

Why Don't We See Poverty Convergence?

Martin Ravallion¹

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

Abstract: Consistently with models of economic growth incorporating borrowing constraints, the analysis of a new data set for 100 developing countries reveals an adverse effect on consumption growth of high initial poverty incidence at a given initial mean. A high incidence of poverty also entails a lower subsequent rate of progress against poverty at any given growth rate (and poor countries tend to experience less steep increases in poverty during recessions). Thus, for many poor countries, the growth advantage of starting out with a low mean ("conditional convergence") is lost due to their high poverty rates. The size of the middle class—measured by developing-country, not Western, standards—appears to be an important channel linking current poverty to subsequent growth and poverty reduction. However, high current inequality is only a handicap if it entails a high incidence of poverty relative to mean consumption.

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 and to Karla Hoff, Aart Kraay, Luis Servén and Dominique van de Walle for comments on an earlier draft.

1. Introduction

Two widely-held “stylized facts” about development are the “advantage of backwardness”—the mean convergence property whereby one sees higher economic growth in countries starting out with a low mean—and that economic growth is the main driving force of poverty reduction in developing countries. There is evidence supporting the convergence property though it is sometimes only evident if one adds controls for initial conditions, giving “conditional convergence.” There is also evidence suggesting that higher rates of growth in the mean bring higher rates of poverty reduction, such as measured by the rate of decline in the proportion of the population living in poverty (the “headcount index”). There are caveats here too, such as that the poverty line is fixed in real terms (or at least not proportional to the mean) and that the same rate of growth can yield very different outcomes for the evolution of aggregate poverty measures. (Measurement errors are also a problem in testing both stylized facts.)

It appears to have gone unnoticed in the literature that, when taken together, these stylized facts imply “poverty convergence:” a catching up process whereby the poorest countries should experience a higher rate of progress against poverty. In fact, under certain conditions, the mean and the poverty rate should have the same speed of convergence. To see this in a simple expository model, let μ_{it} denote the mean consumption or income for country i at date t and let the growth model for the mean be $\Delta \ln \mu_{it} = \alpha_i + \beta_i \ln \mu_{it-\tau} + \varepsilon_{it}$ where $\beta_i (< 0)$ is the convergence parameter.² (And α is a country-specific effect, while τ is the length of the time period and ε_{it} is a zero-mean error term.) Next let H_{it} denote the headcount index which is related to the mean as $\ln H_{it} = \delta_i + \eta_i \ln \mu_{it} + \nu_{it}$. (The elasticity of the poverty rate to the mean, $\eta < 0$, can also vary across countries.) The implied growth model for poverty is $\Delta \ln H_{it} = \alpha_i^* + \beta_i \ln H_{it-\tau} + \varepsilon_{it}^*$ (for which it is readily verified that $\alpha_i^* = \alpha_i \eta_i - \beta_i \delta_i$ and $\varepsilon_{it}^* = \varepsilon_{it} \eta_i + \nu_{it} - (1 + \beta_i) \nu_{it-\tau}$). For any given country, the speed of convergence will be the same for its mean as its poverty measure.

However, that poses a puzzle since, as this paper will show, there is no sign of overall (unconditional) poverty convergence, though there is evidence of convergence in the mean. The overall incidence of poverty is falling in the developing world, but no faster in the poorest

² To see what this means, consider two countries that start with different means, but are identical in all other respects. Then the log of the ratio of their means at date t will be $1 + \beta_i$ of its value at date $t - \tau$. Thus $1 + \beta_i$ can be thought of as the “speed of convergence.”

countries. It seems that something is missing from these two stylized facts. Intuitively, it appears that either the growth process in the mean, or the impact of growth on poverty, must depend directly on the initial poverty rate, in a way that largely nullifies the mean convergence effect.

This echoes some past ideas from the literature. The possibility of dynamic efficiency costs of high current inequality has been one such idea. But is it inequality or something else about the initial distribution that matters, such as poverty or the size of the middle class? Inequality is obviously not the same thing as poverty; inequality can be reduced without reducing poverty 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, as the paper will argue, many of the same theories that have pointed to efficiency costs of high inequality also suggest that high poverty handicaps growth.

To explore these issues empirically, 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, 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 functional form, sample selection by type of survey, and what measures are used for the key variables. A sub-sample with three or more surveys is also used to test robustness to different specification choices, including treating initial distribution as endogenous by treating lagged initial distribution as excludable.

The paper finds that mean-convergence is counteracted by two “poverty effects.” First, there is an adverse direct effect of high initial poverty on growth—working against (conditional) convergence in mean incomes; in terms of the expository model above, the parameter α_i in the growth model for the mean is itself a decreasing function of the initial poverty rate. Second, high initial poverty is dulling the impact of subsequent growth on poverty; the poor enjoy a lower share of the gains from growth in poorer countries. In terms of the above model, the elasticity of poverty to the mean, $-\eta_i$, is itself a decreasing function of the initial level of poverty. These two effects counteract the mean convergence effect, leaving little or no systematic effect of starting out poor on the future 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 a country’s mean consumption.

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.

The following section reviews the literature on distribution-dependent growth and draws out some implications of a simple theoretical model with borrowing constraints. Section 3 describes the data. Section 4 presents the paper’s results on growth and the initial distribution while section 5 focuses on how distribution influences the impact of a given rate of growth on poverty. Bringing together the paper’s main empirical results, Section 6 explains why we do not see poverty convergence. Section 7 concludes.

2. Issues begging from past theories and evidence

There is a large literature, both theoretical and empirical, on growth and distributional change. Here I focus solely on the most relevant strands of the literature, and their implications for the empirical work reported later.

2.1 Theories of distribution-dependent growth

A body of theoretical work has suggested that initial distribution matters to an economy’s aggregate efficiency and (hence) growth prospects.³ One class of models is based on the idea that high inequality restricts efficiency-enhancing cooperation, such that the public goods needed for growth are underprovided or efficiency-enhancing policy reforms are blocked (Bardhan et al., 2000). Political-economy models of redistribution—which argue that high inequality leads democratic governments to implement distortionary redistributive policies—also point to inequality as the key factor; an example is the model of Alesina and Rodrik (1994).

Theories based on credit-market failures also suggest that inequality retards growth. The market failure is typically attributed to information asymmetries, notably that lenders are imperfectly informed about borrowers. The key analytic feature of this class of 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 assumes that credit market failures leave unexploited opportunities for investment in physical and human capital. The information problem means that trades do not occur that would enhance output.

³ There are a number of useful 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).

With diminishing marginal products of capital, the mean future wealth will be a quasi-concave function of the distribution of current wealth; thus higher current inequality implies lower future mean wealth at a given value of current mean wealth, i.e., lower growth. There are various models with such features, including Galor and Zeira (1993), Benabou (1996), Aghion and Bolton (1997) and Banerjee and Duflo (2003).

There is another implication of credit market failures that has received far less attention.⁴ Lopez and Servén (2005) 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. The following section studies the model in Banerjee and Duflo 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 a lower growth at a given level of mean current wealth.

These are not the only arguments that suggest that poverty may be a relevant parameter of the initial distribution. 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 can be expected to mean a higher proportion of impatient consumers and hence lower growth.

Yet another example of models in which poverty rather than inequality matters to aggregate efficiency can be generated by considering how work productivity is affected by past nutritional status 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). 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 (Cunha and Heckman, 2007). Growing up in poverty can thus have lasting impacts on aggregate efficiency.

⁴ The argument that high current poverty retards future growth in a model with credit market failures is made intuitively in Ravallion (2001, 2007).

There are other 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 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.

These models are suggestive of an aggregate efficiency cost of a high incidence of absolute poverty, separately to inequality. But note that the theoretical prediction concerns the level of poverty at a given value of initial mean wealth. Whether initially poorer economies see a higher or lower pace of poverty reduction remains ambiguous. Two opposing effects can be identified. The first is the conditional convergence property regularly found in growth empirics and predicted by standard growth theories (Barro and Sala-i-Martin, 1992). The incidence of absolute poverty is known to be (negatively) correlated with mean income (World Bank, 1990; Ravallion, 1995, 2001). So the convergence property suggests that countries with a lower initial mean will tend to have higher initial poverty incidence and higher subsequent growth. Against this effect, the second category of theories of distribution-dependent growth suggest that the convergence effect is counterbalanced by an adverse distributional effect of higher poverty. Which effect dominates is an empirical question.

A strand of the theoretical literature has also pointed to the possibilities for multiple equilibria in nonlinear dynamic models, whereby one of the equilibria is a poverty trap (“low-level attractor”). Essentially, the nonlinearity in the recursion diagram now comes with 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.

2.2 Implications of a growth model with borrowing constraints

Possibly the simplest illustration of how poverty can handicap growth is found in models incorporating a borrowing constraint, whereby a person cannot borrow more than a fixed ratio of their current wealth. This section draws out this implication from one such model.

Banerjee and Duflo (2003) provide a simple but insightful growth model with a borrowing constraint. Someone who starts her productive life with sufficient wealth will 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, will not be able to do so. Banerjee and Duflo show that higher inequality in such an economy implies lower growth, and they discuss the implications for growth regressions using inequality as a regressor. 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 $k^*/(\lambda + 1)$ will be credit constrained in that $h'(k_t) > r$ and invest all they can, namely $(\lambda + 1)w_t$, while the rest will be free to implement their unconstrained optimum. A share $1 - \beta \in (0,1)$ of current wealth is consumed, leaving β for the next period.

Under these assumptions, the recursion diagram takes the form:

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

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

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

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

By well-known properties of concave functions, we have:

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

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

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

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

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

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:

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

Thus we also have:

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

Note that the inequality effect will still be present, separately to this poverty effect. And the less poverty there is, the less important overall inequality will be to subsequent growth prospects.

⁵ Note that the function ϕ defined by equation (1) is continuous at $k^*/(\lambda + 1)$.

⁶ If there is (unrestricted) first-order dominance, whereby $\partial w_t(p)/\partial H_t^* \leq 0$ for all $p \in [0, 1]$, then $\partial \mu_{t+1}/\partial H_t^* \leq 0$. However, first-order dominance is ruled out by the fact that the mean is held constant in this “though experiment;” there is a redistribution from the “wealth poor” to the “wealth nonpoor”.

2.3 *Past evidence on growth and the initial distribution*

The aspect of distribution that has received almost all the attention in the empirical literature is inequality, as measured by an index that penalizes larger relative disparities across all levels of income. Empirical 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) and Knowles (2005), though not all the evidence has been supportive; also see Li and Zou (1999), Barro (2000) and Forbes (2000). The main reason why these studies have been less supportive appears to be that they have allowed for additive country-level fixed effects; I will return to this point.⁷

There are a number of unresolved specification issues in this literature. The Gini index—half the mean absolute difference between all pairs of incomes normalized by the overall mean—has been (by far) the most popular inequality measure, though this 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 (2005) 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 well 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 controls. The choices made in past empirical work testing for effects of initial distribution have lacked clear justification in terms of the theories predicting such effects. Consider three of 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

⁷ Also see the discussion in Banerjee and Duflo (2003).

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.

channels identified in the theoretical literature (section 2). 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 by Beck et al. (2000, 2007). The theories discussed above based on borrowing constraints imply that the aggregate flow of credit in the economy depends on the initial distribution.

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

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 conducive 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. For example, the main way the middle class expands in a developing country is probably through poverty reduction. Then 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, is likely to be highly correlated with a relative inequality measure, clouding the interpretation. The inter-relationships with other regressors also play a role. 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*, given that one can expect (and I will demonstrate later) that there is a high partial correlation between the poverty rate and inequality holding the mean constant.

2.4 *Growth and poverty reduction*

While the bulk of the literature has focused on growth as the dependent variable, some studies have looked instead at how much impact a given rate of growth has on poverty. The

consensus is that higher growth rates tend to yield more rapid rates of absolute poverty reduction; see World Bank (1990, 2000), Ravallion (1995, 2001, 2007) and Fields (2001).¹⁰ This is consistent with one of the “stylized facts” to emerge from the literature, namely that growth in developing countries tends to be distribution-neutral on average (Ferreira and Ravallion, 2009).

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). 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). Ravallion (1997) found that the following parsimonious specification was found to fit the data very well:

$$g_i(H_{it}) = \eta(1 - G_{it-\tau})g_i(\mu_{it}) \quad (6)$$

where $\eta < 0$ is the elasticity of poverty reduction with respect to the “distribution-corrected” growth rate $(1 - G_{it-\tau})g_i(\mu_{it})$ where $G_{it-\tau} \in [0, 1]$ is the Gini index. At minimum inequality ($G_{it-\tau} = 0$) growth has its maximum effect on poverty while the elasticity reaches zero at maximum inequality ($G_{it-\tau} = 1$). Ravallion (1997) did not, however, find that the growth elasticity of poverty reduction 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

Unlike most past data sets used in the literature on growth empirics, the data set developed for this study is firmly anchored to the household surveys, including their precise dates. This choice is natural given that the focus here is 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

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

data compilations from secondary sources can be eliminated. National accounts (NAS) data are also used, matched as closely as possible to survey dates. However, in common with past literature, there is little choice but to use measures based on household consumption or income rather than the theoretically preferable concept of wealth.

The data set cannot hope to resolve all the issues from the literature on distribution-dependent growth reviewed above. In keeping with the bulk of the literature, the country is the unit of observation here, although 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 postulated by the theories reviewed above would require micro data, and even then the identification problems can be formidable. The aim here is the less ambitious one of describing the (zero and higher-order) correlations found in this new data set, to see whether they are consistent with the patterns found in past studies using country-level data, and to do so within an encompassing framework that can hopefully throw light on what parameter of the initial distribution is best able to predict subsequent growth and poverty reduction.

The data set covers the 99 developing countries with at least two nationally representative surveys over time since about 1980. (For about 70 of these countries there are three or more surveys.) When the poverty measures are used I restrict the sample to the 92 countries in which the earliest sample survey finds that at least some households lived below the poverty line (described below).¹¹ The longest spell between two surveys was used for each country. Both surveys used the same welfare indicator, either consumption or income per person, following standard measurement practices. When both were available, consumption was generally preferred, 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.¹² Three-quarters of the spells use consumption. Comparability problems between surveys remain, such as differences in recall periods and imputation/valuation methods. 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 the interval varies from three to 27 years. All changes between the surveys are annualized. National accounts

¹¹ The data set was constructed from [PovcalNet](#) in December 2008. Seven countries were dropped because the poverty rate was zero in the earliest surveys. Given that log transformations are required this makes little practical difference.

¹² The only exception was Peru, for which a much longer time period could be covered using income-based distributions.

data were mapped as closely as possible to the survey dates, interpolating as need be. All monetary measures are in constant 2005 prices (using country-specific Consumer Price Indices) and all international comparisons are at Purchasing Power Parity (PPP) using the individual consumption PPPs from the 2005 International Comparison Program (World Bank, 2008).

Various measures of distribution are constructed from each survey. Inequality is measured by the usual Gini index. Poverty is mainly measured by the headcount index, given by the proportion of the population living in households with consumption expenditure per capita (or income when consumption is not available) below \$2.00 per day at 2005, which is the median poverty line amongst developing countries.¹³ Coincidentally, in 2005, \$2 a day was also median consumption per person in the developing world, in that roughly half of the population of the developing world lived below this line. This choice 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.¹⁴ The size of the middle class is measured by the proportion of the population living in the interval \$2 to \$13 a day at 2005 purchasing power parity (PPP).¹⁵ The choice of these bounds is 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. 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, as used by Easterly (2001).

Some further notation is useful. Let $F_{it}(z)$ denote the empirical distribution function for country i at date t , giving the proportion of the population living in households with consumption or income per person not exceeding z in \$'s per day at 2005 PPP. So $H_{it} = F_{it}(2)$. The population share of the developing-world middle class is denoted $MC_{it} \equiv F_{it}(13) - F_{it}(2)$ while the share of

¹³ 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).

¹⁴ \$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](#).

¹⁵ See Ravallion (2009) for further discussion of this definition and alternatives. This definition is a variation on that proposed by Banerjee and Duflo (2008), who used the interval \$2 to \$10 a day.

the “Western middle class” is $1 - F_{it}(13)$. The income share of the middle three quintiles (ranked by consumption or income per capita) is MQ_{it} .

The survey means exhibit (unconditional) convergence; the regression coefficient of $g_i(\mu_{it})$ on $\ln \mu_{it-\tau}$ is -0.013 (t=-3.412).¹⁶ But the poverty measures do not. Indeed the proportionate rates of poverty reduction are orthogonal to initial levels; the regression coefficient of $g_i(H_{it})$ on $\ln H_{it-\tau}$ is 0.005 (t=0.542).¹⁷ Figure 1 plots the data, and gives a non-parametric regression line.

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). The signs of convergence in the survey means in Figure 1(a) are clearly stemming in part at least from this problem. It is notable, for example, that the regression coefficient of $g_i(\mu_{it})$ on $\ln \mu_{it-\tau}$ drops to -0.010 using only the consumption surveys, and the coefficient is only significantly different from zero at the 6% level (t=-1.882). Similarly, if one uses mean consumption from the NAS instead, the regression coefficient of $g_i(C_{it})$ on $\ln C_{it-\tau}$ is even lower at -0.007 and only significant at the 8% level (t=-1.743). However, the lack of any sign of poverty convergence in Figure 1(b) cannot be attributed to this problem. Indeed, one might conjecture that correcting for this type of measurement error would reveal divergence; I will return to this issue when considering the subsample with three surveys.

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% (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 over time, from a mean $MC_{it-\tau}$ of 48% to a mean MC_{it} of 53%. However, the changes in the size of the middle class (by developing-country standards) were diverse across countries, as is plain from Figure 2. 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 3, which plots

¹⁶ All t-ratios in this paper are based on White standard errors.

¹⁷ 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).

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.

There are some strong correlations amongst the distributional measures. Table 2 gives the correlation coefficients, focusing on the variables from the earliest surveys that will be used later as regressors. The correlations point to a number of potential concerns about the inferences drawn from past research on distribution-dependent growth. 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).¹⁹ 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 the initial headcount index 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 ($t=13.340$). 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 $\ln H_{it-\tau}$). These observations cast doubt on how one should interpret a regression of growth on the initial mean and inequality if in fact it is the incidence of poverty that is the relevant (omitted) variable; the inequality measure may just be picking up the poverty effect.

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 middle class is accountable to differences in the initial poverty rate. (The bimodality in terms of

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

¹⁹ So it is unclear how much Easterly's (2001) finding that MQ is a significant predictor of growth might in fact reflect inequality.

the size of the middle class in Figure 3 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} .²⁰ 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 4 plots the rate of poverty reduction ($g_i(H_{it})$) against the ordinary growth rate in the survey mean ($g_i(\mu_{it})$). The regression line in Figure 4 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.

4. The relevance of initial distribution to growth in the mean

As discussed in section 2, initial distribution can matter to the rate of poverty reduction via two distinct channels, namely the growth rate and the elasticity of poverty to the mean. This paper postulates a simple triangular model in which the rate of progress against poverty depends on the rate of growth and the initial distribution, while the rate of growth depends in turn on initial distribution. In determining what aspect of initial distribution matters, I will begin with inequality measures, then turn to poverty measures and various robustness tests, including specifications encompassing other measures of initial distribution.

A causal interpretation of these models requires that the initial distribution (in the first of the two surveys used to construct each spell) is exogenous to the subsequent pace of growth and poverty reduction. This can be questioned; for example, there may be some persistent factor in a country's governance or political economy that simultaneously yields high inequality and stifles growth prospects. While one might be unwilling to accept strict exogeneity, it is nonetheless of interest to explore whether the correlations found in the data are at least consistent with the idea that initial distribution matters to an economy's prospects for growth and poverty reduction. I

²⁰ $R^2=0.826$ 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).

shall also provide results for an instrumental variables estimator under widely-used (though still questionable) exclusion restrictions.

The following analysis will focus solely on the relationship between initial distribution and subsequent growth and poverty reduction. One might prefer to expand the regressions to include other covariates commonly found in the literature on growth empirics. However, as noted in section 2, these are arguably channels through which aspects of initial distribution, such as less poverty or a larger middle class, promote growth.

4.1 *Inequality and growth*

These data offer mixed support for the idea that inequality *per se* is harmful for growth. There is no correlation between the initial Gini index and the growth rate in the survey means ($r=-0.043$). Adding a control for the initial mean does not change this fact; for example, using a log transformation for the Gini index (which gives a slightly better fit), one obtains:

$$g_i(\mu_{it}) = 0.092 - 0.013 \ln \mu_{it-\tau} - 0.005 \ln G_{it-\tau} + \hat{\varepsilon}_{it} \quad R^2=0.068, n=92 \quad (7)$$

(1.192) (-2.825) (-0.235)

I tested three variations on these regressions. First, on re-estimating (7) using only the subsample of consumption surveys I found that the Gini index had a larger (more negative) coefficient but still not significantly different from zero at the 10% level ($t=-1.554$; $n=70$). Second, I also tested for an interaction effect between $\ln \mu_{it-\tau}$ and $\ln G_{it-\tau}$, as suggested by a model of with credit constraints that only bind for the poor, which would suggest a positive interaction effect (Banerjee and Duflo, 2003). There was evidence of an interaction effect when $\ln \mu_{it-\tau} \cdot \ln G_{it-\tau}$ was added to (7), but it was negative, and only significant at the 8% level. Third, I also estimated (7) using quintile shares, which were individually and jointly insignificant. Later we will return to inequality as a regressor in an encompassing model.

4.2 *Poverty and growth*

The data offer little support for the simplest idea of an aggregate “poverty trap” as a situation in which countries starting out with a high incidence of absolute poverty by developing country standards tend to experience little or no subsequent growth. Indeed, one finds a small positive correlation between the initial poverty measure ($\ln H_{it-\tau}$) and the subsequent rate of growth in the survey mean ($r=0.25$) although this vanishes when one uses the growth rate of

private consumption per capita from the national accounts ($r=-0.06$). If one divides the sample into countries starting out with more than half the population living below \$2 a day poverty rate ($n=42$) and those with less than half ($n=57$), the average growth rate in the mean is appreciably higher for the first group; the mean growth rate in the survey mean is 2.24% per annum for the countries starting with more than half the population poor, versus 0.77% for the rest.

Such empirical results do not, of course, rule out poverty traps. Indeed, there is clearly considerable persistence in progress against poverty, with most of those countries that started out absolutely poor staying so. What the overall correlations tell us is that high poverty countries did not as a rule see lower subsequent growth; some did and some did not.

The picture changes when one controls for the initial mean, for then a negative and significant partial correlation between initial poverty and growth emerges. The estimated model can be written as:

$$g_i(\mu_{it}) = \alpha + \beta \ln \mu_{it-\tau} + \gamma \ln H_{it-\tau} + \varepsilon_{it} \quad (8)$$

The estimates are given in column (1) of Table 3. The regression is 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; the estimate of β implies a derivative of $\ln \mu_{it}$ with respect to $\ln \mu_{it-\tau}$ that is less than unity for $\tau < 29$ years; the largest value of τ in the data is 27 years.

The estimate of (8) in column (1) of Table 3 suggests that differences in the initial poverty rate have sizeable 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 seeming contradiction between convergence and the hypothesis that poverty is an impediment to growth (as discussed in section 2) can thus be reconciled. We see both effects at work in these regressions, with higher initial poverty incidence emerging as a significant negative factor in growth at a given initial mean on top of the negative effect of the initial mean. The two opposing effects essentially cancel each other out leaving little or no (zero-order) correlation between growth and the initial incidence of poverty.

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 absolute value of the coefficient on the initial mean increases considerably when one adds the initial poverty measure as a regressor. Dropping $\ln H_{it-\tau}$ from (8) 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.

4.3 *Sensitivity to functional form, sample and measures*

This section and the next two test the robustness of the regression in equation (8) to various changes in specification.

It might be conjectured that the effect of $\ln H_{it-\tau}$ in (8) 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 (8) 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 (8) became -0.018 (t=-3.547).

There is, however, a marked nonlinearity in the relationship with the poverty rate, which is being captured by the log transformation of $H_{it-\tau}$ in (8). 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 (8) is to plot $g_i(\mu_{it}) + 0.035 \ln \mu_{it-\tau}$ (from Table 3, column (1)) against $\ln H_{it-\tau}$. Figure 5 gives the results, along with a locally-smoothed (non-parametric) regression line. The relationship is close to linear in the log poverty rate.²¹ The log transformation appears to be the right functional form.

²¹ 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.

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 (8) the poverty rate still had a negative coefficient but it was not significant at the 5% level. At the other extreme, on replacing $\ln H_{it-\tau}$ in (8) by $\ln F_{it-\tau}(13)$, the latter was (highly) insignificant controlling for the initial mean.

The results were also robust to using the poverty gap index instead of the headcount index; the corresponding version of (8) 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 used for estimating (8) used both consumption and income surveys, and that the latter are likely to contain more measurement error. Estimating the regression solely on the subsample of consumption surveys strengthened the result; analogously to (8) one obtains the estimate in column (2) of Table 3. The conditional convergence effect is even stronger as is the poverty effect.

Almost all results discussed above were also robust to using NAS consumption instead. The only notable difference is that the headcount index based on the \$1.25 line was a stronger predictor of the NAS consumption growth; column (3) Table 3 gives the regression.

4.4 *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 poverty convergence more robustly to measurement errors. One way of doing this is to calculate the trend rate of poverty reduction over the three surveys and seeing if this is correlated with the starting value. I estimated the trend for each country by regressing the log of the three headcount indices observed for that country on time. There was no significant correlation between these trends 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 to test for poverty convergence, I regressed $g_i(H_{it})$ (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 $g_i(H_{it})$ on the poverty measure from the earliest survey ($\ln H_{it-\tau_1-\tau_2}$); the result was similar, namely a regression coefficient of 0.027 with $t=1.819$. So there is some sign of divergence, but it is fairly weak.

The second use of the subsample is to use the inter-temporal averages to reduce the attenuation biases due to measurement error in the regression in (8), which can be re-estimated in the form:

$$g_i(\mu_{it}) = \alpha + \beta \ln M_i(\mu_{it-\tau_2}) + \gamma \ln M_i(H_{it-\tau_2}) + \varepsilon_{it} \quad (9)$$

Column (4) of Table 3 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 (8).

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 then 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. I also use NAS consumption as an IV for the survey mean, on the assumption that the measurement errors in these two variables are weakly correlated; however, the results were robust to dropping this IV. Letting τ_i now denote the length of spell i ($=1,2$), the model becomes:

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

The instrumental variables were $\ln \mu_{it-\tau_1-\tau_2}$, $\ln C_{it-\tau_2}$, $\ln C_{it-\tau_1-\tau_2}$, $\ln G_{it-\tau_1-\tau_2}$, $\ln F_{it-\tau_1-\tau_2}$ ($z=1.25, 2.00$) 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. Efficiency gains are possible using Generalized Methods of Moments (GMM); the GMM estimates of (10) are found in Table 3, 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.

Finally, one can use the subsample is to estimate a specification with country-fixed effects, which sweep up any confounding latent heterogeneity at country level. The main results were not 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).²² 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 given 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 4 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.²³ 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 an attenuation bias due to measurement error in the changes in growth rates.

4.5 *Encompassing regressions*

So far I have considered inequality and poverty as separate predictors of the growth rate. Simply adding the log of the initial Gini index to equation (8) does not change the result; the

²² This point is illustrated well by the Monte Carlo simulations found in Hauk and Wacziarg (2009).

²³ 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.

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 (8).

To investigate this further, I added inequality ($\ln G_{it-\tau}$), the income share of the middle three quintiles ($\ln MQ_{it-\tau}$) and the share of the Western middle class ($1 - F_{it-\tau}(13)$) to equation (8).²⁴ The population share of the developing world's middle class is not included given that its value is nearly linearly determined by the poverty rate and share of the Western middle class. Table 4 gives the encompassing regressions using both survey means and consumption per capita from the NAS.

I also tried three variations on the specification in Table 4. In the first, I excluded $MQ_{it-\tau}$, given how highly correlated it is with $\ln G_{it-\tau}$; the results were very similar. In the second, I added an interaction effect between inequality and the initial mean, as discussed above; this was highly insignificant (t-ratio of -0.063) and so were dropped. The third variation added an interaction effect between poverty and inequality, as discussed in section 2. Adding $\ln G_{it-\tau} \cdot \ln H_{it-\tau}$ to the first specification in Table 4 it had a positive coefficient (contrary to the theoretical expectation) but 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; the fit is improved using a log transformation giving:

$$g_i(\mu_{it}) = 0.154 - 0.035 \ln \mu_{it-\tau} + 0.018 \ln [F_{it-\tau}(13) / F_{it-\tau}(2)] + \hat{\varepsilon}_{it} \quad R^2=0.154, n=92 \quad (11)$$

(5.920) (-5.325) (3.863)

On adding $\ln F_{it-\tau}(2)$ ($= \ln H_{it-\tau}$) to (11) one finds that its coefficient is now positive, though only significantly different from zero at the 5% level. The negative (conditional) effect of the poverty rate in Table 3 is clearly transmitted through differences in the size of the middle class. However, the overall explanatory power of these regressions is only slightly better using the

²⁴ The share of the Western middle class was not logged given that 11 observations are lost because of zeros.

specification in (8). The estimate of equation (8) in column (1) of Table 3 and the estimate in (11) are roughly equally data consistent.

The findings that inequality and the relative income share of the middle quintiles do not matter once one controls for the initial poverty rate were robust to using the subsample with three surveys. This subsample 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 in Table 5 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 poverty not inequality that best explains 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 3 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.

5. Initial distribution 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 of the distribution is simply given by (one minus) the elasticity of the poverty rate with respect to the poverty line.²⁵ 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.²⁶ On calculating the elasticity using the initial survey at the \$2 a day poverty line, and denoting that elasticity by $\eta_{it-\tau}$, 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 account for the remaining variance in rates of poverty reduction. Those changes in distribution are clearly not correlated with rates of growth and (hence) the above regression coefficient is 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. What appear to be the most empirically relevant aspect of initial distribution for this second channel?

I find that countries starting with a larger middle-class tended to attain higher subsequent rates of poverty reduction from a given rate of growth. Regressing $g_i(H_{it})$ on both the rate of growth ($g_i(\mu_{it})$) and its interaction with the share of the population in the middle class ($MC_{it-\tau}$) one finds that the interaction effect has a significant negative coefficient:

$$g_i(H_{it}) = -0.011 + (0.043 - 0.029 MC_{it-\tau}) g_i(\mu_{it}) + \hat{\varepsilon}_{it} \quad R^2=0.539, n=91 \quad (12)$$

(-1.749) (0.221) (-4.818)

At the lower mode for $MC_{it-\tau}$ of around 15% (Figure 3), equation (12) implies a growth elasticity of -0.39 (t=-3.13) while at the upper mode, around 75%, it is -2.13 (t=-7.15).

²⁵ 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.

²⁶ On exploiting this fact in a decomposition analysis for a panel of countries (using an earlier version of the same data set used in the present paper) Kraay (2006) concludes that the bulk of the variance in rates of poverty reduction is due to differences in 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) for further discussion.

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

$$g_i(H_{it}) = -0.011 + (0.167 - 0.030 F_{it-\tau}(13) + 0.029 H_{it-\tau}) g_i(\mu_{it}) + \hat{\varepsilon}_{it} \quad R^2=0.539, n=91 \quad (13)$$

(-1.939)
(0.039)
(-0.631)
(3.638)

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 (12).

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.

The picture is clearer on inequality: there is no significant interaction effect between inequality and growth controlling for the interaction effect with either poverty or the size of the middle class.

Building on these observations, Table 6 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. Results are given for both Ordinary Least Squares (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 exists due to common measurement errors (given that the poverty measure and the mean are calculated from the same surveys).

The results in Table 6 indicate that the (absolute) growth elasticity of poverty reduction tends to be lower in countries with a higher initial poverty rate. Poorer countries tend to experience lower proportionate effects on their poverty measures from any given rate of growth. Table 6 also gives homogeneity tests, which 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. These data suggest a parsimonious specification in which the Gini index in equation (6) is replaced by the initial headcount index. Thus, Easterly's (2009) conjecture that the poverty rate is the more relevant predictor of the elasticity is confirmed. At an initial poverty rate of 10%

(about one standard deviation below the mean) the elasticity is about -3 (using the IVE) while it falls to about -0.7 at a poverty rate of 80% (about one standard deviation above the mean).

It is notable that the interaction effect with the poverty rate is stronger than that with the partial elasticity of poverty reduction ($\eta_{it-\tau}$). To test this, I estimated augmented versions of the regressions in Table 6, adding $\eta_{it-\tau}g_i(\mu_{it})$ as an extra regressor. In no case was the coefficient on $\eta_{it-\tau}g_i(\mu_{it})$ significantly different from zero, and the coefficient on $(1-H_{it-\tau})g_i(\mu_{it})$ remained significant in the OLS case, though not for the IVE, although even in that case $(1-H_{it-\tau})g_i(\mu_{it})$ was the better predictor than $\eta_{it-\tau}g_i(\mu_{it})$. (The homogeneity tests also performed well in these augmented regressions.)

I also used the subsample with three survey rounds to implement an IVE using the same instruments as for (10). 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 ($t=-3.092$).

Figure 7 plots the rate of poverty reduction against the poverty-adjusted growth rate in the survey mean (analogous to Figure 4, 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 4. 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.

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 (1) in Table 6 one obtains:

$$g_i(H_{it}) = \underset{(-1.628)}{-0.013} + \underset{(4.246)}{(2.869 H_{it-\tau} - 3.117)} I[g_i(\mu_{it})] g_i(\mu_{it}) + \underset{(3.709)}{(2.218 H_{it-\tau} - 1.984)} (1 - I[g_i(\mu_{it})]) g_i(\mu_{it}) + \hat{\varepsilon}_{it} \quad R^2=0.552, n=91 \quad (14)$$

It can be seen that 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 ($F=2.978$, $prob=0.062$).

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

Based on the above results, my empirically-preferred model takes the form:

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

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

The regressors in (15.2) are not, of course, independent; as we also saw in Section 3, countries with a higher initial mean tend to have a lower poverty rate. (Nonetheless, as we have also seen, the variance across countries in initial relative distribution entails that $\ln \mu_{it-\tau}$ and $\ln H_{it-\tau}$ are not so correlated as to be prohibit disentangling their effects.) In what follows I allow for this non-independence by assuming that $\ln H_{it-\tau}$ varies linearly as a function of $\ln \mu_{it-\tau}$ consistently with the data.

Armed with this model, let us now return to the question posed at the outset as to why we do not find poverty convergence. Define the “poverty convergence elasticity” as $\partial g_i(H_{it}) / \partial \ln H_{it-\tau}$. As was seen in section 3, this is very close to zero. Equations (15.1) and (15.2) motivate a three-way decomposition of the poverty convergence elasticity as follows (with the labels for each term in parentheses):

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

On evaluating all variables at their sample means and using the estimates in column (1) of Table 3 and column (3) from Table 6, 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 out by the combined effect of the two “poverty effects” in (16), which are roughly equal in size, giving an overall poverty convergence elasticity of 0.0055.

Naturally, different data points and parameter estimates from Table 3 and 6 give different magnitudes for the decomposition in (16), though all share the feature that the two poverty effects work in opposition to the mean convergence effect. Evaluating the decomposition at a higher initial headcount index tends to increase 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.

7. Conclusions

The stylized facts of mean convergence and the poverty-reducing power of growth ignore the important role played by initial distribution. The analysis of this paper confirms the intuition that something about the initial distribution is offsetting the “growth advantage of backwardness,” such that we do not observe poverty convergence. That something turns out to be poverty itself. The paper finds that there are three distinct effects of being a poor country on subsequent progress against poverty. The usual conditional convergence effect entails that countries with a lower initial mean, and so (typically) a higher initial poverty rate, grow faster and (hence) enjoy faster poverty reduction. Against this, there is evidence of an adverse direct effect of poverty on growth, such that countries with a higher initial incidence of poverty tend to experience a lower subsequent rate of growth, controlling for initial mean consumption. Additionally a high poverty rate makes it harder to achieve a given proportionate impact on poverty through economic growth. (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 incidence of poverty and the subsequent rate of progress against poverty.

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 the overall level of initial inequality, as measured by the Gini index, matters to 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 strikingly clear if one calculates the predicted rate of poverty reduction for each country, given its initial conditions (combining the paper’s results on the distributional determinants of growth with its results on how much impact

a given rate of growth has on poverty).²⁷ 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 indices in %, so it is plain 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).

Thus these findings confirm that initial inequality matters to subsequent progress against poverty, but they 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, would bring 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 growth and 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. Nonetheless, the cross-country empirical relationships reported here do at least point to the importance of better understanding these deeper handicaps faced by poor countries in their efforts to become less poor.

²⁷ For the following calculation I used specification (1) in Table 3, though other estimates of equation (8) give similar results; I then substituted this equation into the regressions for poverty reduction with homogeneity imposed in Table 6 (rankings are identical whether one uses the OLS or IVE results).

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.
- 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

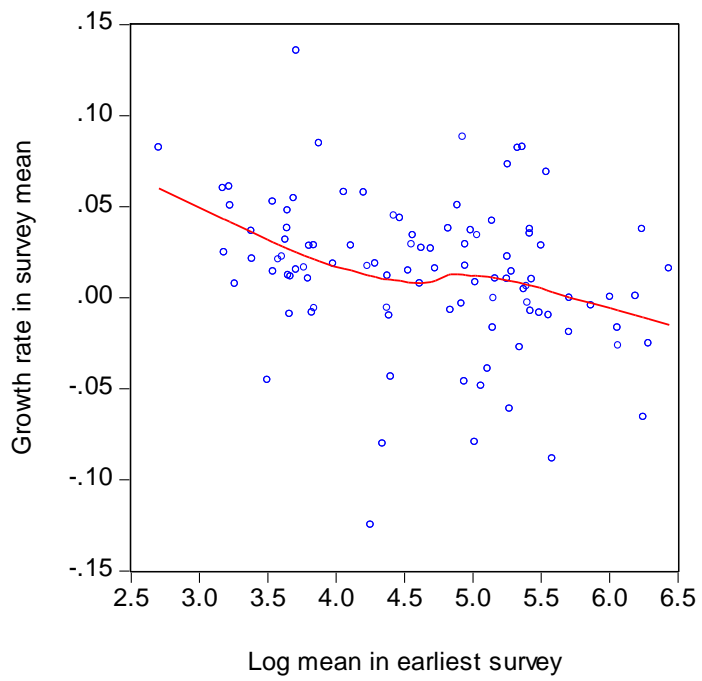
- Growth,” *Journal of Development Economics* 57(2): 259-87.
- Doepke, Matthias and Fabrizio Zilibotti, 2005, “Social Class and the Spirit of Capitalism,” *Journal of the European Economic Association* 3(2-3): 516-524.
- Easterly, William, 2001, “The Middle Class Consensus and Economic Development,” *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*, forthcoming.
- 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, 2005, “Too Poor to Grow,” mimeo, Development Research Group, World Bank.
- _____ and _____, 2006, “A Normal Relationship? Poverty, Growth and Inequality,” Policy Research Working Paper 3814, World Bank.
- 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 an 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*, forthcoming..
- 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.
- 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: Growth rates plotted against initial values

1(a): Survey means



1(b): Headcount indices of poverty

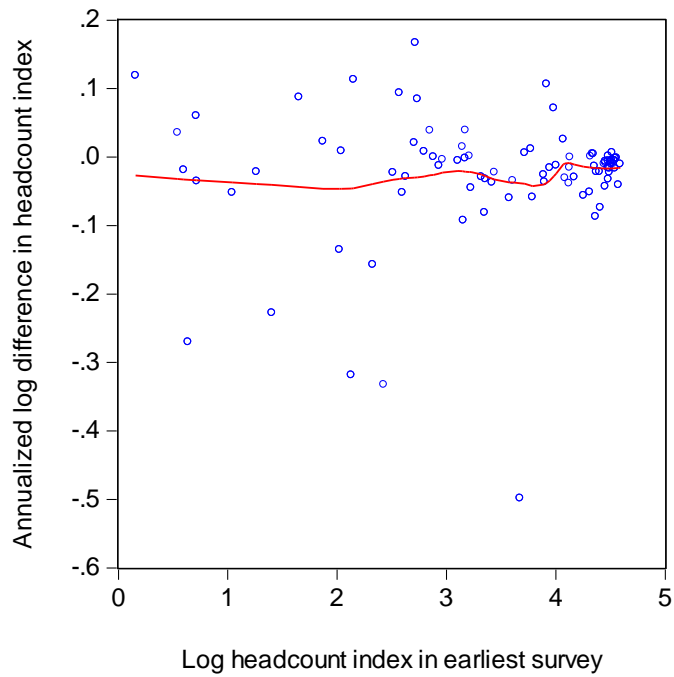


Figure 2: Developing world's middle class across two surveys rounds

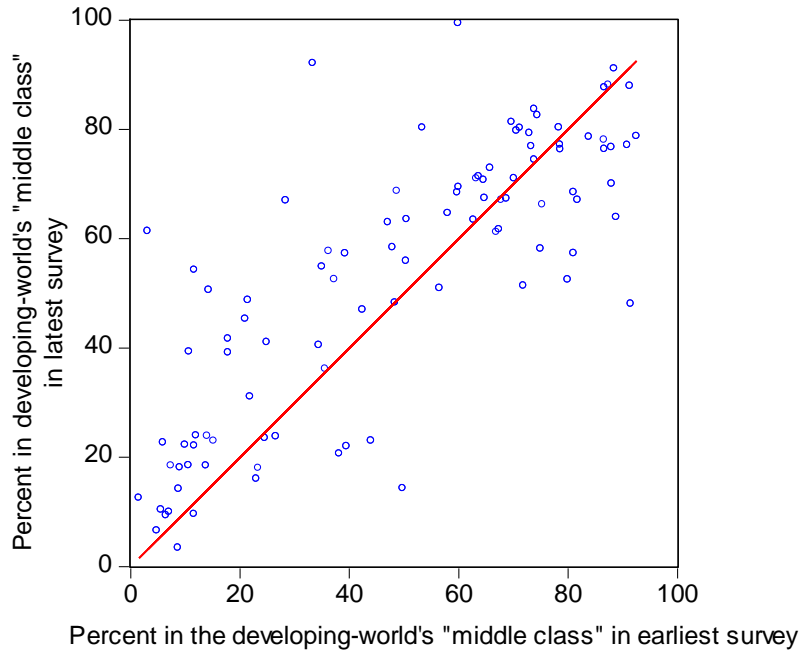


Figure 3: Densities of middle-class population shares

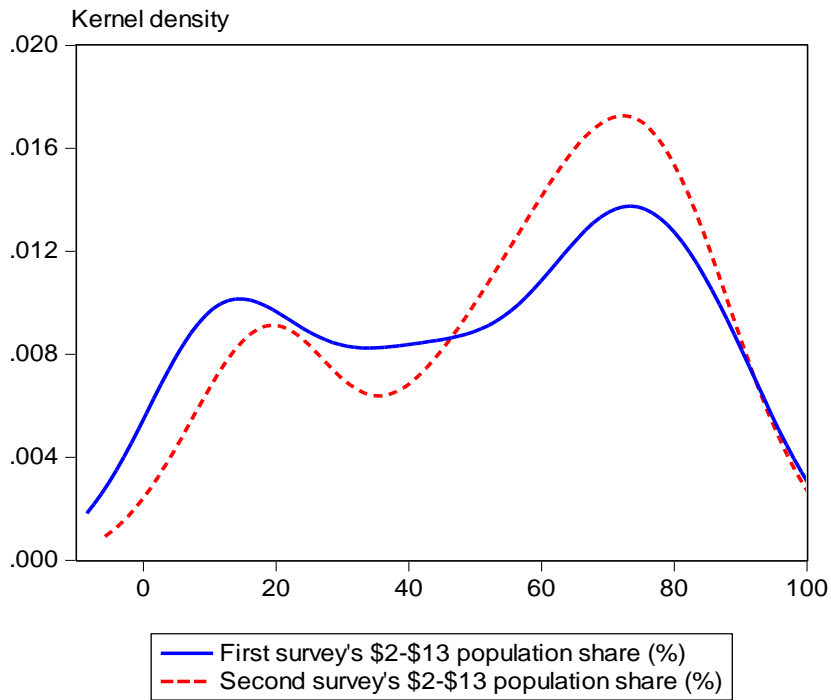


Figure 4: Rate of poverty reduction plotted against rate of growth in survey mean

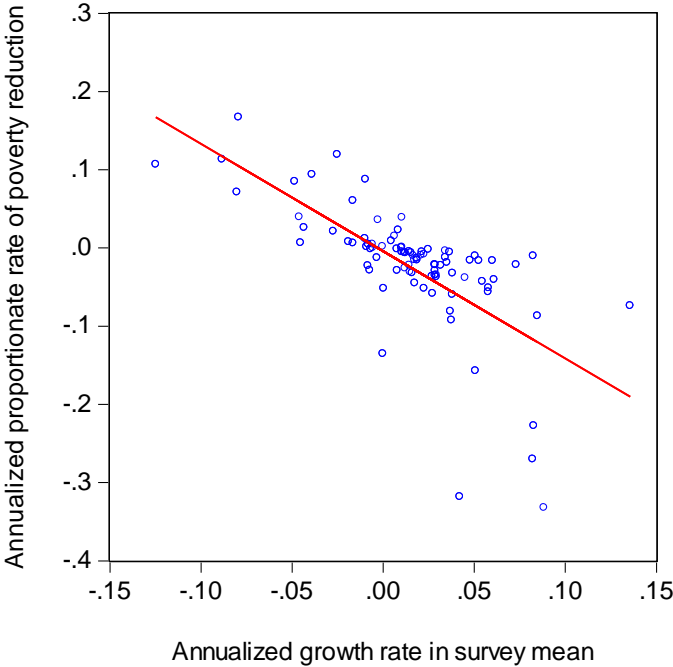


Figure 5: Growth rate with a control for the initial mean plotted against the initial poverty rate

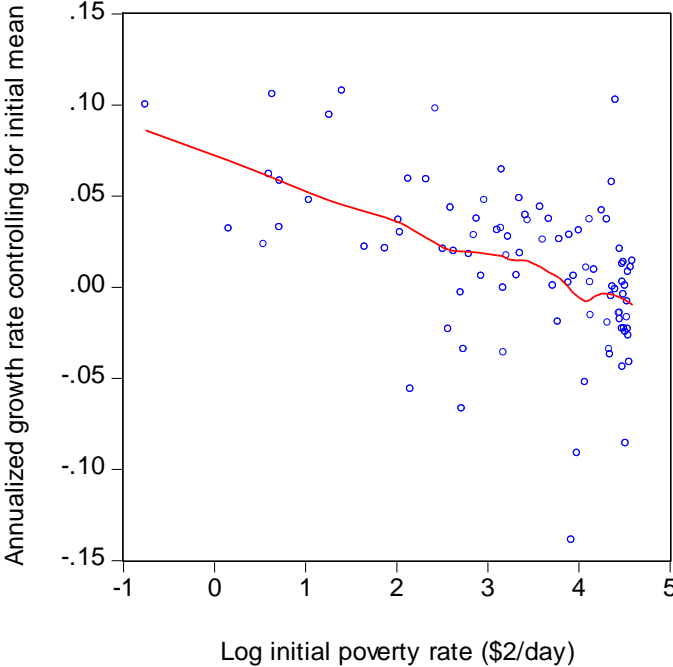


Figure 6: Rate of poverty reduction plotted against distribution-corrected rate of growth

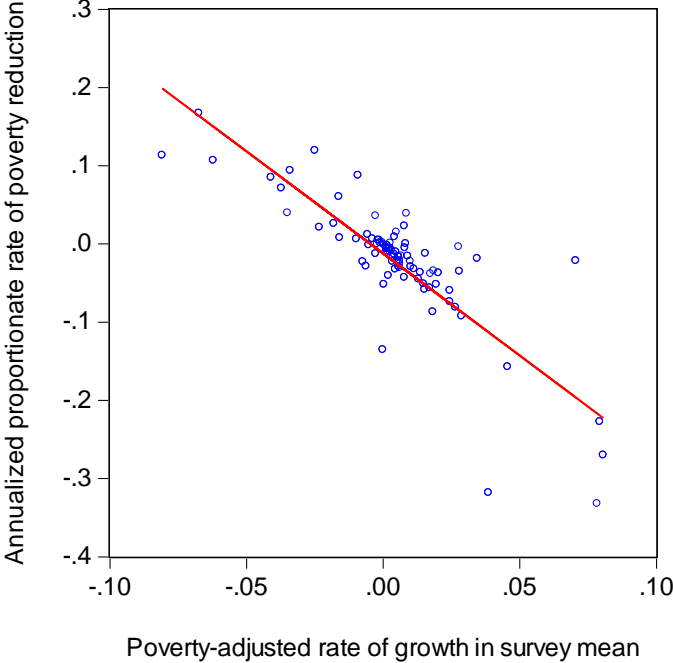


Table 1: Summary statistics

		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: 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	0.024** (5.183)	0.300** (5.850)	0.151** (3.705)	0.235** (4.569)	0.180** (2.772)	0.169** (3.517)
Log initial mean	-0.035** (-5.131)	-0.044** (-5.318)	-0.020** (-3.037)	-0.029** (-3.264)	-0.020* (-1.994)	-0.014* (-2.017)
Log initial headcount index	-0.017** (-3.626)	-0.025** (-4.845)	-0.011** (-2.711)	-0.022** (-6.305)	-0.020** (-3.381)	-0.022** (-4.749)
R ²	0.147	0.201	0.128	0.133	n.a.	n.a.
N	92	70	81	77	64	58

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 4: 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 5: Encompassing regressions for consumption growth rates as a function of initial distribution

	(1) Survey Means	(2) Consumption from NAS	(3) Survey Means	(4) Consumption from NAS
Intercept	0.576 (0.919)	1.069 (2.053)	0.220** (4.988)	0.142** (3.423)
Initial mean ($\ln \mu_{it-\tau}$ for (1) and (3) and $\ln C_{it-\tau}$ for (2) and (4))	-0.036** (-3.813)	-0.014 (-1.765)	-0.030** (-4.054)	-0.015* (-2.077)
Initial poverty rate ($\ln F_{it-\tau}^*(z)$)	-0.028** (-3.643)	-0.016* (-3.012)	-0.019** (-4.030)	-0.013** (-3.468)
Initial Gini index ($\ln G_{it-\tau}$)	0.003 (0.037)	-0.083 (-1.551)	0	0
Income share of middle three quintiles ($\ln MQ_{it-\tau}$)	-0.080 (-0.788)	-0.162 (-1.866)	0	0
Share of population in Western middle class ($1 - F_{it-\tau}(13)$)	-0.158** (-2.775)	-0.176** (-2.405)	-0.099* (-2.012)	-0.168** (-2.950)
N	89	78	92	81
R ²	0.211	0.192	0.164	0.171

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 poverty rate is \$2.00 for survey means and \$1.25 for NAS consumption. The t-ratios in parentheses are based on robust standard errors; * denotes significant at the 5% level; ** denotes significant at the 1% level.

Table 6: Regressions for proportionate change in poverty rate as a function of the growth rate and initial poverty rate

	(1)	(2)	(3)	(4)
	OLS	IVE	OLS	IVE
Intercept	-0.012 (-1.908)	-0.005 (0.607)	-0.012** (-2.175)	-0.008 (-1.365)
Growth rate (annualized change in log survey mean, $g_i(\mu_{it})$)	-2.615** (-6.608)	-3.323** (-4.560)	0	0
Growth rate interacted with initial poverty rate ($g_i(\mu_{it}).H_{it-\tau}$)	2.621** (4.915)	3.101 (3.746)	0	0
(1-Poverty rate) times growth rate ($g_i(\mu_{it}).(1-H_{it-\tau})$)	0	0	-2.613 (-7.273)	-3.294 (-4.585)
N	91	86	91	86
R ²	0.535	0.458	0.535	0.466
Homogeneity test	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 $g_i(\mu_{it}).(1-H_{it-\tau})$.