\tau})$)	-	-	-	-	-2.468*** (-7.367)	-3.091*** (-4.808)
N	89	84	89	84	89	84
R ²	0.680	0.550	0.674	0.487	0.671	0.529
Homogeneity test	1.549	-0.391	0.877	-0.837	n.a.	n.a.

Notes: The dependent variable is the annualized change in log poverty rate for \$2 a day ($g_i(H_{it})$); t-ratios based on robust standard errors in parentheses; * denotes significant at the 10% level; ** denotes significant at the 5% level; *** denotes significant at the 1% level. The homogeneity test is the t-test for the sum of the coefficients on the growth rate $g_i(\mu_{it})$ and the growth rate interacted with initial poverty rate $g_i(\mu_{it}).H_{it-\tau}$; if the relationship is homogeneous then the coefficients sum to zero and the regressor becomes $g_i(\mu_{it}).(1-H_{it-\tau})$. The hyphen indicates that the variable is dropped.