Poverty Accounting

A fractional response approach to the decomposition and prediction of poverty rates[*]

Richard Bluhm[†] Denis de Crombrugghe[‡] Adam Szirmai[§]

This version: 31 July 2017

Revised Manuscript

Abstract

This paper proposes a new empirical framework for poverty accounting consistent with the fractional nature of poverty rates. Using a large collection of household surveys from 124 countries, we estimate the impact of income growth and redistribution on poverty rates, as well as their respective historical contributions to poverty reduction. Both the $2 and $1.25 a day poverty lines are considered over different time periods. We find that initial inequality is a strong moderator of the impact of growth and there has been a shift towards more pro-poor growth around the turn of the millennium. We project poverty rates until 2030 and show that the end of extreme poverty is unlikely to materialize within a generation.

Keywords: poverty, inequality, growth, fractional response models

JEL Classification: I32, C25, O10, O15

[*]This paper and previous versions have been presented at the World Bank, ECINEQ, Agence Française de Développement, Center for Global Development, IARIW, UGöttingen, OECD, a workshop on panel data methods (UMainz) and several workshops in Paris and Maastricht. We gratefully acknowledge financial support from the Agence Française de Développement (AFD). All remaining errors are unfortunately ours.

[†]*Corresponding author.* Leibniz University Hannover, Institute of Macroeconomics, and UNU-MERIT, Maastricht University; e-mail: *bluhm@mak.uni-hannover.de*

[‡]Maastricht University; e-mail: *d.decrombrugghe@maastrichtuniversity.nl*

[§]UNU-MERIT, Maastricht University; e-mail: *szirmai@merit.unu.edu*

1 Introduction

The capacity of economic growth to eradicate poverty is at the heart of the ongoing debate about inclusive growth and equitable development. Okun's equality-efficiency trade-off, which for decades seemed to explain the sacrifice of equity in favor of 'efficiency', has recently lost ground to a renewed focus on 'pro-poor growth' (World Bank, 2005) and 'shared prosperity' (World Bank, 2015). The shift in the policy discussion is underscored by an increasingly large empirical literature that analyzes the impact of changes in incomes and inequality on poverty, and their respective contributions towards poverty reduction.[1] Collectively, these studies established not only that income growth is crucial to achieving sustained decreases in poverty, but also that the benefits of income growth strongly depend on initial levels of income *and* inequality.[2]

In this paper we present a new unified framework for a set of empirical exercises we refer to as 'poverty accounting'. Analogous with growth and development accounting, poverty accounting is the decomposition of (changes in or levels of) poverty into its proximate sources. It is concerned with answering several important questions such as: What is the impact of a one percent change in income growth or in inequality on poverty rates? How much of the historical variation in poverty is due to economic growth, how much to redistribution? Just as in the decomposition of growth, there is uncertainty concerning the correct functional form of the underlying relationship. Yet, there is no ambiguity about the fact that the poverty rate in any country or region is entirely determined by the average income level and the distribution of incomes.

The key insight our analysis builds on is that the poverty headcount ratio is a fraction. This fact alone allows us to derive a natural model of the expected poverty rate, $\mathrm{E}[H]$, given a fixed (absolute) poverty line. The model incorporates two crucial features: first, $\mathrm{E}[H]$ is bounded on the unit interval; and second, $\mathrm{E}[H]$ converges to unity (zero) if mean income becomes arbitrarily small (large) relative to the poverty line. It follows that elasticities or semi-elasticities of poverty with respect to income or inequality must be non-linear. So far, the literature has tried to address this inherent non-linearity within log-linear models, leading to *ad hoc* specifications which are poor approximations and tend to produce unstable or implausible estimates for anything but overall mean effects. To policy makers though, cross-country averages of elasticities are of limited interest. Even within countries, evidence points towards substantially different impacts of growth on poverty across regions or ethnic groups (Aaron, 1967, Hoover et al., 2008).

Our principal contribution is to propose a fractional response approach (à la Papke and Wooldridge, 1996, 2008) to the decomposition and prediction of poverty rates. This

[1]See, for example, Ravallion and Chen (1997), Dollar and Kraay (2002), Besley and Burgess (2003), Kraay (2006), Kalwij and Verschoor (2007), and Dollar et al. (2016).

[2]In fact, this dependence arises mechanically, since poverty is functionally linked to average incomes and inequality (Datt and Ravallion, 1992, Kakwani, 1993, Bourguignon, 2003).

approach dispenses with the constant or linear elasticities assumed by much of the cross-country literature. Most importantly, it allows us to *(i)* estimate credible elasticities and semi-elasticities of poverty with respect to mean income and inequality over the entire range of the data, *(ii)* recover the conditional expectation function of the poverty headcount ratio, *(iii)* estimate the counterfactual quantities needed for computing the historical contributions of growth and redistribution to poverty reduction, and *(iv)* perform out-of-sample forecasting – all within a *single framework*.

In practice, estimating any poverty decomposition using summary household survey data entails a number of difficulties. The second contribution of this paper consists in adapting and extending the fractional response model to simultaneously deal with three problems encountered in our application: unobserved heterogeneity due to persistent measurement differences between surveys, endogeneity due to time-varying measurement errors in incomes, and unbalanced panel data due to infrequently undertaken surveys.

Our third contribution is to apply this new framework to a large data set covering 124 countries over a 30-year period. In their most basic form, the data only contain three variables per country-survey-year: the poverty headcount ratio at a fixed international poverty line, average income or consumption expenditures, and the Gini coefficient.[3] We show that our approach closely approximates the shape of the Lorenz curve near the poverty line using this limited information set.

Our main findings are as follows. A one percent increase in mean income or consumption expenditures reduces the proportion of people living below the poverty line by about two percentage points; a one percent increase in inequality raises the poverty rate by one and a half percentage points. We provide differentiated regional and temporal estimates, often at odds with earlier studies. Regarding historical contributions, we provide evidence that there has been a shift in the poverty reducing pattern of growth around the turn of the millennium. Before 2000 about 90% of poverty reduction was due to income growth, and inequality tended to *rise* with higher growth rates. Since 2000, changes in inequality are responsible for almost a third of all poverty reduction. Finally, we present projections of the poverty headcount ratio for \$2 a day and \$1.25 a day poverty lines (in 2005 PPPs) until 2030. The \$2 a day poverty rate may halve from about 40% in 2010 to below 20% in 2030. This implies another billion people could be lifted out of poverty. The bad news is that the pace of poverty reduction at \$1.25 a day is bound to slow down considerably in the near future.

Let us point out at the outset what this paper is *not*. Just as growth accounting leaves open the question what ultimately drives growth, poverty accounting will not identify the

[3] With micro data, it would be possible to directly estimate the Lorenz curve and calculate the relevant counterfactual quantities to estimate the contributions (see e.g. Datt and Ravallion, 1992). Hence, a viable alternative to our method is to apply the micro-level approach to all countries by creating synthetic data as in Kraay (2006), even though Lorenz curve estimation based on grouped data has its own problems (see Chotikapanich et al., 2007, Bresson, 2009, Krause, 2014).

ultimate determinants of poverty. The causal effect of any specific policy on aggregate household income and its distribution cannot be discerned from the decomposition of poverty into its proximate sources of change. What the framework does reveal is how *given* changes in aggregate income and distribution translate into poverty outcomes.

The remainder of the paper is organized as follows. Section 2 reviews how the existing literature decomposes poverty rates and estimates poverty elasticities. Section 3 explains our approach and discusses the econometrics of fractional response models. Section 4 briefly outlines the data used in this paper. Section 5 presents the estimation results, elasticities, contributions, and poverty projections until 2030. Section 6 concludes.

2 Poverty decompositions and elasticities

With micro-level data it is straightforward to decompose changes in poverty into changes in the average level and the distribution of income (Datt and Ravallion, 1992, Kakwani, 1993). A key problem for cross-country studies of poverty is that we typically do not have access to micro-data of incomes or consumption expenditures but have to estimate poverty using only grouped data. To overcome this limitation, Bourguignon (2003) suggests approximating the entire income distribution of each country using a two-parameter log-normal distribution – an approach that is theoretically grounded, simple and popular but not without its critics (e.g. Bresson, 2009).

Bourguignon (2003) assumes that income, y_t, is a log-normal random variable, such that $\ln y_t \sim \mathcal{N}(\mu_t, \sigma_t^2)$, and mean income can be written as $\bar{y}_t = \mathrm{E}[y_t] = \exp(\mu_t + \sigma_t^2/2)$. Then the poverty headcount ratio at time t may be defined as

$$H_t = H(\bar{y}_t/z, \sigma_t) = \Phi\left(\frac{-\ln(\bar{y}_t/z)}{\sigma_t} + \frac{1}{2}\sigma_t\right) \equiv \Pr\{y_t \leq z\} \tag{1}$$

where $\Phi(\cdot)$ denotes the standard normal cdf, inequality is measured as the standard deviation (σ_t) of log-income, and $\bar{y}_t/z$ is the (relative) distance of mean income ($\bar{y}_t$) to the fixed poverty line (z). We may speak of a 'shortfall' of income when $\bar{y}_t < z$ and of 'affluence' when $\bar{y}_t \geq z$.

Eq. (1) represents the probability that, at a particular time t, an individual randomly drawn from the population is poor. This formulation gave rise to a large literature deriving the income and inequality elasticities of poverty and estimating econometric models inspired by the analytic formulas (e.g. Bourguignon, 2003, Kalwij and Verschoor, 2007, Klasen and Misselhorn, 2008). To summarize, from eq. (1) we can derive the income elasticity ($\varepsilon_t^{H\bar{y}} = \frac{\partial H_t}{\partial \bar{y}_t}\frac{\bar{y}_t}{H_t}$) and inequality elasticity ($\varepsilon_t^{H\sigma} = \frac{\partial H_t}{\partial \sigma_t}\frac{\sigma_t}{H_t}$) of poverty as

$$\varepsilon_t^{H\bar{y}} = -\frac{1}{\sigma_t}\lambda\left(\frac{-\ln(\bar{y}_t/z)}{\sigma_t} + \frac{1}{2}\sigma_t\right) \tag{2}$$

and

$$\varepsilon_t^{H\sigma} = \left(\frac{\ln(\bar{y}_t/z)}{\sigma_t} + \frac{1}{2}\sigma_t\right) \lambda \left(\frac{-\ln(\bar{y}_t/z)}{\sigma_t} + \frac{1}{2}\sigma_t\right) \tag{3}$$

where we define the inverse Mills ratio ($\lambda(x) \equiv \phi(x)/\Phi(x)$) as the ratio of the standard normal pdf to the standard normal cdf, and we require $H_t > 0$.

The decomposition of the poverty rate is often written in proportional terms:

$$\frac{\mathrm{d}H_t}{H_t} \approx \varepsilon_t^{H\bar{y}} \frac{\mathrm{d}\bar{y}_t}{\bar{y}_t} + \varepsilon_t^{H\sigma} \frac{\mathrm{d}\sigma_t}{\sigma_t} \tag{4}$$

where $\mathrm{d}H_t/H_t$ is a small relative change in the poverty headcount, $\mathrm{d}\bar{y}_t/\bar{y}_t$ is a small relative change in mean incomes, and $\mathrm{d}\sigma_t/\sigma_t$ is a small relative change in the standard deviation of log-incomes. This approximation follows from a Taylor linearization of H_t and is derived in Appendix A.

Under log-normality, the standard deviation is a monotone transformation of the Gini inequality coefficient, denoted G_t, and can be obtained via $\sigma_t = \sqrt{2}\Phi^{-1}(G_t/2 + 1/2)$. So eqs. (2) and (3) can be used to predict the elasticities directly using observed values of income and the Gini coefficient. With a little algebra, we can also derive an expression for the Gini elasticity and rewrite eq. (4) accordingly – see eq. (A-5) in Appendix A.

However, the assumption of log-normality is only an approximation and unlikely to hold exactly. The key observation motivating the econometric models is that both elasticities depend only on the initial levels of mean income and inequality (keeping the poverty line fixed). The literature captures the dependence on initial levels by interacting both mean income and inequality with the ratio of initial mean income to the poverty line and with initial inequality. This model is sometimes called the 'improved standard model' (Bourguignon, 2003) and it is usually formulated in (annualized) differences:

$$\begin{aligned} \Delta \ln H_{it} = \;& \alpha + \beta_1 \Delta \ln \bar{y}_{it} + \beta_2 \Delta \ln \bar{y}_{it} \times \ln(\bar{y}_{i,t-1}/z) + \beta_3 \Delta \ln \bar{y}_{it} \times \ln G_{i,t-1} \\ & + \gamma_1 \Delta \ln G_{it} + \gamma_2 \Delta \ln G_{it} \times \ln(\bar{y}_{i,t-1}/z) + \gamma_3 \Delta \ln G_{it} \times \ln G_{i,t-1} + \epsilon_{it} \end{aligned} \tag{5}$$

where Δ is the first-difference operator, α is a linear time trend and ϵ_{it} is an error term; we also inserted the country subscript i.

Specifications like eq. (5) allow for linear variation in the elasticities through the interaction terms. Suppose eq. (5) is estimated via Ordinary Least Squares (OLS), Instrumental Variables (IV) or the Generalized Methods of Moments (GMM). The estimates, $\hat{\varepsilon}_{it}^{H\bar{y}} = \hat{\beta}_1 + \hat{\beta}_2 \ln(\bar{y}_{i,t-1}/z) + \hat{\beta}_3 \ln G_{i,t-1}$ and $\hat{\varepsilon}_{it}^{HG} = \hat{\gamma}_1 + \hat{\gamma}_2 \ln(\bar{y}_{i,t-1}/z) + \hat{\gamma}_3 \ln G_{i,t-1}$, are sometimes referred to as the 'distribution-neutral' income elasticity and the 'growth-neutral' inequality elasticity, since they identify the partial effect of changing only one variable. They are used as linear approximations to eqs. (2) and (3), which has the virtue of simplicity but is coarse, ignoring meaningful restrictions on the parameter space. Actually, in view of the bounded nature of the dependent variable ($H_{it} \in [0, 1]$), the shape of the elasticities (or semi-elasticities) is predictably non-linear. Any linear approximation is likely to be poor away from the center, and can take on implausible values (e.g., $\hat{\varepsilon}_{it}^{H\bar{y}} > 0$) for certain combinations of income and inequality. Yet for the estimates to be policy-relevant, we are precisely interested in temporally and/or regionally differentiated elasticities and not just cross-country averages.

Specifications like eq. (5) have a number of additional disadvantages. First of all, eq. (5) completely disregards any information contained in poverty *levels*. Second, it most likely induces negative serial correlation, and compounds pre-existing measurement error.[4] Third, discarding poverty spells ending or starting with a zero, or 'winsorizing' them by slightly incrementing zero poverty rates, may cause inconsistency. More subtly, it is unclear which level relationship eq. (5) might derive from. Whereas differencing removes time-constant unobserved effects, the inserted interaction terms reintroduce the unobserved effects through the lagged levels. As a result, if there are systematic measurement differences between countries, estimated coefficients will be biased.[5]

Finally, the log transformation on the left-hand side of eq. (5) makes it impossible to recover the conditional expectation of the poverty rate without imposing strong independence assumptions. This means that eq. (5) is not generally suitable for the purpose of estimating the contributions of growth and redistribution to poverty reduction, even though in theory this should be straightforward. Moreover, as Santos Silva and Tenreyro (2006) pointed out for the case of gravity equations, heteroskedasticity in the level equation also introduces bias in the parameter estimates after taking logs.

Poverty elasticities can paint a distorted picture of poverty dynamics. The income elasticity, for example, gives the impression that richer countries become ever better at poverty reduction because a drop in the poverty headcount from 2% to 1% is treated just the same as a drop from 50% to 25%. Klasen and Misselhorn (2008) suggest to focus on absolute rather than relative changes in poverty. Removing the log from the headcount in eq. (5) achieves this. The coefficients and partial effects become semi-elasticities, measuring *percentage point* changes in the number of poor expected for a given percentage change in income or inequality. Likewise, eqs. (2) and (3) provide semi-elasticities if we replace the inverse Mills ratio with the standard normal pdf. In contrast with elasticities, the semi-elasticities converge to zero as mean income becomes large.

[4] Let the original data generating process be represented by $y_{it} = \alpha + \beta(x_{it} + \nu_{it}) + \epsilon_{it}$, where ν_{it} is uncorrelated with x_{it}, and ϵ_{it} is serially uncorrelated. The first-difference transformation induces negative serial correlation, since $\mathrm{E}[\Delta\epsilon_{it}\Delta\epsilon_{i,t-1}] = \mathrm{E}[-\epsilon_{i,t-1}^2] = -\sigma_\epsilon^2$; furthermore, any attenuation bias is magnified: $\mathrm{plim}(\hat{\beta}_{FD}) = \beta\sigma_{\Delta x}^2/(\sigma_{\Delta x}^2 + \sigma_{\Delta\nu}^2)$ where typically $\sigma_{\Delta\nu}^2 = 2\sigma_\nu^2$ but $\sigma_{\Delta x}^2 < 2\sigma_x^2$. Autocorrelation in the mismeasured variables further reduces the signal-to-noise ratio and increases the attenuation bias.

[5] Kalwij and Verschoor (2007) present a model in differences with unobserved effects removed, but later estimate interaction models with unobserved effects again present though not accounted for.

3 Fractional response models of poverty

3.1 Fractional probit

We now outline a new approach to poverty accounting that does not suffer from the above mentioned problems. In a seminal paper, Papke and Wooldridge (1996) suggested modeling proportions using non-linear, parametric, fractional response models, known since as fractional logit, fractional probit and the like. To the best of our knowledge, such an approach has never been applied to poverty decompositions. The models are of the form $\mathrm{E}[y_i|\mathbf{x}_i] = F(\mathbf{x}_i'\boldsymbol{\beta})$, where $y_i \in [0,1]$ and $F(\cdot)$ is an invertible cdf, usually standard normal or logistic. Applying this idea to our problem, we may approximate eq. (1) with

$$\mathrm{E}[H_{it}|\bar{y}_{it}, G_{it}] = \Phi(\alpha + \beta \ln \bar{y}_{it} + \gamma \ln G_{it}) \quad \text{for} \quad i = 1, \ldots, N;\ t = 1, \ldots, T \tag{6}$$

where $\ln z$ is absorbed into the constant and both explanatory variables are expressed in logs to linearize the index and ease interpretation.[6] Naturally, we expect $\beta < 0$ and $\gamma > 0$. We temporarily ignore econometric complications such as unobserved heterogeneity, endogeneity and unbalanced data; they will be tackled in the next subsection.

The inverse of $F(\cdot)$ can serve as a 'link function' in the spirit of the Generalized Linear Model (GLM) literature. Thus, eq. (6) may be written as $\Phi^{-1}(\mathrm{E}[H_{it}|\bar{y}_{it}, G_{it}]) = \alpha + \beta \ln \bar{y}_{it} + \gamma \ln G_{it}$. Figures C-1 and C-2 in Appendix C plot $\Phi^{-1}(H_{it})$ against $\ln \bar{y}_{it}$ and $\ln G_{it}$ for each region, including a regression line. The result is striking: the inverse normal cdf successfully linearizes both relationships. Rather than inverting the cdf, a better way to estimate eq. (6) is Quasi Maximum Likelihood (QML) based on a Bernoulli likelihood. QML is as robust as non-linear least squares but potentially more efficient; it only requires correct specification of the conditional mean, no matter the true distribution of H_{it} (Gourieroux et al., 1984).

It is now straightforward to define the estimated income and Gini elasticities as

$$\hat{\varepsilon}_{it}^{H\bar{y}} = \frac{\partial \hat{\mathrm{E}}[H_{it}|\bar{y}_{it}, G_{it}]}{\partial \bar{y}_{it}} \times \frac{\bar{y}_{it}}{\hat{\mathrm{E}}[H_{it}|\bar{y}_{it}, G_{it}]} = \hat{\beta} \times \lambda\left(\hat{\alpha} + \hat{\beta} \ln \bar{y}_{it} + \hat{\gamma} \ln G_{it}\right) \tag{7}$$

and

$$\hat{\varepsilon}_{it}^{HG} = \frac{\partial \hat{\mathrm{E}}[H_{it}|\bar{y}_{it}, G_{it}]}{\partial G_{it}} \times \frac{G_{it}}{\hat{\mathrm{E}}[H_{it}|\bar{y}_{it}, G_{it}]} = \hat{\gamma} \times \lambda\left(\hat{\alpha} + \hat{\beta} \ln \bar{y}_{it} + \hat{\gamma} \ln G_{it}\right). \tag{8}$$

Contrary to log-linear approximations, eqs. (7) and (8) closely mimic the structure and properties of the analytical elasticities derived from the log-normality assumption in

[6] Simulations analogous to those presented in Table B-1 of Appendix B suggest that using logs greatly improves the ability of the model to accurately recover known (semi-)elasticities.

eqs. (2) and (3). They will approach zero whenever the inverse Mills ratio does and remain plausible even for extreme values of the covariates. The non-linearity arises from the bounded nature of the headcount and not from interaction terms. This secures a number of advantages: the information contained in poverty levels is not wasted, the model will predict poverty headcount ratios strictly within the unit interval, and elasticities as well as semi-elasticities can be estimated consistently within the same model. Using $\hat{\mathrm{E}}[H_{it}|\bar{y}_{it}, G_{it}]$ together with counterfactual values for $\bar{y}_{it}$ or G_{it}, we can even estimate the respective contributions of the two proximate determinants of poverty. Finally, given growth and inequality scenarios, we may predict the future path of poverty rates.

Note that we *do not require log-normality*. Any two-parameter distribution will do as long as the poverty headcount remains a smooth function of mean income and inequality up to statistical error. In fact, it is possible to nest the log-normal case within the GLM.[7]

3.2 Heterogeneity, endogeneity, and unbalanced data

We now turn our attention to three problems we will encounter in our application: unobserved heterogeneity, measurement error, and unbalanced data. For the purpose and duration of this subsection, we stack the *time-varying* covariates and their coefficients in the vectors $\mathbf{x}_{it} = (x_{it,1}, \ldots, x_{it,K})'$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K)'$, and form the matrices $\mathbf{X}_i = (\mathbf{x}_{i1}, \ldots, \mathbf{x}_{iT})'$. The ideal model we would like to estimate is

$$\mathrm{E}[H_{it}|\mathbf{X}_i, \mu_i] = \mathrm{E}[H_{it}|\mathbf{x}_{it}, \mu_i] = \Phi(\mathbf{x}_{it}'\boldsymbol{\beta} + \mu_i) \quad \text{for} \quad i = 1, \ldots, N;\ t = 1, \ldots, T \tag{9}$$

where the covariates are strictly exogenous conditionally on unobserved country-level effects μ_i, and the panel is balanced. The unobserved effects are meant to capture time-persistent differences in measurement or deviations from a two-parameter distribution, which may be arbitrarily correlated with the elements in $\mathbf{X}_i$. The problem, in a context where T is kept small while $N \to \infty$, is that the μ_i are incidental parameters; they cannot be estimated consistently and this affects the estimation of $\boldsymbol{\beta}$ as well as all partial effects.[8]

As a solution, Papke and Wooldridge (2008) propose a *correlated random effects* (CRE) structure, using the Mundlak (1978) and Chamberlain (1984) device: $\mu_i|(\mathbf{x}_{i1}, \ldots, \mathbf{x}_{iT}) \sim \mathcal{N}(\varphi + \bar{\mathbf{x}}_i'\boldsymbol{\theta}, \sigma_u^2)$ where $\bar{\mathbf{x}}_i = T^{-1}\sum_{t=1}^{T}\mathbf{x}_{it}$ contains time averages of *all* time-varying regressors $\mathbf{x}_{it}$. Defining $u_i \equiv \mu_i - \varphi - \bar{\mathbf{x}}_i'\boldsymbol{\theta}$, it follows that $u_i|(\mathbf{x}_{i1}, \ldots, \mathbf{x}_{iT}) \sim \mathcal{N}(0, \sigma_u^2)$. Plugging this into eq. (9) and marginalizing with respect to μ_i we obtain what

[7]To nest the log-normal case in a GLM we specify the conditional expectation function $\mathrm{E}[H|\bar{y}_t, \sigma_t] = \Phi\left(\frac{\alpha+\beta\ln\bar{y}_t+\gamma_1\sigma_t^2}{\exp(\gamma_2\ln\sigma_t^2)}\right)$, where under log-normality we should find $\alpha = \ln z$, $\beta = -1$, $\gamma_1 = \gamma_2 = 1/2$. Testing these restrictions amounts to testing log-normality in the data. A simple Monte-Carlo simulation suggests that such a test is accurately sized in samples comparable to ours.

[8]The biases decrease as T gets larger, but there are no benchmark simulations for the fractional probit case that we know of and in our application the sample sizes are definitely small.

is known in the GLM literature as a 'population-averaged model'

$$\mathrm{E}[H_{it}|\mathbf{X}_i] = \mathrm{E}[\Phi(\varphi + \mathbf{x}_{it}'\boldsymbol{\beta} + \bar{\mathbf{x}}_i'\boldsymbol{\theta} + u_i)|\mathbf{X}_i] = \Phi\left(\varphi_u + \mathbf{x}_{it}'\boldsymbol{\beta}_u + \bar{\mathbf{x}}_i'\boldsymbol{\theta}_u\right) \tag{10}$$

where the coefficients have been scaled by the factor $(1+\sigma_u^2)^{-1/2}$, as indicated by the subscript u. The scaled coefficients and average partial effects (APEs) of all time-varying covariates are identified and can be estimated consistently.

A second issue is the endogeneity due to measurement error in income or consumption expenditures. Classical measurement error is likely to attenuate the income coefficient, while survey-specific (non-classical) measurement error could work in the opposite direction (Ravallion and Chen, 1997). How inference will be affected depends on which type of error is stronger and how this spills over into other variables.[9] Hence we treat the observed mean incomes ($\ln \bar{y}_{it}$) as endogenous.

Building on Rivers and Vuong (1988), Papke and Wooldridge (2008) suggest a two-step control function estimator for such endogeneity problems. For this solution, we require $m \geq 1$ time-varying external instruments, relevant but strictly exogenous conditionally on the unobserved effects, which we arrange in vectors $\mathbf{z}_{it}$, with time averages $\bar{\mathbf{z}}_i$, and matrices $\mathbf{Z}_i = (\mathbf{z}_{i1}, \ldots, \mathbf{z}_{iT})'$. The first step is to estimate a reduced form $\ln \bar{y}_{it} = \pi_{0t} + \mathbf{x}_{1it}'\boldsymbol{\pi}_1 + \mathbf{z}_{it}'\boldsymbol{\pi}_2 + \bar{\mathbf{x}}_{1i}'\boldsymbol{\pi}_3 + \bar{\mathbf{z}}_i'\boldsymbol{\pi}_4 + \nu_{it}$, where $\mathbf{x}_{1it}$ and $\bar{\mathbf{x}}_{1i}$ are $\mathbf{x}_{it}$ and $\bar{\mathbf{x}}_i$ excluding the mismeasured $\ln \bar{y}_{it}$, and where we allow the intercept π_{0t} (and possibly other coefficients as well) to vary over time. In the second step we estimate a population-averaged model conditioned on the exogenous variables and the reduced-form errors, *viz.* $\mathrm{E}[H_{it}|\ln \bar{y}_{it}, \mathbf{X}_{1i}, \mathbf{Z}_i, \nu_{it}] = \Phi\left(\mathbf{x}_{it}'\boldsymbol{\beta}_r + \bar{\mathbf{x}}_{1i}'\boldsymbol{\theta}_r + \bar{\mathbf{z}}_i'\boldsymbol{\zeta}_r + \rho_r\nu_{it}\right)$. Both steps include the Mundlak-Chamberlain device with all strictly exogenous variables involved. The subscript r denotes a new scale factor $(1+\sigma_r^2)^{-1/2}$. The term ν_{it} is replaced by the reduced-form residual $\hat{\nu}_{it}$, and testing $\rho_r = 0$ amounts to testing exogeneity (see Hausman, 1978). Asymptotic standard errors can be approximated via the delta method or the panel bootstrap.

Accounting for unbalanced data adds another layer of complication. Let T_i denote the sample size for country i, with $T = \max_i T_i$. We assume that the selection of observations, as recorded by a selection indicator s_{it}, is conditionally independent (ignorable). Unbalancedness results in the heterogeneity depending on T_i, thereby jeopardizing the consistency of the estimates. Wooldridge (2010a) argues that this dependence can be captured by allowing both the mean and the variance function of the CRE (and therefore of H_{it}) to shift for every different sample size. This is achieved by including dummy variables $\delta_{T_i,n}$ and their interactions with the time averages ($\delta_{T_i,n}\bar{\mathbf{x}}_i'$ and $\delta_{T_i,n}\bar{\mathbf{z}}_i'$) among the covariates, where $\delta_{T_i,n}$ is the Kronecker delta ($\delta_{T_i,n} = 1$ if $T_i = n$

[9] Classical measurement error in income should also induce upward measurement error in the (level of the) Gini coefficient (Chesher and Schluter, 2002); and non-classical error may affect the Gini too. We reach no simple conclusions about the direction of biases, but note that the correlation between the Gini coefficient and average income in our sample is practically zero.

and 0 otherwise). Since the covariates determine both the mean and the variance, we end up with a heteroskedastic model of the following type

$$\mathrm{E}[H_{it}|\{s_{it}, s_{it}\mathbf{x}_{it}, s_{it}\mathbf{z}_{it}; t=1,...,T\}] = \\ \Phi\left(\frac{\mathbf{x}_{1it}'\boldsymbol{\beta}_h + \psi_h \ln \bar{y}_{it} + \rho_h \hat{\nu}_{it} + \sum_{n=2}^{T} \delta_{T_i,n}\left(\varphi_{hn} + \bar{\mathbf{x}}_i'\boldsymbol{\theta}_{hn} + \bar{\mathbf{z}}_i'\boldsymbol{\zeta}_{hn}\right)}{\exp\left(\sum_{n=2}^{T-1}\delta_{T_i,n}\omega_n\right)}\right) \quad (11)$$

where the ω_n are unknown variance parameters, and the subscript h denotes a new scale factor; the reduced-form residuals $\hat{\nu}_{it}$ are obtained from a similarly augmented first-stage regression, as $\hat{\nu}_{it} = \ln \bar{y}_{it} - \mathbf{x}_{1it}'\hat{\boldsymbol{\pi}}_1 - \mathbf{z}_{it}'\hat{\boldsymbol{\pi}}_2 - \sum_{r=2}^{T}\delta_{T_i,r}(\hat{\pi}_{0r} - \bar{\mathbf{x}}_i'\hat{\boldsymbol{\pi}}_{3r} - \bar{\mathbf{z}}_i'\hat{\boldsymbol{\pi}}_{4r})$, where $\mathbf{x}_{1it}'$ includes time dummies and $\bar{\mathbf{x}}_i'$ their time averages.[10] The specification nests the balanced case. If the panel is balanced, the numerator in eq. (11) has only one set of time averages and a constant in addition to the time-varying covariates, while the denominator is unity.[11] Since our estimates of eq. (11) will be quasi maximum likelihood (QML), asymptotic standard errors based on the inverse information matrix will be too conservative and need to be adjusted for clustering at the country level (for details see Papke and Wooldridge, 1996, 2008, Wooldridge, 2010b).

3.3 Elasticities and semi-elasticities

To paraphrase Wooldridge (2010b, p. 602), the challenge is now to intelligently use the estimates. By plugging QML estimates into eq. (11) and averaging over the cross-sectional dimension, we obtain what Blundell and Powell (2004) and Wooldridge (2010a,b) call the *Average Structural Function* (ASF). Since the surveys are irregularly spaced and often cover only parts of a given year, it is not straightforward to track the evolution of the ASF over time. Hence, we also average over time in order to obtain a single scale factor and a single ASF for each variable of interest.

For time-varying covariates, the ASF will deliver estimates of *Average Partial Effects* (APEs). The APE of a continuous variable at time t is the derivative of the ASF with respect to that variable, at some particular $\mathbf{x}_{it}$. In our model (where the covariates of interest are in logs and do not show up in the variance equation) the APEs correspond to semi-elasticities. Elasticities can obtained analogously by dividing the derivative by the estimated value of the cdf. Formulas are provided in Appendix A.

Appendix B reports several Monte Carlo experiments studying the magnitude of the approximation errors. Simulating data from various known distributions we find small approximation errors for the GLM approach, irrespective of whether we focus on

[10] We implement the estimator in a Stata module called `fhetprob` with analytic first and second derivatives, see `www.richard-bluhm.com/data/`.

[11] An interesting extension is to let the conditional variance depend also on inequality. This would imply a marginal proportional rate of substitution $(-\hat{\varepsilon}_t^{H\bar{y}}/\hat{\varepsilon}_t^{HG})$ that is variable rather than constant.

elasticities or predicted poverty rates. The traditional linear models make much larger errors with biases up to 100% away from the overall mean.

4 Data

We compile a data set covering 124 countries over 30 years (1981-2010), based on household surveys from the World Bank's *PovcalNet* database.[12] Subsamples of this data have been used in previous studies (e.g. Adams, 2004, Kalwij and Verschoor, 2007, Chambers and Dhongde, 2011) and the World Bank's methodology is described in detail in Chen and Ravallion (2010). Here we will only mention its main features.

The information provided by *PovcalNet* consists of poverty headcount ratios (H_{it}), per capita monthly income or consumption expenditures ($\bar{y}_{it}$), the corresponding Gini coefficient of inequality (G_{it}), and the total population (pop_{it}). We consider poverty headcount ratios under two different poverty lines (z) widely used for international comparisons: z = \$2 a day (\$60.80 a month) and z = \$1.25 a day (\$38 a month). The latter is typically used to assess *extreme* poverty. In addition, we define dummy variables indicating whether welfare is measured by income or consumption expenditures, and whether it is reported at the level of individual households or groups of households (deciles or finer quantiles). Reported poverty is typically lower in income surveys than in expenditure surveys, and the availability of individual versus grouped data may capture other systematic survey differences. About 63% of the data come from expenditure surveys and about 74% are estimated from grouped data. All monetary quantities are in constant international dollars at 2005 PPP-adjusted prices.

Most included household surveys are nationally representative, but three large countries (China, India and Indonesia) only report separate urban and rural surveys. To construct national series in those cases we weigh the poverty and income data using the urban and rural population shares. As to the Gini coefficient, since it is not subgroup-decomposable, we estimate a national Gini coefficient via an approximation due to Young (2011), based on a mixture of two log-normal distributions.[13] If only one urban or rural survey is available in a given year we drop that survey, except in the case of Argentina where urbanization is near or above 90% for most of the sample period and we thus consider the urban series as nationally representative. We end up with an unbalanced panel of 124 countries spanning 30 years, with an average time series length ($\bar{T}$) of about

[12] The data is publicly available at http://iresearch.worldbank.org/PovcalNet (we use the 2005 PPP version). Per capita averages are simple averages without equivalence scaling.

[13] *PovcalNet* omits weighting some recent data. To use a single consistent method, we apply Young's formula in all cases where separate urban and rural surveys are combined. The approximation is very accurate. The formula is $G = \sum_{i=1}^{K}\sum_{j=1}^{K} \frac{s_i s_j \bar{y}_i}{\bar{Y}} \left(2K\left[\frac{\ln \bar{y}_i - \ln \bar{y}_j + \frac{1}{2}\sigma_i^2 + \frac{1}{2}\sigma_j^2}{(\sigma_i^2+\sigma_j^2)^{1/2}}\right] - 1\right)$ where K is the total number of subgroups, s_i is the population share of the i-th subgroup, $\bar{y}_i$ is mean income, σ_i^2 is the variance, and $\bar{Y}$ is the population-weighted mean income across all subgroups.

6.5 surveys and a total of 809 observations. Table 1 provides summary statistics for the entire panel. In Appendix C, Table C-1 presents summary statistics by region, and List C-1 reports the number of surveys per included country.

Table 1 – Summary statistics

	Mean	Std. Deviation	Min	Max	N
Main variables (2005 PPP $)					
H_{it} – Headcount ($2)	0.303	0.286	.0002	.9845	809
H_{it} – Headcount ($1.25)	0.182	0.219	.0002	.9255	789
G_{it} – Gini coefficient	0.424	0.102	.2096	.7433	809
$\bar{y}_{it}$ – Mean income or expenditure in $ per month	194.59	125.90	14.93	766.78	809
PCE^P_{it} – Consumption (PWT) in $ per month	338.64	234.59	14.39	1231.21	795
Survey type dummies					
Consumption expenditures, grouped	0.611	0.488	0	1	809
Consumption expenditures, household	0.015	0.121	0	1	809
Income, grouped	0.132	0.339	0	1	809
Income, household	0.242	0.429	0	1	809

The standard linear approximations require data in differences or log-differences at the country level. Differences between surveys may involve, or coincide with, switches between income and expenditure surveys or changes in the level of aggregation. To avoid spurious differences, we calculate them only between surveys of the same type: income or expenditure based, and available at the household level or grouped. We also annualize all differences to mitigate any biases arising from estimating elasticities over time periods of different lengths. Since differencing requires $T_i \geq 2$ and we only use comparable surveys, we obtain 648 observed differences from 104 countries. To estimate the contributions we will reduce this data set further, retaining only the longest consecutive runs of the same survey type (yielding 123 observations); and we will also consider two sub-periods, preceding and following the year 2000 (yielding 87 observations in each).

Apart from the survey-based data, we collect consumption data from national accounts, which will later serve both as instruments for survey income and as a basis for the projections. Per capita *personal consumption expenditures* (PCE) are retrieved from both the World Development Indicators (WDI) and the Penn World Table 7.1 (PWT).[14] We construct a 'merged' series for the projections using the WDI as the default but replace it with PWT data if coverage over 1981-2010 is better. We will also make use of population projections covering the period 2010-2030 from the World Bank's Health, Nutrition and Population Statistics database.

[14]Monthly PCE^P_{it} is computed as $(kc_{it}/100 \times rgdpl_{it}/12)$, where kc_{it} is the PWT consumption share and $rgdpl_{it}$ is GDP per capita (Laspeyres) in 2005 constant prices. PCE^W_{it} is WDI household final consumption expenditures in 2005 prices divided by population and converted to monthly figures.

5 Results

5.1 Regressions

Table 2 presents our main estimation results, with each specification addressing an additional issue: unobserved effects, unbalancedness and measurement error, in that order. All include time averages à la Mundlak to capture measurement differences across countries (unobserved effects), and survey type dummies (expenditures or income, grouped or household data) to capture measurement differences across surveys. In addition, a full set of year dummies allows for unspecified common time effects.

Table 2 – Fractional probit models (QMLE) – Dependent variable: H_{it}, \$2 a day

	(1) Regular CRE		(2) Unbalanced		(3) Unbalanced + Two-Step	
	H_{it}	APEs	H_{it}	APEs	H_{it}	APEs
$\ln \bar{y}_{it}$	-1.263	-0.284	-0.880	-0.281	-1.049	-0.339
	(0.054)	(0.012)	(0.048)	(0.011)	(0.198)	(0.035)
$\ln G_{it}$	1.032	0.232	0.786	0.251	0.775	0.251
	(0.114)	(0.026)	(0.098)	(0.026)	(0.163)	(0.032)
$\hat{\nu}_{it}$					0.133	
					(0.113)	
CRE (Corr. Rand. Effects)	Yes		Yes		Yes	
Survey type dummies	Yes		Yes		Yes	
Time dummies	Yes		Yes		Yes	
Panel size dummies	No		Yes		Yes	
Panel size dummies × CRE	No		Yes		Yes	
Variance equation	No		Yes		Yes	
Scale factor	0.225		0.319		0.323	
$N \times \bar{T}$	789		789		775	
N	104		104		103	
pseudo R^2	0.984		0.993		0.993	
$\ln \mathcal{L}$	-219.3		-315.6		-313.4	
$\sqrt{MSE}$	0.0355		0.0238		0.0235	

Notes: The table reports fractional response QMLE estimates. The dependent variable is the poverty rate at \$2 a day (in 2005 PPPs). 20 observations with $T_i = 1$ are not used during estimation. The panel structure is country-survey-year. In models (1) and (2), the standard errors of the coefficients are robust to clustering at the country level and the standard errors of the APEs are computed via the delta method. We include the time averages of the survey type and time dummies in (2) and (3), but constrain their coefficients to be equal across the panel sizes. The standard errors of the coefficients and the APEs in model (3) account for the first stage estimation step with a panel bootstrap using 999 bootstrap replications. The linear projection in the first stage uses $\ln PCE_{it}^P$ as an instrument for $\ln \bar{y}_{it}$. The first-stage cluster-robust F-statistic in (3) is 28.05. Model (3) also excludes West Bank and Gaza entirely (2 observations) and 12 observations from ECA countries pre-1990 for lack of PCE data.

Column (1) includes correlated random effects but ignores unbalancedness. The coefficient on average income is negative and the coefficient on inequality is positive, as they should. Since they are arbitrarily scaled, the adjacent column reports average partial effects (APEs); the scale factor is reported separately in the bottom panel. The APEs in column (1) can be interpreted as average *semi-elasticities*: income growth of one *percent* results in a reduction in the number of poor by 0.284 *percentage points*; and a Gini increase of one *percent* corresponds to an increase in the number of poor by 0.232 *percentage points*. The average income *elasticity* across the entire estimation sample is about -1.83 (SE = 0.084) and the average Gini elasticity is about 1.5 (SE = 0.167). For instance, one *percent* income growth would lead to about a 1.83 *percent* reduction in poverty. These elasticities are close to the lower end (in absolute value) of the range reported in the literature.[15]

Our first specification could be biased due to the strong unbalancedness of the panel and the presence of time-varying measurement error in income and inequality. Column (2) addresses unbalancedness by including panel size dummies, interactions of the time averages with the panel size dummies, and a separate variance equation. We consider this our best specification *without* correcting for measurement error. The substantive conclusions change very little. The APE of income is virtually unchanged and the APE of inequality increases by less than one standard error. Unbalancedness causes little bias on average, but it may still affect the (semi-) elasticities at particular points in time.

Our preferred specification, column (3), is the empirical counterpart of the two-step estimator presented in eq. (11). To account for measurement error in income, we instrument survey mean incomes or expenditures with per capita consumption expenditures from the PWT national accounts (PCE_{it}^{P}). The main identifying assumption is that any measurement error in per capita consumption from the national accounts is uncorrelated with survey-based measurement error in income or expenditures (also see Ravallion, 2001). Figure C-3 in Appendix C shows a partial regression plot highlighting the strength of the first stage relationship.

The evidence of measurement error in income is weak (panel bootstrap t-stat ≈ 1.18; ignoring first-stage sampling error, cluster-robust t-stat ≈ 1.69). The APE and elasticity of income are larger in absolute value than in the previous two specifications ($\bar{\varepsilon}^{H\bar{y}} \approx -2.21$, SE = 0.156). Income growth of one percent leads to a 0.339 percentage point or 2.21 percent reduction in the number of poor, on average. We tentatively conclude that the coefficient of income in columns (1) and (2) is moderately attenuated. This would suggest that classical attenuation bias is more of a problem than systematic survey bias, but we cannot rule out that more complex error structures are at play. The APE and elasticity of inequality remain practically unchanged ($\bar{\varepsilon}^{HG} \approx 1.64$, SE = 0.188).[16]

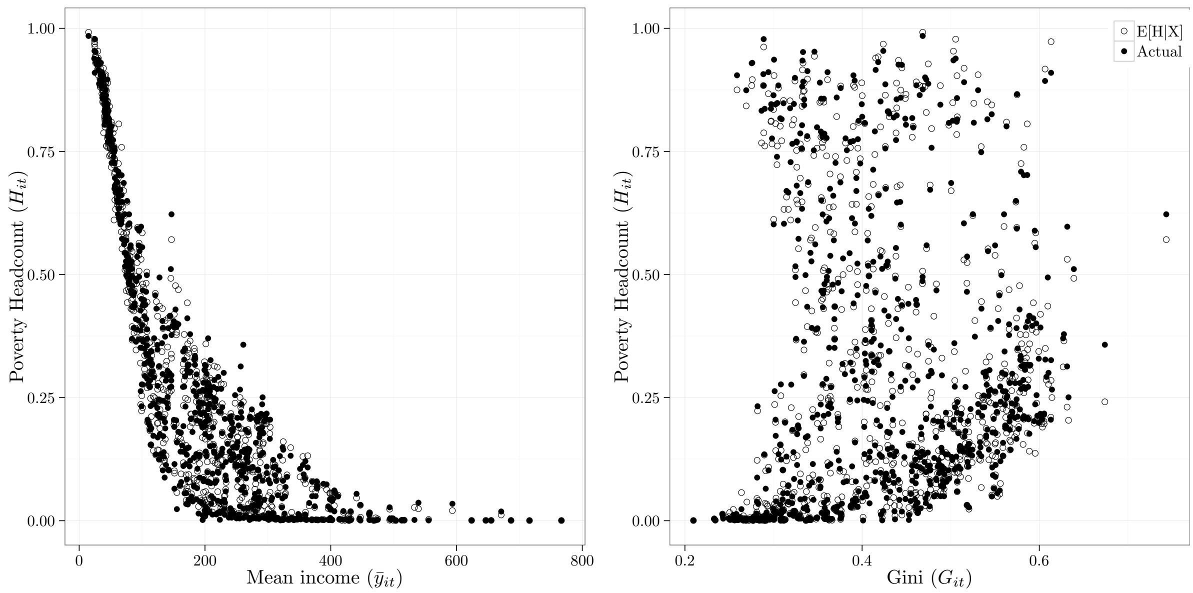
The last row of Table 2 reports the square root of the mean squared residual ($\sqrt{MSE}$). The poverty headcount is predicted with about three and a half percentage

[15] The typical range for the income elasticity in earlier studies is from about -2 to -5, while for the Gini elasticity it is much wider. In newer studies income elasticities are closer to -2.

[16] We recognize that inequality as recorded in the Gini may be affected by measurement error too, but we lack an adequate instrument for the Gini coefficient.

points accuracy in the first model, and with better than two and a half percentage points accuracy in the next two. A close fit is what one would expect from a well-defined decomposition. A simple pseudo-R^2 measure, the squared correlation between the observed and fitted values, with values of 0.98 to 0.99, tells the same story. Figure 1 illustrates this point. Using our preferred specification, we plot both the observed headcount and the predicted headcount over the range of observed mean income or expenditures (left panel) and inequality (right panel). The non-linear fit is very close at either bound (near unity or near zero). The model predicts no nonsensical values, and the entire range of variation of the observed values is covered by the model predictions.

Figure 1 – Data versus fitted values, preferred specification, $2 a day



Notes: Illustration of the model fit from the estimates presented in column (3) of Table 2. The left panel shows the observed poverty rates (actual) and predicted poverty rates ($\hat{\mathrm{E}}[H|X]$) over mean incomes; the right panel plots both over the Gini coefficient.

A comparison with the linear approaches common in the literature is reported in Table C-3 of Appendix C. As is typical for such comparisons, linearly and non-linearly estimated elasticities are very similar *on average*. However, even with interaction terms, the linear equations do not fit nearly as well as the fractional probit models and many coefficients are insignificant. The step from within-transformed data in column (1) to annualized differences in the following columns seems to worsen the impact of measurement error and attenuate the income effect. The two-step GMM estimates of the interaction models are unstable and unable to convincingly reverse the attenuation effect. The last model, which mimics the preferred specification of Kalwij and Verschoor (2007), even implies a negative Gini elasticity. In sum, the traditional linear models perform poorly and are unlikely to produce reliable estimates over a wide range of circumstances.

Before analyzing the impacts in further detail, we check the evidence on log-normality in our data. Table C-4 reports the results of the test outlined in footnote 7. Log-normality is rejected in our three specifications; the added variance parameter in particular diverges substantially from its log-normal value. The overall fit is slightly improved but the improvement is tiny once unbalancedness is accounted for. The estimated income semi-elasticity barely moves so that practical implications are limited anyway. The results strongly favor our simpler and less assumption-laden approximation.

5.2 Impacts

The strength of the fractional response approach lies in its ability to deliver precise and unbiased estimates of effects other than the overall mean response. Table 3 and Table 4 illustrate this point by estimating income and Gini elasticities (Panel a) and semi-elasticities (Panel b) over different time periods for six large geographic regions. Elasticities and semi-elasticities are computed according to eqs. (A-6) and (A-7) by plugging in time-period and region-specific averages of mean income ($\ln \bar{y}_{it}$) and inequality ($\ln G_{it}$), and then averaging over the entire subsample. Standard errors are computed via a panel bootstrap to take into account the sampling uncertainty of the first stage.

There are considerable regional and temporal differences in the estimated income elasticities. As Section 2 showed, heterogeneity of elasticities has a mechanical origin: it is a consequence of heterogeneity in incomes and inequality. More affluent regions (Eastern Europe and Central Asia, Latin America and the Caribbean, and the Middle East and North Africa) have higher income elasticities than poorer regions (East Asia and Pacific, South Asia, and Sub-Saharan Africa). Income dynamics over time are also clearly visible. In Eastern Europe and Central Asia, for example, income is comparatively high before the post-communist transition, sharply collapses through the 1990s, and recovers during the 2000s. Compared to earlier results (e.g. Kalwij and Verschoor, 2007), we find markedly higher average income elasticities in more affluent regions and lower elasticities in poorer regions. All standard errors in Table 3 are small compared to the point estimates and *remain* small for regions with extreme values.

The picture is reversed when we look at the semi-elasticities in Panel b). Comparatively affluent regions have fewer people near the poverty line, and thus the poverty reduction potential from an equivalent increase in incomes is much smaller in terms of people lifted out of poverty. This pattern is (again) best visible in Eastern Europe and Central Asia, where absolute poverty at the $2 a day poverty line is almost non-existent just before the post-communist transition, but rises sharply in the 1990s as incomes decline. Correspondingly, the semi-elasticity is close to zero in the 1980s but then increases as more people fall into poverty. Likewise, the biggest poverty reduction potential in 2005-2010 was in East Asia, South Asia, and Sub-Saharan Africa. This highlights an important point. For development policy, we really care about the share of

Table 3 – Income elasticities and semi-elasticities, $2 a day, by region

	Time period				
	1981–1989	1990–1994	1995–1999	2000–2004	2005–2010
	Panel a) Regional income elasticities				
East Asia and Pacific	-0.991	-1.029	-1.237	-1.139	-1.578
	(0.030)	(0.033)	(0.055)	(0.043)	(0.101)
Eastern Europe and Central Asia	-4.358	-2.892	-2.700	-2.846	-3.304
	(0.555)	(0.309)	(0.277)	(0.304)	(0.384)
Latin America and Caribbean	-2.284	-2.374	-2.425	-2.349	-2.985
	(0.243)	(0.257)	(0.271)	(0.258)	(0.366)
Middle East and North Africa	-2.176	-2.116	-2.024	-1.966	-2.501
	(0.203)	(0.188)	(0.168)	(0.161)	(0.246)
South Asia	-0.548	-0.629	-0.810	-1.024	-1.192
	(0.053)	(0.048)	(0.030)	(0.032)	(0.046)
Sub-Saharan Africa	-0.831	-0.437	-0.436	-0.592	-0.632
	(0.027)	(0.039)	(0.040)	(0.035)	(0.033)
	Panel b) Regional income semi-elasticities				
East Asia and Pacific	-0.568	-0.573	-0.585	-0.583	-0.552
	(0.034)	(0.036)	(0.046)	(0.042)	(0.051)
Eastern Europe and Central Asia	-0.031	-0.214	-0.260	-0.225	-0.134
	(0.008)	(0.015)	(0.020)	(0.015)	(0.010)
Latin America and Caribbean	-0.374	-0.348	-0.334	-0.355	-0.194
	(0.028)	(0.025)	(0.024)	(0.026)	(0.013)
Middle East and North Africa	-0.405	-0.422	-0.447	-0.463	-0.313
	(0.034)	(0.037)	(0.042)	(0.043)	(0.024)
South Asia	-0.418	-0.458	-0.526	-0.572	-0.585
	(0.023)	(0.019)	(0.022)	(0.036)	(0.044)
Sub-Saharan Africa	-0.532	-0.354	-0.353	-0.440	-0.459
	(0.024)	(0.020)	(0.020)	(0.015)	(0.015)

Notes: The table reports regional income elasticities in panel a) and regional income semi-elasticities in panel b). The estimates are computed by plugging period and region-specific averages of mean income and inequality into eq. (A-6) or its semi-elasticity counterpart eq. (A-7) and then averaging over the entire sample. Standard errors are obtained via a panel bootstrap using 999 replications.

the population lifted out of poverty rather than the percent change in the poverty rate.

The region and time specific Gini elasticities in Panel a) of Table 4 show where the potential of redistributive policies in terms of relative reductions of the poverty rate has been largest in the last three decades. Unsurprisingly, these regions are Eastern Europe and Central Asia, Latin America and the Caribbean, and the Middle East and North Africa – all of which have above average inequality. Sub-Saharan Africa starts out with high inequality (the population-weighted mean Gini in the 1980s is 0.4608) and very low incomes, so that the Gini elasticity is small. This is the flip side of the dependency on initial levels: countries can be so poor and unequal that the immediate effects of equalization and income growth on *relative* changes in the poverty rate are small. Again, though, the semi-elasticities presented in Panel b) reverse the picture: poorer and richer countries swap positions. The potential for *absolute* reductions in poverty rates through redistribution was larger in poorer regions throughout the entire period.

Table 4 – Inequality elasticities and semi-elasticities, \$2 a day, by region

	Time period				
	1981–1989	1990–1994	1995–1999	2000–2004	2005–2010
	Panel a) Regional inequality elasticities				
East Asia and Pacific	0.732	0.760	0.914	0.841	1.165
	(0.105)	(0.101)	(0.113)	(0.108)	(0.144)
Eastern Europe and Central Asia	3.219	2.136	1.994	2.102	2.440
	(0.510)	(0.307)	(0.283)	(0.296)	(0.353)
Latin America and Caribbean	1.687	1.753	1.791	1.735	2.205
	(0.186)	(0.198)	(0.199)	(0.189)	(0.269)
Middle East and North Africa	1.607	1.563	1.495	1.452	1.847
	(0.197)	(0.198)	(0.196)	(0.185)	(0.253)
South Asia	0.405	0.464	0.598	0.756	0.880
	(0.093)	(0.097)	(0.095)	(0.107)	(0.127)
Sub-Saharan Africa	0.614	0.322	0.322	0.437	0.467
	(0.087)	(0.055)	(0.060)	(0.066)	(0.069)
	Panel b) Regional inequality semi-elasticities				
East Asia and Pacific	0.419	0.423	0.432	0.431	0.408
	(0.053)	(0.053)	(0.055)	(0.054)	(0.053)
Eastern Europe and Central Asia	0.023	0.158	0.192	0.166	0.099
	(0.007)	(0.015)	(0.019)	(0.017)	(0.012)
Latin America and Caribbean	0.276	0.257	0.247	0.262	0.143
	(0.046)	(0.043)	(0.043)	(0.045)	(0.029)
Middle East and North Africa	0.299	0.311	0.330	0.342	0.231
	(0.041)	(0.040)	(0.041)	(0.044)	(0.025)
South Asia	0.309	0.338	0.389	0.423	0.432
	(0.056)	(0.055)	(0.052)	(0.054)	(0.055)
Sub-Saharan Africa	0.393	0.261	0.261	0.325	0.339
	(0.050)	(0.037)	(0.040)	(0.042)	(0.043)

Notes: The table reports regional inequality elasticities in panel a) and regional inequality semi-elasticities in panel b). The estimates are computed by plugging period and region-specific averages of mean income and inequality into eq. (A-6) or its semi-elasticity counterpart eq. (A-7) and then averaging over the entire sample. Standard errors are obtained via a panel bootstrap using 999 replications.

Two features stand out in these regional and temporal comparisons. First is the contrast between the two impact measures, elasticities and semi-elasticities; the latter have the higher policy relevance. Second is the substantial heterogeneity in the estimated impacts across time and space, mainly due to their dependence on the prevailing levels of income and inequality.[17] As argued by Bourguignon (2003), an improvement in distribution may deliver a 'double dividend', in the sense that it will not only reduce the poverty rate directly, but also increase the beneficial impact of any further income growth. Figure 2 illustrates this by graphing the estimated poverty elasticities and semi-elasticities. The functions are computed according to eqs. (A-6) and (A-7) by

[17]Incomes (or expenditures) have increased substantially in all regions between 1981 and 2010 (see Table C-2 in Appendix C). By contrast, inequality shows no systematic trend over the sample period from 1981 to 2010. In a simple regression of the Gini coefficient on time, we fail to reject the null hypothesis that the time trend is zero (cluster-robust t-stat ≈ 0.07 and $p > 0.94$).

plugging in various combinations of values for per capita income ($\ln \bar{y}_{it}$) and inequality ($\ln G_{it}$), and then averaging over the entire sample. As Figure 2a shows, on top of the direct poverty alleviating effect of income redistribution, a lower level of inequality also raises the income elasticity in absolute value at every point. The magnitude of both elasticities is steeply increasing in the level of income; that is, the returns to either income growth or equalization are bigger, the higher the income level, and the gap between the functions evaluated at different inequality levels keeps widening, inviting the conclusion that redistribution has most impact in affluent societies. That, precisely, is the misleading feature of poverty elasticities.

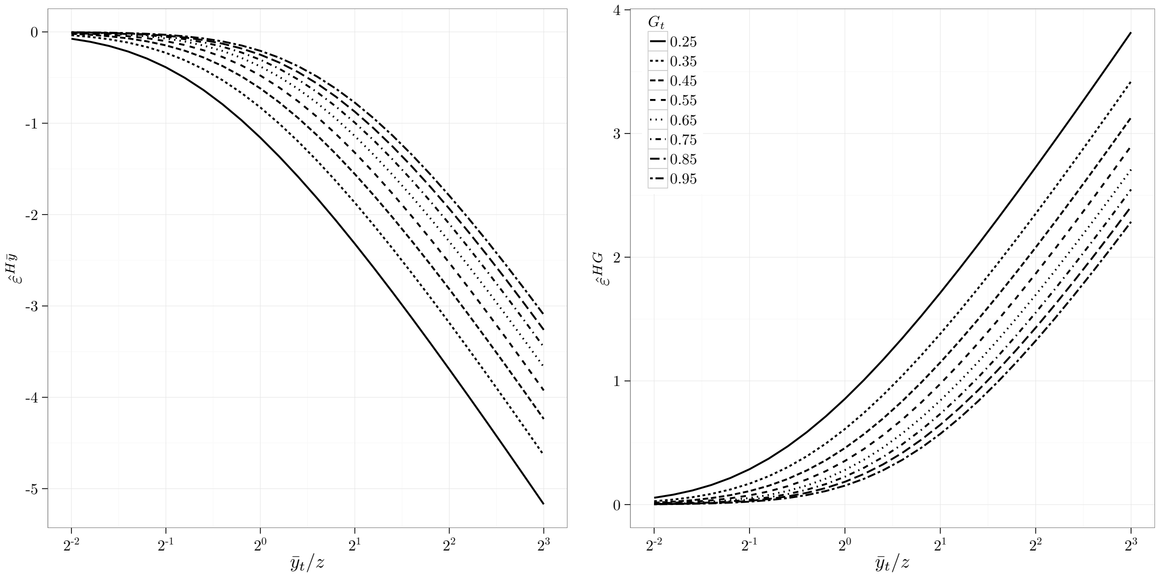
Figure 2b shows the predicted income and Gini *semi*-elasticities of poverty. The picture is very different and in many ways more intuitive. If the income shortfall is too large – the mass of the income distribution is too far below the poverty line – then both the income and the Gini semi-elasticities approach zero. However, if the country is affluent – the mass of the income distribution is far to the right of the poverty line – then both semi-elasticities also approach zero. In between those two extremes, improvements in the income distribution can make a large difference in the proportion of people lifted out of poverty, both directly through redistribution and indirectly through growth. When mean income is at the poverty line ($\bar{y}_t/z = 1$) and the Gini at 0.25, for example, one percent growth will deliver a 0.584 percentage point reduction in the poverty headcount; but if the Gini is at 0.55, one percent growth will deliver only a 0.378 percentage point reduction in the poverty headcount. Especially at very low (or very high) average income levels the initial income distribution is decisive; it practically determines whether there is potential for poverty alleviation through growth at all. Moreover, improvements in the income distribution will have a larger poverty reducing effect at lower (initial) levels of inequality in poor countries (where $\bar{y}_t/z \leq 1$). This pattern leads us to the conclusion that poverty reduction strategies should focus on both income growth *and* equalization, especially in low-income countries where the total returns to redistribution are large.[18]

Could the decomposition be improved by allowing for other, 'more ultimate' determinants of poverty? Under log-normality mean income and the Gini fully describe the distribution of incomes and expenditures, and there is no scope for additional determinants. Yet we deliberately do not rely on log-normality and our data actually reject it. More realistic distributions (see, e.g., Bresson, 2009) usually have more than one shape parameter in order to better capture skewness, a long tail, or the existence of multiple modes. 'Ultimate factors' could thus be proxies for systematic deviations from equiproportional shifts in the distribution of incomes and expenditures. Weak institutions, for example, might explain the fact that the rich capture more of the gains

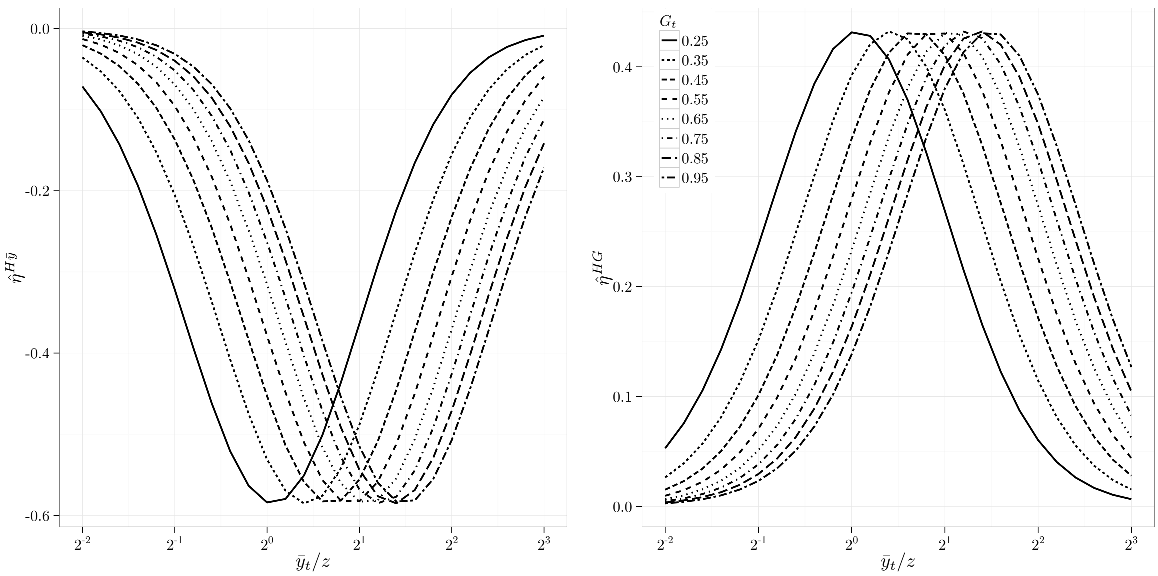
[18]This finding should stand with the new 2011 PPPs as well. Different PPPs would shift incomes, the poverty line and the poverty rates, and so countries would be located differently on the graph; but the position of the curves in terms of relative incomes, $\bar{y}_t/z$, would be comparable.

Figure 2 – Predicted income and Gini elasticities and semi-elasticities of poverty, \$2 a day

(a) elasticities ($\hat{\varepsilon}^{H\bar{y}}$ and $\hat{\varepsilon}^{HG}$)



(b) semi-elasticities ($\hat{\eta}^{H\bar{y}}$ and $\hat{\eta}^{HG}$)



Notes: Illustration of non-linear nature of the poverty-growth-inequality relationship based on the estimates presented in column (3) of Table 2. Panel a) shows the estimated income elasticities of poverty (on the left) and estimated inequality elasticities of poverty (on the right) plotted over the ratio of average incomes to the poverty line. Panel b) shows the estimated income semi-elasticities of poverty (on the left) and estimated inequality semi-elasticities of poverty (on the right) plotted over the ratio of average incomes to the poverty line. The various curves correspond to estimates based on different Gini coefficients.

of growth. Table C-7 in Appendix C extends the heteroskedastic fractional probit models with data on institutions, human capital, access to credit and trade openness. The APEs of income and inequality are not affected by the inclusion of the additional covariates and the APEs of the latter are virtually zero. Thus we conclude that with only two variables, some dummies and correlated random effects, these specifications are essentially saturated. Contrary to linear approximations, the fractional response approach leaves little room for misspecification of the decomposition.

5.3 Contributions

We now turn to the historical contributions of growth and redistribution to poverty reduction. We will argue that there has been a marked shift in the distributional pattern of growth towards pro-poor growth around the turn of the millennium.

The contribution of growth or redistribution to poverty changes in some predefined period is the product of their impact (partial effect, semi-elasticity) and their actual variation, usually expressed as a proportion of the total change in the poverty rate over the same period. Together with the stylized fact that growth tends to be distribution-neutral (see Figure C-4), the relative contributions of growth and redistribution have been used previously to highlight the primacy of growth for poverty reduction (see e.g. Dollar and Kraay, 2002, Kraay, 2006, Kraay et al., 2014, Dollar et al., 2016). Kraay (2006), for example, concludes that long-run differences in poverty reduction can largely be explained by growth in average incomes, as opposed to changes in inequality, which seem to matter only in the short run. A good deal of the primacy of growth is due to the fact that there is a lot more long-run variation in growth than in inequality. Even so, the poverty reducing quality of growth is of legitimate concern. While the cross-country correlation between growth and changes in inequality is near zero, this correlation varies across countries, regions, and time periods (see also Ravallion and Chen, 2003). In fact, the variability of the (semi-)elasticities emphasized above, together with the variability of average incomes and inequality, can give rise to all sorts of relative contributions.

Some additional notation will prove useful. Following Datt and Ravallion (1992), a discrete-time, non-logarithmic version of the decomposition of poverty changes is

$$\Delta H_t \approx \Big[H(\bar{y}_t/z, L_{t-1}) - H(\bar{y}_{t-1}/z, L_{t-1})\Big] + \Big[H(\bar{y}_{t-1}/z, L_t) - H(\bar{y}_{t-1}/z, L_{t-1})\Big], \qquad (12)$$

where L_t is the Lorenz curve describing the income distribution; decomposing absolute differences is in line with the fractional response approach, in which the poverty rate *level* is the natural metric. The first brackets on the right-hand side contain the 'growth component', the second the 'distributional component' of poverty reduction; we will refer to them as Y and D, respectively. Two of the four terms inside the brackets are counterfactual: $H(\bar{y}_t/z, L_{t-1})$ is the poverty rate when average income has evolved but the Lorenz curve is kept unchanged; $H(\bar{y}_{t-1}/z, L_t)$ is the poverty rate when the distribution has shifted but average income is kept constant. The decomposition is not exact: it is 'path dependent' and generates a residual. In the absence of micro-data it is natural to approximate $H(\bar{y}_t/z, L_t)$ by our earlier function $H(\bar{y}_t/z, G_t)$, and replace the unknown quantities by predicted counterparts $\hat{H}(\bar{y}_t/z, G_{t-1})$ and $\hat{H}(\bar{y}_{t-1}/z, G_t)$.[19]

We carry out this decomposition over different time periods based on three different data sets: one using the longest possible spells between surveys of the same type in 1981-2010; and two for the sub-periods 1981-2000 and 2000-2010. Since the availability of surveys before 2000 is more limited, the average spell length in the two sub-periods is comparable; both have a mean and median duration between the initial and final surveys of about seven years. The break point is not chosen by accident. The turn of the millennium marked a qualitative change in the growth performance of Sub-Saharan Africa, and was picked by a formal structural change test (see Figure C-5 and Table C-2 in Appendix C). The developing world collectively grew faster than the developed world for the subsequent decade. What we are interested in here is whether this development coincided with a shift in the quality of growth.

Like Kraay (2006), we borrow from the growth accounting literature to define the shares of growth and inequality (see e.g. Klenow and Rodriguez-Clare, 1997, Caselli, 2005). Table 5 reports a regional variance decomposition of all spells occurring within a particular region and period using the \$2 a day poverty line. Within each of the three periods we compute the variance of the growth component, VAR(Y), the variance of the distribution component, VAR(D), and their covariance, COV(Y, D). From these we obtain the share of the growth component as $s_Y = [\text{VAR}(Y) + \text{COV}(Y, D)]/[\text{VAR}(Y) + \text{VAR}(D) + 2\text{COV}(Y, D)]$, and the share of the distribution component as $s_D = 1 - s_Y$.[20] We also report the root mean squared error ($\sqrt{MSE}$) of the observed versus the predicted values to assess the importance of the residual including model error. Last but not least, we report the number of spells for each region and period. Table C-9 in Appendix C reports comparable results calculated at the \$1.25 a day poverty line.

Our results for all developing countries over the entire 1981-2010 period compare well with the stylized facts from the literature. Overall, the share of growth is about 82%, redistribution accounting for about 18% only. The variance of changes in the poverty rate due to changes in average income is much greater than the variance due to changes in inequality, and the covariance between the two components is virtually zero, consistent with distribution-neutral growth. The residual is less than three quarters of a percentage point, indicating that even with added model error the decomposition works well. The results for the \$1.25 a day poverty line (Table C-9) are qualitatively similar and broadly match those reported in Kraay (2006).

The breakdown of the results by region also tells a familiar story. Growth dominated poverty reduction in Sub-Saharan Africa (where incomes first declined, stagnated and

[19] To get rid of the path dependency eq. (12) may be averaged with a second decomposition swapping t and $t-1$, delivering what Shorrocks (2013) calls the 'Shapley decomposition'. We do not attempt this since in our case the decomposition residual also includes prediction errors.

[20] Note that the share of one factor can exceed one, but both always add up to one. Caselli (2005) discusses a set of alternative measures, but they do not change the qualitative implications of our results.

Table 5 – Decomposition at $2 a day poverty line, by region

	VAR(Y)	VAR(D)	COV(Y,D)	s_Y	s_D	$\sqrt{MSE}$	N
	Panel a) Spells from 1981 to 2010						
East Asia and Pacific	1.166	0.385	0.001	75.16	24.84	1.01	12
Europe and Central Asia	3.429	0.414	0.164	86.14	13.86	0.57	39
Latin America and Caribbean	3.073	0.666	-0.619	98.12	1.88	0.79	28
Middle East and North Africa	0.425	0.452	0.304	49.10	50.90	1.38	8
South Asia	0.659	2.743	0.191	22.46	77.54	0.36	7
Sub-Saharan Africa	1.482	0.134	-0.196	105.08	-5.08	0.48	29
All developing	2.576	0.579	0.002	81.61	18.39	0.73	123
	Panel b) Spells from 1981 to 1999						
East Asia and Pacific	0.951	0.151	0.017	85.23	14.77	0.86	9
Europe and Central Asia	14.568	1.912	-0.706	91.99	8.01	0.92	25
Latin America and Caribbean	3.750	1.229	-0.633	83.94	16.06	0.89	26
Middle East and North Africa	0.372	0.365	0.207	50.31	49.69	0.78	6
South Asia	0.244	0.021	0.004	91.11	8.89	0.33	4
Sub-Saharan Africa	2.217	0.702	-0.309	82.89	17.11	0.92	17
All developing	6.852	1.073	-0.368	90.19	9.81	0.87	87
	Panel c) Spells from 2000 to 2010						
East Asia and Pacific	1.579	1.143	-0.134	58.89	41.11	1.18	10
Europe and Central Asia	3.544	0.498	0.826	76.75	23.25	0.57	26
Latin America and Caribbean	0.365	0.150	0.010	70.10	29.90	0.35	19
Middle East and North Africa	0.780	0.741	0.629	50.69	49.31	1.76	5
South Asia	0.642	1.122	0.731	42.57	57.43	0.40	6
Sub-Saharan Africa	2.051	0.560	0.188	74.97	25.03	0.87	21
All developing	2.049	0.559	0.365	72.31	27.69	0.81	87

Notes: The table reports the results of the decomposition of the observed changes in the poverty rate at $2 a day into its growth and distribution components at the regional level. Panels a) to c) run this decomposition over different sub-samples as denoted in the table. We predict the counterfactual quantities using the first and last available data for the longest runs of survey of the same type within the sample period.

then recovered again), in Latin America (where inequality was persistently high for most of the period in question), and in Europe and Central Asia (where the post-communist transition was accompanied by rapid movements in average incomes). Distributional change played a larger role in East Asia and Pacific, Middle East and North Africa, and South Asia, where it was responsible for about one, two and three quarters of the poverty evolution, respectively. Some of these results come with a caveat, though. In South Asia, although we have all but one countries represented, the sample size is very small.[21] In the Middle East and North Africa, survey coverage is notoriously bad, we have less than half of all countries in the sample, and the residual is more than twice the average.

So what's new? A very different picture emerges once we split the sample at the turn of the millennium. The corresponding results are reported in Panel b) and Panel c) of Table 5. Growth was responsible for nearly all of the changes in the poverty rate before

[21] The large share of the distribution component is driven by the Maldives and Bhutan, where nearly all or about half of poverty reduction can be attributed to changes in distribution.

2000, but its share fell to 72% in the following decade. Since the mean spell lengths are about the same, this is not due to a long run vs. short run difference. Instead, this development seems to be the result of a shift in the poverty reducing quality of growth. Panel b) shows that before 2000 the correlation between the growth and distribution component was negative for all developing countries as a whole and for half of the regions. Panel c) demonstrates that after 2000 the correlation between the two components turned positive for the sample as a whole and in all regions except East Asia: growth coincided with reductions in inequality more often. Table C-9 in Appendix C confirms that this pattern also holds at the $1.25 a day poverty line.

Figure 3 helps to make this point more tangible. The upper panel shows how the correlation between the estimated growth and distribution components changed over the turn of the millennium. Before 2000, reductions in poverty through growth often coincided with adverse distributional change (lower right quadrant), and negative economic growth often coincided with pro-poor distributional change (upper left quadrant). Since 2000, this is no longer true. In a majority of cases, poverty reduction through growth was reinforced by pro-poor distributional change (lower left quadrant). In other cases, growth was not accompanied by pro-poor redistribution (lower right quadrant). Only in a handful of cases did growth fail to reduce poverty (both upper quadrants). The lower panel of the figure shows that the estimated growth components closely track the observed change in the poverty rate in both subsamples, whether the change is positive or negative; but after 2000, growth becomes almost everywhere a net contributor to poverty *reduction*.

Overall, our interpretation of these results challenges the conventional wisdom that growth is distribution-neutral. The evidence suggests that growth has become substantially more pro-poor since the turn of the millennium, in two ways: *(i)* in the relative sense, that is, it benefits the poor disproportionately, and *(ii)* in the absolute sense, that is, it reduces poverty more often than not. If this trend can be maintained this is good news for the goal of ending absolute poverty in the coming decades.

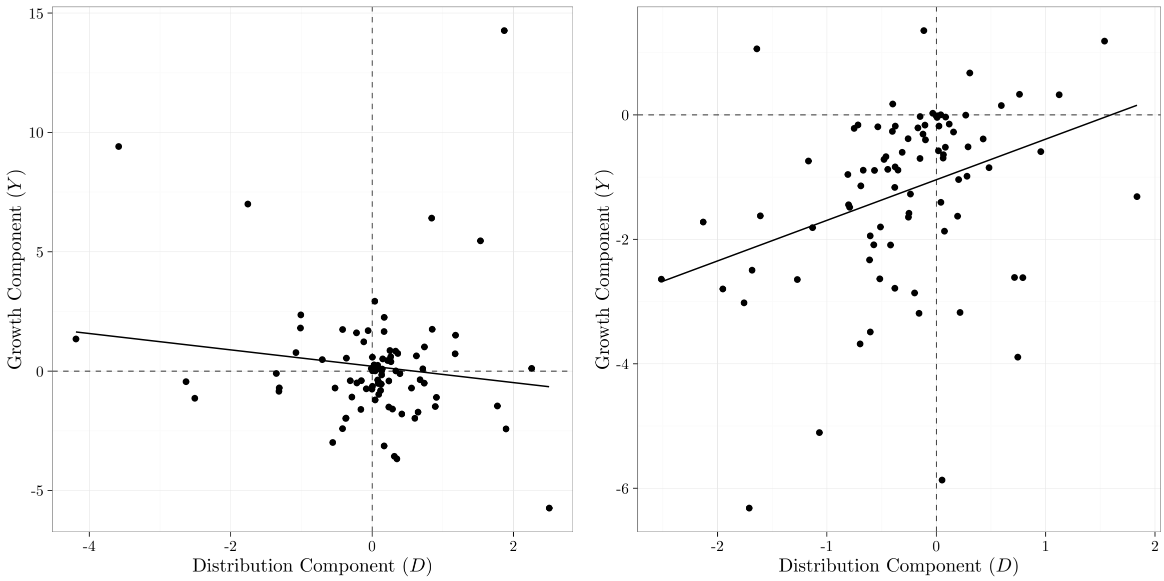
5.4 Projections

So what are the prospects for poverty reduction? In 2015, all member states of the United Nations agreed to 'eradicate extreme poverty for all people everywhere, currently measured as people living on less than $1.25 a day' by 2030. The key question we ask now is whether this new goal is attainable, or if the global goalpost has been set so high, that it may fail to promote a sensible allocation of development assistance.

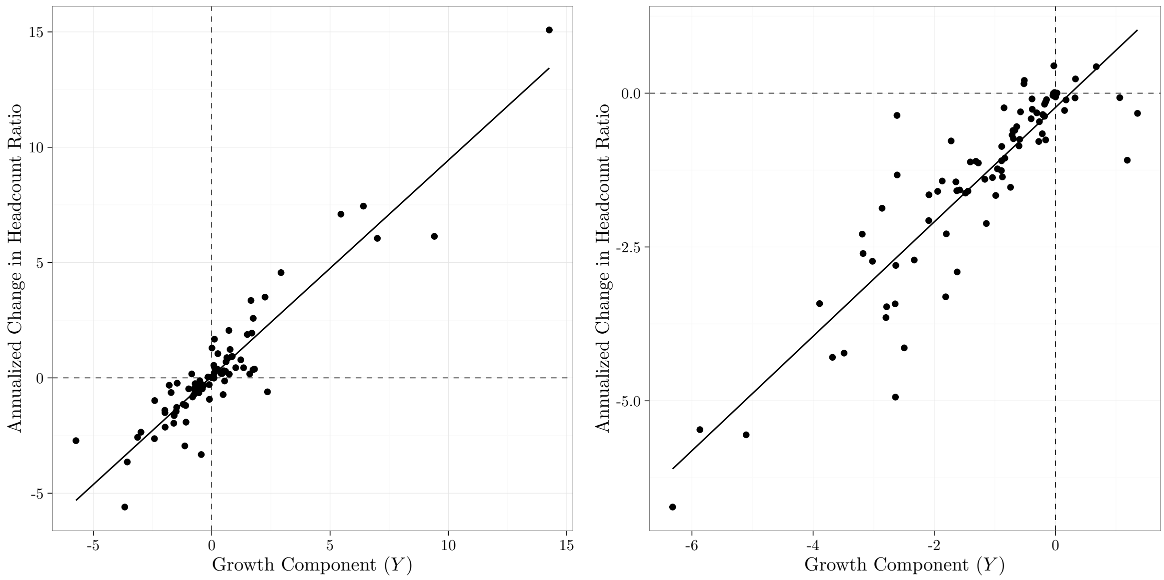
We begin with comparing the in-sample predictions of our model for 2010 to the official World Bank data to show that they match up well. Then, we extrapolate poverty until 2030. Since our model can predict poverty responses to *any* combination of shifts in mean income and inequality, we can consider a variety of scenarios.

Figure 3 – Estimated components from poverty accounting, \$2 a day

(a) Growth versus distribution component, pre and post 2000



(b) Observed change versus growth component, pre and post 2000



Notes: Illustration of the results from the poverty decomposition. Panel a) plots the growth component against the distribution component for the pre-2000 sample (on the left) and for the post-2000 sample (on the right). Panel b) plots the change in the poverty rate against the distribution component for the pre-2000 sample (on the left) and for the post-2000 sample (on the right). The number of dots in a graph may exceed the number of countries, since some countries have both a consumption and an income survey spell.

As most household surveys are not repeated annually, the official World Bank regional poverty figures involve a considerable amount of interpolation and extrapolation (see for details Chen and Ravallion, 2004). Our method is similar in spirit. It involves four steps. First, we extrapolate the last available survey income to 2010 using actually observed country growth rates in personal consumption expenditures per capita (PCE_{it}) from the national accounts.[22] Second, we project mean income into the future, using one of three constant-growth scenarios and one of three distributional scenarios. Third, we predict each country's poverty rate in 5-year intervals from 2010 to 2030 using the estimates from column (2) in Table 2.[23] We may neglect measurement errors in income or inequality for forecasting purposes, implicitly assuming that their influence remains stable over time. Fourth, we calculate regional and global aggregates as population-weighted averages of the country estimates using World Bank population projections.

We base our three different constant-growth scenarios on the historical PCE_{it} growth rates in three predefined periods. An 'optimistic' scenario uses the average PCE_{it} growth rate of each country over 2000-2010, a decade characterized by fast growth. A 'moderate' growth scenario uses the average PCE_{it} growth rate of each country over 1980-2010, the long-run average over the entire sample period. Finally, a 'pessimistic' growth scenario uses the 1980-2000 growth rates, a period during which mean consumption in Sub-Saharan Africa, in particular, shrank at a rate of about 0.82% per year.[24] Table C-2 in Appendix C reports population-weighted regional growth rates over these periods to illustrate the implied regional income dynamics.

For each growth scenario, we also simulate three different inequality patterns. 'Pro-poor growth' implies an annual decline in the Gini coefficient of approximately -0.5%; 'distribution-neutral growth' keeps inequality constant at the level prevailing in 2010, in line with the observed zero correlation between changes in inequality and income growth over the entire sample; and 'pro-rich growth' implies an annual increase in the Gini coefficient of approximately 0.5%. For illustration consider the following example. If a country's Gini coefficient is 0.40 in 2010 and we apply the pro-poor pattern, then by 2030 we project a Gini coefficient of about 0.36; if we apply the pro-rich pattern, then the Gini coefficient is about 0.44 in 2030. Changes of this magnitude are in line with the population-weighted regional trends obtained from the surveys.

Our in-sample estimates for 2010 compare well with the 'official' World Bank figures. We can almost perfectly match the World Bank's results for the world total. We estimate a poverty rate of 40.37% in 2010 at the $2 a day poverty line, whereas the World Bank reports 40.67%.[25] Using the same population data, our estimates imply about 2.378 billion people under the $2 line versus 2.395 billion as reported by the World Bank. For

[22] The term 'national accounts' refers to data from the World Development Indicators or the Penn World Table 7.1, whichever has more data over the 30 year horizon.

[23] Although this specification is only estimated on the sub-sample where $T_i \geq 2$, we can use the estimates to predict poverty for the entire sample ($T_i \geq 1$). We only lack estimates of the panel size effects for $T_i = 1$, so we assign these observations to the adjacent group ($T_i = 2$).

[24] Owing to the post-communist transition, consumption and incomes in Europe and Central Asia were shrinking over the same period. However, given the small number of poor in 2010, the influence of that region on the global poverty headcount in 2030 is minimal.

[25] This was the official number until the Oct. 9, 2014 update of *PovcalNet*.

three regions, our estimates are within one percentage point of the official figures; for the other three, they are within 3 percentage points.[26]

To see how reliable our approach is out of sample, we conduct two cross-validation exercises in Appendix C.[27] Table C-5 uses the data before 2005 to predict regional poverty rates in the 2005–2010 period. The absolute errors are below 2 percentage points in each region and trivially small for the entire developing world. Table C-6 conducts a 10-fold cross-validation where 10% of the time periods are deleted in each fold but the panel is otherwise kept intact. We find a cross-validation error of about 3.6 percentage points and a mean absolute error (MAE) of about 2.5 percentage points – less than twice the in-sample error. Hence, our model has a decent out-of-sample fit for individual observations and very accurately forecasts regional or developing world aggregates.

Table 6 shows our projections for 2030 under the $2 poverty line. Our moderate growth scenario predicts that about 1.87 billion people (26%) live on less than $2 a day in 2030, vs. 2.4 billion (40.37%) in 2010. However, much greater gains are possible. Global poverty under the $2 line falls by half to less than 20% of the developing world's population in the optimistic growth scenario with distribution-neutral or pro-poor growth. If this happens by 2030, then more than one billion people will have left poverty at the $2 line – undeniably a remarkable achievement.

Examining the regional distribution, we find that $2 poverty in East Asia is likely to fall to around 5% by 2030, down from 29.7% in 2010. Nearly everyone in East Asia will have entered the middle class by developing-country standards. Progress in South Asia is also likely to be rapid. According to our moderate growth estimate the expected poverty rate is 35.9% in 2030, meaning about 716 million poor, down from 66.7% or 1.1 billion poor in 2010. In the optimistic pro-poor growth case, the headcount ratio falls further to less than 20% and the number of poor to less than 400 million. In stark contrast, the $2 a day poverty rate in Sub-Saharan Africa is expected to remain very high. Our moderate growth scenario predicts a poverty rate of about 66%, down from 69.9% in 2010, which at current population projections implies almost one billion poor in Sub-Saharan Africa alone. Even in the optimistic and pro-poor growth scenario, we project a poverty rate of over 50% and more than 700 million poor. The bulk of the consumption distribution is too far below the $2 a day poverty line in 2010 for most of the subcontinent. Hence, poverty alleviation in Sub-Saharan Africa remains the primary development challenge of the first half of the 21st century.

These observations can in part be explained by a process of 'bunching up above $1.25 a day and just below $2 a day' occurring in East Asia and, to a lesser extent, in South Asia over the last two decades (Chen and Ravallion, 2010). These two regions have a

[26] For Sub-Saharan Africa, for instance, we estimate a poverty rate of 69.36% and the World Bank reports 69.87%. For East Asia, we estimate 26.78%, whereas the World Bank figure is 29.73%.

[27] We are indebted to an anonymous referee for suggesting this to us.

Table 6 – Projected poverty headcount ratios and poor population at $2 a day in 2030, by region

	Average PCE Growth								
	Optimistic (2000-2010)			Moderate (1980-2010)			Pessimistic (1980-2000)		
	Change in Inequality (Gini)								
	pro-poor	neutral	pro-rich	pro-poor	neutral	pro-rich	pro-poor	neutral	pro-rich
Panel (a) – Headcount at $2 a day in 2030 (in percent)									
East Asia and Pacific	3.79	4.61	5.55	4.03	4.90	5.90	4.39	5.31	6.36
Europe and Central Asia	0.45	0.56	0.69	2.80	3.13	3.49	9.35	10.28	11.31
Latin America and Caribbean	4.00	4.73	5.59	6.29	7.39	8.66	8.57	9.99	11.60
Middle East and North Africa	2.85	3.55	4.39	7.20	8.62	10.25	12.86	14.88	17.12
South Asia	19.83	23.12	26.74	31.60	35.88	40.39	40.35	45.00	49.73
Sub-Saharan Africa	51.62	54.56	57.46	63.67	66.36	68.98	70.93	73.32	75.63
Total	17.36	19.23	21.24	23.73	25.94	28.26	28.71	31.07	33.53
Panel (b) – Poor population at $2 a day in 2030 (in millions)									
East Asia and Pacific	82.42	100.20	120.61	87.49	106.39	128.23	95.44	115.26	138.11
Europe and Central Asia	2.12	2.63	3.26	13.24	14.79	16.49	44.25	48.63	53.50
Latin America and Caribbean	28.39	33.63	39.73	44.71	52.52	61.50	60.90	70.99	82.42
Middle East and North Africa	12.64	15.74	19.47	31.94	38.22	45.46	57.02	65.95	75.91
South Asia	395.33	461.03	533.21	629.97	715.46	805.25	804.56	897.12	991.46
Sub-Saharan Africa	723.19	764.47	805.00	892.01	929.76	966.48	993.71	1027.19	1059.59
Total	1249.19	1383.61	1528.08	1707.20	1866.01	2033.35	2065.12	2235.43	2412.30

Notes: The table reports forecasts of the $2 a day poverty rate in 2030. The forecasts are based on the estimates reported in Column (2) of Table 2 and the different growth/ distribution scenarios outlined in the text. Population projections are from the World Bank's Health, Nutrition and Population Statistics database. The survey data are from the World Bank's *PovcalNet* database.

relatively large population near the poverty line and hence most of the advances are projected to occur there. Latin America and the Caribbean, as well as the Middle East and North Africa, are richer and require stronger income growth to continuously reduce poverty. Sub-Saharan Africa, on the other hand, has a large proportion of poor far below the $2 a day line in 2010 (and almost 50% below $1.25 a day). It is facing a lower income elasticity and an income semi-elasticity that is below its peak. Hence, the subcontinent needs exceptionally strong income growth to make significant strides against poverty.

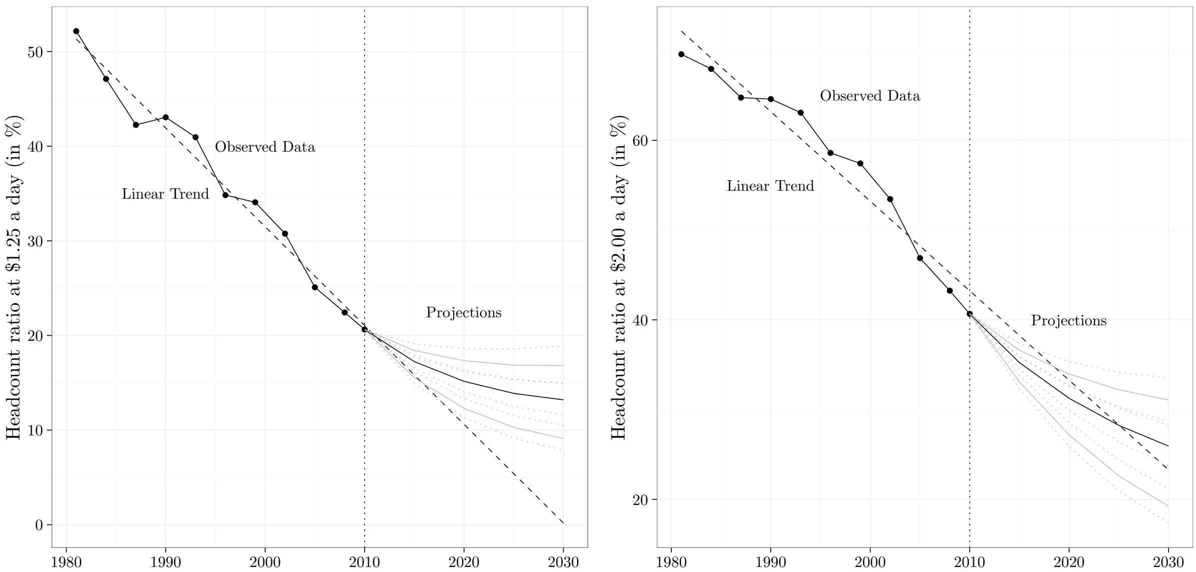
We repeat this exercise using the $1.25 a day poverty line. Table C-10 in Appendix C shows the results based on the estimates of Table C-8. Under moderate growth we project a global poverty rate of 13.2% in 2030, meaning about 950 million extremely poor people versus 1.2 billion in 2010. The pace of poverty reduction will have slowed considerably, both in terms of relative changes and in terms of numbers of poor people. About 70% of the world's poor will live in Sub-Saharan Africa and about 23% in South Asia by 2030. Under optimistic, distribution-neutral growth, we find a global poverty rate of 9.11% in 2030, with about 655 million people remaining extremely poor. About 76% of those live in Sub-Saharan Africa and about 17% in South Asia. Under a pessimistic, distribution-neutral growth scenario, there is almost no progress at all.

What do these results imply for the post-2015 development agenda? In a nutshell, 2030 is unlikely to mark the end of extreme poverty, even under very optimistic assumptions. Figure 4 plots the historical evolution of world poverty from 1981 to 2010, a trend fitted through the observed data and extrapolated until 2030, and our different scenarios. The linear trend serves as a reference for the non-linear projections.

Several points are worth noting. First, only the linear extrapolation predicts an extreme poverty rate in the vicinity of zero by 2030. Regressing the global $1.25 a day poverty rate on time yields a downward slope of about one percentage point per year (as in Ravallion, 2013); at that rate, extreme poverty could be vanquished around 2030. Second, at the $1.25 a day line, all projections show a decelerating trend in poverty reduction, although in the optimistic scenarios the slowdown becomes noticeable later. Third, most scenarios suggest a $1.25 poverty rate higher than 10% in 2030; the optimistic pro-poor and distribution-neutral scenarios project 7.9% and 9.1% respectively. Fourth, most scenarios also exhibit a decelerating pace of poverty reduction at the $2 a day line, but the slowdown tends to occur relatively late.

The changing composition of global poverty has profound implications for the medium-term future. Going forward, fast growing East Asia will contribute less and less to global poverty reduction, especially at the $1.25 a day line, while the share of the global poor residing in Sub-Saharan Africa and South Asia will continue to rise. None of our nine scenarios predict an extreme poverty rate near zero by 2030. This stands in stark contrast to earlier studies and the '3% by 2030' target of the World Bank which became enshrined in the Sustainable Development Goals of the UN. Such an outcome would require an

Figure 4 – Global poverty: observed data, trends and projections, $1.25 and $2 a day



Notes: Illustration of the projections. The left panel shows the projections at the $1.25 a day poverty line; the right panel shows the projections at $2 a day poverty line. The dashed line is a linear trend based on the data until 2010. The black line after 2010 displays the moderate growth scenario, the two gray lines above and below stand for the pessimistic and optimistic growth scenarios with distribution-neutral growth, respectively. Small dashed curves above and below each growth scenario are the associated pro-rich and pro-poor scenarios.

implausible acceleration of growth in many of the poorest countries in the world, coupled with a deceleration of population growth. Even if growth rates in Sub-Saharan Africa were to *double* relative to the post-2000 trend, given the persistent African population growth, the global extreme poverty rate would still be well above 3% in 2030.

Two caveats are in order when it comes to comparing our estimates to recent figures from the World Bank. First, the World Bank switched from reporting global poverty in percent of the *developing world* population to percent of the *global* population. This shaves a few percentage points off their estimates and makes the goal seem more attainable. Second, although the recently released 2011 PPPs suggest that the poorest regions are a little less poor and middle-income regions a little poorer, the global poverty rate in 2010 only changed by about one tenth of a percentage point. This is trivial in comparison to the fundamental uncertainties involved in estimating such a global number.

6 Concluding remarks

In this paper we develop a new, unified poverty accounting framework, apply this framework to study the pace of global poverty reduction using a large cross-country data set, and show that it generates interesting and important findings.

Our approach has several desirable features lacking in the linear models with interactions used so far in the literature. Above all, it accommodates the fractional nature of the poverty headcount ratio and closely approximates the shape of the Lorenz curve near the poverty line using only two summary measures of the distribution: mean income and the Gini coefficient. It does not rely on log-normality, is not sensitive to small variations of poverty rates near zero, and achieves a high degree of data congruence. The approximation errors of the regional aggregates are generally small and negligible in comparison to those of the traditional linear models. The approach is also general enough to be applicable in other contexts than global poverty; it can, for example, be used to decompose national poverty rates across geographical areas or population groups.

Applying the new framework to our data, we show that differences in income and inequality levels across countries induce strong regional heterogeneity in the estimated poverty responses. The model helps us to highlight that poverty semi-elasticities rather than elasticities are the relevant parameters for policy makers, and that these quantities are distinctly non-linear. We confirm earlier reports that growth was chiefly responsible for poverty reduction in the pre-2000 period, but find that changes in inequality play a much larger role since the turn of the millennium. The evolution of inequality matters a lot in the battle against poverty, and not just in the short run. To further demonstrate the flexibility of our framework, we project global and regional poverty rates from 2010 until 2030 based on several scenarios about how historical consumption growth and patterns of inequality extend into the medium-term future. We find that another billion people may be lifted out of $2 a day poverty by 2030, but that at the $1.25 a day poverty line progress is bound to slow down and the goal of ending extreme poverty within a generation will be all but impossible to achieve.

Our results should not be very sensitive to the use of 2005 rather than the new 2011 PPPs, provided we shift the international poverty lines accordingly to $1.90 and $3.10 a day. The new poverty lines were re-drawn with the explicit aim of keeping the global yardstick constant (see Ferreira et al., 2015).

Finally, while economic growth remains essential in the developing world, there is a potentially large 'double dividend' to be reaped if growth can be achieved in combination with simultaneous reductions in inequality. Pro-poor redistribution diminishes poverty by definition, but it can also make economic growth much more effective at reducing poverty. However important these mechanics, it should be remembered that while poverty accounting can attribute poverty changes to its proximate determinants it does not constitute a full causal model of poverty; it is only the last link in a long chain of events.

References

Aaron, H. (1967). The foundations of the "war on poverty" reexamined. *American Economic Review 57*(5), 1229–1240.

Adams, R. H. (2004). Economic growth, inequality and poverty: Estimating the growth elasticity of poverty. *World Development 32*(12), 1989–2014.

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica 61*, 821–856.

Barro, R. J. and J. W. Lee (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics 104*, 184–198.

Beck, T., A. Demirgüç-Kunt, and R. Levine (2010). Financial institutions and markets across countries and over time: The updated financial development and structure database. *World Bank Economic Review 24*(1), 77–92.

Besley, T. and R. Burgess (2003). Halving global poverty. *Journal of Economic Perspectives 17*(3), 3–22.

Blundell, R. W. and J. L. Powell (2004). Endogeneity in semiparametric binary response models. *Review of Economic Studies 71*(3), 655–679.

Bourguignon, F. (2003). The growth elasticity of poverty reduction: Explaining heterogeneity across countries and time periods. In T. S. Eicher and S. J. Turnovsky (Eds.), *Inequality and Growth: Theory and Policy Implications*, pp. 3–26. Cambridge, MA: MIT Press.

Bresson, F. (2009). On the estimation of growth and inequality elasticities of poverty with grouped data. *Review of Income and Wealth 55*(2), 266–302.

Caselli, F. (2005). Accounting for cross-country income differences. Volume 1, Part A of *Handbook of Economic Growth*, Chapter 9, pp. 679–741. Elsevier.

Chamberlain, G. (1984). Panel data. Volume 2 of *Handbook of Econometrics*, pp. 1247–1318. Elsevier.

Chambers, D. and S. Dhongde (2011). A non-parametric measure of poverty elasticity. *Review of Income and Wealth 57*(4), 683–703.

Chen, S. and M. Ravallion (2004). How have the world's poorest fared since the early 1980s? *The World Bank Research Observer 19*(2), 141–169.

Chen, S. and M. Ravallion (2010). The developing world is poorer than we thought, but no less successful in the fight against poverty. *Quarterly Journal of Economics 125*(4), 1577–1625.

Chesher, A. and C. Schluter (2002). Welfare measurement and measurement error. *Review of Economic Studies 69*(2), 357–378.

Chotikapanich, D., W. E. Griffiths, and D. S. P. Rao (2007). Estimating and combining national income distributions using limited data. *Journal of Business & Economic Statistics 25*(1), 97–109.

Datt, G. and M. 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*(2), 275–295.

Dollar, D., T. Kleineberg, and A. Kraay (2016). Growth still is good for the poor. *European Economic Review 81*, 68–85.

Dollar, D. and A. Kraay (2002). Growth is good for the poor. *Journal of Economic Growth 7*(3), 195–225.

Ferreira, F. H., S. Chen, Y. Dikhanov, N. Hamadeh, D. Jolliffe, A. Narayan, E. B. Prydz, A. Revenga, P. Sangraula, U. Serajuddin, and N. Yoshida (2015, October). A global count of the extreme poor in 2012: data issues, methodology and initial results. Policy Research Working Paper 7432, The World Bank.

Gourieroux, C., A. Monfort, and A. Trognon (1984). Pseudo maximum likelihood methods: Theory. *Econometrica 52*(3), 681–700.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica 46*(6), 1251–1271.

Hoover, G. A., W. Enders, and D. G. Freeman (2008). Non-white poverty and macroeconomy: The impact of growth. *American Economic Review 98*(2), 398–402.

Kakwani, N. (1993). Poverty and economic growth with application to Côte d'Ivoire. *Review of Income and Wealth 39*(2), 121–139.

Kalwij, A. and A. Verschoor (2007). Not by growth alone: The role of the distribution of income in regional diversity in poverty reduction. *European Economic Review 51*(4), 805–829.

Klasen, S. and M. Misselhorn (2008). Determinants of the growth semi-elasticity of poverty reduction. Working paper, Ibero America Institute for Economic Research.

Klenow, P. and A. Rodriguez-Clare (1997). The neoclassical revival in growth economics: Has it gone too far? In B. S. Bernanke and J. J. Rotemberg (Eds.), *NBER Macroeconomics Annual 1997*, Volume 12, pp. 73–103. Cambridge, MA: MIT Press.

Kraay, A. (2006). When is growth pro-poor? Evidence from a panel of countries. *Journal of Development Economics 80*(1), 198–227.

Kraay, A., T. Kleineberg, and D. Dollar (2014). Growth, inequality, and social welfare: Cross-country evidence. Working paper, Centre for Economic Policy Research.

Krause, M. (2014). Parametric lorenz curves and the modality of the income density function. *Review of Income and Wealth 60*(4), 905–929.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica 46*(1), 69–85.

Papke, L. E. and J. M. Wooldridge (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics 11*(6), 619–632.

Papke, L. E. and J. M. Wooldridge (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics 145*(1–2), 121–133.

Ravallion, M. (2001). Growth, inequality and poverty: Looking beyond averages. *World Development 29*(11), 1803–1815.

Ravallion, M. (2013). How long will it take to lift one billion people out of poverty? *World Bank Research Observer 28*(2), 139–158.

Ravallion, M. and S. Chen (1997). What can new survey data tell us about recent changes in distribution and poverty? *World Bank Economic Review 11*(2), 357–382.

Ravallion, M. and S. Chen (2003). Measuring pro-poor growth. *Economics Letters 78*(1), 93–99.

Rivers, D. and Q. H. Vuong (1988). Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of Econometrics 39*(3), 347–366.

Santos Silva, J. and S. Tenreyro (2006). The log of gravity. *Review of Economics and Statistics 88*(4), 641–658.

Shorrocks, A. F. (2013). Decomposition procedures for distributional analysis: a unified framework based on the Shapley value. *Journal of Economic Inequality 11*(1), 99–126.

Wacziarg, R. and K. H. Welch (2008). Trade liberalization and growth: New evidence. *World Bank Economic Review 22*(2), 187–231.

Wooldridge, J. M. (2010a, May). Correlated random effects models with unbalanced panels. Manuscript.

Wooldridge, J. M. (2010b). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). The MIT Press.

World Bank (2005). World development report 2006: Equity and development. Technical report, The World Bank.

World Bank (2015). A measured approach to ending poverty and boosting shared prosperity. Technical report, The World Bank.

Young, A. (2011, December). The Gini coefficient for a mixture of ln-normal populations. Manuscript. The London School of Economics and Political Science, London, UK.

For Online Publication: Appendix A

Let the poverty line (z) be fixed and assume incomes follow a two-parameter distribution, so that the poverty headcount may be written as $H(\bar{y}_t/z, \sigma_t) = H(\bar{y}_t, \sigma_t) = H_t$. A Taylor linearization of $H(\cdot)$ about $(\bar{y}_t, \sigma_t)$ gives

$$H(\bar{y}_t + \mathrm{d}\bar{y}_t, \sigma_t + \mathrm{d}\sigma_t) = H(\bar{y}_t, \sigma_t) + \frac{\partial H_t}{\partial \bar{y}_t}\mathrm{d}\bar{y}_t + \frac{\partial H_t}{\partial \sigma_t}\mathrm{d}\sigma_t + \xi_t \tag{A-1}$$

where dx denotes a differential of x, and ξ_t is a second-order remainder. This approach can be extended to allow for a vector of Lorenz curve parameters as in Kakwani (1993), making a very general treatment possible.

Subtracting $H(\bar{y}_t, \sigma_t)$ from both sides, dropping the remainder by approximation, dividing through by H_t (assuming $H_t > 0$), and multiplying the first (second) term by $\bar{y}_t/\bar{y}_t$ (σ_t/σ_t), we arrive at eq. (4) from the main text:

$$\frac{\mathrm{d}H_t}{H_t} \approx \left(\frac{\partial H_t}{\partial \bar{y}_t}\frac{\bar{y}_t}{H_t}\right)\frac{\mathrm{d}\bar{y}_t}{\bar{y}_t} + \left(\frac{\partial H_t}{\partial \sigma_t}\frac{\sigma_t}{H_t}\right)\frac{\mathrm{d}\sigma_t}{\sigma_t} = \varepsilon_t^{H\bar{y}}\frac{\mathrm{d}\bar{y}_t}{\bar{y}_t} + \varepsilon_t^{H\sigma}\frac{\mathrm{d}\sigma_t}{\sigma_t}. \tag{A-2}$$

If we omit the division by H_t, we get a decomposition of dH_t, the (non-relative) change of the poverty rate, in terms of income and inequality *semi*-elasticities.

Similar steps starting from $H(\bar{y}_t, G_t)$ lead to a decomposition in terms of mean income and Gini. Using the chain rule, we have

$$\frac{\mathrm{d}H_t}{H_t} \approx \varepsilon_t^{H\bar{y}}\frac{\mathrm{d}\bar{y}_t}{\bar{y}_t} + \varepsilon_t^{HG}\frac{\mathrm{d}G_t}{G_t} = \varepsilon_t^{H\bar{y}}\frac{\mathrm{d}\bar{y}_t}{\bar{y}_t} + \varepsilon_t^{H\sigma}\left(\frac{\mathrm{d}G_t}{\mathrm{d}\sigma_t}\frac{\sigma_t}{G_t}\right)^{-1}\frac{\mathrm{d}G_t}{G_t}. \tag{A-3}$$

Under log-normality, eqs. (2) and (3) of the main text give $\varepsilon_t^{H\bar{y}}$ and $\varepsilon_t^{H\sigma}$, and σ_t is linked to G_t by $\sigma_t = \sqrt{2}\Phi^{-1}(G_t/2 + 1/2)$. Hence

$$\frac{\mathrm{d}G_t}{\mathrm{d}\sigma_t} = \frac{\mathrm{d}[2\Phi\left(\sigma_t/\sqrt{2}\right) - 1]}{\mathrm{d}\sigma_t} = \sqrt{2}\phi\left(\frac{\sigma_t}{\sqrt{2}}\right), \tag{A-4}$$

giving us an explicit formula for the Gini elasticity:

$$\begin{aligned}\varepsilon_t^{HG} &= \varepsilon_t^{H\sigma}\left(\frac{\mathrm{d}G_t}{\mathrm{d}\sigma_t}\frac{\sigma_t}{G_t}\right)^{-1} \\ &= \left(\frac{\ln(\bar{y}_t/z)}{\sigma_t} + \frac{1}{2}\sigma_t\right)\left(\frac{\sigma_t}{G_t}\sqrt{2}\phi\left(\frac{\sigma_t}{\sqrt{2}}\right)\right)^{-1}\lambda\left(\frac{-\ln(\bar{y}_t/z)}{\sigma_t} + \frac{1}{2}\sigma_t\right).\end{aligned} \tag{A-5}$$

Clearly, the Gini elasticity is a highly non-linear function, and the same is true of the Gini semi-elasticity, as illustrated in Figure 2.

Let us now put aside log-normality and return to the fractional response approach, allowing for any two-parameter income distribution in which headcount rates can be approximated by smooth functions of mean income and inequality. As in Section 3, we allow for heterogeneity, unbalanced data, and possibly measurement error in income. The conditional expectation function of the headcount rate is given by eq. (11) of the main text. Define the linear predictors inside the cumulative normal as $\mathbf{m}'_{it1}\hat{\boldsymbol{\xi}}_1$ for the numerator (main equation) and $\mathbf{m}'_{it2}\hat{\boldsymbol{\xi}}_2$ for the denominator (variance equation). Then the estimated elasticities and semi-elasticities with respect to a variable x_k (appearing in logs) are, respectively,

$$\hat{\varepsilon}_t^{Hx_k} = \hat{\beta}_k \times N^{-1} \sum_{i=1}^{N} \exp\left(-\mathbf{m}'_{it2}\hat{\boldsymbol{\xi}}_2\right) \lambda\left(\mathbf{m}'_{it1}\hat{\boldsymbol{\xi}}_1/\exp(\mathbf{m}'_{it2}\hat{\boldsymbol{\xi}}_2)\right) \tag{A-6}$$

and

$$\hat{\eta}_t^{Hx_k} = \hat{\beta}_k \times N^{-1} \sum_{i=1}^{N} \exp\left(-\mathbf{m}'_{it2}\hat{\boldsymbol{\xi}}_2\right) \phi\left(\mathbf{m}'_{it1}\hat{\boldsymbol{\xi}}_1/\exp(\mathbf{m}'_{it2}\hat{\boldsymbol{\xi}}_2)\right). \tag{A-7}$$

We may plug in any interesting values for incomes/expenditures and the Gini, and average over all panel sizes T_i.

For Online Publication: Appendix B

Approximation errors in the estimated elasticities

Our benchmark is a balanced panel of average incomes, Gini coefficients and poverty rates derived from a log-normal distribution. The log-normal panel DGP is defined as follows. For each replication we generate $\mu_{it} = 4 + e_{it}$ with $e_{it} \sim \mathcal{U}[-1, 1]$ and $\sigma_{it} = 1 + u_{it}$ with $u_{it} \sim \mathcal{U}[-.25, .25]$ where $i = 1, \dots, N$ and $t = 1, \dots, T$. The corresponding formulas for mean income, the Gini coefficient and the headcount ratio are given in Table B-2.

The observed headcount is the average of a sequence of Bernoulli draws, i.e. $H_{it} \sim \mathcal{B}(B, H^*_{it})/B$, where B is the number of trials and H^*_{it} is the poverty rate drawn from the specified distribution. R is the number of Monte Carlo replications. We typically set $N = 100$, $T = 10$, $B = 1000$ and $R = 1000$. This mimics key features of the data analyzed in the main text but abstracts from measurement errors and endogeneity.

Table B-1 – Monte Carlo – GLM versus OLS

	Actual	*Fractional response approach*			*Improved standard model*		
		Estimate	MRE	$\sqrt{MSE}$	Estimate	MRE	$\sqrt{MSE}$
		Panel a) Income elasticity of poverty $\varepsilon^{H\bar{y}}$					
Mean	-1.1091	-1.0875	0.0195	0.0225	-1.0978	0.0103	0.0172
At percentile of $\bar{y}$							
$\dots P_5$	-0.4924	-0.4793	0.0267	0.0132	-0.9396	0.9081	0.4477
$\dots P_{25}$	-0.7534	-0.7238	0.0392	0.0297	-1.0133	0.3450	0.2606
$\dots P_{50}$	-1.1115	-1.0597	0.0465	0.0521	-1.0978	0.0123	0.0240
$\dots P_{75}$	-1.5174	-1.4421	0.0496	0.0760	-1.1823	0.2208	0.3358
$\dots P_{95}$	-1.8947	-1.7990	0.0505	0.0967	-1.2553	0.3375	0.6401
		Panel b) Income semi-elasticity of poverty $\eta^{H\bar{y}}$					
Mean	-0.3277	-0.3280	0.0008	0.0018	-0.3466	0.0577	0.0191
At percentile of $\bar{y}$							
$\dots P_5$	-0.3489	-0.3370	0.0341	0.0120	-0.3716	0.0650	0.0235
$\dots P_{25}$	-0.4020	-0.3927	0.0232	0.0095	-0.3599	0.1047	0.0422
$\dots P_{50}$	-0.3664	-0.3716	0.0143	0.0057	-0.3466	0.0539	0.0203
$\dots P_{75}$	-0.2599	-0.2747	0.0572	0.0149	-0.3333	0.2827	0.0737
$\dots P_{95}$	-0.1624	-0.1735	0.0684	0.0111	-0.3218	0.9815	0.1595

Notes: The table reports the results of 1000 Monte Carlo simulations based on log-normal data with 1000 observations (100 countries over 10 years). Panels a) an b) show the results from applying the proposed panel GLM estimator and OLS with interactions to the simulated data. The former reports elasticities, the latter reports semi-elasticities.

To benchmark the performance of our model against the standard approach, we then compute *(i)* the average of the actual income (semi-)elasticities at each point, *(ii)* the average of the estimated (semi-)elasticities from a simplified fractional response model, and *(iii)* the estimated (semi-)elasticities from an OLS regression in differences with interactions (with and without logs on the left hand side). The corresponding formulas are eqs. (A-6) and (A-7) in Appendix A and eq. (5) in the main text.

Table B-1 reports the results. As expected, both models perform very well if we only focus on averages. Panel a) shows that the mean relative error (MRE) of the income elasticity is below two percent in both cases. However, the fractional response model performs uniformly better once we focus on predictions away from the overall mean (median). The MRE is always below or near five percent even in the tails of the distribution of $\bar{y}_{it}$. The linear model with interactions is strongly biased (e.g. up to 90% at the fifth percentile of $\bar{y}_{it}$). Panel b) focuses on income semi-elasticities – the natural parameter of the fractional response model. Now the fractional response approach also outperforms the linear model on average and remains accurate for semi-elasticities in the tails of the income distribution. Note that while we restricted our attention to income effects, the results for the Gini elasticities and semi-elasticities are comparable.

Approximation errors in the estimated headcount ratios

We now show that our approach also works well when incomes are drawn from a variety of two and three-parameter distributions. We exclusively focus on the ability of our model to recover the cross-country distribution of poverty rates. The rationale behind this choice is simple. Elasticities and semi-elasticities are point-derivatives at the estimated poverty rates. If our model does a good job recovering the latter, the former are likely to be closely approximated as well. There is also a pragmatic reason. More complicated distributions often do not have closed-form solutions for the income or inequality elasticities (see Bresson, 2009, for an alternate approach using Lorenz curves). We do not compare these results to the standard linear approach, since it is formulated in differences and not designed to predict poverty rates directly.

We proceed in line with the simulations presented above. Table B-2 lists the DGPs and characteristics of the five chosen functional forms. The panel is still balanced with $N = 100$ and $T = 10$. We generate random data by averaging over 1000 Bernoulli draws and run 1000 simulations. Table B-2 reports the corresponding results. For each of the five distributions, we compute the average poverty rate and five percentiles of the distribution. Our estimates are usually within a percentage point of the generated data. Noticeable discrepancies only occur in the lowest and highest percentiles. They are also somewhat larger in the lower percentiles of the three-parameter distributions than in the two-parameter cases. There, the MRE is sizable, although the percentage point differences remain moderately small.

We have also computed an error which focuses on the individual observations in each replication data set. Specifically, we compute the square root of the average MSEs over

Table B-2 – Alternate distributions and DGPs

Distribution	Parameters (DGP)	Mean	Gini coefficient	Headcount ratio
			Two parameter distributions	
Log-normal	$\mu = 4 + \mathcal{U}(-1,1)$ $\sigma = 1 + \mathcal{U}(-.25,.25)$	$\exp\left(\mu + \frac{1}{2}\sigma^2\right)$	$2\Phi\left(\frac{\sigma}{\sqrt{2}}\right) - 1$	$\Phi\left(\frac{-\ln(\bar{y}/z)}{\sigma} + \frac{1}{2}\sigma\right)$
Weibull	$a = 1.5 + \mathcal{U}(-.75,.75)$ $b = 75 + \mathcal{U}(-30,30)$	$b\Gamma\left(1 + \frac{1}{a}\right)$	$1 - 2^{-\frac{1}{a}}$	$1 - \exp\left[-\left(\frac{z}{b}\right)^a\right]$
Fisk	see Dagum with $p = 1$			
			Three parameter distributions	
Dagum	$a = 4 + \mathcal{U}(-2,2)$ $p = 1 + \mathcal{U}(-.5,.5)$ $b = 46.5 + \mathcal{U}(-10,10)$	$\frac{b\Gamma(p+1/a)\Gamma(1-1/a)}{\Gamma(p)}$	$\frac{\Gamma(p)\Gamma(2p+1/a)}{\Gamma(2p)\Gamma(p+1/a)}$	$\left[1 + \left(\frac{z}{b}\right)^{-a}\right]^{-p}$
Singh-Maddala	$a = 1.5 + \mathcal{U}(-.5,.5)$ $q = 4 + \mathcal{U}(-2.5,2.5)$ $b = 175 + \mathcal{U}(-25,25)$	$\frac{b\Gamma(1+1/a)\Gamma(q-1/a)}{\Gamma(q)}$	$1 - \left[\frac{\Gamma(q)\Gamma(2q-1/a)}{\Gamma(2q)\Gamma(q-1/a)}\right]$	$1 - \left[1 + \left(\frac{z}{b}\right)^a\right]^{-q}$

Notes: The table summarizes the income distributions and parameters used in the Monte Carlo simulations. $\Phi(\cdot)$ is the cdf of the standard normal distribution and $\Gamma(\cdot)$ is the gamma function. Scale parameters are typically denoted by a (or μ for the log-normal). All other parameters define the shape of the distribution. The poverty line is fixed at $z = 38$ in all simulations. The parameters of the DGPs have been chosen such that they generate an average headcount in the neighborhood of 0.35.

Table B-3 – Monte Carlo – GLM with various income distributions

	Estimate	Actual	Bias	MRE	$\sqrt{MSE}$
Panel a) Log-normal DGP					
Mean of H_{it}	0.3771	0.3770	-0.0001	0.0003	0.0001
$\ldots P_5$	0.0919	0.1010	0.0091	0.0990	0.0096
$\ldots P_{25}$	0.1943	0.1916	-0.0027	0.0139	0.0043
$\ldots P_{50}$	0.3580	0.3564	-0.0016	0.0044	0.0032
$\ldots P_{75}$	0.5538	0.5547	0.0009	0.0016	0.0042
$\ldots P_{95}$	0.7055	0.7064	0.0010	0.0014	0.0047
Panel b) Weibull DGP					
Mean of H_{it}	0.3304	0.3302	-0.0002	0.0006	0.0002
$\ldots P_5$	0.1465	0.1610	0.0145	0.0990	0.0146
$\ldots P_{25}$	0.2377	0.2424	0.0047	0.0198	0.0050
$\ldots P_{50}$	0.3305	0.3179	-0.0126	0.0380	0.0128
$\ldots P_{75}$	0.4196	0.4081	-0.0115	0.0273	0.0119
$\ldots P_{95}$	0.5200	0.5403	0.0203	0.0390	0.0206
Panel c) Fisk DGP					
Mean of H_{it}	0.3275	0.3273	-0.0002	0.0006	0.0002
$\ldots P_5$	0.1403	0.1639	0.0235	0.1677	0.0236
$\ldots P_{25}$	0.2373	0.2377	0.0004	0.0015	0.0026
$\ldots P_{50}$	0.3258	0.3146	-0.0112	0.0343	0.0115
$\ldots P_{75}$	0.4205	0.4109	-0.0096	0.0228	0.0101
$\ldots P_{95}$	0.5118	0.5254	0.0136	0.0267	0.0143
Panel d) Dagum DGP					
Mean of H_{it}	0.3448	0.3442	-0.0006	0.0017	0.0006
$\ldots P_5$	0.1065	0.1429	0.0364	0.3422	0.0366
$\ldots P_{25}$	0.2261	0.2275	0.0013	0.0059	0.0032
$\ldots P_{50}$	0.3400	0.3220	-0.0179	0.0527	0.0182
$\ldots P_{75}$	0.4556	0.4384	-0.0172	0.0378	0.0178
$\ldots P_{95}$	0.6066	0.6271	0.0205	0.0338	0.0218
Panel d) Singh-Maddala DGP					
Mean of H_{it}	0.3268	0.3266	-0.0002	0.0006	0.0002
$\ldots P_5$	0.1202	0.1377	0.0175	0.1458	0.0177
$\ldots P_{25}$	0.2125	0.2113	-0.0012	0.0056	0.0022
$\ldots P_{50}$	0.3048	0.2950	-0.0097	0.0320	0.0100
$\ldots P_{75}$	0.4275	0.4218	-0.0057	0.0133	0.0064
$\ldots P_{95}$	0.5999	0.6123	0.0124	0.0206	0.0130

Notes: The table reports the results of 1000 Monte Carlo simulations based on the DGPs given in Table B-2 with 1000 observations (100 countries over 10 years). Panels a) to c) show the results from applying the proposed panel GLM estimator to the income distribution specified in the heading.

all replications. This error varies from about 2 percentage points (Singh-Maddala) to 4 percentage points (Dagum), suggesting that even at the level of individual observations we are not far off the actual data. This is comparable to the error of our preferred specification in the main text.

For Online Publication: Appendix C

Figure C-1 – Transformed headcount ($2 a day) and log-mean income, by region

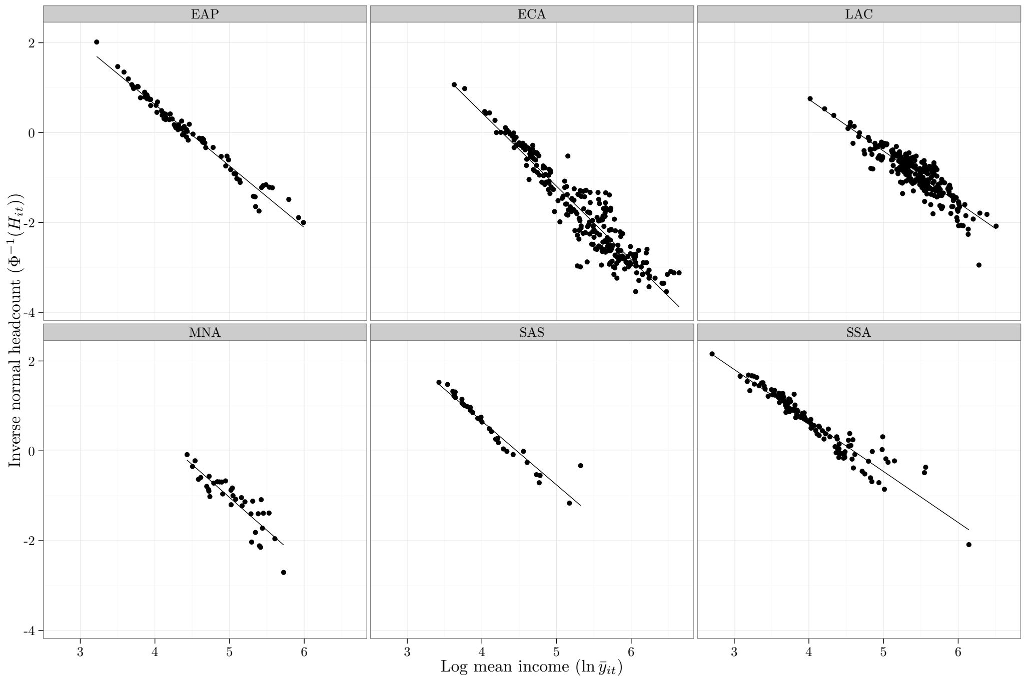


Figure C-2 – Transformed headcount ($2 a day) and log-Gini, by region

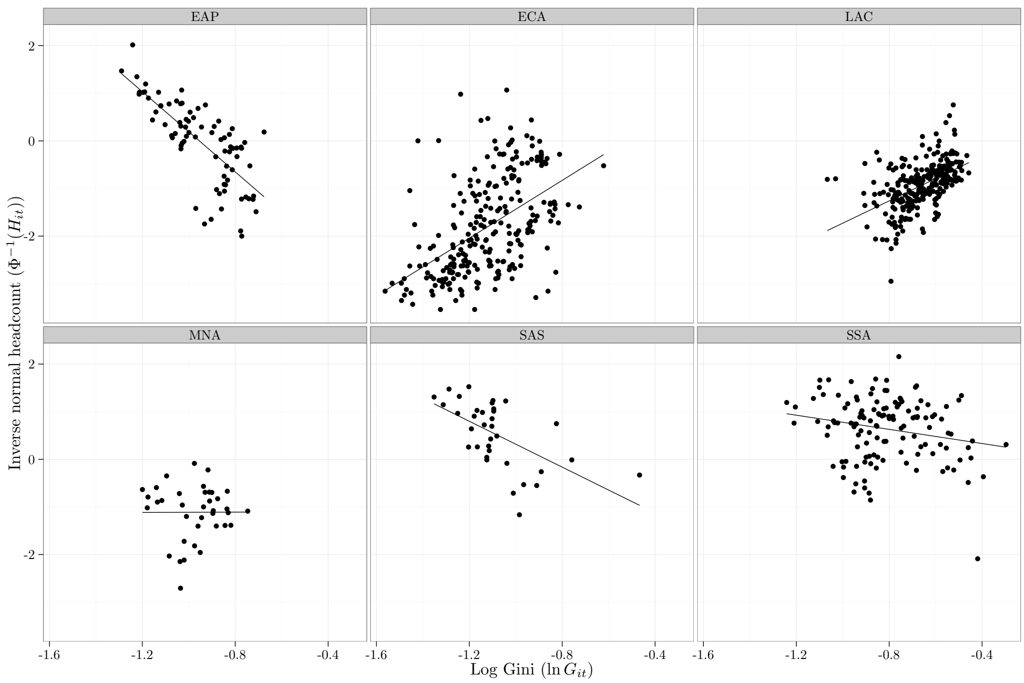
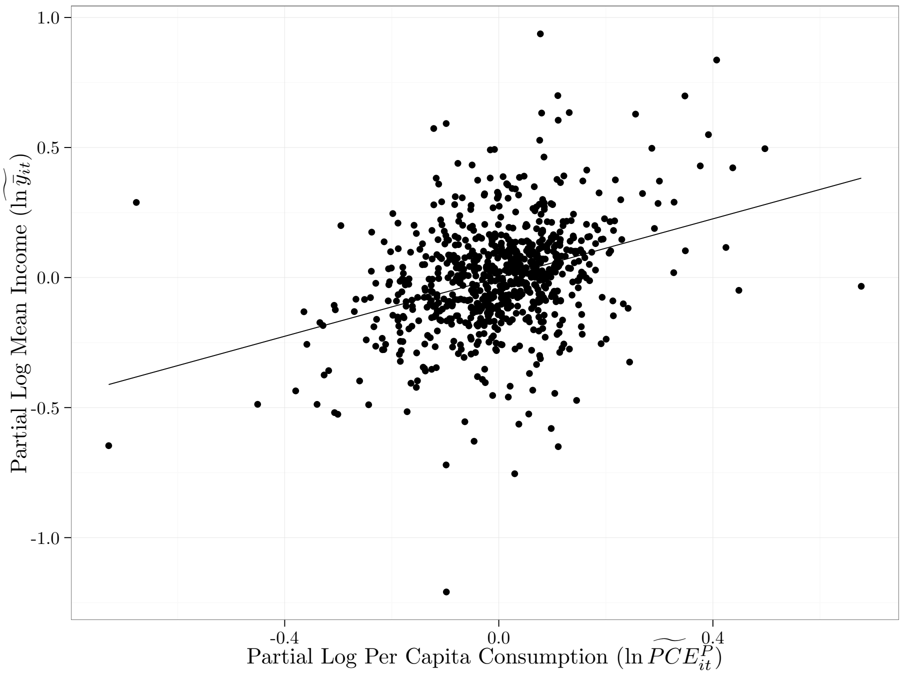


Figure C-3 – Partial regression plot – first stage



Notes: The figure plots two residual series, so that the plotted slope is identical to the slope of $\ln PCE^P_{it}$ in the first stage of the results presented in Table 2. On the x-axis: $\ln \widetilde{PCE^P_{it}} = \ln PCE^P_{it} - \mathbf{x}'_{1it}\hat{\boldsymbol{\beta}}_1 - \sum_{r=1}^{T} \delta_{T_i,r}\hat{\varphi}_{1r} - \sum_{r=1}^{T} \delta_{T_i,r}\bar{\mathbf{x}}'_i\hat{\boldsymbol{\theta}}_{1r}$. On the y-axis: $\widetilde{\ln \bar{y}_{it}} = \ln \bar{y}_{it} - \mathbf{x}'_{1it}\hat{\boldsymbol{\beta}}_1 - \sum_{r=1}^{T} \delta_{T_i,r}\hat{\varphi}_{1r} - \sum_{r=1}^{T} \delta_{T_i,r}\bar{\mathbf{x}}'_i\hat{\boldsymbol{\theta}}_{1r}$. In both cases, $\mathbf{x}'_{1it}$ includes only the log of Gini but $\bar{\mathbf{x}}'_i$ contains the time averages of $\ln G_{it}$ and $\ln PCE^P_{it}$. Both regressions also contain survey type and time dummies, as well as their time averages.

Figure C-4 – Inequality changes and income growth, 1981–2010

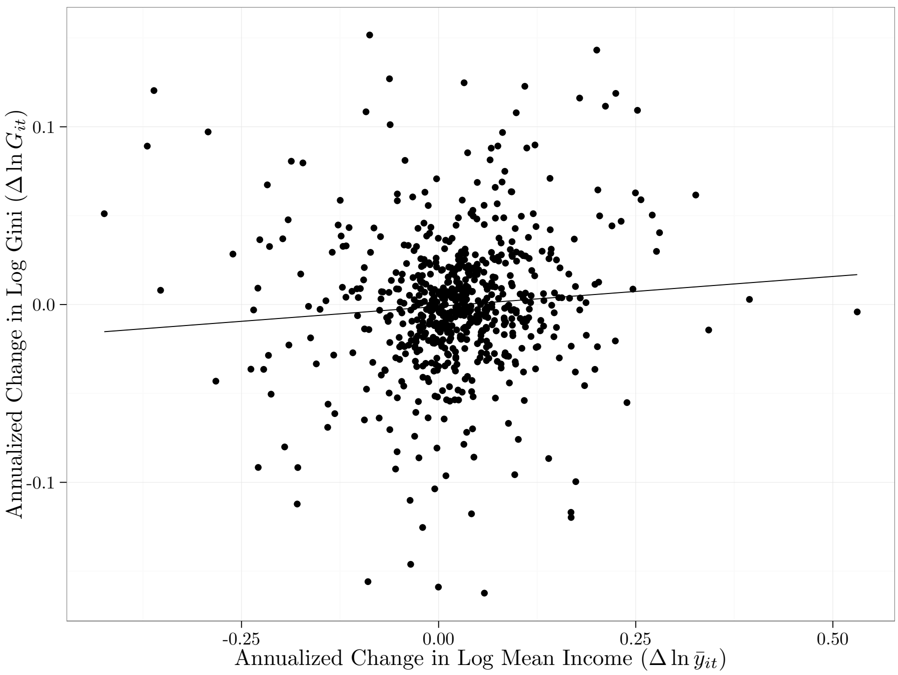
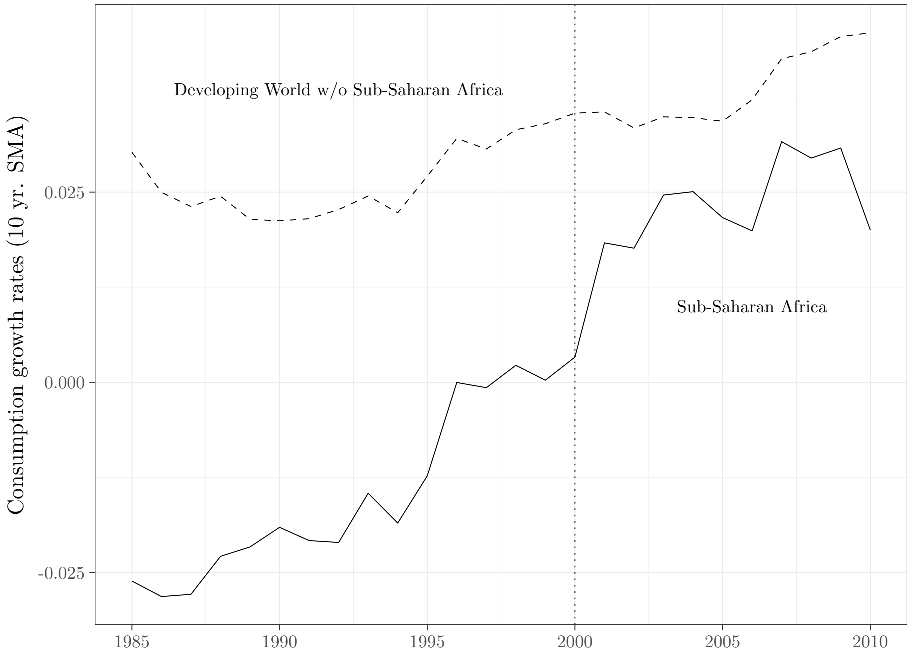


Figure C-5 – Long-run trends in PCE growth – SSA versus developing world, 1985–2010



Notes: The figure plots the long-run trend in per capita consumption expenditure growth in Sub-Saharan Africa (SSA) and the developing world without SSA. The time series have been smoothed using 10 year simple moving averages (SMA). The dotted vertical line marks the year 2000 where the long run trend in SSA experienced a structural break, as indicated by a formal structural change test (Andrews, 1993). The sup-F statistic is 97.43 with an approximate p-value < 0.001.

List C-1 – Included countries (number of surveys)

Albania (5), Algeria (2), Angola (2), Argentina (21), Armenia (11), Azerbaijan (3), Bangladesh (8), Belarus (14), Belize (7), Benin (1), Bhutan (2), Bolivia, Plurinational State of (11), Bosnia and Herzegovina (3), Botswana (2), Brazil (26), Bulgaria (7), Burkina Faso (4), Burundi (3), Cambodia (5), Cameroon (3), Cape Verde (1), Central African Rep. (3), Chad (1), Chile (10), China (16), Colombia (14), Comoros (1), Congo, Dem. Rep. of (1), Congo, Rep. of (1), Costa Rica (23), Cote D'Ivoire (9), Croatia (7), Czech Rep. (2), Djibouti (1), Dominican Rep. (16), Ecuador (12), Egypt (5), El Salvador (14), Estonia (9), Ethiopia (4), Fiji (2), Gabon (1), Gambia (2), Georgia (14), Ghana (5), Guatemala (8), Guinea (4), Guinea-Bissau (2), Guyana (2), Haiti (1), Honduras (20), Hungary (10), India (5), Indonesia (13), Iran, Islamic Rep. of (5), Iraq (1), Jamaica (7), Jordan (7), Kazakhstan (11), Kenya (4), Kyrgyzstan (10), Lao People's Dem. Rep. (4), Latvia (11), Lesotho (4), Liberia (1), Lithuania (9), Macedonia, Rep. of (10), Madagascar (6), Malawi (3), Malaysia (9), Maldives (2), Mali (4), Mauritania (6), Mexico (11), Micronesia, Federated States of (1), Moldova, Rep. of (15), Montenegro (4), Morocco (5), Mozambique (3), Namibia (2), Nepal (4), Nicaragua (4), Niger (4), Nigeria (5), Pakistan (8), Palestinian Territory, Occupied (2), Panama (13), Papua New Guinea (1), Paraguay (14), Peru (16), Philippines (9), Poland (18), Romania (15), Russian Federation (13), Rwanda (3), Saint Lucia (1), Sao Tome and Principe (1), Senegal (4), Serbia (9), Seychelles (1), Sierra Leone (1), Slovakia (7), Slovenia (6), South Africa (5), Sri Lanka (6), Sudan (1), Suriname (1), Swaziland (3), Syrian Arab Rep. (1), Tajikistan (5), Tanzania, United Rep. of (3), Thailand (14), Timor-Leste (2), Togo (1), Trinidad and Tobago (2), Tunisia (6), Turkey (11), Turkmenistan (3), Uganda (7), Ukraine (13), Uruguay (7), Venezuela, Bolivarian Rep. of (13), Vietnam (6), Yemen (2), Zambia (7).

Table C-1 – Summary statistics by region (unweighted)

Variable	Mean	Standard Deviation	Min	Max
	East Asia and Pacific (N=80)			
H_{it} – Headcount ($2)	0.502	0.267	0.023	0.978
G_{it} – Gini coefficient	0.392	0.058	0.275	0.509
$\bar{y}_{it}$ – Mean income or expenditure	107.86	78.39	25.02	399.76
	Eastern Europe and Central Asia (N=254)			
H_{it} – Headcount ($2)	0.110	0.169	0.000	0.857
G_{it} – Gini coefficient	0.330	0.056	0.210	0.537
$\bar{y}_{it}$ – Mean income or expenditure	251.99	136.11	37.66	766.78
	Latin America and Caribbean (N=274)			
H_{it} – Headcount ($2)	0.204	0.122	0.002	0.775
G_{it} – Gini coefficient	0.523	0.054	0.344	0.633
$\bar{y}_{it}$ – Mean income or expenditure	246.63	90.55	55.53	671.04
	Middle East and North Africa (N=37)			
H_{it} – Headcount ($2)	0.166	0.111	0.003	0.466
G_{it} – Gini coefficient	0.380	0.042	0.301	0.474
$\bar{y}_{it}$ – Mean income or expenditure	165.26	56.59	84.02	306.33
	South Asia (N=35)			
H_{it} – Headcount ($2)	0.672	0.226	0.122	0.936
G_{it} – Gini coefficient	0.343	0.067	0.259	0.627
$\bar{y}_{it}$ – Mean income or expenditure	67.78	39.20	30.71	204.98
	Sub-Saharan Africa (N=129)			
H_{it} – Headcount ($2)	0.708	0.202	0.018	0.985
G_{it} – Gini coefficient	0.453	0.087	0.289	0.743
$\bar{y}_{it}$ – Mean income or expenditure	67.62	54.04	14.93	465.80

Notes: The table reports regional summary statistics. Mean income or expenditure is reported in 2005 PPP international dollars per month. 809 observations, 124 countries in total, unbalanced sample from 1981 to 2010.

Table C-2 – Growth in personal consumption expenditures per capita (in %), by region

	Time Period				
	2000-2010	1990-2010	1980-2010	1990-2000	1980-2000
East Asia and Pacific	5.906	5.772	5.598	5.608	5.377
	(0.813)	(0.653)	(0.725)	(0.508)	(0.677)
Europe and Central Asia	6.085	2.755	2.558	-1.225	-0.769
	(0.989)	(0.412)	(0.411)	(1.027)	(0.916)
Latin America and the Caribbean	2.444	2.219	1.445	1.931	0.677
	(0.239)	(0.140)	(0.098)	(0.337)	(0.171)
Middle East and North Africa	3.495	2.532	1.851	1.253	0.495
	(0.443)	(0.440)	(0.293)	(0.648)	(0.545)
South Asia	4.448	3.612	3.179	2.511	2.173
	(0.489)	(0.388)	(0.351)	(0.294)	(0.284)
Sub-Saharan Africa	2.382	1.419	0.698	0.016	-0.818
	(0.689)	(0.470)	(0.472)	(0.688)	(0.540)
N	123	123	123	122	122
$\bar{T}$	10.99	20.64	27.16	9.730	16.30
$N \times \bar{T}$	1352	2539	3341	1187	1989

Notes: The table reports population-weighted estimates of regional PCE growth. Cluster robust standard errors are reported in parentheses.

Table C-3 – Linear models – Dependent variable: $\ln H_{it}$, \$2 a day

	OLS			Two-Step GMM		
	(1) Within R+C '97	(2) Differences R+C '97	(3) Differences Bourg. '03	(4) Differences R+C '97	(5) Differences Bourg. '03	(6) Differences K+V '07
$\Delta \ln \bar{y}_{it}$		-1.895	-0.268	-2.028	2.046	-0.362
		(0.170)	(0.617)	(0.271)	(1.043)	(3.216)
$\Delta \ln \bar{y}_{it} \times \ln(\bar{y}_{i,t-1}/z)$			-0.552		-0.995	-0.517
			(0.179)		(0.258)	(0.785)
$\Delta \ln \bar{y}_{it} \times \ln G_{i,t-1}$			1.108		3.445	2.097
			(0.671)		(1.192)	(2.315)
$\Delta \ln G_{it}$		2.336	-0.527	1.664	1.257	-8.222
		(0.311)	(1.449)	(1.008)	(4.127)	(11.185)
$\Delta \ln G_{it} \times \ln(\bar{y}_{i,t-1}/z)$			1.261		-0.315	-1.382
			(0.427)		(1.172)	(1.996)
$\Delta \ln G_{it} \times \ln G_{i,t-1}$			-1.769		-1.416	-8.164
			(1.586)		(3.929)	(8.296)
$\ln(\bar{y}_{it}/z)$	-2.114					
	(0.204)					
$\ln G_{it}$	3.024					
	(0.409)					
$\ln(\bar{y}_{i,t-1}/z)$						-0.023
						(0.037)
$\ln G_{i,t-1}$						-0.129
						(0.134)
$\bar{\varepsilon}^{H\bar{y}}$	-2.114	-1.895	-1.755	-2.028	-1.905	-2.684
$\bar{\varepsilon}^{HG}$	3.024	2.336	2.201	1.664	2.206	-2.345
$N \times \bar{T}$	648	648	648	641	641	641
N	104	104	104	102	102	102
Hansen's J (p-val.)	–	–	–	0.0418	0.579	0.639

Notes: The table reports OLS and GMM estimates of the model suggested in the previous literature. The dependent variable is the log (difference) of the poverty rate at \$2 a day (in 2005 PPPs). The panel structure is country-survey-year. All standard errors are robust to clustering at the country-level. The GMM results are estimated using two-step efficient GMM. Column (4) uses as instruments ΔPCE_{it}, $PCE_{i,t-1}$, $\ln \bar{y}_{i,t-1}$ and $\ln G_{i,t-1}$. Column (5) uses as instruments ΔPCE_{it}, $PCE_{i,t-1}$, $\Delta PCE_{it} \times \ln G_{i,t-1}$, $\Delta PCE_{it} \times \ln(\bar{y}_{i,t-1}/z)$, $\ln \bar{y}_{i,t-1}$, $\ln \bar{y}_{i,t-1} \times \ln G_{i,t-1}$, $\ln \bar{y}_{i,t-1} \times \ln(\bar{y}_{i,t-1}/z)$, $\ln G_{i,t-1}$ and $\ln G_{i,t-1} \times \ln G_{i,t-1}$. Column (6) uses the same instruments as column (5) but $\ln \bar{y}_{i,t-1}$ and $\ln G_{i,t-1}$ instrument for themselves. All models include a constant (not shown) and column (1) includes a time trend (not shown). Columns (2) and (4) are similar to Ravallion and Chen (1997) (R+C '97) but we update their approach by also including the Gini as in Adams (2004); columns (3) and (5) are similar to the 'improved standard model 2' in Bourguignon (2003) (Bourg. '03); and column (6) is in the spirit of the preferred specification in Kalwij and Verschoor (2007) (K+V '07). The latter also use the annualized log difference of the population size ($\Delta \ln pop_{it}$) as an instrument and rely on real GDP per capita instead of real per capita consumption. A first-stage F-test shows that $\Delta \ln pop_{it}$ is an extremely weak instrument. Kalwij and Verschoor (2007) also use interactions of lagged inequality and lagged income with regional dummies as instruments. However, first stage diagnostics suggest a weak IV problem (the F-stat with regional dummy interactions is always lower than without) and thus we opt for a simpler instrument set. Further, in column (5) and equation (5) we do not include the lagged levels of income and inequality. Column (6) includes them for comparison with Kalwij and Verschoor (2007).

Table C-4 – LN fractional probit models (QMLE) – Dependent variable: H_{it}, $2 a day

	(1) Regular		(2) Unbalanced		(3) Unbalanced + Two-Step	
	H_{it}	APEs	H_{it}	APEs	H_{it}	APEs
	Main equation					
$\ln \bar{y}_{it}$	-1.059	-0.284	-1.134	-0.281	-1.186	-0.299
	(0.020)	(0.005)	(0.040)	(0.005)	(0.364)	(0.021)
σ^2_{it}	0.518	-	0.595	-	0.574	-
	(0.020)		(0.030)		(0.182)	
$\hat{\nu}_{it}$					0.048	
					(0.065)	
	Variance equation					
$\ln \sigma^2_{it}$	0.430		0.372		0.375	
	(0.022)		(0.023)		(0.031)	
	Joint test of log-normality					
p-value	0.003		0.000		0.001	
	Summary statistics					
Scale Factor	0.260		0.248		0.252	
$N \times \bar{T}$	789		789		775	
N	104		104		103	
pseudo-R^2	0.997		0.998		0.998	
$\ln \mathcal{L}$	-314.9		-314.4		-312.2	
$\sqrt{MSE}$	0.0157		0.0112		0.0115	

Notes: The table reports fractional response QMLE estimates. The dependent variable is the poverty rate at $2 a day (in 2005 PPPs). 20 observations with $T_i = 1$ are not used during estimation. The panel structure is country-survey-year. σ^2_{it} is the variance of the log-normal distribution which we obtained by inverting the formula provided in the main text. The joint test of normality reports the result of a Wald test of the null hypothesis that the coefficient on $\ln \bar{y}_{it}$ is unity and the coefficients on the two σ^2_{it} terms are both one-half. Otherwise the specifications mimic the corresponding columns in Table 2. In models (1) and (2), the standard errors of the coefficients are robust to clustering at the country level and the standard errors of the APEs are computed via the delta method. The standard errors of the coefficients and the APEs in model (3) account for the first stage estimation step with a panel bootstrap using 999 bootstrap replications. The linear projection in the first stage uses $\ln PCE^P_{it}$ as an instrument for $\ln \bar{y}_{it}$. The first-stage cluster-robust F-statistic in (3) is 28.05. Model (3) also excludes West Bank and Gaza entirely (2 observations) and 12 observations from ECA countries pre-1990 for lack of PCE data.

Table C-5 – Predicting the 2005–2010 period with data from before

Region (r)	Predicted $\hat{\bar{H}}_r$	True $\bar{H}_r$	Absolute Error ($\lvert\hat{\bar{H}}_r - \bar{H}_r\rvert$)
East Asia and Pacific	0.35589	0.35458	0.00131
Eastern Europe and Central Asia	0.03008	0.02612	0.00397
Latin America and Caribbean	0.12292	0.11684	0.00608
Middle East and North Africa	0.13103	0.13535	0.00431
South Asia	0.68712	0.66714	0.01998
Sub-Saharan Africa	0.68642	0.70236	0.01594
Developing world	0.32925	0.32777	0.00148

Notes: The table reports pseudo out-of-sample forecasts where the data before 2005 are used as training data to predict the observed data in the 2005 to 2010 period (testing data). The $\sqrt{MSE}$ for these six regional data points is 0.011 and the pseudo-R^2 is 0.998. The dependent variable is the poverty rate at $2 a day (in 2005 PPPs). The results have not been population-weighted and are not comparable to the 2010 baseline used in the projections.

Table C-6 – 10-fold cross-validation – preferred specification

Fold	$\sqrt{MSE}$	MAE	Pseudo-R^2	% Training	% Testing
1	0.03354	0.02649	0.991	0.962	0.038
2	0.02779	0.01958	0.992	0.892	0.108
3	0.04142	0.02814	0.985	0.911	0.089
4	0.03623	0.02765	0.988	0.865	0.135
5	0.03838	0.02471	0.986	0.855	0.145
6	0.04448	0.03260	0.984	0.932	0.068
7	0.03550	0.02464	0.988	0.900	0.100
8	0.03134	0.02427	0.989	0.871	0.129
9	0.03000	0.01957	0.989	0.907	0.093
10	0.04397	0.02533	0.982	0.904	0.096
$CV_{(10)}$ - Averages	0.03668	0.02530	0.987	0.900	0.100

Notes: The table reports the results of a 10-fold cross-validation exercise for panel data. In each fold, about 10% of the time series data are deleted while all countries remain in the sample. The remaining 90% of the data (training data) are then used to predict these 10% of observations (testing data). $\sqrt{MSE}$ is the root mean squared error, MAE is the mean absolute out-of-sample error, and the pseudo-R^2 is the squared correlation of the out-of-sample predictions with the testing data. The dependent variable is the poverty rate at $2 a day (in 2005 PPPs).

Table C-7 – Fractional probit models (QMLE) – Dependent variable: H_{it}, \$2 a day

	(1) Institutions		(2) Human Capital		(3) Credit		(4) Trade	
	H_{it}	APEs	H_{it}	APEs	H_{it}	APEs	H_{it}	APEs
$\ln \bar{y}_{it}$	-0.888	-0.285	-0.878	-0.284	-0.950	-0.289	-0.708	-0.302
	(0.050)	(0.012)	(0.060)	(0.011)	(0.036)	(0.009)	(0.032)	(0.012)
$\ln G_{it}$	0.779	0.250	0.805	0.261	0.765	0.233	0.581	0.248
	(0.107)	(0.028)	(0.104)	(0.027)	(0.102)	(0.027)	(0.097)	(0.033)
Executive Constraints	0.005	0.001						
	(0.005)	(0.001)						
Years of Schooling			-0.002	-0.001				
			(0.017)	(0.006)				
Private Credit / GDP					-0.007	-0.002		
					(0.040)	(0.012)		
Trade Openness							0.005	0.002
							(0.017)	(0.007)
Scale factor	0.321		0.324		0.304		0.426	
$N \times \bar{T}$	678		705		697		385	
N	85		87		93		81	
AIC	894.8		914.1		887.6		552.5	
$\ln \mathcal{L}$	-276.4		-286.1		-282.8		-163.2	
$\sqrt{MSE}$	0.0203		0.0211		0.0201		0.0233	

Notes: The table reports fractional response QMLE estimates. The dependent variable is the poverty rate at \$2 a day (in 2005 PPPs). The panel structure is country-survey-year. The estimation samples are reduced due to less data coverage of the covariates. Observations with $T_i = 1$ are not used in estimation. All models include time averages (CRE), time dummies, survey dummies, panel size dummies and interactions between the panel size dummies and the time averages (CRE). The time averages are recomputed for each sample size. The coefficients of the time average of the survey dummies and time effects are constrained to be equal across the panel sample sizes. The variance equation depends on the sample size. The standard errors of the coefficients are robust to clustering at the country level and the standard errors of the APEs are computed via the delta method. Data on *Executive Constraints* is from the Polity IV database. Human capital is measured as *Total Years of Schooling* from Barro and Lee (2013). We linearly interpolate the five-yearly data to an annual series. *Private Credit / GDP* measures financial development and is from Beck et al. (2010). *Trade Openness* is the *de jure* binary measure from Wacziarg and Welch (2008).

Table C-8 – Fractional probit models (QMLE) – Dependent variable: H_{it}, \$1.25 a day

	(1) Regular		(2) Unbalanced		(3) Unbalanced + Two-Step	
	H_{it}	APEs	H_{it}	APEs	H_{it}	APEs
$\ln \bar{y}_{it}$	-1.212	-0.216	-0.668	-0.218	-0.800	-0.263
	(0.056)	(0.010)	(0.038)	(0.008)	(0.180)	(0.034)
$\ln G_{it}$	1.238	0.221	0.726	0.237	0.714	0.235
	(0.121)	(0.022)	(0.074)	(0.020)	(0.180)	(0.032)
$\hat{\nu}_{it}$					0.104	
					(0.104)	
CRE (Corr. Rand. Effects)	Yes		Yes		Yes	
Survey type dummies	Yes		Yes		Yes	
Time dummies	Yes		Yes		Yes	
Panel size dummies	No		Yes		Yes	
Panel size dummies × CRE	No		Yes		Yes	
Variance equation	No		Yes		Yes	
Scale factor	0.179		0.326		0.329	
$N \times \bar{T}$	768		768		754	
N	103		103		102	
pseudo R^2	0.975		0.990		0.990	
$\ln \mathcal{L}$	-172.4		-244.7		-243.7	
$\sqrt{MSE}$	0.0339		0.0214		0.0220	

Notes: The table reports fractional response QMLE estimates. The dependent variable is the poverty rate at \$1.25 a day (in 2005 PPPs). 21 observations with $T_i = 1$ are not used during estimation. The panel structure is country-survey-year. The \$1.25 a day sample is smaller as for 20 observation we only have data at the \$2 a day line. In columns (1) and (2), the standard errors of the coefficients are robust to clustering at the country level and the standard errors of the APEs are computed via the delta method. We include the time averages of the survey type and time dummies in columns (2) and (3), but constrain their coefficients to be equal across the panel sizes. The standard errors of the coefficients and the APEs in model (3) account for the first stage estimation step with a panel bootstrap using 999 bootstrap replications. The linear projection in the first stage uses $\ln PCE_{it}^{P}$ as an instrument for $\ln \bar{y}_{it}$. The first-stage cluster-robust F-statistic in column (3) is 24.40. Column (3) also excludes West Bank and Gaza entirely (2 observations) and 12 observations from ECA countries pre-1990 for lack of PCE data.

Table C-9 – Decomposition at $1.25 a day poverty line, by region

	VAR(Y)	VAR(D)	COV(Y, D)	s_Y	s_D	$\sqrt{MSE}$	N
			Panel a) Spells from 1981 to 2010				
East Asia and Pacific	2.145	0.716	0.260	71.12	28.88	0.88	12
Europe and Central Asia	1.856	0.677	0.438	67.30	32.70	0.31	39
Latin America and Caribbean	1.905	0.642	-0.787	114.89	-14.89	0.78	28
Middle East and North Africa	0.074	0.045	0.024	58.50	41.50	0.32	8
South Asia	0.683	2.594	-0.023	20.42	79.58	0.45	7
Sub-Saharan Africa	2.021	0.394	-0.271	93.45	6.55	0.47	29
All developing	2.140	0.704	0.075	73.99	26.01	0.56	123
			Panel c) Spells from 2000 to 2010				
East Asia and Pacific	2.821	0.399	0.270	82.21	17.79	1.00	9
Europe and Central Asia	8.059	1.468	-0.215	86.22	13.78	0.46	25
Latin America and Caribbean	2.267	1.040	-0.794	85.67	14.33	0.92	26
Middle East and North Africa	0.124	0.014	0.001	89.29	10.71	0.34	6
South Asia	0.865	0.223	0.102	74.89	25.11	0.46	4
Sub-Saharan Africa	4.770	1.826	-0.736	78.73	21.27	1.26	17
All developing	5.138	1.120	-0.268	85.11	14.89	0.86	87
			Panel c) Spells from 2000 to 2010				
East Asia and Pacific	1.410	1.261	0.314	52.26	47.74	0.76	10
Europe and Central Asia	1.405	0.213	0.373	75.20	24.80	0.41	26
Latin America and Caribbean	0.243	0.104	0.090	63.19	36.81	0.47	19
Middle East and North Africa	0.058	0.026	0.022	62.31	37.69	0.38	5
South Asia	0.239	1.146	0.340	28.05	71.95	0.43	6
Sub-Saharan Africa	2.978	1.690	0.706	60.59	39.41	1.02	21
All developing	1.664	0.728	0.454	64.17	35.83	0.66	87

Notes: The table reports the results of the decomposition of the observed changes in the poverty rate at $1.25 a day into its growth and distribution components at the regional level. Panels a) to c) run this decomposition over different sub-samples as denoted in the table. We predict the counterfactual quantities using the first and last available data for the longest runs of survey of the same type within the sample period.

Table C-10 – Projected poverty headcount ratios and poor population at $1.25 a day in 2030, by region

	Average PCE Growth								
	Optimistic (2000-2010)			Moderate (1980-2010)			Pessimistic (1980-2000)		
	Change in Inequality (Gini)								
	pro-poor	neutral	pro-rich	pro-poor	neutral	pro-rich	pro-poor	neutral	pro-rich
Panel (a) – Headcount at $1.25 a day in 2030 (in percent)									
East Asia and Pacific	0.65	0.93	1.31	0.76	1.07	1.48	0.94	1.29	1.74
Europe and Central Asia	0.12	0.16	0.21	1.21	1.45	1.71	5.17	5.74	6.44
Latin America and Caribbean	2.27	2.74	3.28	3.46	4.12	4.91	4.59	5.46	6.48
Middle East and North Africa	0.48	0.66	0.91	1.54	2.07	2.75	3.72	4.77	6.05
South Asia	4.19	5.54	7.24	8.48	10.89	13.79	12.76	15.99	19.77
Sub-Saharan Africa	32.09	35.69	39.37	43.62	47.17	50.70	51.75	55.12	58.47
Total	7.88	9.11	10.49	11.63	13.20	14.96	14.96	16.82	18.89
Panel (b) – Poor population at $1.25 a day in 2030 (in millions)									
East Asia and Pacific	14.05	20.23	28.56	16.59	23.29	32.22	20.44	28.00	37.87
Europe and Central Asia	0.59	0.76	0.97	5.72	6.86	8.11	24.47	27.16	30.45
Latin America and Caribbean	16.15	19.44	23.32	24.61	29.30	34.86	32.61	38.77	46.05
Middle East and North Africa	2.12	2.94	4.04	6.84	9.18	12.18	16.47	21.15	26.83
South Asia	83.47	110.38	144.35	169.13	217.08	275.00	254.40	318.89	394.16
Sub-Saharan Africa	449.54	499.97	551.61	611.15	660.85	710.32	725.03	772.26	819.22
Total	567.20	655.36	754.95	836.34	949.49	1076.37	1076.35	1209.95	1359.23

Notes: The table reports forecasts of the $1.25 a day poverty rate in 2030. The forecasts are based on the estimates reported in Column (2) of Table 2 and the different growth/ distribution scenarios outlined in the text. Population projections are from the World Bank's Health, Nutrition and Population Statistics database. The survey data are from the World Bank's *PovcalNet* database.