WPS3592

Which Inequality Matters?

Growth Evidence Based on Small Area Welfare Estimates in Uganda

Youdi Schipper and Johannes G. Hoogeveen*

Abstract

Existing empirical studies on the relation between inequality and growth have been criticized for their focus on income inequality and their use of cross-country data sets. This paper uses two sets of small area welfare estimates – often referred to as poverty maps – to estimate a model of rural per capita expenditure growth for Uganda between 1992 and 1999. We estimate the growth effects of expenditure and education inequality while controlling for other factors such as initial levels of expenditure and human capital, family characteristics and unobserved spatial heterogeneity. We correct standard errors to reflect the uncertainty due to the fact that we use estimates rather than observations. We find that per capita expenditure growth in rural Uganda is affected positively by the level of education as well as by the degree of education inequality. Expenditure inequality does not have a significant impact on growth.

World Bank Policy Research Working Paper 3592, May 2005

* Schipper is with the Vrije Universiteit Amsterdam and consultant to the World Bank. Hoogeveen is with the World Bank. Please send correspondence to both yschipper@feweb.vu.nl and jhoogeveen@worldbank.org. We are grateful to the Uganda Bureau of Statistics, Entebbe, for their help with the provision of the survey and census data. For useful comments and other forms of help we would like to thank Chris Elbers and seminar participants at the Vrije Universiteit.

1. Introduction

Poverty reduction is fully determined by the rate of growth of mean per capita expenditure and (changes in) the distribution of expenditure (Bourguignon, 2004; Ravallion, 1997). This puts the empirical relation between inequality and growth at the heart of poverty reduction strategies. Despite decades of theoretical and empirical research on this relation, the aggregate evidence on the effect of inequality on growth is inconclusive on the sign and robustness of the effect. Recent debate has focused on (1) the nature of the inequality measure (income versus physical and human capital inequality, e.g. Castello and Domenech, 2002; Elbers and Gunning, 2004) and (2) problems of measurement and inference in macro data sets (Atkinson and Brandolini, 2001; Banerjee and Duflo, 2003). A further, remarkable feature of the literature in this area is an apparent geographical mismatch: although the effect of inequality on growth has important implications for poverty and African levels of poverty incidence remain persistently high, empirical evidence on the inequality-growth relation is virtually absent for Africa.[1]

Taking into account points (1) and (2), this paper provides empirical evidence for Uganda. Regarding (1), recent studies have questioned the focus on income inequality as a determinant of growth. A number of theoretical papers show that inequality in human capital determines both income inequality and income growth (Benabou, 1996; Galor and Tsiddon, 1997; Elbers and Gunning, 2004). On the empirical side,

[1] A number of recent cross-country growth studies focus on sub-Sahara African countries (Block, 2001; Bloom and Sachs, 1998). There are also growth studies for Africa using household level data (Deininger and Okido, 2003, for Uganda and Dercon, 2001, for Ethiopia). However, these analyses do not consider the effect of inequality on growth. Mbabazi et al. (2002) estimate a cross-country inequality growth regression for a set of developing countries, a quarter of which are in Sub-Sahara Africa.

Deininger and Squire (1998) find that the coefficient on income inequality in growth regressions is not robust to the inclusion of regional dummies. Birdsall and Londono (1997) show that once land and human capital inequality are entered in a cross-country growth regression, income inequality no longer has a significant effect on growth. Similarly, Castello and Domenech (2002) find a negative and robust growth effect of human capital inequality, but no robust income inequality effect. We construct a human capital Gini coefficient using census data and estimate the growth effect of inequality in human capital too. Our results indicate that it is human capital inequality rather than income inequality that affects growth; however, the effect we find for Uganda is positive.

On point (2), we present a study based on micro data for Uganda, which allows us to avoid data comparability problems that affect cross-country studies. While there is a small number of inequality-growth studies that use micro (meso) data, we believe this is the first such study for a relatively small country and the first for an African country. Our analysis has been made possible by recent advances in the field of small area welfare estimation (Hentschel et al., 2000). The data set thus consists partly of imputed variables or so-called small area welfare estimates. These are obtained by deriving expenditure estimates for a complete population, combining information from a census and a survey (Elbers et al., 2003).

The paper is organized as follows. Section 2 briefly reviews data problems encountered in empirical inequality-growth studies and introduces small area welfare estimates as an alternative data source. In Section 3 we present our growth model and descriptive statistics. Section 4 presents a discussion of econometric issues that need to be addressed; in particular, we discuss a variance correction that is required because we use imputed data. We present results in Section 5 and conclude in Section 6.

2. Data: macro, micro and small area welfare estimates

The effect of income inequality on economic growth is the subject of a large literature. Aghion et al. (1999) and Thorbecke and Charumilind (2002) review this literature and show that theory does not provide firm predictions of the sign of the effect.[2] Empirical studies in the 1990s have been ".. impressively unambiguous .." (Aghion et al., 1999, p1617) in concluding that the growth effect of inequality is negative, but more recently some authors have obtained contrasting results (e.g. Forbes, 2000; Banerjee and Duflo, 2003). The most common denominator in these studies is the nature of the data used: the empirical inequality-growth literature is largely based on cross-country data.

Cross-country inequality-growth studies, while providing the bulk of existing empirical evidence, have been criticized for a number of reasons. First and foremost, the quality and internal consistency of datasets, in particular inequality series, have been questioned. Deininger and Squire (1996) challenge the quality of inequality data in growth regressions and offer an improved 'high quality' dataset. However, the consistency of inequality data in this set has been challenged by Atkinson and

[2] Positive inequality-growth effects can be attributed to a positive effect on savings, to the existence of investment indivisibilities or to positive incentive effects of inequality. A negative inequality-growth effect can be explained by political tension, instability and demands for redistribution due to inequality, by reduced investment opportunities for the poor, worsened borrowers' incentives and by higher macro-economic volatility. A 'unified' model that aims to reconcile these conflicting effects is presented in Galor (2000): this paper predicts that the effect of inequality on growth is non-linear, with a positive effect at an 'early stage of economic development' and a negative effect at a 'later stage'.

Brandolini (2001), who show that national statistical agencies differ in their income measures so that cross-country comparability of income inequality is questionable. Moreover, changes in definitions or ruptures within country series may suggest structural shifts in inequality without real significance. Banerjee and Duflo (2003) argue that this type of measurement error may seriously distort causal inference in inequality-growth models. Brock and Durlauf (2001) reject causal interpretations in cross-country studies except under quite exceptional conditions. Their main argument is that causal interpretation requires that estimated parameters can be assumed constant, which is not plausible given the importance of country-specific unobserved information (e.g. regarding policy). Deininger and Okidi (2003) also argue that data used in cross-country studies are national aggregates that are likely to lose valuable region or gender specific information; as a result, they question the relevance of cross-country evidence for national policy formation – even in case of perfect data.

A related, but often ignored, measurement issue affecting growth and inequality regressions is related to the way the dependent variable is defined. Consider growth over a period *t* for a country or region *i*, usually specified as

$$gr_i = \frac{y_{i,t} - y_{i,0}}{t} \tag{1}$$

where *y* is a measure of income or expenditure. This measure is often specified as the logarithm of the mean of per capita expenditure over households *h* for country/region *i*, i.e.

$$\ln M(y_i) = \ln\left(\frac{\sum_{h=1}^{H} y_{h,i}}{H}\right) \tag{2}$$

where $M(.)$ denotes an average. Ravallion (1998) points out that the use of the logarithm of mean expenditure introduces a measure of the change in inequality in the error term of a regression with income growth as the dependent variable. The argument is that if I is general measure of inequality

$$I(y_i) = \ln M(y_i) - M(\ln y_i)$$

then we find after rearranging terms:

$$\ln M(y_i) = M(\ln y_i) + I(y_i) \tag{3}$$

The LHS of (3) is the same as (2), and implicitly comprises a measure of inequality. Therefore, rather than the log of household income we should use

$$y_{i,t} = \frac{\sum_{n=1}^{N} \log(y_n)}{N} \tag{4}$$

which is the first term on the RHS of (3). We have addressed this point by calculating growth using (4). Comparing two separate growth regressions, one of which has (3) as dependent variable, we find a clear difference between the estimates (results available, not reported). The (absolute) inequality coefficient in the model that uses definition (3) is much larger and more significant than the alternative that uses (4). This points to a possible spurious effect due to the definition of the variable.

Considering the various data issues, there appear to be two ways forward: higher quality cross-country datasets (Bourguignon 2004) or country specific micro data (Banerjee and Duflo 2003). A problem with the latter is that only surveys for very large countries provide sufficient data points to meaningfully include inequality indicators in a regression while census data typically do not provide the income or wealth variables and covariates needed in a growth regression. As a result, there is only a very small number of inequality-growth studies that use micro or regional data. Ravallion (1998) estimates a household level growth model with local externalities and finds a significant negative effect of asset inequality for rural China. Balisacan and Fuwa (2003) find a positive effect of land inequality on provincial level growth for the Philippines. An important advantage of regional or household data, apart from the usually large number of observations, is that comparability problems are much less severe than in cross-country datasets: the definitions of variables or phrasing of survey questions are generally uniform across regions for a given dataset.

The unavailability of inequality data has long precluded the study of the inequality-growth relation for smaller countries. However, application of techniques of welfare estimation for small area target populations (see Elbers et al., 2003) has recently provided expenditure estimates for all households in Uganda for 1992 and 1999 (Hoogeveen et al., 2003). It is the availability of these estimates that allows us to estimate the effect of income inequality on growth for a smaller country such as Uganda.

Estimates of expenditure and inequality are available for 719 rural sub-counties. Based on comparable household expenditure data, they represent one of the first data sets for Africa with comparable inequality estimates for a substantial number of observations; see Emwanu et al. (2004), for details. A summary of the welfare estimates used in this paper is presented in Table 1. The table confirms that on average poverty fell over the 1990s but that the decrease in poverty was not distributed uniformly. The decline in poverty was lowest in the North as was the growth rate. Initial expenditure inequality was highest in the West region, whereas initial inequality in the three other regions is almost the same.

Table 1

Region	Expenditure Growth 92-99	Expenditure Inequality 92	Poverty ratio 92	99	Poverty change 92-99
Central	0.057	0.302	0.543	0.245	-0.297
East	0.058	0.299	0.661	0.362	-0.299
North	0.019	0.300	0.768	0.678	-0.091
West	0.040	0.327	0.557	0.326	-0.230
National	0.043	0.308	0.633	0.403	-0.229

Note: column entries are (unweighted) region means of sub-county estimates.

3. An empirical inequality – growth model

We estimate, at the sub-county level, yearly per capita expenditure growth over the period 1992 – 1999 as a function of 1992 per capita expenditure, expenditure inequality, human capital inequality, male and female human capital and household demographics. The model can be represented as

$$g_{i,d} \equiv (\tilde{y}_{i,99} - \tilde{y}_{i,92})/7 = \tilde{y}_{i,92} b_1 + \tilde{I}^{\exp}_{i,92} b_2 + I^{\mathrm{edu}}_{i,92} b_3 + \mathbf{X}_{i,92} ? + a_d + u_i \qquad (5)$$

All variables, except for the inequality measures and dummies, are averages by sub-county i: g is the annual expenditure growth rate between 1992 and 1999; y is the logarithm of per capita expenditure; I^{exp} is the Gini coefficient for per capita household expenditure; I^{edu} is the Gini coefficient for the number of years of formal education of the household head. **X** is a matrix of other covariates consisting of human capital (number of years of formal education entered separately for household heads and for spouses), household head age, gender of the household head, adult equivalent household size, and the fraction of own children in the household. Given our approach, we are limited in our choice of covariates in **X** to what the census has to offer. We use district fixed effects, represented by α_d, to control for unobserved spatial heterogeneity; u_i is an error term. The specification includes most variables that are common in (cross-country) inequality-growth studies. We add demographics to account for differences in production technology and fertility, as many theoretical models require. We do not have a measure of 'market distortions', but do not expect the value of this variable to vary much between sub-counties.

A non-standard econometric issue lies in the fact that some of our variables are not observed but imputed as described in the previous section. The imputed variables, expenditure, growth and expenditure inequality, are denoted using tildes. See Table 2 for definitions and summary statistics and Section 4 for a discussion of the estimation issues involved.

Table 2

Variable	Definition*	Mean	Standard error	Minimum	Maximum
gr	Annual growth of log per capita exp, 92-99: $Ln(pcx_{99}) – Ln(pcx_{92})/7$	0.04	0.04	-0.15	0.15
y	Log expenditure pc 92: $Ln(pcx_{92})$	9.46	0.23	8.74	10.20
gini	Expenditure Inequality: Gini coefficient wrt pcx	0.31	0.03	0.20	0.50
gi_hyredu	Education Inequality: Gini coefficient wrt household head education years	0.58	0.09	0.34	0.99
n_hyredu	Household head's education, number of years	3.59	1.00	0.04	8.41
n_spyredu	Spouse's education, number of years	1.36	0.54	0.00	3.67
hage	Age of household head	42.17	2.18	30.89	47.75
hfem	Gender, equals 1 if household head female	0.28	0.08	0.11	0.75
aesize	Adult equivalent household size	3.90	0.55	2.16	8.09
pchild	Percentage of children in household	0.44	0.06	0.12	0.60

*Note: all observations are sub-county means of the household values of the variables mentioned, with the exception of the Inequality measures. No. of observations: 719

4. Estimation

The properties of estimators obtained from downstream[3] regressions using imputed values for welfare indicators are investigated in Elbers et al. (2005). Their main proposition is that coefficients from regressions involving imputed welfare indicators which have been derived with small area estimation techniques, either on the LHS or on the RHS, do not differ systematically from regressions with 'real data'. The intuition for this consistency result is that imputed variables can be regarded a special kind of instrumental variables and may therefore be safely used in estimation. We briefly explore the issues involved in estimation for the general case with imputed values on both the LHS and the RHS of a regression equation.

We consider a simple version of our downstream regression model (omitting inequality measures)

$$g_i = y_i b + \mathbf{x}_i \,?+ u_i \tag{6}$$

The dependent g and the independent y are obtained from upstream imputation; in what follows, imputed variables have tildes in order to distinguish them from 'true' values or observations. Writing imputed values as the difference between true values and an error term, $\tilde{g} \equiv g - w$ and $\tilde{y} \equiv y - x$, we obtain

$$\tilde{g}_i = \tilde{y}_i b + (x_i b - w_i) + \mathbf{x}_i \,? + u_i \tag{7}$$

The β coefficient can be consistently estimated provided that (a) the imputed values $\tilde{g}$ and $\tilde{y}$ are consistent estimators of the conditional expectation of the true welfare measures and (b) the error terms x and w are uncorrelated with the regressors $\tilde{y}$ and $\mathbf{x}$.

Elbers et al. (2005) show that when small area welfare estimates are used (a) is satisfied and (b) is likely to be satisfied. To see the latter, first note that $\tilde{y}$ is imputed per capita expenditure (pcx) or a non-linear measure calculated from pcx, e.g.

[3] It is convenient to refer to our inequality-growth regression as a 'downstream' model so as to distinguish it from the 'upstream' expenditure model which has been used to generate the imputed values.

inequality.[4] Both x and w are prediction errors and are thus orthogonal to the predicted values $\tilde{y}$ and $\tilde{g}$, respectively. Moreover, since $\tilde{y}$ and $\tilde{g}$ are based on the same prediction model, the prediction errors should be orthogonal with respect to both $\tilde{y}$ and $\tilde{g}$.[5] The prediction errors should also be uncorrelated with regressors in $\mathbf{x}$: since the upstream modelling process makes use of as many available instruments as possible, these regressors will have been considered as instruments in the upstream pcx prediction model, ruling out the presence of any remaining correlation.

However, a correction of the estimated standard errors of the coefficients is necessary because the (upstream) imputation process creates correlation between the welfare estimates. Following Elbers et al. (2005), the prediction error of imputed variables, e.g. expenditure, can be decomposed as

$$x \equiv y - \tilde{y} = [y - E(y)] + [E(y) - \tilde{y}] \tag{8}$$

where E(y) is the conditional expectation of expenditure. The first term on the RHS of (8) is termed the idiosyncratic error, which is due to unobserved factors that determine expenditure, and the second part is the model error, which reflects uncertainty about the upstream model's parameters. Applying this error decomposition to both g and y (7) can be written as

$$\begin{aligned} \tilde{g}_i = & [\tilde{y}_i b + \mathbf{x}_i \mathbf{?}] + [(E(y_i) - \tilde{y}_i) b - (E(g_i) - \tilde{g}_i)] \\ & + [(y_i - E(y_i)) b - (g_i - E(g_i)) + u_i] \end{aligned} \tag{9}$$

The RHS of the equation consists of three parts, each in square brackets. First we have a structural part consisting of imputed and non-imputed regressors and their

[4] Other variables could in principle be imputed or predicted as well; however, we consider pcx imputations.
[5] This holds *a fortiori* when either y or z is a non-linear transformation of pcx or its distribution, such as a poverty or inequality measure.

respective coefficients. The second part represents the model error, the third part the sum of upstream idiosyncratic error and downstream error.

We simplify notation by rewriting these three parts as $\tilde{g}_i = \mathbf{z}_i^{*}\mathbf{?} + j_i + e_i$ where $\mathbf{z}^{*}$ = ($\tilde{y}$,**x**) represents all regressors, both observed and imputed, and l = (β,g); φ represents the 'model part' of the error and *e* the idiosyncratic part. Assuming that the idiosyncratic part of the error is i.i.d., the variance matrix of the OLS coefficient estimates of (9) is

$$V(\mathbf{?}) = s_e^2(\mathbf{Z'Z})^{-1} + (\mathbf{Z'Z})^{-1}\mathbf{Z'}V(j)\mathbf{Z}(\mathbf{Z'Z})^{-1} \tag{10}$$

where the model part variance is

$$V(j) = b^2 V(E(y) - \tilde{y}) + V(E(g) - \tilde{g}) - 2b\,Cov[(E(y) - \tilde{y}),(E(g) - \tilde{g})] \tag{11}$$

Equation (10) shows that, compared to OLS variance estimates, variance has to be adjusted upwards. As (11) shows, this adjustment depends on the variance in the model error. The more imputed variables are used the more terms will have to be added: with *n* imputed variables, the number of terms on the RHS of (11) equals *n* variance terms plus *n*(*n*-1)/2 covariance terms. For example, if one imputed variable is used on the RHS only, the adjustment is limited to the first term. In our regression model (equation (5)), two imputed variables are used on the RHS, one on the LHS.

Two additional econometric problems affect our growth model. First, Caselli et al. (1996) show that estimating a cross-section growth model using a fixed effects estimator will lead to substantial bias when the number of periods is small, especially on the coefficient for initial expenditure (y92). The empirical growth literature suggests a number of solutions to this problem, most notably the Arrelano-Bond estimator. Such estimators, however, need at least three periods to estimate the model, using the first period to instrument for the initial conditions of the second period which explain growth between period two and three. Since we have only two periods, we cannot follow this approach. However, although the bias on the 'convergence coefficient' may be significant, Monte Carlo experiments indicate that the bias on the other RHS coefficients tends to be small (Forbes, 2000).

The second problem is endogeneity. Even though our model does not contain 'flow' variables but only beginning-of-period 'stocks', initial expenditure y(92) has been used to construct the growth variable and is thus correlated with the error term. Initial inequality may also be an endogenous variable, as the literature suggests that growth affects inequality (e.g. Aghion et al., 1999; Lundberg and Squire, 2003). One would expect this to be more problematic for changes in inequality (which we do not use) rather than for initial inequality. Put to scrutiny, a Hausman test rejects exogeneity of expenditure inequality, but cannot reject exogeneity of education. Consequently we deal with the endogeneity of initial expenditure and expenditure inequality.

Since we do not have lagged values, e.g. *y*(t-1), to use as instruments, we have to find instruments amongst the (few) available sub-county census means. We have chosen as instruments for income a variable that measures the 'education deficit' (the number of school years missed) of children below the age of 13. The (initial) education deficit for children in this age group is strongly negatively correlated with initial income, but, arguably, does not affect growth in the period analysed. The instruments for expenditure inequality are the maximum education deficit for children below 13 and ‘ethnic fractionalization’, which is the probability that any two citizens randomly chosen from a sub-county population are from different ethnic groups. The latter variable has itself been used to explain growth (Easterly and Levine, 1997) and is thus possibly a less suited instrument. We tested its validity by including ethnic fractionalization in the model. It does not alter the other coefficient estimates in any significant way; moreover, an overidentifying restrictions test rejects endogeneity. Finally, we note that the instrumentation also affects the calculation of the model’s variance: imputed endogenous variables have to be instrumented first and then instrumented values are used in the calculation of the variance-covariance matrix $V(j)$.

5. Results

The estimated standard errors in all our regressions are adjusted to account for prediction errors following the approach outlined in the previous section. The adjustments – illustrated for the baseline equation in Table 3 – result in an increase in estimated standard errors for all coefficients. The last column of Table 3 gives the ratio of the adjusted standard error estimate to the standard 2SLS estimate. The increase varies over coefficients between a factor 1.2 and 1.7. This is the typical trade-off when analysing small area welfare estimates: the gain in the number of ‘observations’ obtained by using imputed variables is partly offset by the loss in precision due to (downstream) model prediction errors.

The results in Table 3, which includes both inequality variables and their squares, illustrates the general decrease in significance when taking into account the fact that estimates or predictions, not data, are used. In the case of education inequality the adjustment even ‘destroys’ a significant result, that is, causes the significance level to increase to over ten percent.

Table 3 – Variance adjustments

Dependent: growth	*coef*	*t-val(adj)*	*t(2SLS)*	*incr(se)*
y	-0.1723	-6.6374	-8.7981	1.3255
gini	-1.1909	-0.7139	-1.0176	1.4254
gini2_	1.8082	0.7405	1.0748	1.4515
gi_hyredu	0.2377	1.3318	2.2749	1.7081
gi_hyredu2	-0.1151	-0.9057	-1.5812	1.7458
n_hyredu	0.0086	1.7077	2.6470	1.5500
n_spyredu	0.0175	4.1999	4.9129	1.1698
hage	-0.0008	-1.0305	-1.6072	1.5597
hfem	0.0487	2.7689	3.7914	1.3693
aesize	-0.0065	-3.1557	-4.5537	1.4430
pchild	-0.0463	-1.9806	-2.5756	1.3004
_cons	1.7947	3.9631	5.5353	1.3967

Our main findings are presented in Table 4 in a series of six regressions. Conditional convergence is pronounced in all specifications: the coefficient on initial income is negative, highly significant and has a value of around –0.17 in all specifications. Apparently, sub-counties with lower mean per capita expenditure in 1992 have grown faster over the 1990s, ceteris paribus. However, note that the coefficient estimate is biased so we should not attach significance to its exact value. The main variable of interest, inequality, has been entered using expenditure inequality and education inequality; these variables have been entered in linear and quadratic form in alternative specifications. The results show that expenditure inequality (gini) does not have a significant effect on growth in any of the specifications. In contrast, education inequality (gi_hyredu) has a significant positive effect on growth in all specifications we tried. This effect is robust to the inclusion of expenditure inequality (columns 4 and 5). When only education inequality is entered – without expenditure inequality, columns 5 and 6 – the effect is significant at the one percent level (note again that this

Table 4 – Regression results

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
	gr	gr	Gr	gr	gr	gr
y	-0.172	-0.161	-0.171	-0.171	-0.186	-0.175
	(6.64)***	(5.78)***	(5.23)***	(6.02)***	(7.46)***	(7.64)***
gini	-1.191	0.126	-2.048	0.096		
	(0.71)	(1.13)	(1.17)	(0.84)		
gini2	1.808		3.094			
	(0.74)		(1.22)			
gi_hyredu	0.238			0.089	0.099	0.349
	(1.33)			(2.46)**	(2.90)***	(3.24)***
gi_hyredu2	-0.115					-0.192
	(0.91)					(2.53)**
n_hyredu	0.009	0.001	0.003	0.008	0.011	0.010
	(1.71)*	(0.21)	(0.61)	(1.51)	(2.54)**	(2.38)**
n_spyredu	0.018	0.016	0.015	0.017	0.015	0.018
	(4.20)***	(3.91)***	(3.31)***	(4.01)***	(3.94)***	(4.56)***
hage	-0.001	-0.000	-0.001	-0.000	-0.001	-0.001
	(1.03)	(0.52)	(0.88)	(0.64)	(1.71)*	(1.87)*
hfem	0.049	0.057	0.055	0.046	0.047	0.053
	(2.77)***	(3.45)***	(2.73)***	(2.90)***	(2.97)***	(3.37)***
aesize	-0.006	-0.005	-0.006	-0.006	-0.007	-0.007
	(3.16)***	(2.60)***	(2.42)**	(3.03)***	(3.35)***	(3.49)***
pchild	-0.046	-0.042	-0.035	-0.044	-0.056	-0.063
	(1.98)**	(1.91)*	(1.48)	(2.02)**	(2.39)**	(2.70)***
Constant	1.795	1.550	2.034	1.593	1.785	1.604
	(3.96)***	(5.08)***	(3.96)***	(5.20)***	(7.13)***	(7.03)***
Observations	719	719	719	719	719	719
R-squared	0.92	0.90	0.91	0.91	0.91	0.92
p(Hausman Chi-sq)	0.07	0.00	0.01	0.00	0.15	0.44
p(Sargan Chi-sq)	0.10	0.06	0.16	0.05	0.04	0.20

Notes: Coefficient estimates on district dummies omitted. Absolute value of t statistics in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%

is after variance correction related to the imputation of the growth and expenditure variables). The last specification (column 6) includes education inequality squared: the coefficient has a negative sign and is significant at the five percent level. We conclude that education inequality should be entered in both linear and quadratic form. Although growth thus appears to be an inverted U-shaped function of education inequality, ceteris paribus, we should not overstate the non-monotonicity: the maximum of the parabola is found where gi_hyredu equals 0.91. In our data, more than 95 percent of sub-counties has lower education inequality. In other words, only for extremely high values of education inequality is the effect on growth negative.

All other variables in our growth regressions except age of the household head (hage) are significant at the five percent level or better in the last specification (column 6). Also, most of them are reasonably insensitive to variations in specification. Only the coefficient on household head's education level becomes large and significant when inequality (squared) with respect to this variable is added. Most other coefficients have expected signs. The effect of human capital levels – measured by the years of education completed by the household head and spouse – is positive: investments in human capital levels of household members pay-off in higher growth. Interestingly, the effect of an additional year of education for spouses is nearly twice as large as for household heads. Although on average 28 percent of Ugandan households has a female head in 1992, the effect of spouse human capital is striking, both in relative size and significance. Moreover, sub-counties with larger shares of households that are headed by females (hfem) grow faster. The latter effect is significant at the one percent level in all specifications.[6]

The effect of the age of household head is small, negative and marginally significant (ten percent level). Adult equivalent household size and the percentage of own children in the household both have a negative effect on growth. In other words, given age structure, sub-counties with larger households experience lower per capita growth. Moreover, sub-counties with households with a larger number of own children relative to total household size grow more slowly than others, ceteris paribus.

[6] We note that some of our findings contrast with evidence presented in Deininger and Okidi (2003): these authors find a non-linear, U-shaped growth effect of education levels and a negative effect of the female household head dummy. One possible explanation for this difference is that we consider rural households only.

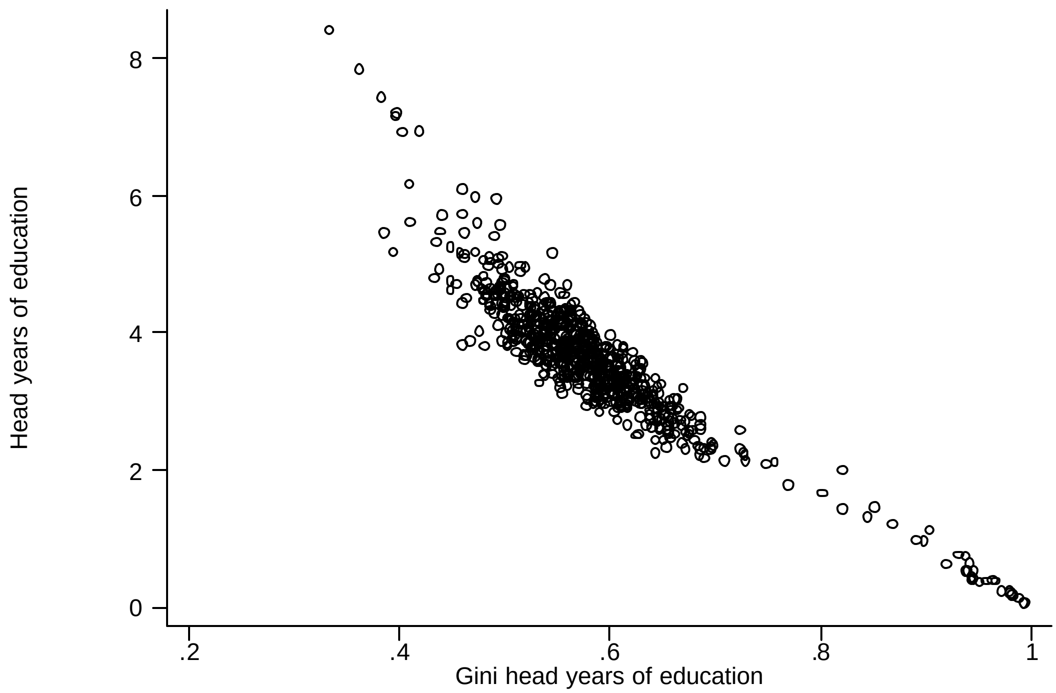
Discussion

The two most important findings of this study are (1) income inequality does not have a statistically significant effect on growth and (2) education (human capital) inequality has a positive effect on growth. The first of these findings is in line with cross-country evidence in Birdsall and Londono (1997) and Castello and Domenech (2002), while the second finding contrasts with findings in these papers.

This second finding may appear somewhat counter intuitive at first sight: growth is enhanced when human capital (or access to it) of the household head is more unequally distributed. The key to understanding what is going on is the fact that we control for mean level of education: this means that our conclusion is that *at a given mean level* of human capital, a more unequal distribution of this capital is good for growth. As noted before, a number of authors have addressed this point theoretically. In particular, Elbers and Gunning (2004) show that our result is to be expected in a Ramsey growth model: under the condition that the production function is convex in human capital, a mean-preserving spread in human capital results in higher long-run output growth. For instance suppose we were to redistribute one year of education from someone with low educational attainment to someone who is reasonably well educated. This would make the distribution of human capital more unequal while keeping the mean constant. But if the increase in output by the well-educated person exceeds the decline for the less well educated person, then the increased spread in education has a positive effect on growth – as long as the mean level of education is kept constant.

Mean preserving spreads in human capital are not possible within a given population; they only exist in thought experiments or in the long run, that is, over generations. In reality, mean level of education and inequality change simultaneously. In Uganda, for instance, education inequality – as measured by the Gini coefficient – is highly (positively) correlated with the percentage of households whose head has never attended school. The average years of head education is also highly (negatively) correlated with the percentage of heads who have never attended school. Consequently there is a strong negative correlation between education inequality and the average level of education (see Figure 1): both are largely determined by the number of household heads who never attended school. The implication of such a correlation is that while raising the general level of education through policies like universal primary education will be good for growth its positive effects will be partly offset by the associated decline in the education inequality. This effect is substantial. A 10% increase in education is associated with a 3% decline in the Gini which, in turn, offsets about ? of the growth impact of an increase in the level of education. Put differently: if the additional education years had been distributed unequally, e.g., in such a way that the Gini would have remained constant, the growth effect would have been larger.

Figure 1



6. Conclusion

We estimated the effect of income and education inequality on growth, using imputed data on expenditure inequality and growth for small administrative units in Uganda (sub-counties), along with census data for education inequality. Analysing this relation for a specific country has important benefits: first, it avoids data comparability problems that typically affect cross-country growth regressions. Moreover, by identifying the effects of inequality on growth for a given country, country specificity is taken into account. This enhances the relevance of our results for local policy makers.

In the empirical section we adjusted the standard errors of variable coefficients for the fact that some regressors are imputed–in our case initial expenditure levels and expenditure inequality, and therefore associated with a standard error. The adjustments are considerable; they typically increase standard errors by a factor 1.2 to 1.7.

Our results show that higher levels of education enhance growth. Controlling for the level of educational attainment, larger variation in education is good for growth. Our results also indicate that income inequality does not affect growth.

7. References

Aghion, P., Caroli, E. and Garcia-Penalosa, C. (1999), Inequality and economic growth: The perspective of the new growth theories, *Journal of Economic Literature,* **37** (4)**,** 1615-60.

Atkinson, A. B. and Brandolini, A. (2001), Promise and pitfalls in the use of "secondary" data-sets: Income inequality in oecd countries as a case study, *Journal of Economic Literature,* **39** (3)**,** 771-799.

Balisacan, A. M. and Fuwa, N. (2003), Growth, inequality and politics revisited: A developing-country case, *Economics Letters,* **79** 53-58.

Banerjee, A. V. and Duflo, E. (2003) "Inequality and growth: What can the data say?", mimeo.

Benabou, R. (1996), Heterogeneity, stratification, and growth: Macroeconomic implications of community structure and school finance, *American Economic Review,* **86** (3)**,** 584-609.

Birdsall, N. and Londono, J. L. (1997), Asset inequality matters: An assessment of the world bank's approach to poverty reduction, *American Economic Review, Papers and Proceedings,* **87** (2)**,** 32-37.

Bourguignon, F. (2004) "The poverty-growth-inequality triangle", mimeo, Paper presented at the Indian Council for Research on International Economic Relations, New Delhi.

Brock, W. A. and Durlauf, S. N. (2001), Growth empirics and reality, *World Bank Economic Review,* **15** (2)**,** 229-272.

Caselli, F., Esquivel, G. and Lefort, F. (1996), Reopening the convergence debate: A new look at cross-country growth empirics, *Journal of Economic Growth,* **1** (3)**,** 363-390.

Castello, A. and Domenech, R. (2002), Human capital inequality and economic growth: Some new evidence, *Economic Journal,* **112** (478)**,** C187-C200.

Deininger, K. and Okidi, J. (2003), Growth and poverty reduction in uganda, 1999-2000: Panel data evidence, *Development Policy Review,* **21** (4)**,** 481-509.

Deininger, K. and Squire, L. (1996), A new data set measuring income inequality, *World Bank Economic Review,* **10** (3)**,** 565-91.

Deininger, K. and Squire, L. (1998), New ways of looking at old issues: Inequality and growth, *Journal of Development Economics,* **57** (2)**,** 259-287.

Easterly, W. and Levine, R. (1997), Africa's growth tragedy: Policies and ethnic divisions, *Quarterly Journal of Economics,* **112** (4)**,** 1203-1250.

Elbers, C. and Gunning, J. W. (2004) "Transitional growth and income inequality: Anything goes", mimeo, Vrije Universiteit, Amsterdam.

Elbers, C., Lanjouw, J. O. and Lanjouw, P. (2003), Micro-level estimation of poverty and inequality, *Econometrica,* **71** (1)**,** 355-364.

Elbers, C., Lanjouw, J. O. and Lanjouw, P. (2005), Imputed welfare estimates in regression analysis, *Journal of Economic Geography,* **forthcoming**.

Emwanu, T., Okwi, P., Kristjensen, P., Hoogeveen, H. and Kuppa, R. (2004, forthcoming), Regional poverty estimates in Uganda, 1992 and 1999, UBOS/ILRI/WB/DFID.

Forbes, K. J. (2000), A reassessment of the relationship between inequality and growth, *American Economic Review,* **90** (4)**,** 869-887.

Galor, O. and Tsiddon, D. (1997), The distribution of human capital and economic growth, *Journal of Economic Growth,* **2** (March)**,** 93-124.

Hentschel, J., Lanjouw, J. O., Lanjouw, P. and Poggi, J. (2000), Combining census and survey data to trace the spatial dimensions of poverty: A case study of ecuador, *World Bank Economic Review,* **14** (1)**,** 147-65.

Hoogeveen, H., Emwanu, T. and Okiira Okwi, P. (2003) "Updating small area welfare indicators in the absence of a new census", mimeo, Washington / Kampala.

Lundberg, M. and Squire, L. (2003), The simultaneous evolution of growth and inequality, *Economic Journal*, **113** ((April))**,** 326-344.

Mbabazi, J., Morrisey, O. and Milner, C. (2002) "Inequality, trade policy and growth: An explanation for the africa dummy", mimeo, Nottingham.

Ravallion, M. (1997), Can high-inequality developing countries escape absolute poverty?, *Economics Letters,* **56** 51-57.

Ravallion, M. (1998), Does aggregation hide the harmful effects of inequality on growth?, *Economics Letters,* **61** 73-77.

Thorbecke, E. and Charumilind, C. (2002), Economic inequality and its socioeconomic impact, *World Development,* **30** (9)**,** 1477-1495.