Why Don't We See Poverty Convergence?

Martin Ravallion[1]

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

Abstract: We see signs of convergence in average living standards amongst developing countries and of greater progress against poverty in faster growing economies. Yet we do not see poverty convergence; the poorest countries are not enjoying higher rates of poverty reduction. The paper tries to explain why. Consistently with some growth theories, 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. Starting with a high incidence of poverty also entails a lower rate of progress against poverty at any given growth rate (and conversely 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 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.

Keywords: Poverty trap, middle class, inequality, economic growth

JEL: D31, I32, O15

---

[1] These are the views of the author and should not be attributed to the World Bank or any affiliated organization. Address: mravallion@worldbank.org. 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 Karla Hoff, Aart Kraay, 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 and the Institute of Social Sciences, Cornell University.

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 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 (reflecting a lower mean) should enjoy a higher subsequent growth rate and (hence) higher proportionate rate of poverty reduction. Indeed, as will be demonstrated later, the mean and the poverty rate will have the same speed of convergence in widely-used log-linear models.

That poses a puzzle. As this paper will also show, there is no sign of poverty convergence amongst developing countries, let alone a similar speed of convergence to that found for the mean. The overall incidence of poverty is falling in the developing world, but no faster (in proportionate terms) in its poorest countries.[2] Clearly something is missing from the story. Intuitively, one hypothesis is that either the growth process 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." Later I will point to a number of theoretical arguments as to how this can happen.

To test this hypothesis, a new data set was constructed for this paper from household surveys for almost 100 developing countries, each with two or more surveys over time. These data are used to estimate a model in which the rate of progress against poverty depends on the rate of growth in the mean and various parameters of the initial distribution—encompassing those identified in the literature—while the rate of growth depends in turn on initial distribution as well as the initial mean. The model is subjected to a number of tests, including alternative functional forms, sample selection by type of survey, and alternative measures are used for the key variables. A sub-sample with three or more surveys is also used to test robustness to

[2] Note that poverty convergence is defined in proportionate rather than absolute terms, in keeping with usage the in the growth literature. The absence of poverty convergence by this definition implies that poorer countries tend to see larger absolute reductions in their poverty rate.

different specification choices, including treating initial distribution as endogenous by treating lagged initial distribution as excludable.

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 incomes. Second, high initial poverty dulls the impact of growth on poverty; the poor enjoy a lower share of the gains from growth in poorer countries. On balance there is little or no systematic effect of starting out poor on the rate of poverty reduction. Other aspects of the initial distribution play no more than a secondary role. 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. 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 though this too is largely accountable to differences in the incidence of poverty.

2. Past theories and evidence

Growth theories incorporating credit-market failure suggest that high inequality reduces an economy's aggregate efficiency and (hence) growth rate.[3] The market failure is typically attributed to information asymmetries—that lenders are poorly informed about borrowers. The key analytic feature of such models is a suitably nonlinear relationship between an individual's initial wealth and her future wealth (the "recursion diagram"). The economic rationale for a nonlinear recursion diagram is that the credit market failure leaves unexploited opportunities for investment in physical and human capital and that there are diminishing marginal products of capital. Then 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. Models with such features include Galor and Zeira (1993), Benabou (1996), Aghion and Bolton (1997) and Banerjee and Duflo (2003).

But is it inequality that matters, or something else, such as poverty or the size of the middle class? Inequality is obviously not the same thing as poverty; inequality can be reduced

[3] There are a number of surveys including Perotti (1996), Hoff (1996), Aghion et al. (1999), Bardhan et al. (2000), Banerjee and Duflo (2003), Azariadis (2006) and World Bank (2006, Chapter 5). Borrowing constraints are not the only way that inequality can matter to growth. Another class of models is based on the idea that high inequality restricts efficiency-enhancing cooperation, such that key public goods are underprovided or efficiency-enhancing policy reforms are blocked (Bardhan et al., 2000). Other models argue that high inequality leads democratic governments to implement distortionary redistributive policies, as in Alesina and Rodrik (1994).

without a lower poverty measure by redistributing income amongst the non-poor, and poverty can be reduced without lower inequality. (Similarly, efforts to help the middle-class may do little to relieve current poverty.) In fact there is another implication of credit market failures that has received very little attention.[4] The following section studies one theoretical model from the literature more closely and shows that the simple fact of a credit constraint implies that unambiguously higher current poverty incidence—defined by any poverty line up to the minimum level of initial wealth needed to not be liquidity constrained in investment—yields lower growth at a given level of mean current wealth.

This is not the only argument suggesting that poverty is a relevant parameter of the initial distribution. 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. Another example can be found in the theories that have postulated impatience for consumption (high time preference rates possibly associated with low life expectancy) and hence low savings and investment rates by the poor (see, for example, Azariadis, 2006). Here too, while the theoretical literature has focused on initial inequality, it can also be argued that a higher initial incidence of poverty means a higher proportion of impatient consumers and hence lower growth.

Yet another example is found by considering how work productivity is likely to be 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 to work 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.

There are also theoretical arguments involving market and institutional development, though this is not a topic that has so far received as much attention in this literature. While past theories have often taken credit-market failures to be exogenous, poverty may well be a deeper causative factor in financial development (as well as an outcome of the lack of financial

[4] Ravallion (2001, 2007) argues intuitively that poverty retards growth when there are credit market failures.

development). For example, given fixed cost of lending (both for each loan and for setting up the lending institution), liquidity constraints can readily emerge as the norm in very poor societies.

A strand of the theoretical literature has also pointed to the possibilities for multiple equilibria in nonlinear dynamic models, whereby the lowest equilibrium is a poverty trap ("low-level attractor"). Essentially, the recursion diagram now has a low-level non-convexity, whereby a minimum level of current wealth is essential before any positive level of future wealth can be reached. In poor countries, the nutritional requirements for work can readily generate such dynamics, as illustrated by the model of Dasgupta and Ray (1986). Such a model predicts that a large exogenous income gain may be needed to attain a permanently higher income and that seemingly similar aggregate shocks can have dissimilar outcomes; growth models with such features are also discussed in Day (1992) and Azariades (1996, 2006) amongst others. Sachs (2005) has invoked such models to argue that a large expansion of development aid would be needed to assure a permanently higher average income in currently poor countries.

2.1 A model of aggregate growth with micro borrowing constraints

I now explore one of these models more fully. 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 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:

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

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

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

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

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:

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

Now consider the growth effect of a mean-preserving increase in the poverty rate. I assume that H_t^* increases and that no individual with wealth less than $k^*/(\lambda+1)$ becomes better off, implying that $\partial w_t(p)/\partial H_t^* \leq 0$ for all $p \leq H_t^*$. If this holds then I will say that poverty is unambiguously higher. It is readily verified that:[5]

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

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

The sign of (4) cannot be determined under the assumptions so far.[6] However, on imposing a constant initial mean $\mu_t = \bar{\mu}$, equation (4) simplifies to:

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

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 <u>given</u> 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 distributional effect of higher poverty (Proposition 2). Which effect dominates is an empirical question.

2.2 Past evidence on growth and the initial distribution

Following Barro and Sala-i-Martin (1992), cross-country regressions for GDP growth rates have found a significant negative coefficient on initial GDP once one controls for initial conditions. A subset of the literature has used inequality as one such initial condition. 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), Knowles (2005) and Voitchovsky (2005). 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.

---

[6] If there is (unrestricted) first-order dominance, whereby $\partial w_t(p)/\partial H_t^* \le 0$ for <u>all</u> $p \in [0, 1]$, then $\partial \mu_{t+1}/\partial H_t^* \le 0$. 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".

There are a number of unresolved specification issues in this literature. The aspect of initial distribution that has received almost all the attention in the empirical literature is inequality, as typically measured by the Gini index. Wealth inequality is arguably more relevant though this has rarely been used due to data limitations.[7]

The popularity of the Gini index appears to owe more to its availability in secondary data compilations than any intrinsic relevance to the economic arguments.[8] In the only paper I know of in which a poverty measure was used as a regressor for aggregate growth across countries, Lopez and Servén (2009) find evidence that a higher initial poverty rate retards growth. As Lopez and Servén observe, the significance of the Gini index in past studies may reflect an omitted variable bias, given that one expects (and I will later verify empirically) that inequality will be highly correlated with poverty at a given mean.

There are also issues about the relevant control variables when studying the effect of initial distribution on growth. The specification choices in past work testing for effects of initial distribution have lacked clear justification in terms of the theories predicting such effects. Consider three popular predictors of growth, namely human development, the investment share, and financial development. On the first, basic schooling and health attainments (often significant in growth regressions) are arguably one of the channels linking initial distribution to growth. Indeed, that is the link in the original Galor and Zeria (1993) model.[9] Turning to the second, one of the most robust predictors of growth rates is the share of investment in GDP (Levine and Renelt, 1992); yet arguably one of the main channels through which distribution affects growth is via aggregate investment and this is one of the channels identified in the theoretical literature. Finally, consider private credit (as a share of GDP), which has been used as a measure of "financial sector development" in explaining growth and poverty reduction (Beck et al., 2000, 2007). The theories discussed above based on borrowing constraints suggest that the aggregate flow of credit in the economy depends on the initial distribution.

Another set of specification issues concerns interaction effects. As Banerjee and Duflo (2003) point out, while liquidity constraints stemming from credit-market failures imply that the

[7] An exception is Ravallion (1998), who studies the effect of geographic differences in the distribution of wealth on growth in China.

[8] The compilation of Gini indices from secondary sources (and not using consistent assumptions) in Deininger and Squire (1996) led to almost all the tests in the literature since that paper was published.

[9] More recently, Gutiérrez and Tanaka (2009) show how high initial inequality in a developing country can yield a political-economy equilibrium in which there is little or no public investment in basic schooling; the poorest families send their kids to work, and the richest turn to private schooling.

growth rate depends on the extent of inequality in the initial distribution, they also suggest that there will be an interaction effect between the initial mean and inequality. However, as the further analysis of the Banerjee-Duflo model in the last section suggests, the more relevant interaction effect may well be that between poverty and inequality.

Some of the literature has focused instead on testing the assumptions of these theories. The empirical evidence on poverty traps is mixed. At least some of the theoretical models of poverty traps appear to be hard to reconcile with the aggregate data; see, in particular, the discussion in Kraay and Raddatz (2007) of poverty traps that might arise from low savings (high time preference rates) in poor countries. There are also testable implications for micro data. An implication of a number of the models based on credit-market failures is that individual income or wealth at one date should be an increasing concave function of its own past value. This can be tested on micro panel data. Lokshin and Ravallion (2004) provide supportive evidence in panel data for Hungary and Russia while Jalan and Ravallion (2004) do so using panel data for China. These micro studies suggest seemingly sizeable efficiency costs of inequality. The same studies do not, however, find the properties in the empirical income dynamics that would be needed for a poverty trap. There is also evidence of nonlinear wealth effects on new business start ups in developing countries, though with little sign of a non-convexity at low levels due to lumpiness in capital requirements (Mesnard and Ravallion, 2006). Similarly, McKenzie and Woodruff (2006) find no sign of non-convexities in production at low levels amongst Mexican microenterprises. However, Hoddinott (2006) and Barrett et al (2006) find evidence of wealth-differentiated behaviors in addressing risk in rural Zimbabwe and Kenya (respectively) that are consistent with the idea of poverty traps.

Micro-empirical support for the claim that there are efficiency costs of poor nutrition and health care for children in poor families has come from a number of studies. In a recent example, an impact evaluation by Macours et al. (2008) of a conditional cash transfer scheme in Nicaragua found that randomly assigned cash transfers to poor families improved the cognitive outcomes of children through higher intakes of nutrition-rich foods and better health care. This echoes a number of findings on the benefits to disadvantaged children of efforts to compensate for family poverty; for a review see Currie (2001).

While the theories and evidence reviewed above point to inequality and/or poverty as the relevant parameters of the initial distribution, yet another strand of the literature has pointed to various reasons why the size of a country's middle class can matter to the fortunes of those not (yet) so lucky to be middle class. It has been argued that a larger middle class promotes economic growth, such as by fostering entrepreneurship, shifting the composition of consumer demand, and making it more politically feasible to attain policy reforms and institutional changes conducing to growth. 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.

So we have three contenders for the distributional parameter most relevant to growth: inequality, poverty and the size of the middle class. The fact that very few encompassing tests are found in the literature, and that these different measures of distribution are not independent, leaves one in doubt about what aspect of distribution really matters. As already noted, 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*. Similarly, the main way the middle class expands in a developing country is probably through poverty reduction, so it is unclear whether it is a high incidence of poverty or a small middle class that impedes growth. Similarly, a relative concept of the "middle class," such as the income share of middle quintiles, will probably be highly correlated with a relative inequality measure, clouding the interpretation.

2.3 Growth and poverty reduction

The consensus in the literature is that higher growth rates tend to yield more rapid rates of absolute poverty reduction; see World Bank (1990, 2000), Ravallion (1995, 2001, 2007), Fields (2001) and Kraay (2006).[10] This is implied by another common finding in the literature, namely that growth in developing countries tends to be distribution-neutral on average, meaning that changes in inequality are roughly orthogonal to growth rates in the mean (Ravallion, 1995, 2001; Ferreira and Ravallion, 2009). Distribution-neutrality in the growth process implies that the

[10] Also see the review of the arguments and evidence on this point in Ferreira and Ravallion (2009).

changes in any standard measure of absolute poverty (meaning that the poverty line is fixed in real terms) will be negatively correlated with growth rates in the mean.

There is also evidence that inequality matters to how much a given growth rate reduces poverty (Ravallion, 1997, 2007; World Bank, 2000, 2006; Bourguignon, 2003; Lopez and Servén, 2006). Intuitively, in high inequality countries the poor will tend to have a lower share of the gains from growth. Ravallion (1997, 2007) examined this issue empirically using household survey data over time (earlier versions of the data set used here). Ravallion (1997) found that the following parsimonious specification fits the data for developing countries well:

$$\Delta \ln H_{it} = \eta(1 - G_{it-1})\Delta \ln \mu_{it} + \upsilon_{it} \tag{6}$$

where H_{it}, G_{it} and μ_{it} are the headcount index, the Gini index and the mean respectively for country *i* at date *t*, $\eta < 0$ is the elasticity of poverty reduction to the "distribution-corrected" growth rate $(1 - G_{it-1})\Delta \ln \mu_{it}$ and υ_{it} is a zero mean error term (uncorrelated with the growth rates). At minimum inequality ($G_{it-1} = 0$) growth has its maximum effect on poverty (in expectation) while the elasticity reaches zero at maximum inequality ($G_{it-1} = 1$). Ravallion (1997) did not find that the elasticity varied systematically with the mean, although Lopez and Servén (2006) showed that if incomes are log-normally distributed then such a variation is implied theoretically. Easterly (2009) conjectured that the initial poverty rate is likely to be the better predictor of the elasticity than initial inequality, though no evidence was provided.

3. Data and descriptive statistics

In keeping with the bulk of the literature, the country is the unit of observation.[11] 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 the initial distribution, which is measured from surveys. By calculating the distributional statistics directly from the primary data, some of the inconsistencies and 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.

[11] 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.

I found 99 developing and transition countries with at least two suitable household surveys since about 1980. (For about 70 of these countries there are three or more surveys.) For the bulk of the analysis I restrict the sample to the 92 countries in which the earliest available survey finds that at least some households lived below the average poverty line for developing countries (described below).[12] 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.[13] 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" effect associated with the transition. The experience of these countries is clearly not typical of the developing world.

The longest 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;[14]three-quarters of the spells use consumption.

Naturally the time periods between surveys are not uniform across countries. 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 household survey for date t_i in country *i* and the earliest available survey for date $t_i - \tau_i$, the proportionate annualized difference ("growth rate") for the variable *x* is denoted $g_i(x_{it}) \equiv \ln(x_{it} / x_{it-\tau}) / \tau$ (dropping the *i* subscript on *t* and τ for brevity). National accounts (NAS) data and social indicators are also used, matched as closely as possible

12 The data set was constructed from *PovcalNet* in December 2008. Seven countries were dropped because the poverty rate was zero in the earliest surveys.

13 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 (2008).

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

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 \$2.00 per day at 2005 PPP, which is the median poverty line amongst developing countries.[15] Let $F_{it}(z)$ denote the distribution function for country *i* at date *t*; then $H_{it} = F_{it}(2)$. In 2005, \$2 a day was also very close to the median consumption per person in the developing world. 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 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.[16]

Inequality is measured by the usual Gini index (G_{it})—half the mean absolute difference between all pairs of incomes normalized by the overall mean.

The size of the middle class (MC_{it}) is measured by the proportion of the population living between \$2 and \$13 a day (following Ravallion, 2009); so $MC_{it} \equiv F_{it}(13) - F_{it}(2)$.[17] These bounds are also somewhat arbitrary, although this definition appears to accord roughly with the idea of what it means to be "middle class" in China and India (Ravallion, 2009). By contrast, those living above \$13 a day can be thought of as the "middle class" by Western standards; the share of the "Western middle class" is $1 - F_{it}(13)$. These are interpretable as absolute measures of the middle class. I also calculated a relative definition of the middle class, namely the consumption or income share controlled by the middle three quintiles (MQ_{it}), as used by Easterly (2001).

Table 1 provides summary statistics for both the earliest and latest survey rounds. The mean Gini index stayed roughly unchanged at about 42%. The initial index ranged from 19.4%

---

15 This is based on the compilation of national poverty lines presented in Ravallion et al. (2009). The methods used in measuring poverty and inequality using these data are described in Chen and Ravallion (2008).

16 The \$1.25 line is the mean of the poorest 15 countries in terma of consumption per person. \$13 per person per day corresponds to the official poverty line in the US for a family of four; see Department of Health and Human Services.

17 Similarly Banerjee and Duflo (2008) used the interval \$2 to \$10 a day to define the middle class.

(Czech Republic) to 62.9% (Sierra Leone), both around 1990. In the earliest surveys, about one quarter of the sample had a Gini index below 30% while one quarter had an index above 50%.

The average size of the middle class increased, from a mean $MC_{it-\tau}$ of 48% to a mean MC_{it} of 53%. The middle-class expanded in 64 countries and contracted in 35. There is also a marked bimodality in the distribution of countries by the size of their middle class, as is evident in Figure 1, which plots the kernel densities of MC_{it} and $MC_{it-\tau}$. Taking 40% as the cut-off point, 30 countries are in the lower mode and 69 are in the upper one for the most recent survey; the corresponding counts for the earliest surveys are 42 and 57.[18] The relative measure of the size of the middle class behaved differently; there was little change in the mean *MQ* over time (Table 1) and the density function was unimodal in both the earliest and latest surveys.

Table 2 gives the correlation coefficients, focusing on the main regressors used later. The correlations point to a number of potential concerns about the inferences drawn from past research. The Gini index is highly (negatively) correlated with the income share of the middle three quintiles (r=-0.971 for the earliest surveys and -0.968 for the latest). The poverty measures are also strongly correlated with the survey means; $\ln H_{it-\tau}$ and $\ln \mu_{it-\tau}$ have a correlation of -0.851 (while it is -0.836 for $\ln F_{it-\tau}(1.25)$ and $\ln \mu_{it-\tau}$). The least-squares elasticity of $\ln H_{it-\tau}$ with respect to the initial survey mean (i.e., the regression coefficient of $\ln H_{it-\tau}$ on $\ln \mu_{it-\tau}$) is -1.305 (t=13.340). (All t-ratios in this paper are based on White standard errors.) There is a very high correlation between the poverty measures using \$1.25 a day and \$2.00 a day (r=0.974). There are weaker correlations between the two poverty measures and the initial Gini index (r= 0.241 and 0.099 for z=1.25 and z=2.00). However, there is also 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=0.802. The log Gini index also has a strong partial correlation with the log of the poverty rates holding the log mean constant (t=4.329 for $\ln H_{it-\tau}$).

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 -0.975; 95% of the variance in the initial size of the

[18] For further discussion of the developing world's rapidly expanding middle class, and the countries left behind in this process, see Ravallion (2009).

middle class is accountable to differences in the initial poverty rate. (The bimodality in terms of the size of the middle class in Figure 1 reflects a similar bimodality in terms of the \$2 a day poverty rate.) Across countries, 80% of the variance in the changes over time in MC_{it} can also be attributed to the changes in H_{it}.[19] The absolute and relative measures of the size of the middle class are positively correlated but not strongly so.

There is a strong correlation between the rate of poverty reduction and the ordinary growth rate in the survey mean (confirming the studies reviewed in section 2). Figure 2 plots the rate of poverty reduction ($g_i(H_{it})$) against $g_i(\mu_{it})$. The regression line in Figure 2 has a slope of -1.372 (t=-5.948) with R^2=0.363.

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 the later results. Table 2 also gives the correlation coefficients between τ and the various measures of initial distribution. The correlations are all small.

While this paper focuses mainly on the developing world as a whole, one region stands out: Sub-Saharan Africa (SSA). By the \$2.00 a day line, the mean of $H_{it-\tau}$ for SSA is 76.04% as compared to 29.51% for non-SSA countries; the difference is significant (t=8.84). Similarly, in terms of the size of its middle class, SSA is more concentrated in the lower mode in Figure 1. Two-thirds (20 out of 29) of SSA countries are in the lower mode for the earlier survey round; the corresponding means of $MC_{it-\tau}$ were 22.89% (s.e.=3.62%) and 59.07% (3.01%) for SSA and non-SSA countries respectively and the difference is statistically significant at the 1% level. Inequality too is higher in SSA; the mean Gini index in the earliest surveys is 0.474 (0.018) for SSA versus 0.390 (0.017) in non-SSA countries, and the difference is significant (t=7.68). There is clearly a “SSA effect” in both growth and poverty reduction, though we will see that this is accountable to the other variables in the estimated models.

4. Convergence?

Virtually all of the papers in the empirical literature reviewed in section 2 have assumed that the parameters of the dynamic processes for growth and poverty reduction are independent

---

19 R^2 =0.826 for the regression of $MC_{it} - MC_{it-\tau}$ on $F_{it}(2) - F_{it-\tau}(2)$; the regression coefficient is -0.896 (t=-25.496;n=92), which is significantly different from -1 (t=2.946).

of the initial level of poverty. The easiest way to see that this assumption cannot be right is to show that the standard models imply something that is not supported by the data.

Consider the most common empirical specification for the growth process in the mean:

$$\Delta \ln \mu_{it} = \alpha_i + \beta_i \ln \mu_{it-1} + \varepsilon_{it} \tag{7}$$

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:

$$\ln H_{it} = \delta_i + \eta_i \ln \mu_{it} + \nu_{it} \tag{8}$$

where $\eta_i < 0$ and ν_{it} is a zero-mean error term. This assumes that relative distribution fluctuates around a stationary mean, with changes in distribution orthogonal to growth rates in the mean. The implied growth model for poverty is then:

$$\Delta \ln H_{it} = \alpha_i^* + \beta_i \ln H_{it-1} + \varepsilon_{it}^* \tag{9}$$

for which it is readily verified that $\alpha_i^* = \alpha_i \eta_i - \beta_i \delta_i$ and $\varepsilon_{it}^* = \varepsilon_{it}\eta_i + \nu_{it} - (1+\beta_i)\nu_{it-1}$. The parameters of (7) and (8) ($\alpha_i, \beta_i, \delta_i, \eta_i$) can vary across counties but (for the sake of this argument) suppose they do so independently of $H_{it-\tau}$. Then the "speed of convergence" for the mean, $\partial \Delta \ln \mu_{it} / \partial \ln \mu_{it-1} = \beta_i$, is the same as that for poverty: $\partial \Delta \ln H_{it} / \partial \ln H_{it-1} = \beta_i$. 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.*

However, this is not borne out by the data. Table 3 gives convergence tests for both the mean and the poverty measures, with and without controls.[20] The 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. The survey means exhibit convergence with a β coefficient of -0.013 (t=-

[20] The test is the regression coefficient of $g_i(\mu_{it})$ on $\ln \mu_{it-\tau}$. Alternatively one can estimate the nonlinear regression $g(\mu) = \alpha - [(1-e^{\beta\tau})/\tau]\ln \mu_{-\tau} + \varepsilon$. This gave a very similar result to (1) in Table 4, namely $\hat{\beta} = -0.012$. (t=-2.865). Clearly, the approximation that $e^{\beta\tau} = 1 + \beta\tau$ works well.

3.412) without the controls and -0.042 (t=-7.435) with them. But this not true of the poverty measures. Indeed the proportionate rates of poverty reduction are orthogonal to initial levels.[21] Figure 3 plots the data, and gives a non-parametric regression line.

Clearly these results do not support the idea that the mean and the poverty measure have the same speed of convergence; indeed, there is no convincing sign of poverty convergence. The rest of this paper will try to explain why. In terms of the model above, it will be shown that the parameter α_i is a decreasing function of the initial poverty rate while the elasticity of poverty to the mean, $-\eta_i$, is a decreasing function of the initial level of poverty.

5. The relevance of initial poverty to growth in the mean

As discussed in section 2, initial distribution can matter to the rate of poverty reduction through two distinct channels, namely the growth rate and the elasticity of poverty to the mean. I postulate a simple triangular model in which rate of growth depends in turn on initial distribution while the rate of progress against poverty depends on the interaction between the growth rate and the initial distribution. This section focuses on the first relationship; section 6 turns to the second.

The section begins with benchmark regressions of growth on the initial mean and initial poverty rate. A causal interpretation of these regressions requires that the initial distribution (in the earliest survey used to construct each spell) is exogenous to the subsequent pace of growth. This can be questioned. I shall test encompassing models with controls for other factors. I also provide results for an instrumental variables estimator under widely-used (though still questionable) exclusion restrictions.

5.1 Benchmark regression for growth

Table 4 gives estimates of the following regression:[22]

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

[21] For the $1.25 line the corresponding regression coefficient was -0.005 with t=-0.393; at the other extreme, for the $13 line it was -0.009 (t=-0.480). Again, the nonlinear specification gave a very similar result.

[22] 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 4 implies a derivative that is less than unity for $\tau < 29$ years; the largest value of τ in the data is 27 years.

The estimates in column (1) suggest that differences in the initial poverty rate have sizeable negative impacts on the growth rate at a given initial mean. A one standard deviation increase in $\ln H_{it-\tau}$ would come with 0.021 (2% points) decline in the growth rate for the survey mean.

The fact that a significant (partial) correlation with the initial poverty rate only emerges when one controls for the initial mean is suggestive of an adverse distributional effect of high poverty. However, it is not simply a "relative poverty" effect, stemming from the variance in absolute poverty attributable to differences in relative distribution. This is evident in the fact that the convergence parameter increases considerably when one adds the initial poverty measure as a regressor. Dropping $\ln H_{it-\tau}$ from (10) the coefficient on $\ln \mu_{it-\tau}$ falls to -0.013 (t=-3.413). The presence of the poverty rate as a regressor magnifies the convergence parameter, suggesting that the fact that the absolute poverty rate depends on the mean is also playing a crucial role in determining its significance in these regressions—working against the convergence effect.

It might be conjectured that the effect of $\ln H_{it-\tau}$ in (10) 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 (10) using cubic functions of $\ln \mu_{it-\tau}$ to control for the initial mean. While I found some sign of higher-order effects of $\ln \mu_{it-\tau}$, these made very little difference to the regression coefficient on the poverty rate in the augmented regression; the coefficient on $\ln H_{it-\tau}$ in column (1) in Table 4 became -0.018 (t=-3.547).

There is, however, a marked nonlinearity in the relationship, which is being captured by the log transformation of $H_{it-\tau}$ in (10). If one uses $H_{it-\tau}$ rather than $\ln H_{it-\tau}$ on the same sample, the negative effects are still evident but they are much less precisely estimated, with substantially lower t-ratios—a t-ratio of -1.292 for the coefficient on $H_{it-\tau}$—though in both cases the effects come out somewhat more strongly if one adds a squared term in $H_{it-\tau}$ to pick up the nonlinearity, with both the linear and squared terms significant at the 10% level or better.

A simple graphical test for misspecification of the functional form in (10) is to plot $g_i(\mu_{it}) + 0.035 \ln \mu_{it-\tau}$ (from column (1) in Table 4) against $\ln H_{it-\tau}$. Figure 4 gives the results, along with a locally-smoothed (non-parametric) regression line. The relationship is close to linear in the log poverty rate.[23] The log transformation appears to be the right functional form.

The more relevant poverty line is that using the \$2.00 a day line. On replacing $\ln H_{it-\tau}$ by $\ln F_{it-\tau}(1.25)$ in (10) the poverty rate still had a negative coefficient but it was not significant at the 5% level. I also estimated the following specification:

$$g_i(\mu_{it}) = \alpha + \beta \ln \mu_{it-\tau} + \gamma_1[\ln H_{it-\tau} - \ln F_{it-\tau}(1.25)] + \gamma_2 \ln F_{it-\tau}(1.25)) + \varepsilon_{it} \quad (11)$$

The estimate of $\gamma_1 - \gamma_2$ was -0.010, but was not significantly different from zero (t=-0.801), suggesting that (10) is the correct specification.

The results were also robust to using the poverty gap index instead of the headcount index; the corresponding version of (10) was similar, with a coefficient on the log of the poverty gap index of -0.011, with t-ratio of -2.338. However, the fit is better using the headcount index.

Recall that the sample in estimating (10) used both consumption and income surveys, and that the latter may have more measurement error. Estimating the regression solely on consumption surveys strengthened the result; analogously to (10) one obtains column (2) of Table 4. The conditional convergence effect is even stronger, as is the poverty effect.

The results are robust to using NAS consumption growth instead (Table 4). The notable differences are that the convergence parameter in (10) is lower, $-\hat{\beta} = 0.02$ (column 3, Table 4) and that the headcount index based on the \$1.25 line is a slightly stronger predictor of the NAS consumption growth. (The results using NAS consumption growth were less sensitive to the choice of poverty line between \$2.00 and \$1.25 a day.)

Another way to use NAS consumption is as a control for other initial conditions influencing the long-run value of the survey mean. Augmenting (10) with this extra control variable gives, for the full sample:

$$g_i(\mu_{it}) = \underset{(3.942)}{0.181} - \underset{(-6.817)}{0.050} \ln \mu_{it-\tau} - \underset{(-2.382)}{0.011} \ln H_{it-\tau} + \underset{(3.682)}{0.022} \ln C_{it-\tau} + \hat{\varepsilon}_{it} \quad R^2=0.288;\ n=87 \qquad (12)$$

And for the sample of consumption surveys:

$$g_i(\mu_{it}) = \underset{(4.868)}{0.232} - \underset{(-6.507)}{0.061} \ln \mu_{it-\tau} - \underset{(-3.624)}{0.017} \ln H_{it-\tau} + \underset{(3.346)}{0.025} \ln C_{it-\tau} + \hat{\varepsilon}_{it} \quad R^2=0.317;\ n=66 \qquad (13)$$

The results are consistent with the expectation that $\ln C_{it-\tau}$ is picking up long-run differences. The poverty effect remains evident, though with a lower coefficient.

---

[23] In both cases I have scaled the vertical axis to accord with the sample mean growth rate by using the deviation of the log initial mean from its sample mean value.

5.2 *Further tests on the subsample with three surveys*

One can form a subsample of about 70 countries with at least three household surveys. When there were more than three surveys I picked the one closest to the midpoint of the interval between the latest survey and the earliest.

There are at least four ways one can exploit the extra round of surveys. The first is to test for convergence more robustly to measurement errors.[24] One way of doing this is to calculate the trend over the three surveys and test if this is correlated with the starting value. I estimated the trend for each country by regressing the logs of the three (date-specific) means for that country on time and similarly for the headcount indices. Convergence in the mean was still evident; the regression coefficient of the estimated trend on the log mean from the earliest survey was -0.009 (t=-2.052), which is significant at the 4% level. And again there was no significant correlation between these trends in poverty reduction and the initial poverty measures; the regression coefficient of the estimated trend on the log headcount index from the earliest survey was 0.007 (t=0.805). Another method is to form means from the first two surveys and look at their relationship with the changes observed between the last survey and the middle one. Define the mean from the first two surveys as $M_i(x_{it-\tau_2}) \equiv (x_{it-\tau_2} + x_{it-\tau_1-\tau_2})/2$ while the growth rate is $g_i(x_{it}) \equiv \ln(x_{it}/x_{it-\tau_2})/\tau_2$. Using this method, unconditional mean convergence was no longer evident (though conditional convergence was still found) but there was an indication of poverty <u>divergence</u>; regressing $g_i(H_{it})$ (the proportionate change in the poverty measure between the middle and final rounds) on $M_i(H_{it-\tau_2})$; the coefficient was 0.029, which is significant at the 6% level (t=1.901). There is still some contamination due to measurement error in these tests. Yet another method is to regress $g_i(x_{it})$ on the measure from the earliest survey ($\ln \mu_{it-\tau_1-\tau_2}$); the result was similar, namely little sign of (unconditional) mean convergence but mild divergence for poverty (a β coefficient of 0.027 with t=1.819).

24 As is well known, measurement errors can create spurious signs of convergence; if the initial mean is over-(under-) estimated then the subsequent growth rate will be lower (higher).clearly stems in part at least from this problem. It is notable that the β coefficient drops using only the consumption surveys (Table 4) or NAS consumption. However, significant conditional convergence in the means (including those only from consumption surveys) and NAS consumption is still evident (Table 4).

Secondly, the subsample can be used to form inter-temporal averages, to reduce the attenuation biases in the benchmark regression due to measurement error; equation (10) can be re-estimated in the form:

$$g_i(\mu_{it}) = \alpha + \beta \ln M_i(\mu_{it-\tau_2}) + \gamma \ln M_i(H_{it-\tau_2}) + \varepsilon_{it} \qquad (14)$$

Column (4) of Table 4 gives the results. The regression coefficients are larger (in absolute value), consistent with the presence of attenuation bias in the earlier regressions. The standard errors also fall noticeably. This strengthens the earlier results based on equation (10).

The third way of using the extra survey rounds is as a source of instrumental variables (IVs). Growth rates between the middle and last survey rounds were regressed on the mean and distributional variables for the middle round but treating the latter as endogenous and retaining the data for the earliest survey round as a source of IVs. Letting τ_i now denote the length of spell *i* (=1,2), the model becomes:

$$g_i(\mu_{it}) = \alpha + \beta \ln \mu_{it-\tau_2} + \gamma \ln H_{it-\tau_2} + \varepsilon_{it} \qquad (15)$$

The instrumental variables were $\ln \mu_{it-\tau_1-\tau_2}$, $\ln C_{it-\tau_1-\tau_2}$, $\ln G_{it-\tau_1-\tau_2}$, $\ln F_{it-\tau_1-\tau_2}(z)$ (*z*=1.25, 2.00) and τ_1. The first-stage regressions for $\ln \mu_{it-\tau_2}$ and $\ln H_{it-\tau_2}$ had R^2=0.884 (F=61.06) and R^2=0.796 (F=31.30) respectively. The Generalized Methods of Moments (GMM) estimates of (15) are found in Table 4, Column (5). (I also give the corresponding result using NAS consumption in column (6).) We see that the finding that a higher initial poverty rate implies a lower subsequent growth rate (at given initial mean) is robust to allowing for the possible 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. Analogously to equation (12.1), on adding $\ln C_{it-\tau_2}$ to specification (5) in Table 4, and treated it as exogenous, one obtains (using the same set of instruments):

$$g_i(\mu_{it}) = \underset{(3.731)}{0.235} - \underset{(-4.407)}{0.066} \ln \mu_{it-\tau} - \underset{(-3.954)}{0.022} \ln H_{it-\tau} + \underset{(4.481)}{0.035} \ln C_{it-\tau} + \hat{\varepsilon}_{it} \qquad (16)$$

Dropping $\ln C_{it-\tau_2}$ from the set of IVs gives instead:

$$g_i(\mu_{it}) = \underset{(1.815)}{0.144} - \underset{(-2.206)}{0.041} \ln \mu_{it-\tau} - \underset{(-2.634)}{0.016} \ln H_{it-\tau} + \underset{(3.288)}{0.025} \ln C_{it-\tau} + \hat{\varepsilon}_{it} \qquad (17)$$

Finally, one can use the subsample is to estimate a specification with country-fixed effects, which sweep up any confounding latent heterogeneity in growth rates at country level.

The main results were <u>not</u> robust to this change. Regressing the change in annualized growth rates ($g_i(\mu_{it}) - g_i(\mu_{it-\tau_2})$) 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 remained significant but the poverty rate ceased 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. Simulation studies have found that the coefficients on growth determinants are heavily biased toward zero in fixed-effects growth regressions (Hauk and Wacziarg, 2009).[25] 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 a low signal-to-noise ratio 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}). Table 5 gives the levels regression in logs, which implies an elasticity of μ_{it} to C_{it} of 0.75 (R^2=0.82) for the latest survey rounds.[26] Using a country-fixed effects specification in the levels, the elasticity drops to 0.46 while with fixed-effects in the growth rates (using the subsample with at least three surveys) it 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.

5.3 Encompassing regressions

It might be conjectured that the poverty measures (at given initial means) are picking up other aspects of the initial distribution, such as inequality (the variable identified in almost all the empirical literature, as discussed in section 2). Simply adding the log of the initial Gini index to equation (10) does not change the result; the coefficient on the Gini index is not significantly different from zero; the coefficient on $\ln H_{it-\tau}$ remains (highly) significant in the augmented version of (10). To investigate this further, I added inequality ($\ln G_{it-\tau}$), the income share of the

25 This point is illustrated well by the Monte Carlo simulations found in Hauk and Wacziarg (2009).

26 Including the seven developing countries with zero initial poverty ($F_{it-\tau}(2) = 0$) increases the elasticity in the levels to 0.750 (t=21.543) but makes little difference to the fixed effects estimates.

middle three quintiles ($\ln MQ_{it-\tau}$), the share of the Western middle class ($1-F_{it-\tau}(13)$) and three commonly used variables from the literature on growth empirics mentioned above, namely the primary school enrollment rate, life expectancy at birth, and the price index of investment goods. The population share of the developing world's middle class was not included given that its value is nearly linearly determined by the poverty rate and share of the Western middle class.

Table 6 gives the encompassing regressions using both survey means and consumption from the NAS. The table also gives restricted forms that passed comfortably. The initial poverty rate remains a (highly) significant predictor of growth in these encompassing models. Furthermore, its coefficient falls only modestly in the encompassing regressions (comparing columns (1) and (3) in Table 4 with (1) and (2) respectively in Table 6); this suggests that a large share of its explanatory power is independent of these extra variables. The size of the Western middle class, life expectancy and the price of investment are also significant predictors. The relative share of the middle quintiles is also significant for the growth rates in the survey means (but not NAS consumption), though with a negative sign. (That was also true if one replaced $MQ_{it-\tau}$ with $\ln G_{it-\tau}$.)

The two regional effects that have been identified in the literature on growth empirics are for Sub-Saharan Africa (negatively) and East Asia (positively). I tested augmented versions of the regressions in Table 6 with dummy variables for these two regions. There was no sign of an SSA effect in any specification. There was a negative East Asia effect though only (mildly significant (at the 8% level). Of course, there are unconditional effects on growth in both regions. But these are largely captured within the model, particularly for Africa.

I also tried adding two interaction effects. In the first, I added an interaction effect between inequality and the initial mean, as discussed above; this was highly insignificant (t-ratio of -0.063). Second, adding $\ln G_{it-\tau}.\ln H_{it-\tau}$ I found that it had a positive coefficient (contrary to the theoretical expectation discussed in section 2) though it was not significantly different from zero at even the 15% level.

Inequality and the income share of the middle quintiles are insignificant when one controls for initial poverty (though, of course, inequality is one factor leading to higher poverty), but the population share of the Western middle class emerges with a significant negative coefficient. The jointly negative coefficients on the poverty rate and the share of the Western middle class imply that a higher population share in the developing-world middle class is growth enhancing. Thus the data can also be well described by a model relating growth to the share of the developing world's middle class. (As one would expect, replacing $\ln H_{it-\tau}$ and $1-F_{it-\tau}(13)$ by $\ln[F_{it-\tau}(13)/H_{t-\tau}]$ gave very similar overall fit, though not quite as good as Table 6.) The negative (conditional) effect of the poverty rate may well be transmitted through differences in the size of the middle class.

The subsample with three surveys also allows one to test for the distributional effect reported by Banerjee and Duflo (2003), who argued that 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. To test for this, I repeated the regressions above using the 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. (Other variables were the same except for the middle survey.) The coefficients on the initial poverty rate (now the poverty rate for the middle survey) remained significant at the 1% level and the "Western middle class effect" remained evident but with reduced significance. However, the 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.

While the above results appear to be convincing 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. As already noted, the between-country differences in the incidence of poverty at a given mean reflect differences in relative distribution. While those differences are not simply a matter of "inequality" as normally defined, they are correlated with inequality. The predicted values of the growth rates from the regression in column (1) of Table 4 are significantly correlated with inequality; r=-0.442. Since higher inequality tends to imply higher poverty at a given mean (section 3), it also implies lower growth prospects.

6. Initial poverty and the growth elasticity of poverty reduction

I turn now 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 rate of poverty reduction, as distinct from the indirect effect via the rate of growth. Again I focus on the \$2 line, although the \$1.25 line gave similar results.

For any given relative distribution the elasticity of the poverty rate to the mean is simply given by (one minus) the elasticity of the poverty rate with respect to the poverty line.[27] This can be calculated at any given poverty line. The interaction effect between this elasticity and the growth rate is then an obvious predictor of the rate of poverty reduction.[28] On calculating the elasticity for the \$2 a day poverty line using the initial survey for each country, and denoting that elasticity by $\eta_{it-\tau}$, one finds that the regression coefficient of $\ln(H_{it}/H_{it-\tau})$ on $\eta_{it-\tau}\ln(\mu_{it}/\mu_{it-\tau})$ is not significantly different from unity; the coefficient is 1.062 with a standard error of 0.198 and R^2=0.389. Of course there are also changes in relative distribution, which presumably account for the bulk of the remaining variance in rates of poverty reduction. Consistently with past findings in the literature,[29] the changes in relative distribution are virtually orthogonal to rates of growth and (hence) the above regression coefficient is very close to unity. (If higher growth was systematically associated with worsening distribution then the regression coefficient would be biased downward, and so below unity.) However, there may well be relevant correlations with the properties of the initial distribution. Additionally, the elasticity is itself a function of the initial mean and initial distribution. These observations motivate a reduced form model in which the rate of poverty reduction depends on both the rate of growth and its interaction effects with relevant aspects of the initial distribution.

Table 7 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 regression:

$$g_i(H_{it}) = \delta_0 + \delta_1 \ln H_{it-\tau_2} + (\eta_0 + \eta_1 H_{it-\tau_2}) g(\mu_{it}) + \upsilon_{it} \tag{18}$$

Results are given for both OLS and IVE; the IVE method uses the growth rate in private consumption per capita from the NAS as the instrument for the growth rate in the survey mean. Following Ravallion (2001), this IV allows for the possibility that a spurious negative correlation

[27] This follows immediately from the aforementioned fact that the poverty rate is homogeneous of degree zero in the poverty line and the mean for a given Lorenz curve.

[28] On exploiting this fact in a decomposition analysis for a panel of countries (using an earlier version of the same data set used here) Kraay (2006) concludes that the bulk of the variance in rates of poverty reduction is due rates of growth. Note that this can be true and yet there is a large difference in the rates of poverty reduction at a given rate of growth between countries with different initial distributions; see Ravallion (2007).

[29] For a recent overview see Ferreira and Ravallion (2009).

exists due to common measurement errors (given that the poverty measure and the mean are calculated from the same surveys).

The results in Table 7 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 7 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 *times* one minus the poverty rate. At an initial poverty rate of 10% (about one standard deviation below the mean) the elasticity is about -3 while it falls to about -0.7 at a poverty rate of 80% (about one standard deviation above the mean). I also used the subsample with three survey rounds to implement an IVE using the same instruments as for (15). The homogeneity restriction was (again) easily accepted (t=-0.447). The IVE of the regression coefficient of $g_i(H_{it})$ on $(1-H_{it-\tau_2})g_i(\mu_{it})$ was -3.478 (t=-3.092).

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

$$g_i(H_{it}) = -\underset{(-1.749)}{0.011} + (\underset{(0.221)}{0.043} - \underset{(-4.818)}{0.029} MC_{it-\tau}) g_i(\mu_{it}) + \hat{\upsilon}_{it} \quad R^2=0.539,\ n=91 \tag{19}$$

At the lower mode for $MC_{it-\tau}$ of around 15% (Figure 1), equation (19) implies a growth elasticity of -0.39 (t=-3.13) while at the upper mode, around 75%, it is -2.13 (t=-7.15). 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:

$$g_i(H_{it}) = -\underset{(-1.939)}{0.011} + (\underset{(0.039)}{0.167} - \underset{(-0.631)}{0.030} F_{it-\tau}(13) + \underset{(3.638)}{0.029} H_{it-\tau}) g_i(\mu_{it}) + \hat{\upsilon}_{it} \quad R^2=0.539,\ n=91 \tag{20}$$

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 (F=0.001), implying that it is the middle-class share that matters, as in equation (19).

Statistically it is a dead heat then 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 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 tested an encompassing model with extra interaction effects with $G_{it-\tau}$, the partial elasticity of poverty reduction ($\eta_{it-\tau}$), 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 to the region's above-average poverty incidence.) These were individually and jointly insignificant (the joint F-test accepted the null with prob.=0.199).

Does the relationship differ according to whether growth is positive or negative? The survey mean decreased over time for about 30% of the spells; the mean $I[g_i(\mu_{it})] = 0.687$ where $I[x]$ is the indicator function ($I[x] = 1$ if x>0 and $I[x] = 0$ otherwise). On stratifying the parameters according to whether the mean is increasing or not, and re-estimating specification (3) in Table 7 one obtains:

$$\begin{aligned} g_i(H_{it}) &= \underset{(-1.628)}{-0.013} + (\underset{(4.246)}{2.869} H_{it-\tau} - \underset{(-5.046)}{3.117}) I[g_i(\mu_{it})] g_i(\mu_{it}) \\ &+ (\underset{(3.709)}{2.218} H_{it-\tau} - \underset{(-5.717)}{1.984})(1 - I[g_i(\mu_{it})]) g_i(\mu_{it}) + \hat{\upsilon}_{it} \end{aligned} \qquad R^2=0.552,\ n=91 \quad (21)$$

The positive interaction effect is found during spells of contraction in the mean ($I[g_i(\mu_{it})] = 0$) as well as expansions ($I[g_i(\mu_{it})] = 1$); the homogeneity restriction passes in both cases (the t-test for contractions is 1.143, versus 1.425 for expansions). Nor can one reject the null that the coefficients are the same for expansions versus contractions (F=2.978, prob=0.062).

So the key proximate determinant of the elasticity is the initial poverty rate. Figure 5 plots the rate of poverty reduction against the poverty-adjusted growth rate in the survey mean (analogous to Figure 2, which used the ordinary growth rate). The slope of the regression line is almost twice as high (a coefficient of -2.613, t=-7.273) and R^2=0.535, as compared to 0.363 for the regression in Figure 2. So allowing for initial distribution, as measured by the $2 a day poverty rate, adds 17% points to the share of the variance in the rate of poverty reduction that can be explained by the rate of growth in the survey mean.

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

Recall that the speed of poverty convergence, $\partial g_i(H_{it}) / \partial \ln H_{it-\tau}$, is very close to zero. We can now combine the main results to help explain why. Based on the various encompassing tests above, my empirically-preferred model takes the form:

$$g_i(H_{it}) = \eta(1 - H_{it-\tau})g_i(\mu_{it}) + \upsilon_{it} \quad (22.1)$$

$$g_i(\mu_{it}) = \alpha + \beta \ln \mu_{it-\tau} + \gamma \ln H_{it-\tau} + \varepsilon_{it} \quad (22.2)$$

The regressors in (22.2) 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.[30] 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 three-way decomposition of the poverty convergence elasticity:

$$\frac{\partial g_i(H_{it})}{\partial \ln H_{it-\tau}} = \underbrace{\eta\beta(1 - H_{it-\tau})\left(\frac{\partial \ln H_{it-\tau}}{\partial \ln \mu_{it-\tau}}\right)^{-1}}_{\text{(Mean convergence effect)}} + \underbrace{\eta\gamma(1 - H_{it-\tau})}_{\text{(Direct effect of poverty)}} - \underbrace{\eta g_i(\mu_{it})H_{it-\tau}}_{\text{(Poverty elasticity effect)}} \quad (23)$$

On evaluating all variables at their sample means and using the estimates in column (1) of Table 4 and column (5) from Table 7, and using the OLS elasticity of elasticity of the initial headcount index with respect to the initial survey mean of -1.305, one finds that the mean convergence effect is -0.038, while the direct effect of poverty is 0.024 and the poverty elasticity effect is 0.0195. 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.

8. Conclusions

Arguably the most interesting thing about the fact that we do not see poverty convergence in the developing 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

---

[30] Nonetheless, as we have also seen, the variance across countries in initial distribution entails that $\ln \mu_{it-\tau}$ and $\ln H_{it-\tau}$ are not so correlated to prevent disentangling their effects.

growth reduces poverty—suggests that something about the initial distribution is offsetting the "advantage of backwardness."

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 with a lower initial 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 (as well as other controls). Additionally a high poverty rate makes it harder to achieve a given proportionate impact on poverty through growth in the mean. (By the same token, the poverty impact of economic contraction 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 "advantage of backwardness" 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, such as human underdevelopment and policy distortions.

The evidence is mixed on the role played by other aspects of distribution. A larger middle class—by developing-country (but not Western) standards—makes growth more poverty-reducing. But this effect is largely attributable to the lower poverty rate associated with a larger middle class. Controlling for the initial incidence of poverty, there is no sign that a higher overall level of initial inequality, as measured by the Gini index, inhibits the pace of poverty reduction via either the rate of growth or the growth elasticity. Nonetheless, initial inequality is empirically important, via its bearing on the extent of poverty. This is plain if one calculates the predicted rate of poverty reduction for each country, given its initial conditions.[31] The five countries with the highest (most negative) predicted rates of poverty reduction are Lithuania (32), Estonia (30), Jordan (36), Belarus (30) and Hungary (25); the numbers in parentheses are their initial Gini

[31] For the following calculation I substituted equation (1) in Table 4 (though other specifications give similar results) into the regressions with homogeneity imposed in Table 7.

indices in %, so it is clear that most of these are relatively low-inequality countries. By contrast, the five countries with the lowest predicted values were all high inequality countries, namely: Venezuela (56), Chile (56), Brazil (57), Colombia (57) and South Africa (59).

While these findings confirm that initial inequality matters to subsequent progress against poverty, they also reveal that the main way it matters is via its bearing on the initial incidence of poverty. There is no sign in this paper's results that lower inequality amongst the non-poor, leaving the incidence of absolute poverty unchanged, brings any longer-term payoff in terms of growth and poverty reduction. And in the minority of cases in which high inequality comes with low absolute poverty at a given mean, it does not imply worse longer-term prospects for growth and poverty reduction.

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, 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 "knock out" the effect of poverty, either on growth or poverty reduction at a given rate of growth. However, the cross-country empirical relationships reported here do point to the importance in future work of better understanding the 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, 1996, "The Economics of Poverty Traps. Part One: Complete Markets," *Journal of Economic Growth* 1: 449-486.

______________, 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.

_____________ and ___________, 2008, "What is Middle Class about the Middle Classes Around the World?," *Journal of Economic Perspectives* 22(2): 3-28.

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.

Barrett, Christopher, Paswel Phiri Marenya, John McPeak, Bart Minten, Festus Murithi, Willis Oluoch-Kosura, Frank Place, Jean Claude Randrianarisoa, Jhon Rasambainarivo and Justine Wangila, 2006, "Welfare Dynamics in Rural Kenya and Madagascar," *Journal of Development Studies* 42(2): 248-277.

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.

Benabou, Roland, 1996, “Inequality and Growth”, in Ben Bernanke and Julio Rotemberg (eds) *National Bureau of Economic Research Macroeconomics Annual*, Cambridge: MIT Press.

Beck, Thorsten, Ross Levine and Norman Loayza, 2000, “Finance and the Sources of Growth,” *Journal of Financial Economics* 58: 261-300.

Beck, Thorsten, Asli Demirguc-Kunt and Ross Levine, 2007, “Finance, Inequality and the Poor,” *Journal of Economic Growth* 12: 27-49.

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, 2008, “The Developing World is Poorer than we Thought, but no less Successful in the Fight against Poverty,” Policy Research Working Paper 4703, World Bank.

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.

Currie, Janet, 2001, “Early Childhood Development Programs,” *Journal of Economic Perspectives* 15(2): 213-38.

Dasgupta, Partha and Debraj Ray, 1986, “Inequality as a Determinant of Malnutrition and Unemployment”, *Economic Journal* 96: 1011-34.

Datt, Gaurav and Martin Ravallion, 1992, “Growth and Redistribution Components of Changes in Poverty Measures: A Decomposition with Applications to Brazil and India in the 1980s,” *Journal of Development Economics*, 38: 275-295.

Day, Richard H., 1992, “Complex Economic Dynamics: Obvious in History, Generic in

Theory, Elusive in Data," *Journal of Applied Econometrics* 7: S9-S23.

Deininger, Klaus and Lyn Squire, 1996 "A New Data Set Measuring Income Inequality," *World Bank Economic Review* 10: 565-591.

______________ and _________, 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.

Gutiérrez, Catalina and Ryuichi Tanaka, 2009, "Inequality and Education Decisions in Developing Countries," *Journal of Economic Inequality* 7: 55-81.

Hauk, William R. and Romain Wacziarg, 2009, "A Monte Carlo Study of Growth Regressions," *Journal of Economic Growth* 14(2): 103-147.

Hoddinott, John, 2006, "Shocks and their Consequences Across and Within Households in Rural Zimbabwe," *Journal of Development Studies* 42(2): 301-321.

Hoff, Karla, 1996, "Market Failures and the Distribution of Wealth: A Perspective from the Economics of Information," *Politics and Society* 24(4): 411-432.

Jalan, Jyotsna and Martin Ravallion, 2004, "Household Income Dynamics in Rural China," in Stefan Dercon (ed) *Insurance Against Poverty,* Oxford University Press.

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.

Lokshin, Michael and Martin Ravallion, 2004, "Household Income Dynamics in Two Transition Economies," *Studies in Nonlinear Dynamics and Econometrics,* 8(3).

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.

Macours, Karen, Norbert Schady and Renos Vakis, 2008, "Cash Transfers, Behavioral Changes and Cognitive Development in Early Childhood," Policy Research Working Paper 4759, World Bank, Washington DC.

McKenzie, David and Christopher Woodruff, 2006, "Do Entry Costs Provide and Empirical Basis for Poverty Traps? Evidence from Mexican Microenterprises," *Economic Development and Cultural Change* 55(1): 3-42.

Mesnard, Alice and Martin Ravallion, 2006, "The Wealth Effect on New Business Startups in a Developing Economy," *Economica,* 73: 367-392.

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, "The Developing World's Bulging (but Vulnerable) 'Middle Class'", Policy Research Working Paper 4816, World Bank, Washington DC (http://econ.worldbank.org/docsearch).

Ravallion, Martin, Shaohua Chen and Prem Sangraula, 2009, "Dollar a Day Revisited," *World Bank Economic Review* 23(2):.163-184.

Sachs, Jeffrey, 2005, *Investing in Development: A Practical Plan to Achieve the Millennium Development Goals*, Millennium Project, United Nations, New York.

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: Densities of middle-class population shares

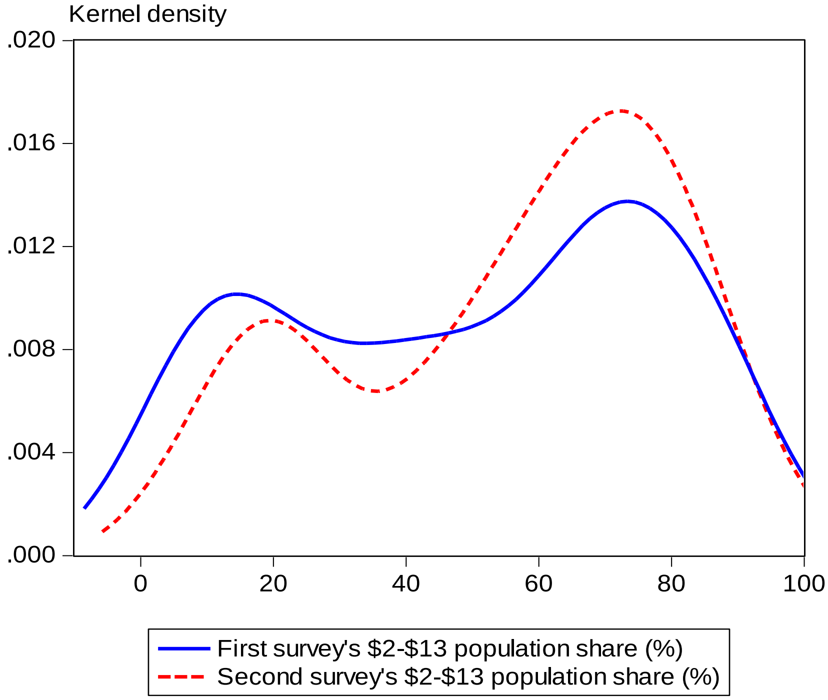


Figure 2: Rate of poverty reduction plotted against rate of growth in survey mean

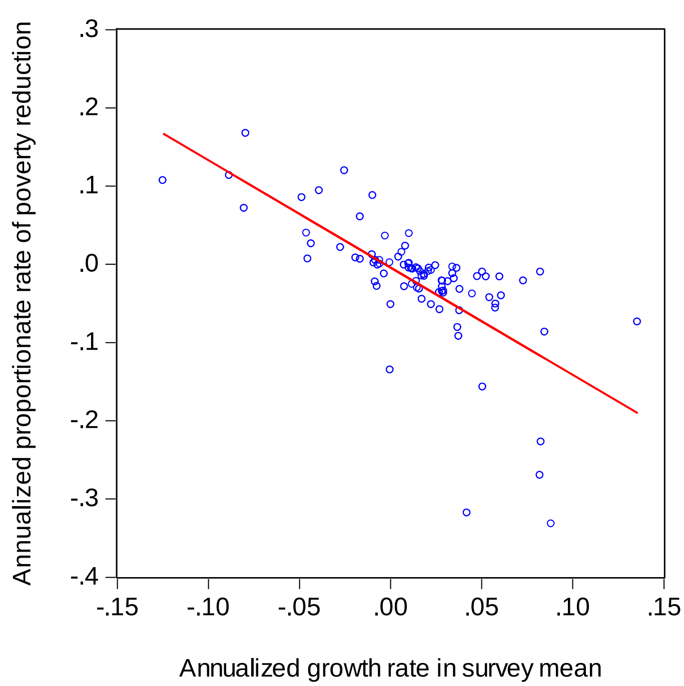
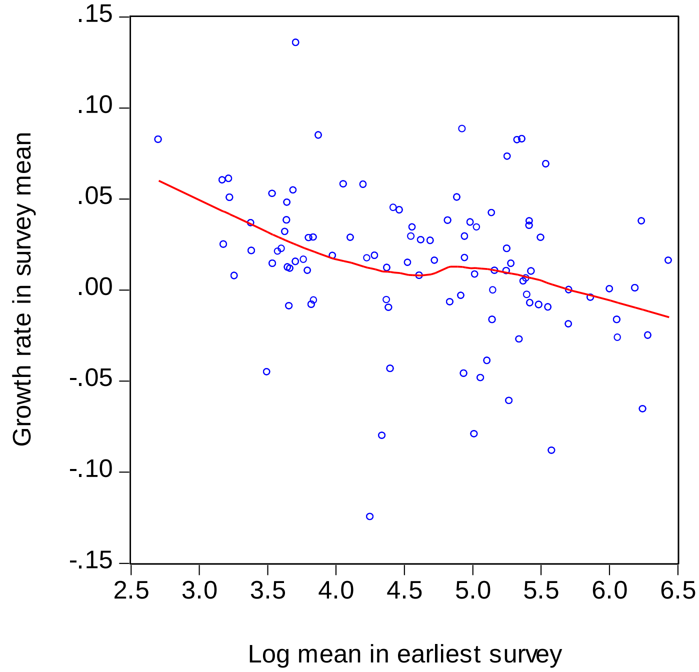


Figure 3: Growth rates plotted against initial values

1(a): Survey means



1(b): Headcount indices of poverty

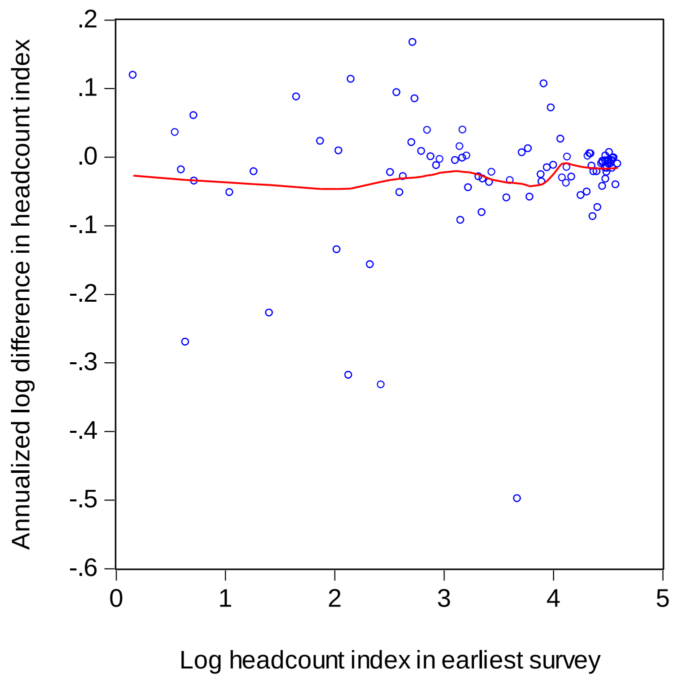


Figure 4: Growth rate with a control for the initial mean plotted against the initial poverty rate

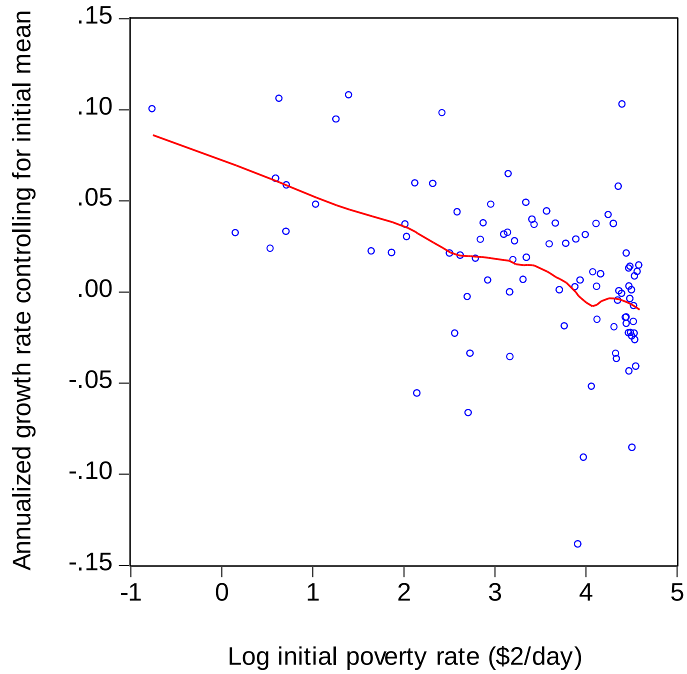


Figure 5: Rate of poverty reduction plotted against distribution-corrected rate of growth

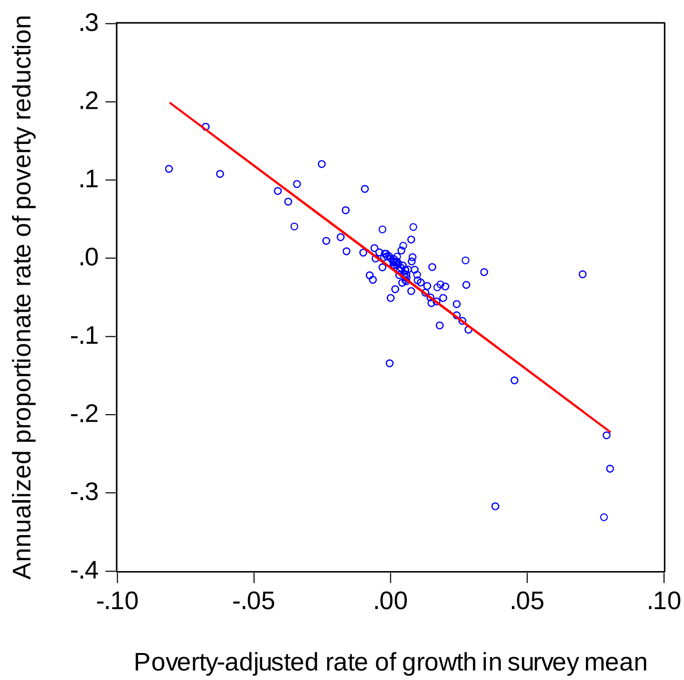


Table 1: Summary statistics

		No. observations	Mean	Standard deviation
Year	Earliest survey	92	1990.23	5.51
	Latest survey	92	2003.34	3.10
Survey mean ($PPP, 2005)	Earliest survey	92	126.88	98.55
	Latest survey	92	151.00	104.84
Annualized rate of growth in survey mean (%/year)		92	1.61	4.01
Gini index (%)	Earliest survey	92	42.49	9.87
	Latest survey	92	41.96	8.33
Poverty rate for $1.25 a day (%)	Earliest survey	92	31.46	24.77
	Latest survey	92	24.12	24.55
Poverty rate for $2 a day (%)	Earliest survey	92	46.42	33.57
	Latest survey	92	39.79	31.04
Share of developing-world middle class (%)	Earliest survey	92	48.43	29.27
	Latest survey	92	52.97	25.17
Share of Western middle class (%)	Earliest survey	92	5.15	8.20
	Latest survey	92	7.23	9.14
Income share of middle three quintiles (%)	Earliest survey	89	45.26	5.96
	Latest survey	89	45.54	4.95

Note: The sample is all Part 2 member countries of the World Bank with adequate nationally-representative household surveys and for which the estimated headcount index for the $2 a day line is positive in the earliest survey.

Table 2: Correlation matrix

Growth of poverty rate for \$2/day ($g_i(H_{it})$)								
Growth rate of survey mean ($g_i(\mu_{it})$)	-0.573							
Survey mean ($\ln\mu_{it-\tau}$)	0.017	-0.355						
Poverty rate for \$2/day ($\ln H_{it-\tau}$)	0.017	0.209	-0.859					
Gini index of inequality ($\ln G_{it-\tau}$)	0.176	-0.048	0.010	0.253				
Middle class population share ($MC_{it-\tau}$)	-0.122	-0.244	0.880	-0.815	-0.199			
Western middle class share ($1-F_{it-\tau}(13)$)	0.189	-0.335	0.742	-0.722	0.136	0.399		
Income share of middle three quintiles ($MQ_{it-\tau}$)	-0.139	-0.017	0.069	-0.324	-0.959	0.279	-0.054	
Time between survey rounds (τ)	0.108	-0.107	0.070	0.083	0.190	-0.046	0.134	-0.146
	$g_i(H_{it})$	$g_i(\mu_{it})$	$\ln\mu_{it-\tau}$	$\ln H_{it-\tau}$	$\ln G_{it-\tau}$	$MC_{it-\tau}$	$1-F_{it-\tau}(13)$	$MQ_{it-\tau}$

Note: Correlation matrix for common sample of complete data for all variables (n-83); pair-wise correlations quoted in text use all available observations for that pair of variables and so may differ from those above.

Table 3: 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)
Unconditional	-0.013**	-0.010	-0.007	0.005	-0.005
	(-3.413; n=99)	(-1.882; n=74)	(-1.743; n=92)	(0.542; n=86)	(-0.393; n=79)
Conditional	-0.042**	-0.040**	-0.026**	-0.015	-0.028
	(-7.435; n=90)	(-4.928; n=68)	(-4.431; n=90)	(-1.035; n=86)	(-1.734; n=79)

Note: The table gives $\hat{\beta}$ in the regression $g_i(\mu_{it}) = \alpha + \beta \ln \mu_{it-\tau} + \gamma X_{it-\tau} + \varepsilon_{it}$. T-ratios based on White standard errors (corrected for heteroskesdicity). The controls (all for earliest survey date) used in testing for conditional convergence were log mean consumption per capita from NAS (for the survey means and poverty measures), log primary school enrollment rate; log life expectancy; log relative price index of investment goods.

Table 4: Alternative estimates of the regression of growth rates on initial mean and initial headcount index of poverty

	(1)	(2)	(3)	(4)	(5)	(6)
	Sample with two surveys				Sample with three surveys	
	Full sample	Consumption surveys only	NAS consumption per capita	Means from first two surveys used as initial conditions	GMM estimator with IVs from earliest survey rounds	As for (5) but using NAS consumption instead of survey means
Intercept	0.234**	0.300**	0.151**	0.235**	0.175*	0.177**
	(5.183)	(5.850)	(3.705)	(4.569)	(2.772)	(3.517)
Log initial mean	-0.035**	-0.044**	-0.020**	-0.029**	-0.019	-0.014
	(-5.131)	(-5.318)	(-3.037)	(-3.264)	(-1.469)	(-1.804)
Log initial headcount index	-0.017**	-0.025**	-0.011**	-0.022**	-0.020**	-0.026**
	(-3.626)	(-4.845)	(-2.711)	(-6.305)	(-3.090)	(-4.468)
R^2	0.147	0.201	0.128	0.133	n.a.	n.a.
N	92	70	81	77	64	59

Notes: The dependent variable is the annualized change in log survey mean ($g_i(\mu_{it})$ for (1), (2), (4) and (5) and annualized change in log private consumption per capita from NAS ($g_i(C_{it})$) 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 NAS consumption (column 2). The t-ratios in parentheses are based on robust standard errors; * denotes significant at the 5% level; ** denotes significant at the 1% level.

Table 5: Alternative estimates of the elasticity of the survey mean to NAS consumption per capita

		$\hat{\beta}$	N	R^2
Levels for latest survey	$\ln \mu_{it} = \alpha + \beta \ln C_{it} + \varepsilon_{it}$	0.747 (21.463)	97	0.823
Levels for earliest survey	$\ln \mu_{it-\tau} = \alpha + \beta \ln C_{it-\tau} + \varepsilon_{it-\tau}$	0.748 (14.082)	92	0.728
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

Table 6: Encompassing regressions for consumption growth rates

	(1) Survey Means	(2) Consumption from NAS	(3) Survey Means	(4) Consumption from NAS
Intercept	0.442 (0.795)	0.617 (1.234)	0.26 (1.279)	-0.275 (-1.914)
Initial mean ($\ln \mu_{it-\tau}$ for (1) and (3) and $\ln C_{it-\tau}$ for (2) and (4))	-0.058** (-5.961)	-0.035** (-3.657)	-0.060** (-6.912)	-0.030** (-3.764)
Poverty rate ($\ln H_{it-\tau}$)	-0.027** (-5.482)	-0.017** (-3.033)	-0.027** (-5.750)	-0.014** (-3.024)
Gini index ($\ln G_{it-\tau}$)	-0.020 (-0.400)	-0.081 (-1.784)	0	0
Income share of middle three quintiles ($\ln MQ_{it-\tau}$)	-0.117 (-1.477)	-0.167* (-2.167)	-0.091** (-3.985)	0
Share of population in Western middle class ($1 - F_{it-\tau}(13)$)	-0.102* (-2.284)	-0.128** (-2.815)	-0.110** (-2.432)	-0.133** (-3.691)
Primary school enrolment rate (log)	0.007 (0.700)	0.003 (0.271)	0	0
Life expectancy (log)	0.117** (2.768)	0.154** (3.653)	0.129** (3.068)	0.139** (3.665)
Price of investment (log)	-0.014** (-2.650)	-0.016** (-3.140)	-0.014** (-2.698)	-0.017** (-3.434)
N	0.434	0.470	0.430	0.453
R^2	88	84	88	87

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 was 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 5% level; ** denotes significant at the 1% level.

Table 7: Regressions for proportionate change in poverty rate as a function of the growth rate and initial poverty rate

	(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 ($\ln H_{it-\tau}$)	-0.004	0.008	0	0	0	0
	(-0.792)	(0.267)				
Growth rate (annualized change	-2.674**	-3.564**	-2.615**	-3.323**	0	0
in log survey mean, $g_i(\mu_{it})$)	(-6.660)	(-4.325)	(-6.608)	(-4.560)		
Growth rate interacted with initial	2.780**	3.492**	2.621**	3.101**	0	0
poverty rate ($g_i(\mu_{it}).H_{it-\tau}$)	(5.206)	(3.650)	(4.915)	(3.746)		
(1-Poverty rate) *times* growth rate	0	0	0	0	-2.613**	-3.294**
($g_i(\mu_{it}).(1-H_{it-\tau})$)					(-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.

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 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 $g_i(\mu_{it}).(1-H_{it-\tau})$.