OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 72, 2 (2010) 0305-9049
doi: 10.1111/j.1468-0084.2009.00576.x

A Micro-Decomposition Analysis of Aggregate Human Development Outcomes*

Sylvie Lambert†, Martin Ravallion‡ and Dominique van de Walle‡

†*Paris School of Economics, 48, Bld Jourdan, 75014 Paris, France*
(e-mail: sylvie.lambert@ens.fr)
‡*The World Bank, 1818 H Street, NW, MSN MC4-400, Washington DC 20433, USA*
(e-mails: dvandewalle@worldbank.org and mravallion@worldbank.org)

Abstract

We show how differences in aggregate human development outcomes over time and space can be additively decomposed into a pure mean income (growth) component, a component attributed to differences in the distribution of income, and components attributed to 'non-income' factors and differences in the model linking outcomes to income and non-income characteristics. The income effect at the micro level is modelled non-parametrically, so as to flexibly reflect potentially complex distributional changes. Our proposed method is illustrated using data for Morocco and Vietnam, and the results offer some surprising insights into the observed aggregate gains in schooling attainments.

I. Introduction

The challenge of reaching the Millennium Development Goals (MDGs) has renewed interest amongst donors, development agencies and developing country governments in the question of how much economic growth affects aggregate human development (HD) outcomes, and what role is played by other factors, such as changes in income inequality.[1] Understanding the extent and nature of that linkage can help inform efforts

*The research assistance of Silvia Redaelli is also gratefully acknowledged as are helpful comments from the journal's anonymous referees. The findings, interpretations and conclusions of this article are those of the authors and should not be attributed to the World Bank, its Executive Directors or the countries they represent.
JEL Classification numbers: D31, I21, O15.

[1]The MDGs are a commitment by the United Nations to achieve a set of poverty and HD targets by 2015. For further details, see http://www.un.org/millenniumgoals/

to monitor and forecast progress in raising aggregate HD attainments as well as to recommend policy reforms.

A large literature has focused on cross-country empirical relationships between average incomes and aggregate HD indicators such as school enrolments, literacy, life expectancy and infant mortality.[2] Unsurprisingly, attainments in basic health and education tend to be higher in higher income countries and (less obviously) economic growth tends to be accompanied by improvements in social indicators. Cross-country comparisons have also been used to try to assess the extent of 'social inefficiency' in HD at given mean macroeconomic and fiscal aggregates.[3] The deviations of education and health indicators from their expected values at given mean incomes (and/or social spending) have been used to infer differences in the efficiency of HD efforts within countries; an early example was Sen (1981) who showed that Sri Lanka's HD attainments were far higher than one would expect given its income.

There is considerable uncertainty about how the empirical relationship between aggregate indicators of HD and mean incomes should be interpreted when discussing HD policy, including progress towards the MDGs. Is it that higher average income allows a society to buy goods and services that promote health and schooling? Or does the empirical relationship found in cross-country comparisons stem from a correlation between average incomes and other country characteristics?[4] Can the cross-country relationship be relied upon in drawing inferences about the macroeconomics of HD in any specific country, or the performance of that country relative to another? The answers to these questions are left begging when the cross-country relationship between aggregate HD outcomes and mean income is used to predict the impacts of future growth (or the income growth needed to attain specific HD targets) or to assess country performance at a given mean income. Indeed, given the uncertainties about what 'mean income' is capturing in cross-country regressions for HD outcomes, and the high level of aggregation in the conditional means based on cross-country regressions, it is little more than a leap of faith to believe that this approach can tell us something useful about how economic growth in any specific country setting will impact on HD.

There is also a large body of work exploring some of these issues using micro data.[5] Typically household- or individual-level data on health or education attainments are regressed on a range of individual and household variables, including income or

[2]Contributions to the literature on the cross-country relationship between average incomes and average HD attainments include: Sen (1981, 1988), Bhalla and Glewwe (1986), Anand and Ravallion (1993), World Bank (1993), Aturupane, Glewwe and Isenman (1994), Pritchett and Summers (1996), UNDP (1996), Bidani and Ravallion (1997), Moore *et al.* (1999), WHO (1999), Ranis, Stewart and Ramirez (2000).

[3]Examples include: World Bank (1993), UNDP (1996), Moore *et al.* (1999), WHO (1999), Wang *et al.* (1999), Evans *et al.* (2000), Gupta and Verhoevan (2001), Jayasuriya and Wodon (2003). Ravallion (2005) reviews the literature using cross-country comparisons to infer 'social efficiency'.

[4]For example, Anand and Ravallion (1993) argue that what might really be driving the relationship is that richer countries tend to have lower income poverty and/or higher public spending on health and education, and it is these channels that hold the key lessons for policy.

[5]For a survey, see Strauss and Thomas (1995). Recent examples of micro approaches to studying HD outcomes include: Bhargava (1999), Filmer and Pritchett (1999) and Glewwe and Jacoby (2004).

wealth indicators. This body of micro work has been a rich source of knowledge about the micro determinants of HD attainments. However, the specifications used and the analytics have not allowed this literature to effectively explore the implications for the *aggregate* relationship between HD attainments in a country and economic growth, the distribution of income and non-income factors. What is needed is a set of tools for consistently aggregating the empirical micro relationships to throw light on the macroeconomics of HD.

This article aims to help bridge this gap, by developing and implementing a micro-based decomposition method for investigating the proximate determinants of aggregate HD outcomes. The aim is to measure the relative importance of growth in mean incomes vs. changes in the distribution of income and 'non-income' characteristics of the population, such as maternal schooling. In essence we provide a growth-redistribution decomposition method for HD indicators analogous to the widely used Datt and Ravallion (1992) decomposition for poverty measures. However, in the latter case there is a precise mathematical link between the measure of poverty and the mean and distribution. That is not the case with HD indicators, which makes the decomposition a more difficult task. In investigating income distributional impacts on HD we use non-parametric regression methods for the HD indicators. This allows a high degree of flexibility in representing the underlying nonlinearity in the micro relationship between HD attainments and household incomes, so as to estimate the decomposition. The modelling and decomposition methods developed in this article are available in a user-friendly STATA program (available from the authors).

Although the primary aim of the article is to provide a new methodological tool that will allow a deeper micro-based understanding of the macroeconomic determinants of HD outcomes, it is of interest to also provide some empirical examples. We use our decomposition method in studying the proximate determinants of changes over time in schooling attainments and the inter-group disparities in schooling found in two developing countries. In the context of schooling, there are reasons to think that both mean income and its distribution will matter to aggregate outcomes given the likely nonlinearities involved (whereby marginal income effects on HD outcomes are likely to be lower the higher the level of income). However, past research on micro data has also pointed to the importance of a wide range of 'non-income' characteristics, including characteristics of the household (such as parental education, number of siblings, demographics, ethnicity, household inputs), provider characteristics (such as their quality, distance, relevance of the curriculum, gender of teachers, availability of latrines, etc.) and geographic characteristics (such as average consumptions and consumption inequality within the area of residence and the provision of schooling facilities). Our proposed modelling method allows for a potentially wide range of such non-income characteristics, all of which are potential correlates of incomes, which may well account in part at least for the 'macroeconomic' income effect evident in aggregate data.

The specific case studies focus on school enrolments for boys and for girls in two rather different countries, Morocco and Vietnam. These were chosen both because

of the differences in their growth and HD performances over the 1990s and because suitable survey data are available for our purposes.

II. Empirical model

Our aim is to decompose differences in average schooling attainments between groups, which could be different regions or countries, different population groups such as gender or ethnicities or different dates. In setting up the decomposition we want to explicitly identify a component attributed to average consumption and one attributed to consumption inequality. To isolate the latter we need a suitably flexible nonlinear representation of how schooling varies with consumption. By incorporating such flexibility we are better able to provide a convincing account of the role played by distribution (both changes over time, and differences between groups at one date, such as between urban and rural areas) – unconfounded by *ad hoc* parametric assumptions about the nature of the economic gradient in schooling.

The probability (S_{ij}) of being in school for child i in group j with characteristics x_{ij} and income y_{ij}, is given by the following model:

$$S_{ij} = \alpha_j + \phi_j(y_{ij}) + \pi_j x_{ij} + v_{ij}, \tag{1}$$

where v_{ij} is a zero-mean error term with variance σ_v^2. All that we assume about the function ϕ is that it is smooth and single-valued; in particular, its first derivatives are bounded by constants, $c \geq |\Delta\phi(y)|/|\Delta y|$. The function need not be monotonic, or take any parametric form. The vector x_{ij} includes both family and geographical characteristics, and the latter include both the mean consumption of the area of residence and the inequality of consumption in that area.

For large group sizes we can treat the consumption distribution as continuous, and let $F_j(y)$ denote the distribution function of consumption in group j. Let

$$\bar{\phi}_{jk} \equiv \int_0^1 \phi_j(y)\, \mathrm{d}F_k(y), \tag{2}$$

(so when $j = k$ we get the mean of $\phi_j(y_{ij})$ across all i). The expected value of the school enrolment rate for group j is:

$$\bar{S}_j = \alpha_j + \bar{\phi}_{jj} + \pi_j \bar{x}_j. \tag{3}$$

The enrolment gap is $1 - \bar{S}_j$. The latter can be exactly decomposed as:

$$1 - \bar{S}_j \equiv 1 - \phi_j(\mu_j) + I_j - \alpha_j - \pi_j \bar{x}_j, \tag{4}$$

where μ_j is the mean consumption of group j and

$$I_j \equiv \phi_j(\mu_j) - \int_0^1 \phi_j(y)\,\mathrm{d}F_j(y) \tag{5}$$

is the contribution of consumption inequality to the aggregate enrolment gap.

To estimate the model in equation (1) we draw on the literature on partial linear models (as reviewed in Yatchew, 1998).[6] All observations are ordered in terms of their consumption values. Differences are taken between the data for successive ranked observations, giving the regression:

$$\Delta S_{ij} = \Delta\phi_j(y_{ij}) + \Delta x_{ij}\pi_j + \Delta v_{ij}, \tag{6}$$

where Δx_{ij} is the difference between the values for the ith observation and that for $i-1$ when ranked in ascending order of y. Under our assumption about the function ϕ, the first term on the right-hand side vanishes as the sample size goes to infinity ($\mathrm{plim}[\phi_j(y_{ij}) - \phi_j(y_{i-1,j})] = 0$). So the following parametric regression can be estimated by least squares:

$$\Delta S_{ij} = \Delta x_{ij}\pi_j + \Delta v_{ij}. \tag{7}$$

To estimate the functions $\phi_j(.)$, we can then estimate the non-parametric regression:

$$S_{ij} - (x_{ij} - \bar{x}_j)\hat{\pi}_j = \phi_j(y_{ij}) + v_{ij}. \tag{8}$$

Finally, we estimate the constant terms in the usual way by forcing the estimate of equation (1) through the mean points, so $\hat{\alpha}_j = \bar{S}_j - \bar{\hat{\phi}}_{jj} - \bar{x}_j\hat{\pi}_j$, where $\bar{\hat{\phi}}_{jj}$ is the sample predicted value of $\phi_j(y_{ij})$.

Higher-order differencing allows efficiency gains in this method (Yatchew, 1998). We rewrite equation (8) as:

$$\sum_{k=0}^{m} d_k S_{i-k,j} = \left(\sum_{k=0}^{m} d_k x_{i-k,j}\right)\pi + \sum_{k=0}^{m} d_k v_{i-k}, \tag{9}$$

where $\sum d_k = 0$ [which allows us to drop the non-parametric effect from equation (6)] and the normalization condition $\sum d_k^2 = 1$ (which assures that the transformed residuals have variance σ_v^2). Hall *et al.* (1990) provide the optimal weights up to $m = 10$.

III. Decomposition

Our purpose is to explain the difference in average schooling attainments between two groups, indexed j and k. As is typically the case in decompositions, the values taken by the components are measured relative to a reference group r. In many applications, it will be natural to treat one of the groups j or k as the reference (or 'base'). That is what one would normally do in assessing changes over time between two survey

[6]We have used the PLREG program for STATA by Lokshin (2006).

dates, in which case the base date is typically chosen as the reference. However, when there are more than two dates, or there is a third group that is of interest (e.g. the reference group might be a sub-group of people at one date who have already attained the MDG), one can allow the reference group to be different to either comparator. In general, the values taken by each component will depend on the choice of reference; how much that choice matters is an empirical question.

The proposed decomposition for the difference in means between the two groups is:

$$\bar{S}_k - \bar{S}_j = G^r_{jk} + R^r_{jk} + N^r_{jk} + M_{jk}. \tag{10}$$

The four components of the decomposition are defined as follows.

1. *Mean income* (G^r_{jk}): The contribution of the difference in mean consumption between groups j and k using group r as the reference; this is given by:

$$G^r_{jk} \equiv \phi_r(\mu_k) - \phi_r(\mu_j). \tag{11}$$

For example, if the groups are two dates, t and $t+1$, and the reference is the base date (t) then this is interpretable as the contribution of economic growth:

$$G^t_{t,t+1} = \phi_t(\mu_{t+1}) - \phi_t(\mu_t). \tag{12}$$

Notice that $G^r_{jk} = 0$ if and only if $\mu_k = \mu_j$ and that [for $\phi'_r(.) > 0$] a positive (negative) value of G^r_{jk} implies higher (lower) mean consumption. Although these properties hold for any reference group, choice of the latter will affect the quantitative magnitude of G^r_{jk}. The mean income component can be thought of as the 'pure' effect of economic growth, holding distribution, non-income factors and the model's parameters constant.

2. *Redistribution* (R^r_{jk}): The contribution of changes in the distribution of consumption using group r as the reference. Recall that the contribution of inequality to the level of mean schooling is measured by the difference between the expected value of the ϕ function and its value at the mean consumption. The redistribution component is then the inter-group difference in the contribution of inequality, namely:

$$R^r_{jk} \equiv \bar{\phi}_{rk} - \bar{\phi}_{rj} - [\phi_r(\mu_k) - \phi_r(\mu_j)]. \tag{13}$$

The properties of this component depend on the curvature of the ϕ function and how the distribution of consumption changes. Suppose that distribution k is obtained from j by a mean-preserving redistribution in favour of the poor and that ϕ_r is strictly concave (convex); then by standard properties of concave functions, R^r_{jk} must be positive (negative).

3. *Non-income factors* (N^r_{jk}): This is the contribution of differences in mean non-income characteristics:[7]

$$N^r_{jk} \equiv \pi_r(\bar{x}_k - \bar{x}_j). \tag{14}$$

[7]The vector x can include geographical effects of local-level average income and inequality.

This will be recognized as the 'characteristics' component of the Blinder–Oaxaca decomposition (Blinder, 1973; Oaxaca, 1973), although strictly N as defined here is only that sub-component not attributed to incomes. Plainly, if $\bar{x}_k = \bar{x}_j$ then $N^r_{jk} = 0$.

4. *Model* (M^r_{jk}): This component arises from any structural differences in the model for outcomes between groups j, k and the reference r. This is a nonlinear generalization of the usual 'structure' term in the Blinder–Oaxaca decomposition. More precisely, the model component is given by:

$$M^r_{jk} \equiv \bar{\phi}_{kk} - \bar{\phi}_{rk} - (\bar{\phi}_{jj} - \bar{\phi}_{rj}) + \alpha_k - \alpha_j + (\pi_k - \pi_r)\bar{x}_k - (\pi_j - \pi_r)\bar{x}_j. \qquad (15)$$

This can be thought of as the aggregate contribution of all group-specific attributes that determine the model's functional form (both parametric and non-parametric) for each group. When one is comparing different dates, economy-wide institutional changes will yield a model component in the change in aggregate HD outcomes. When comparing different groups at one date, the model term will reflect differences in how the economy and society as a whole at that date influence the group-specific functional form. For example, if there is factor-market discrimination against one group (an ethnic minority) then this will show up in the model term of our decomposition.

The following remarks are in order about this decomposition.

First, notice that separately identifying N from M requires that the covariates are identical between the models for j and k. When the models are different (such as because of a comparability problem between the two surveys) one can only estimate $N + M$.

Second, the 'mean income' and 'redistribution' components we identified before can be thought of as 'internal' effects at household level, meaning that they are derived from the direct household-level income effect. In principle, we can also allow for 'external' effects operating through geographic effects on HD outcomes. For example, an internal effect would stem from the effect of a higher income on household-level purchases of inputs to schooling. Independently of this, there may be an external effect such as when growth in the local region leads to better schools. In our decomposition, these external effects are identified here as 'non-income' factors (N).

Third, one might also be interested in the *internal* composition of the components of our decomposition. When we come to interpreting our empirical results it will be useful to note that there is no *a priori* reason why the various sub-components will necessarily move in the same direction. For example, within the model component there may well be an economy-wide upward shift in the ϕ function, in combination with (partially or fully) offsetting shifts in the coefficients on some non-income components. Figure 1 illustrates this for one non-income factor X. There are two sub-components to the model component between dates 1 and 2. First, the gradient of the relationship between schooling (S) and X declines in date 2; for example, birth order might become less important in determining a child's school enrolment probability. Secondly, there is a generalized increase in schooling at all levels of X as indicated by the dashed line. The model component is the net effect of this positive upward shift at all levels and the negative shift (more pronounced at higher values of X) that is required to get back to the actual gradient for date 2.

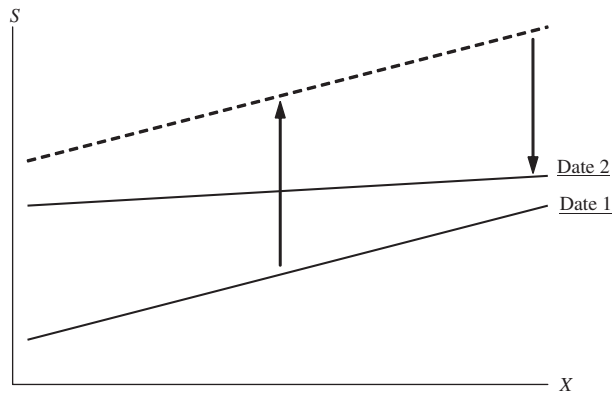

Figure 1. Stylized representation of the components of the model term between dates 1 and 2

IV. Case studies for Morocco and Vietnam

The 1990s in Morocco is an interesting period for implementing the decomposition and studying the proximate determinants of changes in enrolment outcomes. Morocco has long been singled out for its poor social indicators given its per capita income levels (about 1,200 USD per capita in 1998 and 1,700 USD per capita in 2005 in nominal terms) and its severely unequal education outcomes as reflected in large urban/rural differentials and a substantial gender gap. There were improvements during the 1990s – net primary enrolment rates rose from 58% in 1990 to 70% in 1998, and gross enrolment rates from 63% to 85% – reflecting increased social spending and targeted programmes aimed at redressing regional disparities. These efforts focused on education, but also on health, electrification and drinking water. Progress also occurred for rural girls as evidenced by a concomitant change in their net primary enrolment rate from 28% to 47%. The most recent statistics indicate that Morocco achieved 92% net primary enrolment rates (89% for girls) for the 2002/03 school year.

Improvements in social indicators coincided with a series of devastating droughts and poor economic conditions. Slow economic growth, especially in rural areas, resulted in a rise in income poverty from 13.1% of the population in 1990/91 to 19% in 1998 (and 18–27 % in rural areas). During the same period, consumption inequality remained stable at a fairly high level (a Gini of 0.39 is reported in World Bank, 2001). Despite recent amelioration, Morocco's social indicators (including enrolments in postprimary schooling) still lag significantly behind that of other countries in the region with similar levels of income.

The generally poor past state of education outcomes is variously explained by poverty, cultural attitudes towards women's roles, poor access and quality of educational facilities and parents' general pessimism about schooling's net rewards. A recent review of the education system shows that as recently as 1998, less than

45% of Douars (rural communes) had a primary school, for about 25% a facility existed only over 2 km away, whereas 6% provided no possibility of schooling at all. Eighty per cent of rural primary schools had no running water and 87% were without electricity (Commission Spéciale Education Formation, 1999). Access is even worse for higher levels of schooling.

The 1990s saw rapid urbanization in Morocco. In 1994, for example, the annual urban population growth rate was 3.6% compared with 0.7% in rural areas, where the fertility rate is nevertheless twice as high. Forty-six per cent of the population remains rural. Urbanization is driven largely by rural migration. Many such migrants are illiterate and enter low-wage, low-return informal sector activities. A new phenomenon in the Moroccan economy was a surge in women's labour force participation starting around 1993. This reflects migration of rural women who had previously worked on the family farm (and were thus not counted as participating in the labour market) caused by severe droughts in 1992, 1993 and 1995 (World Bank, 1997).

Vietnam presents a very different situation. In stark contrast to Morocco, Vietnam experienced a dramatic reduction in income poverty in just 5 years during the 1990s from 58% of the population in 1993 to 37% in 1998 (Government-Donor–NGO Working Group, 1999). In rural areas, where 80% of the population lived in 1993, poverty fell from 66% to 48% during the same period. Rapid economic growth, accompanied by only a small increase in inequality – the Gini of consumption expenditures rose from 0.33 in 1993 to 0.35 in 1998 – underlies this trend.

Social indicators also improved during this period. In the education area, primary school enrolment rates, already high as a result of the communist regime's emphasis on education, increased from 87% to 91% for girls and from 86% to 92% for boys. More dramatic gains were found for lower secondary enrolment rates, which doubled to 61% for girls and to 62% for boys. Upper secondary enrolment rates have likewise increased substantially. Yet, these high enrolment rates and the lack of a gender gap hide the fact that there are still regions – primarily mountainous and more isolated areas – and groups of people – primarily ethnic minorities – that are particularly disadvantaged in terms of access and quality (van de Walle and Gunewardena, 2001). Ethnic minority children account for half of the children not in school. Enrolments also tend to rise with household consumption expenditures, and increasingly so at higher schooling levels. Glewwe and Jacoby (2004) have shown that there is a strong wealth effect in education. Households complain of high costs, low quality and low relevance of the curriculum. Existing policies such as fee exemptions for the poor and ethnic minority children have low coverage and have been found to account for only a small share of total household outlays on schooling (Behrman and Knowles, 1999; Nguyen, 2004; van de Walle, 2004).

V. Data

Our main aim here is to illustrate the decomposition method. We do so using the sorts of regressions for school enrolments that are common in past work, although (as we

discuss) the causal interpretation of these regressions is not beyond question. We do some robustness checks.

Both Morocco and Vietnam have nationally representative, household consumption expenditure surveys that were implemented with World Bank assistance for two dates during the 1990s. They also collect information on various other aspects of living standards including health, education, work and migration. We use the Morocco Living Standards Survey (MLSS) rounds of 1990/91 and 1998/99 with 3,323 and 5,131 households, respectively, and the Vietnam Living Standards Surveys (VLSS) of 1992/93 and 1997/98, covering about 4,800 and 6,000 households, respectively (World Bank, 1995, 2000). In both countries, the surveys are comparable across the two years and sufficient information on school attendance, household- and community-level variables is available. These are fairly comprehensive data sets, which allow us to include a potentially wide range of controls.

We created two individual-level files of children aged 6 through 11 years, and 12 through 18 years for Vietnam (respectively containing 3,328 and 3,100 observations in 1993, and 3,289 and 4,042 observations for 1998). Six is the official primary school starting age and indeed some 66% of children were in school at that age in 1993, and 83% in 1998. The first file is expected to cover primary school-aged children pretty well. The second group covers all of secondary school and possibly vocational training. The sample includes all children who live in the household (regardless of whether they are the children of the head), as well as children who do not but whose parents do.

For Morocco, the size and cut-off points of the groups are different given the different schooling profile. (As we do not directly compare the two countries this is not a problem for the analysis.) The two sub-samples are for children aged 7–12 years (containing 2,576 observations in 1991 and 3,637 in 1998) and 13–15 years (1,239 observations in 1991 and 1,860 in 1998), the ages corresponding to the legal age ranges for attending primary and lower secondary school, respectively. In 1991, no child below age 7 attends school in the data and in 1998, only 15 do. We drop them from the sample. In the Morocco database, we include only children of the head of household who also reside in the household as the questionnaire did not collect sufficient information on non-residents.

The schooling indicator we use as the dependent variable is enrolment defined as whether a child is currently in school or not. There are many other potential schooling indicators that could be used. We did repeat all the calculations using years of schooling as a proportion of the maximum possible years. Results were qualitatively very similar and hence, are not reported.

As noted in section II, the explanatory variables (which it will be recalled enter linearly) include individual, household and geographical (or community) characteristics. Child characteristics include specific age dummies, gender and birth order, calculated over all children aged 18 years and below in each household. For Vietnam we also include a dummy for whether the child is the head's child. At the household level we include log consumption per capita, log household size, demographic

composition (shares by age group), the gender of the head and dummies for urban and regional location. For Vietnam we enter the father and mother's age, age-squared and total years of education; for Morocco, we have age and age-squared only for the father, and education for both the father and mother is measured as dummy variables for various attainment levels rather than years. In addition, for Vietnam we include a dummy variable for whether the head is from an ethnic minority, a group that tends to be poorer and less likely to be schooled. In Morocco we bring in controls for migration (whether the father has lived somewhere else for the last 12 months, whether he was born in the current residence and number of years father has been living in the current residence). Geographical variables consist of the mean consumption and inequality (as measured by the mean log deviation of consumption per person) of the community of residence, as well as distance to school.

Although we have included a reasonably typical set of regressors from past micro studies of school enrolments in developing countries, one can always question the exogeneity of such regressors; for example, female headship might be endogenous to schooling choices through migration decisions of the male head of household. (Of course, dropping such variables can create new concerns about omitted variable bias; and arbitrary exclusion restrictions for identification using an instrumental variables estimator can always be questioned.)

One variable probably stands out as potentially endogenous, namely household consumption. (For example, the decision to keep a child in school could lower current consumption owing to foregone income from child labour.) Measurement error in consumption can also create a bias. Consumption is believed to be measured well in the surveys, following standard practices.[8] Nevertheless, potential endogeneity clouds a causal interpretation of the regression coefficients. We also test a measure of predicted expenditures, which we calculate from a regression of actual consumption on a wide array of household-level characteristics that are likely to influence household consumption. Results are not significantly affected.

Using these data we calibrate the regression models described before and implement the decompositions. We begin by studying the change in school attainments over time, both in the aggregate and by gender. Next we present a cross-sectional comparison of the sources of the differences in schooling between urban and rural areas. We then focus on other decompositions, specific to each country, namely ethnicity in Vietnam and parents' literacy in Morocco.

VI. Results

Table 1 gives some descriptive statistics. School enrolment rates have risen over time in both countries, but are appreciably higher in Vietnam. Real consumption per person rose over the period in Vietnam (a growth rate of 5.8% per annum), but fell in

[8] It incorporates the value of own production and the imputed value of housing expenditures, and is expressed in real 1998 prices. The reference used is of January 1998 for Vietnam and average 1998 prices for Morocco.

TABLE 1

Summary statistics for children and households with children

		Morocco			*Vietnam*		
		Observations	*Mean*	*SD*	*Observations*	*Mean*	*SD*
Enrolment rate	Date 2	3,637	0.73	0.44	3,289	0.93	0.25
('primary' age group)	Date 1	2,576	0.61	0.49	3,328	0.87	0.34
Enrolment rate	Date 2	1,860	0.52	0.50	4,042	0.65	0.48
('secondary' age group)	Date 1	1,239	0.16	0.37	3,100	0.43	0.50
Real household consumption	Date 2	30,457	8,070.65	7,627.93	39,035	3,027.52	2,510.35
per person, whole sample	Date 1	16,741	8,708.18	9,780.29	26,526	1,960.07	1,388.50
Real household consumption	Date 2	16,715	10,537.4	9,118.98	9,827	4,993.45	3,675.77
per person (urban areas), whole sample	Date 1	7,181	12,237.8	13,020.7	5,180	3,041.86	2,149.76
Real household consumption per	Date 2	13,742	5,070.28	3,378.16	29,208	2,366.08	1,462.43
person (rural areas), whole sample	Date 1	9,524	5,959.78	4,232.15	21,346	1,697.56	959.99
Real household consumption per person,	Date 2	5,497	6,324.90	4,696.63	7,331	2,319.59	1,682.77
school-age children sample	Date 1	3,815	6,905.48	5,096.53	6,428	1,736.18	1,079.07
Real household consumption per person (urban areas),	Date 2	2,766	8,292.99	5,511.94	974	3,918.05	2,467.48
school-age children sample	Date 1	1,563	9,306.09	6,077.86	925	2,740.3	1,574.73
Real household consumption per person (rural areas),	Date 2	2,731	4,331.58	2,395.29	6,339	2,093.69	1,395.74
school-age children sample	Date 1	2,252	5,239.34	3,405.59	5,485	1,577.87	875.51
Gini index of real household consumption	Date 2	30,463	0.40		39,035	0.36	
per person, whole sample	Date 1	16,741	0.41		26,526	0.31	
Gini index of real household consumption	Date 2	16,721	0.38		9,827	0.35	
per person (urban areas), whole sample	Date 1	7,181	0.41		5,180	0.33	
Gini index of real household consumption	Date 2	13,742	0.32		29,208	0.29	
per person (rural areas), whole sample	Date 1	9,524	0.32		21,346	0.27	
Gini index of real household consumption	Date 2	5,497	0.36		7,331	0.31	
per person, school-age children sample	Date 1	3,815	0.36		6,428	0.29	
Gini index of real household consumption	Date 2	2,766	0.33		974	0.32	
per person (urban areas), school-age children sample	Date 1	1,563	0.34		925	0.30	

TABLE 1

(continued)

		Morocco			*Vietnam*		
		Observations	*Mean*	*SD*	*Observations*	*Mean*	*SD*
Gini index of real household consumption	Date 2	2,731	0.29		6,339	0.28	
per person (rural areas), school-age children sample	Date 1	2,252	0.30		5,485	0.25	
CV of real household consumption per person	Date 2	5,497	0.74		7,331	0.73	
	Date 1	3,815	0.74		6,428	0.62	
Father's age	Date 2	5,497	42.25	16.91	7,331	39.44	7.55
	Date 1	3,815	42.63	18.09	6,428	39.01	7.70
Mother's age	Date 2	5,497	39.89	8.16	7,331	36.69	6.38
	Date 1	3,815	39.78	8.68	6,428	36.33	6.40
Father's education	Date 2	5,497	0.69	0.46	7,331	7.23	3.47
	Date 1	3,815	0.74	0.44	6,428	6.20	3.72
Mother's education	Date 2	5,497	0.84	0.37	7,331	6.23	3.63
	Date 1	3,815	0.88	0.33	6,428	5.81	3.78
log household size	Date 2	5,497	1.97	0.32	7,331	1.75	0.28
	Date 1	3,815	2.00	0.33	6,428	1.81	0.29
Distance to primary school	Date 2	5,317	1.76	2.36	7,331	21.5	15.2
	Date 1	3,815	2.04	2.00	6,428	16.7	7.0
Urban	Date 2	5,497	0.50	0.50	7,331	0.20	0.40
	Date 1	3,815	0.41	0.49	6,428	0.17	0.38
Father born in the place of current	Date 2	5,497	0.37	0.48			
residence (Morocco only)	Date 1	3,815	0.32	0.47			

Notes: 'Primary', children of 7–12 years in Morocco and 6–11 years in Vietnam; 'secondary', children of 13–15 years in Morocco and 12–18 years in Vietnam. First date is 1991 in Morocco and 1993 in Vietnam; second date is 1998 in both cases. Parent's education is measured by the share of mothers and fathers without any education in Morocco and by years of education in Vietnam. Consumption is expressed in thousands of real 1998 Vietnamese Dongs for Vietnam and in real 1998 Dirhams for Morocco. Distance to school is in kilometres for Morocco and in minutes for Vietnam.
CV, coefficient of variation; SD, standard deviation.

Morocco. Inequality remained stable in Morocco while it rose in Vietnam, although the change in the Gini index appears to be small.

Figure 2 gives the non-parametric regressions on log consumption per person for each country at both dates. We see that the expected concave relationship is generally evident.[9] In Vietnam in particular, there is notably more nonlinearity for the primary school level so we expect distribution to matter more for the lower grades there. This is also true for Morocco, although less so. (The detailed regression results can be found in a Statistical Addendum available from the authors.)

Note that the location of the estimated ϕ functions depends on the precise specification of the control variables. For example, given our choice of references for the various dummy variables in the linear controls the ϕ function for primary schooling in Morocco should be interpreted as the income effect for a 12-year-old girl in a female-headed rural household in the South, whose father never moved from his present location, and for whom both parents attended primary school. In the case of Vietnam, it is the income effect for an 11-year-old girl who is the child of the head in an ethnic majority female-headed household in the rural Red River Delta.

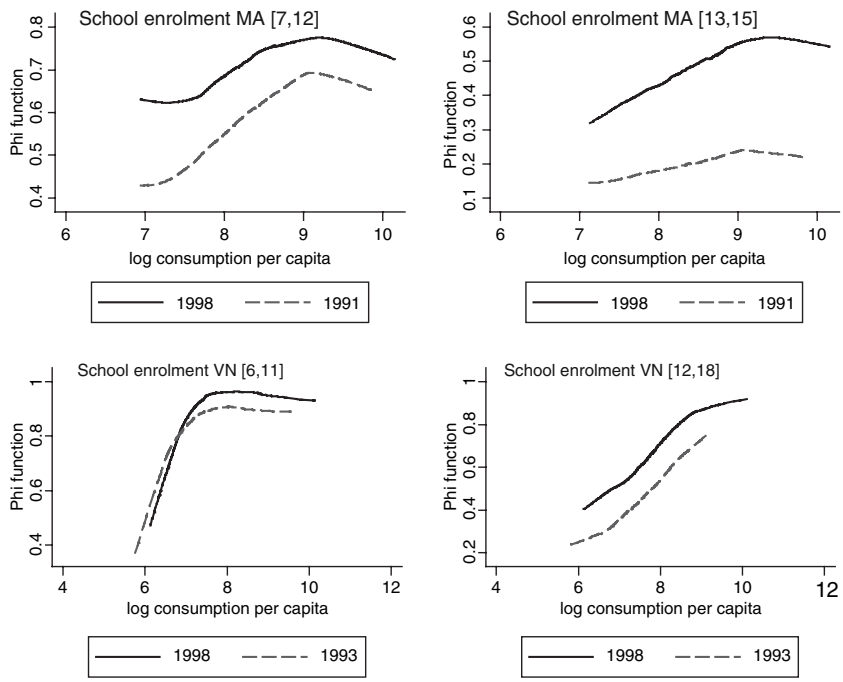

Figure 2. Non-parametric regression functions in partial linear models for school enrolments in Morocco (1991–98) and Vietnam (1993–98) over time
Note: MA, Morocco; VN, Vietnam.

[9]Note that concavity of ϕ in y does not require that the non-parametric regression function is concave in $\ln y$.

We begin with the decomposition of changes over time in school enrolments for boys and girls (Table 2). We give results for both primary and secondary school-age groups and using both the initial and final years as the base. The choice of reference turns out to matter little in this case and we will focus the discussion on the results for the initial year. (Later, we give an example where the choice of reference matters more.)

In Morocco, we see that in 1991, 61% of children (53% of girls and 71% of boys) aged 7–12 years were enrolled in school. By 1998 this had risen to 76%. Given that there was only modest change in mean income or the distribution of income, it is not too surprising that the bulk of the gain is attributed to the structural and non-income factors. However, dominance of the model component is striking. Of the 15% point increase in the enrolment rate, non-income factors account for 4% whereas the changes over time in the model's parameters – the structure

TABLE 2

Decompositions for school enrolment changes over time for boys and girls

	Primary school enrolment rate (%)*			*Secondary school enrolment rate† (%)*		
	Girls + boys	*Girls only*	*Boys only*	*Girls + boys*	*Girls only*	*Boys only*
Morocco						
Baseline value (1991)	61.48	52.99	70.29	16.02	10.58	21.59
Total increase 1991–98	14.77	15.62	13.06	37.48	36.23	38.65
Decomposition using initial year as reference						
Growth	−1.21	−1.18	−1.43	−0.05	−1.32	0.73
Redistribution	0.33	−0.05	1.1	0.03	0.98	−0.52
Non-income factors	4.24	5.75	2.09	1.40	−0.17	3.13
Model	11.39	11.27	11.16	35.95	36.52	35.01
Decomposition using final year as reference (times −1)						
Growth	−0.39	−0.35	−0.28	−0.9	−1.65	−0.83
Redistribution	−0.05	−0.23	−0.22	0.31	1.12	0.02
Non-income factors	3.55	4.95	1.46	5.73	8.17	3.33
Model	11.66	11.42	11.96	32.18	28.37	35.83
Vietnam						
Baseline value (1993)	86.54	86.86	86.52	43.29	36.61	50.10
Total increase 1993–98	6.65	5.28	7.98	21.63	23.10	20.10
Decomposition using initial year as reference						
Growth	1.37	1.40	1.61	5.78	4.78	6.30
Redistribution	0.73	0.30	1.00	−0.69	−0.36	−0.99
Non-income factors	−0.10	0.28	−1.44	0.63	3.12	−1.85
Model	4.54	3.16	6.64	15.72	15.10	16.84
Decomposition using final year as reference (times −1)						
Growth	1.63	2.12	1.09	6.59	6.31	6.88
Redistribution	1.39	0.47	2.37	−1.40	−1.26	−1.51
Non-income factors	0.67	1.85	0.08	2.61	4.55	0.54
Model	2.84	0.71	4.27	13.65	13.05	14.40

Notes: 1, enrolled; 0, not enrolled. *'Primary', children of 7–12 years in Morocco and 6–11 years in Vietnam; †'secondary', children of 13–15 years in Morocco and 12–18 years in Vietnam.

component – accounted for 11%. The 'pure' effect of the (negative) growth in mean consumption was to reduce enrolments; in other words, a distribution-neutral growth process with no other changes (in inequality, other 'non-income' covariates, or in the model's parameters) would have decreased the enrolment rate to 60% in 1998 (down from 61% in 1991), as compared with the actual figure in 1998 of 76%. Changes in distribution account for almost none of the increase in the enrolment rate.

The quantitative importance of the model component is evident in Table 2 in all the decompositions, for both countries. The (large) increase of 37% points in the enrolment rate for Morocco in the 13–15 age group is nearly entirely because of this component. It is worth noting that some children in this age group are likely to be enrolled in primary school despite being of secondary school age. In fact, given that children frequently enter school after age 6, many may remain in primary school past the theoretical age.

What accounts for this model component? The answer is plain from Figure 2, where we see large upward shifts in the ϕ functions. Controlling for the other (non-income) variables in our models, there was a sizeable increase in school enrolment at given household consumption per person. The 'intercept effect', at the lowest consumption level, entailed a 20% point increase in the enrolment rate. Indeed, the average upward shift in the ϕ functions evident in Figure 2 is greater than the model component of 11% in Table 2, implying that changes in the model's parameters had an overall *negative* effect on the enrolment rates. On studying the regression results we find that this is mainly because of changes in the regression coefficients on birth order, the age of the father and whether the head is male. As this upward shift in the ϕ function is found on controlling for the covariates in our model we interpret it as an economy-wide factor. Note, however, that the size of the effect stemming from the shift in the ϕ function vs. changes in the parameters on the linear control variables depends on the model's specification. (Changing the reference for one or more dummy variables, or any additive transformation of the linear controls, would change the location of the estimated ϕ function, similar to any linear regression with dummy variables.)

To help understand these two effects working in opposite directions within the model component consider the changing effect of birth order in Morocco. The detailed regression results indicate that there was a more pronounced positive gradient in enrolments in Morocco in 1991 than in 1998. This 'flattening' of the birth-order gradient reflected a rise over time in the enrolment rates of children who were amongst the first born. At the same time, there was an overall upward shift in enrolment rates at given mean consumption and given values of the controls, including birth order. The negative effect of the change in the coefficient on birth order in the structure component of the decomposition is then needed to compensate for the overestimation of the enrolment rates of the youngest children (highest birth order) that is implied by the overall upward shift in the ϕ function. In other words, the structural change for young children is the net effect of the positive upward shift in enrolments across the board with a partially offsetting negative effect stemming from the fact that first born

children are catching up over time with the youngest. Figure 1 illustrates the two effects.

Although the attribution of this model component is unclear, its pronounced effect on schooling is at least consistent with the policy effort made by Morocco towards education during this period. These efforts, which included the provision of new school infrastructure, are likely to have played a major role (as discussed in section IV).

New infrastructure provision also contributes to the non-income factor component. In fact, the largest changes between 1991 and 1998 in non-income factors included in the regression concern distances to primary school which declined by nearly 15% on average. A far away school is found to be more detrimental to girls than to boys which explains why the non-income factor is larger for girls in the upper left panel of Table 2.

The other factor which plays a large role in the non-income component is a decline in the share of illiterate fathers from 72% to 68%. Here again, because illiterate fathers have a more pronounced negative impact on girls, the change contributes to the larger non-income component in the girls' model.

In Vietnam, economic growth played a more important role, but even here the model component turns out to be the dominant factor (Table 2). Again the shift in the ϕ functions in Figure 2 accounts for the bulk of this component. The importance of the model component suggests economy-wide behavioural changes that could be a sign of the success of the efforts to promote education during the decade.

Urban–rural comparisons naturally figure prominently in a developing country context. We provide two sets of such decompositions. In Table 3, we decompose the change over time in the school enrolment rate separately for urban and rural areas of both countries. At the primary school age, there was only a small increase in the enrolment rate in urban areas of either country, which reflects the high initial base. Here again structure plays the main role. It also played an important role in accounting for the (much larger) gains seen in the primary school enrolment rate for rural areas, notably in Morocco where it dwarfs the strong negative effect of the mean income component, which reflects the sharp decline in real consumption witnessed by rural areas. Although a cautionary note on attribution remains, the model component is at least consistent with the policy emphasis on increasing access to schooling in rural areas during the 1990s (e.g. through rural school construction, and programmes targeted towards rural girls). At the secondary level, growth accounted for a large share (about two-thirds) of the increase in enrolments in urban Vietnam, but was secondary to the model component in rural areas.

As is evident in Table 3, there are large urban–rural disparities in schooling in both countries, although more so in Morocco. Figure 3 gives the non-parametric regressions on log consumption per person separately for the urban and rural areas in 1998. Here nonlinearities are also more pronounced for the lower grades in the rural areas of both Morocco and Vietnam, but only for the urban areas of Vietnam in the case of enrolments beyond the primary level.

TABLE 3

Decompositions for urban–rural differences over time for Morocco and Vietnam

	Primary school enrolment rate (%)*		*Secondary school enrolment rate† (%)*	
	Urban	*Rural*	*Urban*	*Rural*
Morocco				
Baseline value (1991)	87.87	43.31	18.25	14.50
Total increase 1991–98	3.40	17.29	55.9	14.64
Decomposition (initial year as reference)				
Growth	−0.9	−2.94	0.42	0.45
Redistribution	−0.15	1.28	−0.52	−0.35
Non-income factors	1.75	−0.84	3.7	2.33
Model	2.9	19.95	51.70	11.4
Vietnam				
Baseline value (1993)	95.22	85.10	60.43	40.29
Total increase 1993–98	2.12	7.58	17.29	22.67
Decomposition (initial year as reference)				
Growth	1.10	1.86	11.17	4.15
Redistribution	0.66	0.63	−0.71	−0.17
Non-income factors	−1.13	0.21	0.08	1.77
Model	1.57	4.81	5.97	17.06

Notes: 1, enrolled; 0, not enrolled. *'Primary', children of 7–12 years in Morocco and 6–11 years in Vietnam; †'secondary', children of 13–15 years in Morocco and 12–18 years in Vietnam.

Table 4(a) decomposes the urban–rural differences found in 1998, using rural areas as the reference. The large disparity in primary school enrolment rates between urban and rural areas of Morocco – 91% as compared with 60% – is largely attributed to the role of *non-income* factors. Among these, it is worth underlining the average distance to school – nearly twice as much in rural areas – and the share of illiterate parents, also much higher there. These variables are very negative contributors to the enrolment probability in rural areas. The (non-negligible) differences in mean consumption between urban and rural areas are clearly not the reason for the difference in schooling outcomes in Morocco. Indeed, the mean income component turns out to be negative (although small); this is because the function ϕ for rural areas (the reference) turns out to be negatively sloped in the relevant interval (between the urban and rural means), as shown in Figure 3. In contrast, for urban areas, the function is positively sloped in the relevant interval. So this illustrates the potential sensitivity to the choice of reference.

To see this more clearly, Table 4(b)[10] gives the results using urban areas as the reference. In this case, the higher mean consumption in urban areas contributes to

[10]To render results directly comparable between the two panels of Table 4, results in panel (b) have been multiplied by −1.

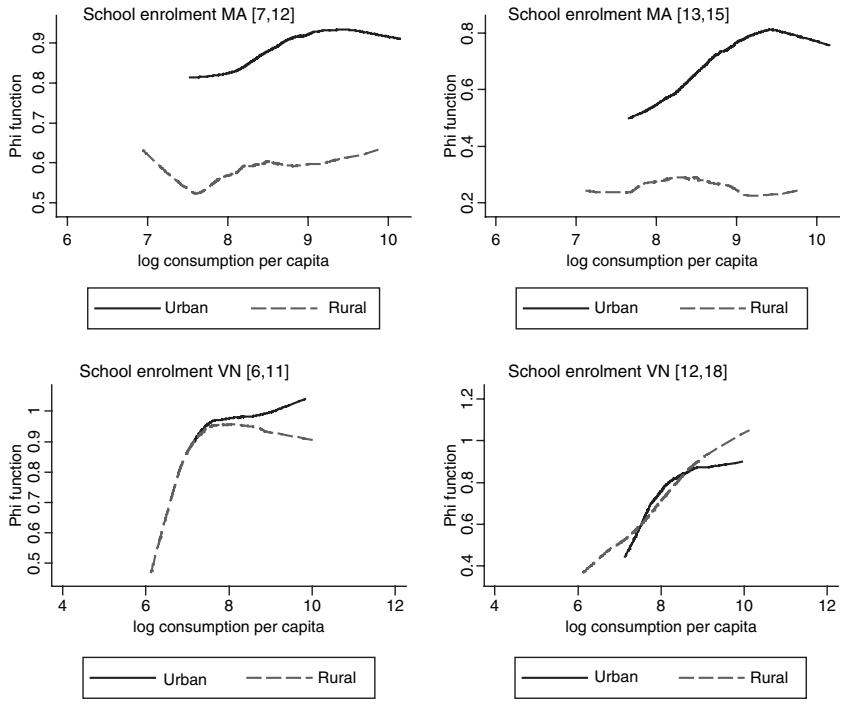

Figure 3. Non-parametric regression functions in partial linear models for school enrolments in urban and rural areas for 1998
Note: MA, Morocco; VN, Vietnam.

the (now negative) difference in school enrolments between urban and rural areas of Morocco, but now the model component is dominant. In this case the apportionment of the total difference between the non-income and structural factors is found to be highly sensitive to the choice of reference, although the aggregate of these two components ($N + M$) is affected rather little. This reversal might be the result of the fact that the linear regression on non-income factors explains as little as 4% of the variance in school enrolment in 1998 for urban areas, whereas it explains 15% of the variance in rural areas. Hence, when urban areas are used as the reference, the decomposition hardly captures any of the changes owing to non-income factors, while everything is captured by the model component that plays a residual role here. In Vietnam, at the secondary school level, the mean income component is by far the most important component of the decomposition, and this remains true when we change the reference.

In Table 5 we apply the decomposition method to the (sizeable) ethnic group disparities in schooling found in Vietnam, whereas in Table 6 we give the corresponding decompositions over time for each ethnic group. In 1993, the primary-level school enrolment rate for minorities was 23% points lower than for the majority (Kinh)

TABLE 4

Decompositions for school enrolment disparities between urban and rural areas in Morocco and Vietnam in 1998

	Primary school enrolment rate (%)*			*Secondary school enrolment rate† (%)*		
	Girls + boys	*Girls only*	*Boys only*	*Girls + boys*	*Girls only*	*Boys only*
(a) Rural areas as reference						
Morocco 1998						
Rural rate	60.59	48.47	72.55	29.15	17.53	40.39
Total difference (urban–rural)	30.68	40.91	20.49	45.00	54.08	37.04
Decomposition						
Growth	0.37	−3.49	2.22	−2.85	−3.35	−4.72
Redistribution	1.73	2.81	1.31	2.49	3.15	4.07
Non-income factors	31.56	44.22	22.49	32.43	37.19	35.13
Model	−3.34	−2.82	−5.70	12.70	16.92	2.15
Vietnam 1998						
Rural rate	92.68	91.41	94.23	62.96	56.68	69.27
Total difference (urban–rural)	4.66	6.52	3.39	14.76	22.09	8.47
Decomposition						
Growth	0.13	−1.10	1.29	14.04	12.85	15.48
Redistribution	2.62	2.97	1.89	−1.83	−1.47	−2.18
Non-income factors	7.53	10.31	5.59	−5.54	6.96	−16.58
Model	−4.83	−5.56	−4.37	6.79	2.42	10.56
(b) Urban areas as reference						
Morocco 1998						
Decomposition (times −1)						
Growth	5.29	6.08	5.40	12.89	12.90	14.16
Redistribution	−1.51	−1.68	−2.16	−5.34	−6.65	−6.63
Non-income factors	1.45	0.59	1.69	3.21	3.61	3.40
Model	25.08	35.71	15.39	34.00	44.05	25.71
Vietnam 1998						
Decomposition (times −1)						
Growth	0.97	4.87	−1.14	12.26	15.10	12.00
Redistribution	1.22	−0.20	−0.36	−1.88	−3.53	−0.97
Non-income factors	3.54	−47.60	5.10	−2.81	−3.63	−4.71
Model	−0.29	49.53	0.81	5.88	12.83	0.95

Notes: 1, enrolled; 0, not enrolled. *'Primary', children of 7–12 years in Morocco and 6–11 years in Vietnam; †'secondary', children of 13–15 years in Morocco and 12–18 years in Vietnam.

group.[11] The enrolment rates for both groups rose over time, and the gap narrowed appreciably to about 15% points (Table 5). At the earlier date we find that over half of the gap is accountable to the difference in mean consumption between the Kinh and

[11]The majority ethnic group comprises the Kinh and Han Chinese. In 1993, our data includes 2,833 and 2,711 majority group 11–16 and 12–18 year olds, respectively, whereas the numbers for 1998 are 2,729 and 3,473. The ethnic minorities represent all 52 other ethnic groups and account for 486 and 380 observations in 1993, and 551 and 560 observations in 1998 for the primary- and secondary-aged groups, respectively.

TABLE 5

Decompositions for school enrolment disparities between ethnic majority and minority groups in Vietnam

	Primary school enrolment rate (%)*			*Secondary school enrolment rate† (%)*		
	Girls + boys	*Girls only*	*Boys only*	*Girls + boys*	*Girls only*	*Boys only*
Vietnam 1993						
Minority rate	66.87	68.44	66.67	34.74	29.21	40.93
Total difference (majority–minority)	23.14	21.62	23.58	9.82	8.53	10.67
Decomposition						
Growth	12.85	10.60	14.56	2.07	−2.34	6.50
Redistribution	−6.38	−6.67	−5.09	3.00	4.91	−0.71
Non-income factors	12.41	12.75	15.25	1.57	2.22	0.04
Model	4.53	4.76	−0.80	3.69	5.13	4.99
Vietnam 1998						
Minority rate	80.58	78.60	84.50	61.43	54.85	68.90
Total difference (majority–minority)	15.21	16.51	12.17	4.13	5.87	1.52
Decomposition						
Growth	0.72	4.07	−3.99	11.58	13.40	10.17
Redistribution	3.98	0.05	5.13	−2.44	−5.44	−4.01
Non-income factors	15.89	26.91	6.51	−13.45	−14.49	−13.04
Model	−4.76	−14.61	5.38	8.05	11.81	7.96

Notes: 1, enrolled; 0, not enrolled. The ethnic majority group includes Kinh and Han; the minority group includes all remaining ethnic groups. *'Primary', children of 6–11 years; †'secondary', children of 12–18 years.

the (poorer) minorities; the proportion is higher for boys than for girls. The difference in distribution tended to narrow the gap. Non-income factors accounted for as high a share as the difference in mean consumption. This pattern changed noticeably by 1998. The non-income differences between the two ethnic groups continue to play an important role, but differences in mean income are no longer an important factor in accounting for primary school enrolment differences. However, this has switched to the higher (secondary) age group. For this group, the non-income factors work in the opposite direction – reducing the enrolment gap. From Table 6 it can be noted that the share of the mean income component in explaining changes over time for each group is larger at the primary level for the minority group, whereas the reverse is true for the secondary level. This suggests that growth benefitted all groups, but that as the Kinh were already close to full primary enrolment, all of the improvement for them is concentrated on the secondary level.

Finally, in Tables 7 and 8, we give the results for a decomposition of the disparities in enrolments between children whose father is literate and those whose father is not in Morocco. In 1991, the enrolment rate of primary school-aged children with illiterate fathers was 25 percentage points below that for other children. Table 7 shows that differences in average income between these two groups plays an important role, notably for boys for whom it accounts for two-thirds of the difference at the primary

TABLE 6

Decompositions for school enrolment changes over time for ethnic majority and minority in Vietnam

	Primary school enrolment rate (%)*			*Secondary school enrolment rate† (%)*		
	Girls + boys	*Girls only*	*Boys only*	*Girls + boys*	*Girls only*	*Boys only*
Vietnam 1993–98 ethnic majority						
Baseline value	90.01	90.07	90.24	44.56	37.75	51.60
Total difference (1998–93)	5.78	5.05	6.43	21.00	22.97	18.82
Decomposition						
Growth	2.06	1.93	1.98	6.02	4.99	6.56
Redistribution	−0.38	−0.19	−0.26	−0.65	−0.60	−0.97
Non-income factors	−1.30	−0.65	−2.90	0.23	4.31	−3.60
Model	5.30	4.09	7.32	14.94	13.50	16.59
Vietnam 1993–98 ethnic minority						
Baseline value	66.87	68.44	66.67	34.74	29.21	40.93
Total difference (1998–93)	13.71	10.15	17.84	26.69	25.64	27.97
Decomposition						
Growth	6.80	6.23	5.73	0.96	0.11	2.21
Redistribution	−2.93	−4.53	1.16	1.75	1.25	0.84
Non-income factors	4.99	5.88	6.97	7.51	1.49	11.71
Model	4.41	2.63	3.17	16.92	23.99	13.56

Notes: 1, enrolled, 0, not enrolled. *'Primary', children of 6–11 years; †'secondary', children of 12–18 years.

school level in 1991. Interestingly, the weight of the mean income component is vastly reduced in 1998 at the primary level, but not at the secondary level. This could well reflect the fact that the aforementioned policy efforts were mainly concentrated on bridging the enrolment gap at the primary level. The large increase in the difference in school enrolments for children of secondary school age between the two years suggests that children of literate fathers were better able to reap the benefits of increases in secondary schooling opportunities.[12] Here, the weight of the difference in average income in accounting for this result is fairly important. The role of policy changes is reflected in Table 8 by the consistently dominant role played by structural factors in accounting for the change in enrolments over time for the two groups separately. Girls from literate parents seem to have particularly benefitted from the structural changes over the 1990s.

We also tested the sensitivity of the decompositions to using predicted consumption per person, using x as a vector of instrumental variables, using the nonlinearity for identification. (As the model is inherently nonlinear the standard identification conditions for linear models do not strictly apply.) The basic pattern of the results in

[12]The negative difference between the enrolment rate of secondary school-age children with a literate father and those without in 1991 is explained by the fact that the latter take more time to finish primary school (or enter school at an older age). This is still true in 1998 but is massively compensated by the actual increase in secondary-level enrolments of children with a literate father.

TABLE 7

Decompositions for school enrolment disparities between children with a literate vs. an illiterate father in Morocco

	Primary school enrolment rate (%)*			*Secondary school enrolment rate† (%)*		
	Girls + boys	*Girls only*	*Boys only*	*Girls + boys*	*Girls only*	*Boys only*
Morocco 1991						
Illiterate rate	54.59	44.80	64.73	17.01	11.42	22.82
Total difference (literate–illiterate)	25.52	31.72	20.70	−4.17	−3.29	−6.02
Decomposition						
Growth	10.52	7.98	13.50	0.92	1.41	2.04
Redistribution	−0.93	−4.05	−0.81	−2.27	1.12	−7.33
Non-income factors	7.04	11.00	−1.37	−9.84	−18.87	9.64
Model	7.91	16.70	7.90	7.20	13.34	−10.84
Morocco 1998						
Illiterate rate	70.85	62.45	78.74	47.29	39.75	54.95
Total difference (literate–illiterate)	16.69	19.63	14.11	22.09	26.14	19.66
Decomposition						
Growth	1.98	3.29	2.21	7.00	3.59	8.28
Redistribution	0.13	0.15	0.22	−2.21	0.61	−2.91
Non-income factors	8.24	11.34	1.96	15.13	16.17	14.60
Model	5.86	4.39	9.58	0.95	4.94	−1.86

Notes: 1, enrolled, 0, not enrolled; *'Primary', children of 7–12 years; †'secondary', children of 13–15 years. Reference, father is literate.

Table 2 was affected little by this change. Table 9 gives examples, comparing selected decompositions for Vietnam from the previous tables with those based on predicted consumptions.

VII. Conclusions

Changes in aggregate HD outcomes can be additively decomposed into four components: (i) a pure 'growth' effect associated with differences in mean income; (ii) a redistribution effect attributed to differences in the distribution of income; (iii) 'non-income' factors; and (iv) a model component reflecting any differences in the model parameters – the 'HD returns' to income or non-income characteristics. The analytical complication over other decompositions (in the Blinder–Oaxaca tradition) is that to convincingly estimate the decomposition we have to use a flexible, non-parametric, representation of the economic gradient in HD outcomes across households. Non-parametric regression methods allow us to implement the decomposition proposed here, and we have programmed it in STATA.

In applying this decomposition tool to data for Morocco and Vietnam, we find that growth and distributional change have played only a modest role in the changes in school enrolments over time observed in both countries. There are a couple of

TABLE 8

Decompositions for school enrolment changes over time for children with a literate vs. an illiterate father in Morocco

	Primary school enrolment rate (%)*			*Secondary school enrolment rate† (%)*		
	Girls + boys	*Girls only*	*Boys only*	*Girls + boys*	*Girls only*	*Boys only*
Morocco 1991–98 literate father						
Baseline value	80.11	76.52	85.43	12.84	8.13	16.80
Total increase (1991–98)	7.43	5.55	7.42	56.54	57.75	57.80
Decomposition						
Growth	−1.39	−1.21	−0.95	−0.02	−0.70	1.26
Redistribution	−1.20	−0.42	−1.68	0.48	−0.47	0.18
Non-income factors	0.76	3.19	0.59	0.13	−1.09	2.67
Model	9.60	3.77	10.42	54.53	58.74	52.60
Morocco 1991–98 illiterate father						
Baseline value	54.59	44.80	64.73	17.01	11.42	22.82
Total increase (1991–98)	16.26	17.64	14.01	30.28	28.33	32.12
Decomposition						
Growth	−0.27	−0.27	−0.04	0.16	−0.52	1.39
Redistribution	−0.28	−0.59	−0.11	−0.28	0.08	−1.12
Non-income factors	6.53	9.51	4.16	2.68	2.34	4.58
Model	10.13	9.13	9.62	27.70	26.27	27.26

Notes: 1, enrolled, 0, not enrolled. *'Primary', children of 7–12 years; †'secondary', children of 13–15 years.

TABLE 9

Sensitivity test using predicted expenditures for Vietnam

	Primary school enrolment rate (%)*				*Secondary school enrolment rate† (%)*			
	Girls + boys expenditure		*Rural expenditure*		*Girls + boys expenditure*		*Rural expenditure*	
	Actual	*Predicted*	*Actual*	*Predicted*	*Actual*	*Predicted*	*Actual*	*Predicted*
Baseline value (1993)	86.54	88.32	85.10	87.06	43.29	44.45	40.29	41.44
Total increase 1993–98	6.65	5.33	7.58	5.83	21.63	26.62	22.67	27.94
Decomposition (initial year as reference)								
Growth	1.37	0.56	1.86	1.41	5.78	3.15	4.15	6.54
Redistribution	0.73	−0.25	0.63	−0.76	−0.69	−0.25	−0.17	−1.25
Non-income factors	−0.10	2.30	0.21	3.05	0.63	0.42	1.77	3.71
Model	4.54	3.06	4.81	2.20	15.72	22.14	17.06	17.97

Notes: 1, enrolled; 0, not enrolled; *'Primary', children of 6–11 years; †'secondary', children of 12–18 years.

notable exceptions to this generalization; in particular, growth emerges as the most important factor in the changes in higher-level school enrolments in urban Vietnam and also for ethnic minority primary school enrolments. However, taken as a whole, our results for these two countries do not suggest that aggregate economic growth or

changes in the distribution of income have played an important role; nor did improvements in relevant non-income factors (such as parental education) account for much of the aggregate gains over time. Rather, the bulk of the changes observed over time are accountable to changes in the structure of the model linking these variables to schooling attainments. As a generalization, structure is the dominant factor in the gains observed in school enrolment rates over time in both countries. This is generally the case nationally as well as within urban and rural areas, for each of the ethnic groups in Vietnam and the literate, illiterate groups in Morocco. Our decomposition cannot tell us what economy-wide policies or other factors drive these structural changes, but it is at least suggestive that there were substantial public policy efforts at increasing enrolments; increases in the overall economic returns to schooling may also have played a role, possibly helped by more widely shared knowledge about those returns.

Possibly it is not too surprising that the distributional effect is so small in our case studies, as there was not much change in distribution in either country. Also, given the lack of growth in Morocco, it is not surprising that growth had so little impact. What is far more telling is the Vietnam case where there was considerable economic growth over the study period, and yet structure was still the dominant factor in the gains over time.

The results look very different when we use our decomposition tool to study the cross-sectional disparities in schooling between socio-economic or geographical groups. Then we find many cases in which structure becomes secondary to both differences in mean incomes and non-income factors and (although less often) to inter-group differences in the distribution of incomes. Differences in higher-level enrolments between Vietnam's urban and rural areas are largely because of differences in mean consumption in 1998. In Morocco, however, non-income factors are dominant in explaining the urban–rural educational differences at both the primary and secondary levels. The disparity in mean consumption between Vietnam's ethnic minority and majority groups is a major factor in educational inequality, although the role of this economic inequality has shifted over time from primary to secondary schooling. Non-income factors also play a big role here, increasing the gap for primary school enrolments at both dates, but reducing ethnic group differences in secondary enrolments in 1998. Finally, income differences also explain a large part of the enrolment gap between children with and without literate fathers in Morocco, although here too, non-income factors explain the largest part.

Final Manuscript Received: October 2009

References

Anand, S. and Ravallion, M. (1993). 'Human development in poor countries: on the role of private incomes and public services', *Journal of Economic Perspectives*, Vol. 7, pp. 133–150.

Aturupane, H., Glewwe, P. and Isenman, P. (1994). 'Poverty, human development and growth:an emerging consensus?' *American Economic Review, Papers and Proceedings*, Vol. 84, pp. 244–249.

Behrman, J. and Knowles, J. (1999). 'Household income and child schooling in viet Nam', *World Bank Economic Review*, Vol. 13, pp. 211–256.
Bhalla, S. and Glewwe, P. (1986). 'Growth and equity in developing countries: a reinterpretation of the Sri Lankan experience', *World Bank Economic Review*, Vol. 1, pp. 35–63.
Bhargava, A. (1999). 'Modeling the effects of nutritional and socioeconomic factors on the growth and morbidity of Kenyan school children', *American Journal of Human Biology*, Vol. 11, pp. 317–326.
Bidani, B. and Ravallion, M. (1997). 'Decomposing social indicators using distributional data', *Journal of Econometrics*, Vol. 77, pp. 125–140.
Blinder, A. (1973). 'Wage discrimination: reduced form and structural estimates', *Journal of Human Resources*, Vol. 8, pp. 436–455.
Commission Spéciale Education Formation (1999). 'Regards sur le Système Education-Formation au Maroc: Réalisations, Problématiques, Dysfonctionnements', Rabat, Maroc.
Datt, G. and Ravallion, M. (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*, Vol. 38, pp. 275–295.
Evans, D., Tandon, A., Murray, C. and Lauer, J. A. (2000). *The Comparative Efficiency of National Health Systems in Producing Health: An Analysis of 191 Countries*, GPE Discussion Paper 29, World Health Organization, Geneva.
Filmer, D. and Pritchett, L. (1999). 'The effect of household wealth on educational attainment: evidence from 35 countries', *Population and Development Review*, Vol. 25, pp. 85–120.
Glewwe, P. and Jacoby, H. (2004). 'Economic growth and the demand for education: is there a wealth effect?' *Journal of Development Economics*, Vol. 74, pp. 33–51.
Government-Donor –NGO Working Group (1999). *Vietnam: Attacking Poverty*, Vietnam Development Report 2000, Hanoi, Vietnam.
Gupta, S. and Verhoeven, M. (2001). 'The efficiency of government expenditure: experience from Africa', *Journal of Policy Modeling*, Vol. 23, pp. 433–467.
Hall, P., Kay, J. W. and Titterington, D. M. (1990). 'Asymptotically optical difference-based estimation of variance in nonparametric regression', *Biometrika*, Vol. 77, pp. 521–528.
Jayasuriya, R. and Wodon, Q. (2003). *Efficiency in Reaching the Millennium Development Goals*, World Bank Working Paper No.9, World Bank, Washington DC.
Lokshin, M. (2006). 'Difference-based semiparametric estimation of partial linear regression models', *STATA Journal*, Vol. 6, pp. 377–384.
Moore, M., Leavy, J., Houtzager, P. and White, H. (1999). *Polity Qualities: How Governance Affects Poverty*, Poverty Research Programme, Institute of Development Studies, University of Sussex.
Nguyen, N. (2004). 'Trends in the education sector', chapter 12, in Glewwe, Agrawal and Dollar (eds), *Economic Growth, Poverty and Household Welfare: Policy Lessons from Vietnam*, Washington, DC: World Bank Regional and Sectoral Studies, World Bank,Washington, DC, pp. 425–465.
Oaxaca, R. (1973). 'Male–female wage differentials in urban labor markets', *International Economic Review*, Vol. 9, pp. 693–709.
Pritchett, L. and Summers, L. (1996). 'Wealthier is healthier', *Journal of Human Resources*, Vol. 28, pp. 197–220.
Ranis, G., Stewart, F. and Ramirez, A. (2000). 'Economic growth and human development', *World Development*, Vol. 28, pp. 197–220.
Ravallion, M. (2005). 'On measuring aggregate "Social Efficiency" ', *Economic Development and Cultural Change*, Vol. 53, pp. 273–292.
Sen, A. (1981). 'Public action and the quality of life in developing countries', *Oxford Bulletin of Economics and Statistics*, Vol. 43, pp. 287–319.
Sen, A. (1988). 'Sri Lanka's achievements: when and how?' in Srinivasan T. N. and Bardhan P. K. (eds), *Rural Poverty in South Asia*, Columbia University Press, New York, pp. 549–556.

Strauss, J. and Thomas, D. (1995). 'Human resources: empirical modeling of household and family decisions', in Behrman J. and Srinivasan T. N. (eds), *Handbook of Development Economics*, Vol. 3, North-Holland, Amsterdam, pp. 1883–2023.

United Nations Development Programme (UNDP), (1996). *Human Development Report*, Oxford University Press, New York.

van de Walle, D. (2004). 'The static and dynamic incidence of Vietnam's Public Safety Net', chapter 6, in Glewwe, Agrawal and Dollar (eds), *Economic Growth, Poverty and Household Welfare: Policy Lessons from Vietnam*, World Bank Regional and Sectoral Studies, World Bank, Washington, DC, pp. 189–227.

van de Walle, D. and Gunewardena, D. (2001). 'Sources of ethnic inequality in Vietnam', *Journal of Development Economics*, Vol. 65, pp. 177–207.

Wang, J., Jamison, D., Bos, E., Preker, A. and Peabody, J. (1999). 'Measuring country performance on health: selected indicators for 115 countries', *Health, Nutrition and Population Series*, World Bank, Washington DC.

World Bank (1993). *World Development Report: Investing in Health*, Oxford University Press, New York.

World Bank (1995). 'Vietnam Living Standards Survey (VLSS), 1992–93: basic information', Mimeo, Research Development Group, World Bank, Washington, DC.

World Bank (1997). 'Kingdom of Morocco: growth and labor markets: an agenda', Report No. 16598-MOR, Volumes I and II: Main Report and Annexes, World Bank, Washington DC.

World Bank (2000). 'Viet Nam Living Standards Survey (VNLSS), 1997–98: basic information', mimeo, Research Development Group, World Bank, Washington, DC.

World Bank (2001). 'Kingdom of Morocco: poverty update', World Bank, Washington DC.

World Health Organization (1999). *The World Health Report:Making a Difference*, World Health Organization, Geneva.

Yatchew, A. (1998). 'Nonparametric regression techniques in economics', *Journal of Economic Literature*, Vol. 36, pp. 669–721.