AGRICULTURAL PRODUCTIVITY AND POVERTY IN DEVELOPING COUNTRIES[1]

Extensions to DFID Report No.7946

Lin Lin[*], Victoria McKenzie[*], Jenifer Piesse[**] and Colin Thirtle[***]

[*] Environmental Policy and Management Group
T.H.Huxley School of Environment, Earth Sciences and Engineering
Imperial College of Science, Technology and Medicine, London, SW 7 2BP

[**] School of Management and Organizational Psychology
Birkbeck College, University of London, London, WC 1E 7HX
and University of Stellenbosch, Stellenbosch, 7602, Republic of South Africa.

[***] Environmental Policy and Management Group
T.H.Huxley School of Environment, Earth Sciences and Engineering
Imperial College of Science, Technology and Medicine, London, SW 7 2BP
and University of Pretoria, Pretoria, 0002, Republic of South Africa

OUTLINE

[1] We thank Lawrence Haddad of IFPRI, Jonathan Kydd and Peter Dorward of Imperial College at Wye, for help and advice and especially Gavin McGillivray of DFID for the original suggestion to investigate the relationship between productivity and poverty in this manner. Michael Lipton suggested that we should also look at the effect of industrial productivity growth.

EXECUTIVE SUMMARY

This paper investigates the empirical relationships between agricultural productivity growth, poverty reduction, nutritional improvement, inequality and GDP per capita growth in some detail. The empirical estimates of the relationship between labour productivity in agriculture and poverty reduction appears to be particularly robust. For all the specifications of the model and for all the different samples, labour productivity in agriculture is found to be a powerful and always significant cause of poverty reduction.

The study begins with simple models, which concentrate on the explaining poverty measures with just the key variables, which are the two components of agricultural labour productivity (value added per unit of land, or yield and the land labour ratio) and the Gini coefficient, which measures inequality. For the latest cross section of 40 counties, from the 2001 World Development report, the two agricultural productivity variables alone explain 63% of the variance in the percentage of the population living on less than $1 per day. The poverty elasticity of the yield is -0.91and is significant at the 99% level of confidence. For the 19 African countries alone, 47% of the variance is explained, the significance level is the same despite the small sample and the elasticity is slightly higher at –0.96. This means that a 1% increase in yields decreases the percentage of the population living on less than $1 per day by 0.96%.

For a larger pooled sample, in which there are 109 observations on 58 countries, where there are two or even three observations for some countries, at different points in time, the yield elasticity falls to 0.621, but the high level of significance is maintained and 50% of the variance is explained. The double counting of some countries could be the cause of the difference, but it is more likely that including the Gini coefficient reduces the elasticity. The Gini measure of inequality has a poverty elasticity of 2.2 and is highly significant. Thus, a 1% increase in inequality increases the percentage of the population living on less than $1 per day by 2.2%, which is clear evidence of the adverse effects of increases in inequality. For the African countries alone, the sample is increased to 38 and although the Gini coefficient has a similar impact, the poverty elasticity of the yield rises to –1.02%, with 55% of the variance explained.

The $1 per day poverty measures begin only in 1985, so to cover the green revolution period, this is replaced by the Human Development Index. The agricultural productivity variables remain significant at high levels of confidence, for both the full sample of 280 observations on 174 countries and for the smaller African sample of 78 observations. The yield elasticity is 0.12 for the full sample, with 76% of the variance explained and 0.094 for the Africa sample, with an explanatory power of 48%. Thus, a 1% increase in yields increases the value of the index by 0.12% for the full ample and by 0.094% for the African countries. Since the units of measurement are totally different, these results are not comparable to those for the $1 per day poverty measure and they are not so easy to interpret, but they do confirm the power and robustness of the agricultural productivity variables.

The final applications of this simple model are to two nutrition variables, which are the per capita dietary energy supply and the percentage of children under five years old that are underweight. For the full sample of 109 observations, the yield elasticity with respect to energy supply is 0.05, which indicates that a 1% increase in yields increases the per capita dietary energy supply by only 0.05%. Again, the result is not comparable to the $1 per day poverty estimates, but also the explanatory power is only 25%, so these two results suggest that nutritional improvement depends on many other variables and is more difficult to achieve. For the Africa sample the yield elasticity is 0.09, which is almost twice as great, but again only 25% of the variance is explained.

The yield elasticity with respect to the percentage of children that are underweight is comparable (as it is again a reduction measured in percent) and it is –0.42, which means that a 1% increase in yields reduces child malnutrition by 0.42%. This is more encouraging, but again only 31%

of the variance is explained. For the African sample the yield elasticity falls to –0.27 and only 28% of the variance is explained. Thus, the yield elasticity with respect to calories is far higher for Africa, but it seems to be more difficult to improve the diet of children.

These results confirm the predominant view in the literature surveyed in Thirtle et al. (2001), that agricultural productivity growth can be expected to have an impact on poverty. However, it is fair to ask how great this is relative to other improvements in economic performance. Thus, the impact of labour productivity growth in agriculture is first compared with the effects of productivity growth in industry and services. These results need to be treated with some caution, because there was even more missing data for industry and services than for agriculture, but regardless of the samples used, agricultural labour productivity had a poverty elasticity of -0.63, which was significant at least the 95% confidence level, whereas labour productivity in industry and services has no significant impact on poverty reduction.

Attempts to explain inequality were less successful. The yield and land labour ratios had significant, but small (-0.05 and 0.06 respectively), elasticities and the only other variables which consistently explained the Gini coefficients were the percentage of the population that were rural and the rate of population growth. The elasticity for the percentage of the population that was rural was –0.20, which indicates more rural societies are less unequal, but this may be because they are poorer and thus have less differentiation. The elasticity of population growth was 0.19, meaning that 1% greater population growth increases inequality by 0.2%. This confirms the notion that rapid population growth tends to increase inequality.

The danger with these simple models is that they may be mis-specified, in the sense of omitting important explanatory variables, so that the poverty elasticity of the yield variable is biased upwards, or possibly even downwards. To overcome this problem, data on all the available variables was collected and then tested by looking for correlations with the poverty and inequality measures. All the combinations of significant explanatory variables were then used in addition to the agricultural productivity variables and tests were used to determine which variables improved the models. Gross fixed investment (GFI), which includes infrastructure investment, was found to be the most consistently significant variable. This is again in keeping with the findings in the literature: see Fan, Hazell and Thorat, 1999, for example. The elasticity of GFI in the full sample was –0.93 and the poverty elasticity of yields was reduced to –0.56, so it does seem that infrastructure is also crucial to poverty reduction. For the Africa sample the yield elasticity was far higher at –0.95 and the GFI elasticity was lower at –0.48, which suggests that infrastructure is so poor in the African sample that it has had less effect. Open economies also appear to have less poverty, as the percentage of GDP that was traded had an elasticity of –0.26 and GDP growth has positive effects, with an elasticity of –0.16. The point on the effect of missing variables biasing the yield elasticity is bourn out, but the problem is not severe. The minimum value of the yield elasticity in these models with five explanatory variables was still –0.54 and the significance levels were maintained.

The difficulties of modelling these relationships become apparent when GDP per capita is included in the equation as well as agricultural productivity. Then, the yield elasticity falls to –0.25 and the elasticity of GDP per capita is –0.47. However, this is not an omitted variable problem. Rather, agricultural productivity is a key variable in explaining differences in GDP per capita. Indeed, the yield and the land labour ratio, which are the two components of agricultural labour productivity, alone explain 84% of the variance in GDP per capita. Similarly, these same two variables explain 22% of the variance in equality, as measured by the Gini coefficient. Thus, there is causality running from agricultural labour productivity to both GDP per capita and the Gini coefficient, so these three variables should not appear in the same equation.

There are two solutions to this problem. If the single equation approach is to be maintained, GDP per capita can be regressed on the agricultural productivity variables and the residual from this equation, which may be termed "agricultural productivity-free GDP per capita can be used instead of the original variable. Similarly, the Gini coefficient is regressed on the agricultural productivity variables and the residual used instead of the original Gini. If this is done, the yield elasticity remains close to its original value, at –0.65, the Gini elasticity is unchanged at 2.21 and the GDP per capita elasticity rises to –0.59, which is within the range suggested by Hanmer and Naschold (2000). The other significant variables are the percentage of GDP traded, with an elasticity of –0.35 and GFI, which has an elasticity of –0.67. These variables explain 60% of the variance in the percentage of the population living on less than $1 per day.

The preferred alternative is to resort to systems of equations. In the first equation, the percentage of the population living on less than $1 per day is explained by the Gini coefficient (elasticity of 1.74), the percentage of GDP traded (elasticity of –0.37), gross fixed investment (elasticity of –0.75) and GDP per capita (elasticity of –0.76). These variables explain 65% of the variance in poverty. In the second equation, the Gini coefficient is explained by the percentage of the population that is rural (elasticity of –0.20) and the rate of population growth (elasticity of 0.33). These variables explain 42% of the variance in inequality. In the third equation, GDP per capita is explained by the yield (elasticity of 0.58), the land labour ratio (elasticity of 0.57) and expenditure per student in primary education (elasticity of –0.14). These variables explain 77% of the variance in GDP per capita. Thus, it is possible to impose some structure on the complex inter-relationships between the nine variables in the system and produce credible results.

This is one of two preferred recursive models, but all the results come to the same basic conclusion, which is that agricultural productivity growth has almost as much effect on poverty reduction as does GDP per capita growth itself. But, the real point is that the gains from agricultural productivity growth are not confined to the poor, since the yield also has an elasticity of 0.58 with respect to growth in GDP per capita. This means that 1% growth in yields both reduces the percentage of the population lining on less than $1 per day by roughly 0.5% and increases GDP per capita by almost 0.6%.

The final step is to calculate the cost of achieving a 1% increase in yields. This can be done within the confines of the recursive system, by adding a fourth equation, in which yield is explained by agricultural R&D and education. For this R&D data must be collected, so for now rough calculation will have to suffice. In 1991, the less developed countries spend $8 billion on agricultural R&D (Alston et al., 2000) and yields have been increasing at 2.4% per annum since the 1960s, so the cost of a 1% yield increase is £3.33. The GDP of the LDCs in 1990 was about $2,850 billion, so with a GDP elasticity of 0.56, the payoff to this R&D investment was $16.13 billion. If the lag from expenditures to gains in output is 5 years, the rate of return is 37%: if it is four years, it rises to 50%, which is surely an attractive investment. This gives a rough estimate of the rate of return to agricultural research and the poverty reduction impact is a bonus on top of this figure. The number of persons living on less than $1 per day was 1,200 million in 1998 and the best estimate of the poverty elasticity of agricultural research is 0.65. Thus a 1% increase in yields reduces this count by 0.65%, which is almost 7.8 million. Viewed in this way, agricultural research may well be a useful and cost effective instrument for reducing poverty, but it is a pretty blunt one as many have argued (Alston et al., 1995). It needs to be sharpened by aiming it at the poorest segment of the population.

1 INTRODUCTION

The previous report on this subject (Thirtle et al, 2001) reviewed the literature on the role of agricultural productivity in alleviating poverty in developing countries and presented some limited empirical results suggesting that there are significant relationships between productivity growth and both poverty and nutrition. The major finding was that the empirical estimates of this relationship appear to be robust. Regardless of differences in data and formulation, the results showed that a 1% increase in yields leads to a reduction in the percentage of people living on less than $1 per day of between 0.6% and 1.2%.

This study extends the previous work in several respects.[2] First, the sample is broken up to show that the relationships hold for Sub-Saharan Africa, which is perhaps the region of greatest interest. Second, the effects of inequality on the productivity-poverty relationship are further investigated and the causes of inequality are also considered.[3] Third, although the relationship between agricultural productivity and poverty appeared to be robust, there are many other omitted variables that may affect the relationship, so these are added to the model. Fourth, adding variables to a single equation is of limited use, since many of the relationships are collinear and the variables tend to cancel each other out. Also there is genuine simultaneity between variables, so it is necessary to resort to multiple equation models to improve explanatory power. Fifth, although agricultural productivity has a significant effect on poverty reduction, it is fair to ask how the impact compares with other options. So, in this study the impact of productivity growth in industry and services is compared with the results for agriculture. Last, the link between agricultural R&D and productivity growth is investigated, with a view to estimating the costs of achieving poverty reduction by means of increasing agricultural productivity. Also, the results show that agricultural productivity growth increases GDP per capita as well as reducing poverty: it is not only the poor that gain.

The paper proceeds as follows. The next section explains the basic methodology and outlines the data and its sources, before presenting an exploratory correlation analysis to determine the variables, which appear to be important in explaining poverty and inequality. Section three describes the basic regression models and the results they generate for the full sample and for the African countries alone. Section four contrasts the impacts of agricultural productivity growth with the effects of labour productivity growth in industry and services. Section five considers the ways in which the specification of the model can be improved, by removing the interactions between the variables in the single equation model and resorting to recursive models. Last, the R&D costs of reducing poverty and increasing GDP per capita by means of yield increases are estimated and the rate of return to agricultural R&D is estimated, before the conclusion summarises the results.

2 METHODOLOGY AND DATA

2.1 Basic Regression Models

The basic approach is ordinary least squares regressions, with the poverty index as the dependent variable. The explanatory variables are alternative measures of agricultural productivity and other variables that may be expected to have an impact on poverty. The intention is to estimate poverty reduction elasticities with respect to agricultural productivity, similar to the poverty elasticities with respect to growth (Chen and Ravallion, 2000 Hanmer and Naschold, 2000, Ravallion, 1997, Ravallion and Datt, 1999.). Since there is no well-defined economic theory of poverty alleviation, the regression are somewhat ad hoc and the possibility of misspecification due to variable omission is obviously a danger.

---

[2] Many of the results from the empirical section of the previous paper are reported here, as they are needed to allow comparisons with the new estimates. Thus, this paper can stand alone, but the previous paper literature review is needed for the literature review, which led to this empirical investigation.

[3] Maxwell (2001) discusses the reinstitution of inequality on the poverty agenda, for instance in the World Bank World Development and refers to work by Howard White suggesting international targets for inequality reduction.

We exploit the link between labour and land productivity used by Hayami and Ruttan (1986), which is stated here in value added terms as an identity:

$$\frac{VALUE\ ADDED}{PER\,UNIT\ LABOUR} \equiv \frac{VALUE\ ADDED}{LAND} \times \frac{LAND}{PER\,UNIT\ LABOUR} \quad \textbf{(1)}$$

Value added is net of the costs of intermediate inputs, which will remove the cost effects of intensification using increasing amounts of modern inputs. This should make this measure closer to total factor productivity (TFP) than the total output-based measure. In this way labour productivity can be decomposed into the product of two components: land productivity, or yield, and the land labour ratio, which can be viewed as an indicator of a country's resource endowment. Thus, the yield contribution to labour productivity can be separated from the relative scarcity of land, which it is not possible to change. This is important because a country such as the USA has several hundred times as much land per unit of labour as a land scarce country, such as Bangladesh, and will have far higher labour productivity as a result.

Thus, the poverty indicator is explained by the productivity indices for land and for labour, by means of cross sectional regressions, in models 1 and 2. Then, the decomposition is used, in equation 3, making the land labour ratio and land productivity the independent variables. In these three basic models all the variables are expressed in logarithms, so that the coefficients can be interpreted as elasticities, the X_i are the other relevant variables (which vary, but usually include the Gini index) and ε is the error term.

$$Model\,1:\ \ln Poverty\ Index = \alpha + \beta \ln\left[\frac{Value\,added}{Land}\right] + \phi_i \ln X_i + \varepsilon \quad \textbf{(2)}$$

$$Model\,2:\ \ln Poverty\ Index = \alpha + \beta \ln\left[\frac{Value\,added}{Labour}\right] + \phi_i \ln X_i + \varepsilon \quad \textbf{(3)}$$

$$Model\,3:\ \ln Poverty\ Index = \alpha + \beta \ln\left[\frac{Value\,added}{Land}\right] + \delta \ln\left[\frac{Land}{Labour}\right] + \phi_i \ln X + \varepsilon \quad \textbf{(4)}$$

These three models form the starting point of the analysis. The more complex models developed later, such as the recursive estimation of the poverty equation and an inequality equation, are described in section five, when they are applied.

2.2 Data

Dependent Variables: Poverty and Nutrition Indicators

The first poverty indicator used was the percentage of the population living on less than $1 per day, taken from the World Development Report 2000/2001. This gives a simple cross section, with a minimum of 40 countries, which are listed later.

To increase the size of the data set it is necessary to pool the results of poverty surveys, so that for some countries there are two or even three observations at different points in time. This may bias the results, but it does increase the sample size to 109, which is far more satisfactory. The observations are for 58 developing countries and range in time from 1985 to 1995. They are from the World Bank and Chen, Datt and Ravallion, (1994) and were used by Hanmer and Naschold (2000), who we thank for making these data available.

The earliest observations for the $1 per day poverty index are for 1985, which does not allow coverage back to the green revolution period. The available index that does go back further, at least to the final stages of the green revolution, is the Human Development Index (HDI), from the UNDP. The HDI is a composite index of development. The three most crucial components of the HDI are measures of longevity, education and income and it may serve as a reasonable poverty index. Educational attainment, income, and life expectancy are all associated with poverty. Thus, the HDI is used as a second poverty indicator, although is obviously not as satisfactory. The HDI covers 174 countries for 1975, 1980, 1985, 1990, 1998, giving a total number of observations without missing data of 280.

The nutrition indicators are from Lawrence Haddad of IFPRI. There are 181 mixed cross section and time series observations, but again this reduces to 109 observations. The dependent variables are per capita dietary energy supply and the percentage of children under five who are underweight.

In later models the Gini coefficient is added as a dependent variable, to give a two equation recursive model. These data are also from Hanmer and Naschold (2000).

Explanatory variables

The data for all the independent variables were obtained from the World Development Indicators (2000) CD. The variables are listed in Table 1, which also reports the correlation of the variables with the $1 per day poverty measure and the Gini coefficients. The correlation coefficients provide some guidance as to which variables are likely to have explanatory power in the regression models. These significant variables are picked out in red. For the percentage of the population living on less than $1 per day poverty index, a negative sign indicates that the variable is likely to be poverty reducing and a positive sign the opposite. The range of the Gini is from zero, which indicates perfect equality, to unity, which would be complete inequality, so again a negative sign suggests that the variable may decrease inequality. Thus, the Gini coefficient has a detrimental effect on poverty reduction, because the positive coefficient means that greater inequality correlates with a higher percentage of the population living on less than $1 per day, which is the result suggested by the literature (see Hanmer and Naschold, 2000, for example). Conversely, the three agricultural productivity variables from equations (1)-(3) are all significantly poverty reducing, whereas productivity growth in industry and services has no significant effect on poverty, but does decrease inequality, as does land productivity.

The other explanatory variables are those available that seemed most likely to affect poverty and inequality, some of which appear in the literature. The combination of the signs on $1 a day poverty and inequality correlations is sometimes interesting. Of the other agricultural variables, the rural population as a percentage of the total and agriculture's share in value added are both positively correlated with poverty, indicating that counties with large agricultural sectors tend to be poorer. However, the negative correlations with the Gini coefficient show that they also lessen inequality. The percentage of the land that is irrigated has no significant effects.

The literature review in Thirtle et al (2001) found that there was evidence that health and education expenditures may reduce poverty, so literacy and health expenditures were included and were found to be significantly poverty reducing, according to the correlation coefficients.[4] Literacy also correlates with less inequality. Educational expenditures per student in primary education correlates negatively with inequality, but has no significant impact on poverty. Educational expenditures per student in secondary education are not significant and neither of these expenditure variables is correlated with literacy. Similarly, infrastructure has been shown to be important in the literature (Fan, Hazell and Thorat, 1999, for example), so gross fixed investment is included to try to capture infrastructure investment.[5] This is also negatively correlated with poverty.

Table 1: Socio-Economic Dataset

4 Note that the variable is illiteracy, which is why the sign appears to be wrong.

5 Gross domestic fixed investment includes land improvements (fences, ditches, drains, and so on); plant, machinery, and equipment purchases; and the construction of roads, railways, and the like, including commercial and industrial buildings, offices, schools, hospitals, and private residential dwellings. Data are in constant 1995 U.S. dollars. For more information, see WDI table 4.10.The World Bank data used to include government expenditure on transport and communications, which would have been preferable, but this has been discontinued.

Variable Name	Correlation with $1 per day poverty	Correlation with GINI Coefficient	Sample Size
1 Dollar per Day (% Total Population)*	1.00**	0.35**	121
Gini Coefficient*	0.35**	1.00**	121
Agriculture value added per worker (constant 1995 US$)	-0.55**	0.12	113
Agriculture value added per hectare of agricultural land (constant 1995 US$)	-0.30**	-0.41**	109
Agriculture land per worker	-0.18**	0.44**	109
Services, value added per worker (constant 1995 US$)	-0.05	-0.23**	57
Industry, value added per worker (constant 1995 US$)	-0.02	-0.23**	54
Rural population (% of total population)	0.42**	-0.17**	121
Agriculture, value added (% of GDP)	0.43**	-0.16**	119
Land use, irrigated land (% of cropland)	-0.12	-0.04	115
Health expenditure per capita, PPP (current international $)	-0.44**	0.19	54
Illiteracy rate, adult total (% of people aged 15 and above)	0.50**	0.21**	115
Expenditure per student, Primary Education (% of GDP)	-0.10	-0.26**	61
Expenditure per student, Secondary Education (% of GDP)	-0.28	-0.09	28
Gross domestic fixed investment (% of GDP)	-0.33**	0.02	118
Trade (% of GDP)	-0.32**	0.09	119
Exports of goods and services (% of GDP)	-0.42**	0.06	119
General government consumption (% of GDP)	-0.07	0.13	119
State-owned enterprises, economic activity (% of GDP)	-0.18**	-0.03	52
Tax revenue (% of GDP)	-0.47**	0.05	87
GDP per capita, (current international PPP $)	-0.61**	-0.01	117
GDP growth (annual %)	-0.02	-0.013	90
Population growth (annual %)	0.29**	0.35**	114
Unemployment, total (% of total labour force)	-0.14	0.02	45
Bread and cereals price in PPP terms (U.S. price = 100)	--	--	0
Inflation, consumer prices (annual %)	0.04	0.06	108

* From Hanmer and Naschold (2000)   ** significant at the 5% level, two-tailed test.

The liberalisation literature argues that openness is important and both the percent of GDP traded and exports as a percentage of GDP are negatively correlated with poverty. The damaging effects of overly-large public sectors is also a part of the liberalisation creed, but government consumption is not significant and the percentage of GDP accounted for by state owner enterprises is negatively correlated with poverty. This may be a function of the sample, but perhaps the countries with larger government sectors do tend to be more egalitarian (there are no formerly communist countries in this sample, which could have been the cause of this result). Inflation is also slated as a bad, but has no discernable effect on poverty or inequality. However, tax revenue also correlates with less poverty, so at least some taxation would seem to be re-distributive.

The key variable in the literature on the poverty-reducing effect of growth (surveyed by Hanmer and Naschold, 2000) is GDP per capita, which also shows a strong negative correlation with poverty, but not with inequality and the GDP growth rate is not significant. Population growth is positively correlated with both poverty and inequality, indicating that countries with more rapid population growth do experience difficulties, but inflation has no discernable impact.

These results are summarised in Tables 2 and 3, which list variables in descending order of correlation with first the $1 per day poverty index and then the Gini coefficient. Thus, Table 2 first lists the poverty reducing variables, showing that GDP per capita has the highest correlation with poverty reduction, followed by agricultural labour productivity. The correlation coefficients are bivariate relationships, which give some indication of the possible impacts, but they take no account of the multiple causes of poverty reduction or the interactions between the variables and correlation implies nothing about the direction of causality. Thus, although the evidence from the literature suggests that GDP per capita growth reduces poverty, it is also true that having more poor people in the population reduces GDP per capita. Despite these caveats, it is worth noting that the effect of agricultural labour productivity has almost as much impact as growth in GDP per capita. This is consistent with the lack of impacts on poverty of productivity growth in industry and services. Also, the costs of increasing agricultural productivity growth are very low, relative to the cost of an equal increase in GDP per capita. It would be surprising that productivity growth in a single sector can have such a big impact, were it not for the extensive literature, which almost all points to the link between agricultural growth and poverty reduction. The actual impacts will be better established with multivariate regressions, in the next section.

The last five variables in Table 2 are those that appear to increase the percentage of the population living on less than $1 per day. Again, these results are well supported by the literature, which has always argued for basic education as a means of improving the position of the poor. Table 3 repeats this summary exercise for the Gini measure of inequality, showing that increasing yields has the greatest impact on inequality reduction. Although there was no direct poverty impact of labour productivity in industry and services, they must have an indirect effect, by way of reducing inequality. Last, the inequality-increasing effect of more land per worker is difficult to interpret. It could well result from the Asian countries with more population pressure on the land have less poor people than the African countries, which have far more land, of much worse quality.

To conclude the data analysis in this section, Tables 2 and 3 also show the summary statistics for all the variables in Table 1, for the full sample and for the African countries alone, so that the two can be compared. The mean values in the first part of Table 2 show that the Africa sample has substantially lower GDP per capita and far lower agricultural land and labour productivity. The higher land labour ratio is not quality adjusted: if the FAO data on land potential were taken into account this difference could even be reversed, as these data show that Kenya is at least as land-scarce as the heavily populated Asian countries, once land quality and rainfall differences are included. Africa has a greater proportion of GDP accounted for by state-owned enterprises and this variable was negatively correlated with poverty.

The lower part of Table 2 shows that literacy rates in Africa are far lower than for the full sample and that agriculture has a higher share of GDP in Africa, which is not surprising, given the much higher proportion of the population that are rural. The Gini coefficients show that there is not much difference in inequality levels, but if anything this Africa sample has greater inequality. Finally, population growth is greater for the Africa sub-sample.

The top part of Table 3 repeats some of the Table 2 results, but also shows that Africa actually spends more per student on primary education, despite, or perhaps because of the higher level of illiteracy. Labour productivity in industry for the Africa sample is only 53% of the level in the full sample and in services the difference is even more pronounced, at only 35%. The lower part of the Table adds no new information.

This completes the exploratory analysis of the data and suggests that the agricultural productivity variables are strongly correlated with poverty and inequality. We now use the information gleaned above in constructing regression models to explain poverty and inequality reduction.

Table 2: Variables Correlated with $1 Day Poverty

Poverty Reducing	Correlation	Mean		Standard Deviation		Minimum		Maximum	
		All	Africa	All	Africa	All	Africa	All	Africa
GDP per capita, (current international PPP $)	-0.61**	2799.8	1579.0	2096.4	1591.9	335.6	335.6	10948.0	8029.3
Agriculture value added per worker (constant 1995 US$)	-0.55**	1441.8	622.1	1363.8	754.8	164.4	164.4	5319.5	3388.8
Tax revenue (% of GDP)	-0.47**	17.6	19.0	8.6	8.0	3.6	8.5	44.7	38.6
Health expenditure per capita, PPP (current international $)	-0.44**	187.9	147.2	154.3	180.2	10.0	10.0	784.0	571.0
Exports of goods and services (% of GDP)	-0.42**	27.2	24.8	16.4	14.3	5.5	5.6	89.4	62.0
Gross domestic fixed investment (% of GDP)	-0.33**	20.6	20.1	8.6	12.4	5.8	5.8	71.4	71.4
Trade (% of GDP)	-0.32**	60.9	61.3	34.6	30.6	13.9	19.7	182.7	151.4
Agriculture value added per hectare of agricultural land (constant 1995 US$)	-0.30**	317.3	104.0	323.2	102.5	4.3	4.3	1444	432.9
Agriculture land per worker	-0.18**	10.7	16.6	17.6	27.6	0.29	0.6	102.5	102.5
State-owned enterprises, economic activity (% of GDP)	-0.18**	10.5	13.8	8.2	11.0	0.6	5.3	35	35
Poverty Increasing									
Illiteracy rate, adult total (% of people aged 15 and above)	0.50**	32.4	48.4	23.1	19.2	0.4	17.4	88.0	88.0
Agriculture, value added (% of GDP)	0.43**	22.2	29.6	13.9	15.7	4.2	4.2	60.8	60.8
Rural population (% of total population)	0.42**	57.3	90.6	20.7	13.7	15.7	42.1	95.0	95
Gini Coefficient*	0.35**	0.433	0.467	0.111	0.102	0.205	0.289	0.634	0.623
Population growth (annual %)	0.29**	2.4	2.9	0.9	0.5	0.1	1.8	6.6	3.7

** Significant at 5% level.

Table 3: Variables Correlated with GINI Coefficient

Inequality Reducing	Coefficient	Mean		Standard Deviation		Minimum		Maximum	
		All	Africa	All	Africa	All	Africa	All	Africa
Agriculture value added per hectare of agricultural land (constant 1995 US$)	-0.41**	317.3	104.0	323.2	102.5	4.3	4.3	1444	432.9
Expenditure on Primary Education (% of GDP)	-0.26**	13.4	17.4	8.5	9.4	2.5	6.3	47.4	47.4
Services, value added per worker (constant 1995 US$)	-0.23**	13641.1	4719.7	20413.4	5518.0	459.4	459.4	91755.2	19898.3
Industry, value added per worker (constant 1995 US$)	-0.23**	14368.3	7678.0	21984.4	9446.8	374.7	374.7	125846.1	37127.2
Rural population (% of total population)	-0.17**	57.3	90.6	20.7	13.7	15.7	42.1	95.0	95
Agriculture, value added (% of GDP)	-0.16**	22.2	29.6	13.9	15.7	4.2	4.2	60.8	60.8
Inequality Increasing									
Agriculture land per worker	0.44**	10.7	16.6	17.6	27.6	0.29	0.6	102.5	102.5
Population growth (annual %)	0.35**	2.4	2.9	0.9	0.5	0.1	1.8	6.6	3.7
Illiteracy rate, adult total (% of people aged 15 and above)	0.21**	32.4	48.4	23.1	19.2	0.4	17.4	88.0	88.0

*From Hanmer and Naschold (2000)

** Significant at 5% level.

3 REGRESSION ANALYSIS AND RESULTS

3.1 Regressions of the Cross Section of $1 per Day Poverty from WDR 2000

The analysis begins by fitting equations (1)-(3) with the available cross section of $1 per day poverty percentages as the dependent variable. The results are reported in Table 4. Model 1 has only 40 observations because the yield data ends at 1995 and many of the 72 poverty estimates are for later dates. However, the 40 countries are a reasonable sample of the developing world. They are Algeria, Botswana, Bulgaria, Burkina Faso, Central African Republic, Chile, Cote d'Ivoire, Ecuador, Egypt, Estonia, Guatemala, Kenya, Korea, Lesotho, Madagascar, Mali, Mauritania, Mexico, Mongolia, Morocco, Namibia, Nepal, Niger, Paraguay, Poland, Portugal, Romania, Rwanda, Senegal, Sierra Leone, Slovenia, South Africa, Sri Lanka, Tanzania, Tunisia, Turkey, Uganda, Uruguay, Uzbekistan and Zimbabwe. The greatest weakness is that the largest countries of Asia, such as China and India, are missing, so a huge proportion of the world's poor people are not included. Also, some would omit the central and eastern European countries.

The adjusted R^2 of 0.20, in model 1of Table 4, means that yields explains only 20% of the variance in poverty, which is not satisfactory, since it suggests that other, omitted variables explain the majority of the differences, but the productivity measure is significantly different from zero at a high level of confidence. The poverty elasticity of – 0.37 means firstly that higher yields result in lower percentages of the population living in poverty and secondly that a 1% increase in yields reduces the percentage of the populations living on less than $1 per day by 0.37%.

Table 4: Dependent Variable is % of Population with Less than $1 per Day: Cross Section

Explanatory Variables	*Expected Sign*	*Estimated Coefficients*		
		Model 1	**Model 2**	**Model 3**
VA/LAND	Negative	-0.37^{**}		-0.91^{**}
VA/LABOUR	Negative		-0.83^{**}	
LAND/LABOUR	Negative			-0.819^{**}
Constant		4.26^{**}	8.06^{**}	8.48^{**}
R square		0.20	0.506	0.625
F Test		13.35^{**}	13.35^{**}	53.42^{**}
Sample Size		40	66	40

** significant at the 1% level, two-tailed test.

Model 2 explains just over 50% of the variance, which is far better and the poverty elasticity, which is again highly significant, rises to – 0.83, so a 1% improvement in labour productivity reduces the poverty count by 0.83%. The sample increases to 66, but the problem with this model is that the effect could all be coming from the land-labour ratio component of the labour productivity index.

Thus, following these preliminary tests, Model 3 separates the two terms. The model explains 62% of the variance in poverty and the large increase in the F statistic indicates that it is statistically preferred to the two previous attempts. A 1% increase in the land labour ratio reduces poverty by 0.82%, which is surprisingly low relative to the effect of the land productivity term, which indicates that **a 1% improvement in yields decreases the percentage of the population living on less than $1 per day by 0.91%.** Again, the variables are highly significant and this is the preferred model. The result can be developed further if an elasticity can be calculated to link R&D expenditure to yield gains. Then, the cost of generating a 1% decrease in poverty could be calculated. Since R&D expenditures are quite modest, our expectation is that this could look like a very cost effective means of reducing poverty.

The exercise is repeated for the 19 African countries, since Africa is regarded as a major problem area. Although the sample size is barely sufficient, the results are still robust. Table 5 shows that yield alone has no explanatory power, but the poverty elasticity of labour productivity changes very little and once this is decomposed into its two elements in model 3, both are significant. Indeed, the poverty elasticity of land productivity rises from –0.91 to –0.96 and the land/labour ratio adjustment also increases slightly. The land/labour ratio is playing an important role in making the yield variable significant for the Africa sample, by controlling for land quality. The countries with poor soil and water have far more land per unit of labour, so this variable is quality adjusting the data to allow the yield effect to show through. Whilst this is an encouraging result, it is also true that a 1% increase in yields is probably harder to achieve in Africa.

Table 5: Dependent Variable is % of Population with Less than $1 per Day: Africa only

Explanatory Variables	*Expected Sign*	*Estimated Coefficients*		
		Model 1	**Model 2**	**Model 3**
VA/LAND	Negative	-0.022		-0.96**
VA/LABOUR	Negative		-0.87**	
LAND/LABOUR	Negative			-0.95**
Constant		3.19**	8.46**	8.92**
R square		0.0005	0.42	0.47
F Test		0.01	14.39**	7.11**
Sample Size		19	22	19

** significant at the 1% level, two-tailed test.

3.2 Regressions using Pooled Data on $1 per Day Poverty

Whereas the WDR data has only single observations for each country, the data used by Hanmer and Naschold (2000) has scattered observations from 1985 to 1995 for 58 countries, which increased sample size to 109 observations.[6] Models 1 to 3 are the same as in the previous section and the results are reported in Table 6.

Table 6: Dependent Variable is % of population with less than $1 per day: Pooled Sample

Variables	*Expected Sign*	*Estimated Coefficients*			
		Model 1	**Model 2**	**Model 3**	**Model 4**
VA/LAND	Negative	-0.299**		-0.72**	-0.621*
VA/LABOUR	Negative		-0.629**		
LAND/LABOUR	Negative			-0.605**	-0.742*
GINI	Positive				2.153*
YEAR DUMMY		-0.014	0.117	0.1066	0.185
Constant		4.498*	7.177*	7.616*	-0.776
R square		0.088	0.3095	0.328	0.50
F-statistic		6.44**	20.2**	14.79**	15.66**
Sample Size		109	113	109	109

* significant at the 5% level, two tailed test. ** significant at the 1% level, two tailed test.

In model 1, the 1$ a day poverty indicator is regressed on land productivity and only 9% of the variance of poverty is explained. Dummy variables were included to allow for the different time periods, but the coefficients were not significant. Again, the productivity measure is significantly different from zero at a high level of confidence. From the poverty elasticity we can infer that an increase of 1% in labour productivity would bring about a 0.3% decrease in the poverty headcount index

In model 2, the poverty indicator is regressed on labour productivity, which explains 30% of the variance and gives a highly significant elasticity of 0.63. Model 3, where the $1 per day poverty indicator is regressed against both land productivity and the labour land ratio, explains over 32% of the variance in poverty. Both components, land productivity and the land labour ratio, are significant at the 5% level. The poverty elasticities indicate that if land productivity were to increase by 1% there would be a 0.72% reduction in the percentage of the population living on less than $1 per day, whilst if the land

[6] We thank them for making these data available.

labour ratio were to increase by 1%, this would bring about 0.6% decrease in the percentage of people living on less than $1 per day.

Model 4 adds the Gini coefficient, which is an index of inequality, varying from zero, which is perfect equality to unity, which would be complete inequality. Throughout the literature review, it was suggested that greater inequality prevented growth from reducing poverty. The adjusted R^2 in Model 4 rises to 50% and all three variables are statistically significant at the 5% level. The results infer that a 1% increase in land productivity would reduce the poverty headcount index by 0.62 and that a 1% increase in the land/labour ratio would reduce poverty the poverty headcount index by 0.62%. However, the most striking effect is that if there were a 1% decrease in the Gini index, there would be a 2.19% decrease in the poverty headcount index. Thus, the larger sample gives a very similar result to the first cross section regressions. The poverty reduction from a 1% increase in yields still appears to be between 0.6% and 0.7% and the relationship is again highly significant.

Again, the exercise is repeated for the African countries in the sample, which now gives a more reasonable count of 38 observations. Again, the effect of yield alone is not significant, due to the huge differences in land quality and water availability in Africa, but the poverty elasticity of labour productivity rises from –0.629 in the full sample to almost exactly –1.0 for the African countries. Thus, productivity growth has a greater impact in the more predominantly agricultural countries of Africa and this effect is just as marked in the preferred models which decompose labour productivity. A one percent increase in yields gives a slightly greater than one percent reduction in the percentage of the population living on less than $1 per day. Note too that although the Gini coefficient is much the same size it is not significant. This may be due to the smaller sample size, but is also affected by inter-relationships with the agricultural productivity variables, which will be pursued later.

Table 7: Dependent Variable is % of population with less than $1 per day: Pooled Africa sample

Variables	*Expected Sign*	*Estimated Coefficients*			
		Model 1	**Model 2**	**Model 3**	**Model 4**
VA/LAND	Negative	-0.11		-01.04**	-1.02**
VA/LABOUR	Negative		-1.00**		
LAND/LABOUR	Negative			-0.96**	-0.99**
GINI	Positive				2.23
YEAR DUMMY		-0.33	0.15	0.15	0.13
Constant		3.82**	9.44**	9.59**	8.64
R square		0.03	0.55	0.55	0.55
F-statistic		0.68	8.26**	5.79**	4.3**
Sample Size		38	38	38	38

** significant at the 1% level, two tailed test.

3.3 Regressions using the Human Development Index

The earliest observations for the $1 per day poverty index are for 1985, whereas the HDI goes back to 1975, so the HDI was used in the place of the $1 per day poverty measure in an attempt to cover the effects of the green revolution period. Only the preferred model 3 is reported as the preliminary tests have been repeated several times. The two agricultural productivity variables alone explain 76% of the variance in the HDI and both are highly significant. Thus, this regression confirms the apparently solid link between agricultural productivity growth and poverty reduction. Raising yields by 1% increases the HDI by 0.12%, which is the right direction, but improving the value of a composite index does not have the obvious and appealing meaning of reducing the $1 per day measure.

Table 8: Explaining the Human Development Index: Full sample

Variables	*Expected Sign*	*Estimated Coefficients*
		Model 3
VA/LAND	Positive	0.1226**
LAND/LABOUR	Positive	0.1011**

Constant ***R square*** ***F-statistic*** ***Sample Size***		-2.39** 0.759 328.48** 280

** significant at the 1% level, two tailed test.

For the Africa only sample, the elasticities smaller but are still significant, but only 48% of the variance is explained. This is a reasonable result in that the green revolution was a predominantly Asian phenomenon and despite some successes in Africa, agricultural technology has had a lesser impact.

Table 9: Explaining the Human Development Index: Africa sample

Variables	***Expected Sign***	***Estimated Coefficients***
		Model 3
VA/LAND	Positive	0.094**
LAND/LABOUR	Positive	0.079**
Constant		-1.01**
R square		0.48
F-statistic		20.86**
Sample Size		79

** significant at the 1% level, two tailed test.

3.4 Regressions using Nutrition Indicators

Here the data are from Lawrence Haddad of IFPRI. There are 181 mixed cross section and time series observations, from 1971-96, but again this reduces to 109 observations in model 3, which is again the most successful regression. Dummy variables were again used to deal with the time difference. Two variables D8089 and D9096 were generated to adjust for the time differences in the data. Thus, there are two time dummies, for the 1980-1989 and the 1990-1996 respectively, which measure differences relative to 1971-79.

In the first case, the dependent variable is per capita dietary energy supply and 22% of the variance is explained by land productivity and the land labour ratio. The elasticities are highly significant and a 1% increase in land productivity increases the energy supply by 5.3%. This seems somewhat low, especially relative to the results for the second nutrition variable.

Table 10: Explaining Per Capita Dietary Energy Supply: Full sample

Variables	***Expected Sign***	***Estimated Coefficients***
		Model 3
VA/LAND	Positive	0.053**
LAND/LABOUR	Positive	0.060**
DUMMY 8089		0.02
DUMMY 9096		0.016
Constant		7.38**
R square		0.22
F-statistic		9.89**
Sample Size		109

**significant at the 1% level, two tailed test.

For the Africa sample, which was limited to 35 observations, the elasticites are considerably larger than for the full sample, so improvements in land productivity may have been lower than for Asia, but their effects on calorie consumption are greater, for this poorer group of countries.

Table 11: Explaining Per Capita Dietary Energy Supply: Africa sample

Variables	***Expected Sign***	***Estimated Coefficients***

		MODEL 3
VA/LAND	Positive	0.09**
LAND/LABOUR	Positive	0.07*
DUMMY 8089		0.03
DUMMY 9096		-0.04
Constant		7.17**
R square		0.25
F-statistic		2.35*
Sample Size		35

**significant at the 5% level, two tailed test.
*significant at the 10% level, two tailed test.

Table 12 shows 30% of the variance in under-weight five year old children is explained by the two variables and that a 1% increase in yields decreases the percentage by 0.42%.

Table 12: Explaining the % of Under-weigh Children below 5 years old: Full sample

Variables	*Expected Sign*	*Estimated Coefficients*
		Model 4
VA/LAND	Negative	-0.42**
LAND/LABOUR	Negative	-0.25**
D8089		-0.21
D9096		-0.3
Constant		5.04**
R square		0.31
F-statistic		7.63**
Sample Size		109

** significant at the 5% level, two tailed test.

For the Africa sample, the results in Table 13 show that the elasticities are still significant, but in contrast to the calorie consumption results, they are smaller than for the full sample, rather than larger. This would seem to indicate that although increases in land productivity have a bigger impact on energy consumption, the share going to mothers and children is lower in Africa.

Table 13: Explaining the % of Under-weigh Children below 5 years old: Africa sample

Variables	*Expected Sign*	*Estimated Coefficients*
		Model 4
VA/LAND	Negative	-0.27**
LAND/LABOUR	Negative	-0.2*
D8089		-0.26*
D9096		0.15
Constant		4.76**
R square		0.28
F-statistic		2.6**
Sample Size		38

** significant at the 1% level, two tailed test.
* significant at the 5% level, two tailed test.

Summary

Agricultural productivity increases have significant effects on all the poverty and nutrition measures, both for the full sample and for the African countries alone. There is also clear evidence that inequality, as measured by the Gini coefficients, has a substantial negative impact on poverty reduction. However, although these results appear to be robust, before the actual magnitudes can be believed, further work is required. Particularly, Tables 1-3 suggested that there are many other variables affecting poverty and inequality, so these must be taken into account. Otherwise, variable omission may bias the elasticities. Specifically if a variable that is positively correlated with the explanatory variables is omitted, this will tend to bias the elasticities upwards, since the contribution of the omitted variable will tend to be picked up by those that are included. The reverse will be true if the omitted variable is negatively correlated with the regressors. However, we first compare the impacts of agricultural productivity growth with the effects of productivity growth in industry and the services sector.

4 DOES PRODUCTIVITY GROWTH IN INDUSTRY AND SERVICES REDUCE POVERTY?

The effect of agricultural productivity on poverty reduction appears to be consistent and substantial, but is agriculture different? It is possible that productivity growth in industry or services may have similar effects. The literature review in Thirtle et al (2001) suggested that agricultural productivity growth should have a greater effect, but this proposition has not yet been tested. Thus, we now extend the analysis to industry and services.

First, the correlations between productivity in the three sectors are reported in Table 14, which suggests that the literature is indeed correct.[7] The first column shows that the correlation coefficient between labour productivity in services and the percentage of the population living on less than $1 per day is only –0.05, which the probability value shows is not significantly different from zero. Similarly, the correlation coefficient for industry is only –0.015, which is again insignificant. However, the correlation coefficient for agricultural productivity is –0.55 and this is significant at the 99% confidence level. This is 11 times greater than the services coefficient and 36 times larger than that for industry. The other essential finding in the correlation matrix is the correlation between labour productivity in industry and services, which is 0.888. This is high enough to suggest that if both are included in the same regression, they will cancel each other out, in the sense of making both elasticities insignificant.

Table 14: Inter-sectoral productivity correlations

	$1 per day Poverty	VA/Labour (service)	VA/Labour (industry)
VA/Labour (service) *(p-value)*	-0.0524 *0.6984*	1	
VA/Labour (industry) *(p-value)*	-0.015 *0.9141*	0.8879 *0*	1
VA/Labour (agriculture) *(p-value)*	-0.5545 *0*	0.0329 *0.8152*	0.085 *0.5571*

Thus, the first regressions fit the productivity variables individually, although this may be a poor specification. Since bivariate regression is very little different from correlation, it is not surprising that the elasticities reported in Table 15 are very similar to the correlation coefficients. The differences that do occur can be attributed to the logarithmic transformation, which makes the coefficients elasticities. The R^2 values for the services and industry regression show that these variables have almost no explanatory power and this is confirmed by the t statistics, which show that both are insignificant in explaining the percentage of the population living on less than $1 per day. In contrast, agricultural labour productivity alone explains 31% of the variance and the poverty elasticity is highly significant. However, the models are fitted to the available samples, which differ for the three sectors.

---

[7] The data on value added for industry and services has missing data before 1992, which required some interpolation.

Table 15: Dependent Variable is % of Population with Less than $1 per Day: Sectoral Productivities

Explanatory Variables	*Expected Sign*	*Estimated Coefficients*		
		Model 1	**Model 2**	**Model 3**
VA/LABOUR (Service)	Negative	-0.054		
VA/LABOUR (Industry)	Negative		-0.014	
VA/LABOUR (Agriculture)	Negative			-0.63*
Constant		3.25**	2.97**	7.2**
R square		0.003	0.0002	0.31
F Test		0.75	0.97	6.44**
Sample Size		57	54	113

** significant at the 1% level, two-tailed test.
* significant at the 5% level, two tailed test.

Since, the test is more reliable if the same sample is used for all three sectoral productivities, this is the next step and now all three variables are included in a single equation. The results, reported in Table 16 as model 4 show that the single equations outcomes are changed. The elasticities are of the same magnitude, but the sign is reversed for industry and only agricultural productivity is significant. However, the high correlation between labour productivity in services and industry makes this result suspect, so we next resort to two stage estimation in which the two are not included together. Still, the results, reported in models 5 and 6 change very little, in that industrial and service sector productivity are insignificant. These data show that there is little doubt that labour productivity in agriculture has a significant impact on poverty reduction, whereas productivity growth in industry and services has no discernable effect.

Table 16: Dependent Variable is % of Population with Less than $1 per Day: Sectoral productivities

Explanatory Variables	*Expected Sign*	*Estimated Coefficients*		
			2-Stage estimation	
		Model 4	**Model 5**	**Model 6**
VA/LABOUR (Service)	Negative	-0.05	(exogenous)	0.06
VA/LABOUR (Industry)	Negative	0.09	0.04	(exogenous)
VA/LABOUR (Agriculture)	Negative	-0.65*	-0.64*	-0.64*
Constant		6.9**	6.98**	6.8**
R square		0.29**	0.29	0.29
F Test		3.29	4.48	3.41**
Sample Size		50	50	50

** significant at the 1% level, two-tailed test.
* significant at the 5% level, two tailed test.

5 EXPLAINING INEQUALITY & IMPROVING THE MODEL SPECIFICATION

5.1 Single equation approach

The results in the previous two sections are an improvement on the simple correlations of section 2, but as Table 1 showed, there are several other variables, which appear to highly correlated with poverty and inequality. Prior knowledge suggests that some of these should be included in models that explain poverty reduction, and misspecification, due to variable omission can bias the estimates of the elasticities that have been reported. Thus, we now attempt to extend the model to include all the variables that explain poverty and inequality reduction. Given the limited sample size, the dubious quality of some of the data and the co linearity of the explanatory variables, not too much can be expected.

Comparisons can be made with other models in which the less than $1 per day poverty index is used as the dependent variable, especially Hanmer and Naschold (2000), who use the same poverty data. The models that are reported were chosen on the basis of goodness of fit (the F statistic is preferred to the R^2, although the latter was not ignored), the significance of the variables and particularly analysis of the residuals. The last takes precedence, since well-behaved, randomly distributed residuals indicate that the model is free from the econometric problems that can bias the estimates. All the significant variables from Table 1 were included in these tests, in various combinations, but the number of significant variables that could be retained in a single equation was extremely limited.

Thus, the only additional variable that improved the model for all three samples was gross fixed investment, which was included to capture the effects of infrastructure. Table 17 shows that the preferred model retains the agricultural productivity variables and the Gini, in addition to gross fixed investment. Model 3 should be compared with model 4 in Table 9, which used the same full sample of 109 observations. Adding the investment variable increases the explanatory power of the model from 50% to 58% and increases the F statistic from 15.66 to 22.31, which is a substantial improvement. The key finding is that the poverty elasticity of the yield variable is reduced only from –0.62 to –0.56. Gross fixed investment has a larger elasticity of –0.93, which indicates its importance. The elasticity for the Gini is far larger, at 2.28, but this variable has a very limited variance, so the actual magnitude could be misleading.

Table 17: Dependent Variable is % of population with less than $1 per day

Explanatory Variables	*Expected Sign*	***Estimated Coefficients***		
		Model 1 (Non-Africa)	**Model 2 (Africa)**	**Model 3 (Combined)**
VA/LAND	Negative	-0.61^{***}	-0.95^{***}	-0.56^{***}
LAND/LABOUR	Negative	-0.81^{***}	-0.92^{***}	-0.71^{***}
GINI	Positive	2.76^{***}	0.72	2.28^{***}
GROSS FIX INV.	Negative	-1.47^{***}	-0.48^{*}	-0.93^{***}
Constant		1.55	7.74^{**}	1.16
R square		0.59^{***}	0.59^{**}	0.58^{**}
F Test		14.61^{***}	7.83^{***}	22.31^{***}
Sample Size		71	38	109

*** significant at the 1% level, two tailed test. ** significant at the 5% level, two-tailed test. * significant at the 10% level, two-tailed test.

For the non-African countries, the results are essentially similar to those for the full sample, but the Africa only estimates are more interesting. These results can be compared with model 4, in Table 7, and now inequality appears to be much less of a problem in Africa. However, this is not the case, according to Table 3, which reported higher inequality for the Africa sample. The lower impact of the Gini coefficient in the regressions is the result of including gross fixed investment. Although the yield elasticity is slightly reduced, yield is still the most powerful variable in this case.

5.2 Explaining inequality

Thus, the single equation approach is somewhat limited, but this is partly as a result of the small sample for Africa. However, before presenting some better results for the full sample, we move on to estimating single equation models to explain inequality. This leads on to a two equation, two-stage model and a recursive model. The first model reported in Table 18 is a simple experiment to see if the agricultural productivity variables have a significant impact on inequality. Although their elasticities are significant and explain 22% of the variance, both are extremely small. Whereas the correlation coefficients for both value added per unit of land and the land labour ratio were slightly greater than 0.4 in absolute value, in the bivariate correlations of Table 1, they had opposite signs and labour productivity was not significant, because the two elements cancelled each other out.

Table 18: Dependent Variable is the Gini coefficient: Full sample

Explanatory Variables	*Expected Sign*	*Estimated Coefficients*				
		Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
VA/LAND	Negative	-0.046*			-0.06***	-0.09***
LAND/LABOUR	Negative	0.06***		0.96**		
RURAL POPUPATION (%)	Negative		-0.20***	0.13**		-0.124***
POPULATION GROWTH	Positive		0.19***		0.11**	
Constant		2.1	4.41	3.53**	4.01**	4.73***
R square		0.22	0.20	0.36	0.18	0.21
F Test		20.3***	11.73***	30.3**	13.24***	16.81***
Sample Size		109	88	102	102	109

*** significant at the 1% level, two-tailed test. ** significant at the 5% level, two-tailed test. * significant at the 10% level, two-tailed test.

The two other explanatory variables that were significant in these regressions were the percentage of the population that was rural, which had a negative effect on inequality, and the rate of population growth, which increased inequality. In model 2, on their own, these two variables explained less of the variance than did the agricultural productivity variables and the F statistic is far lower. This model is reported here as a single equation, because it appears later in as one half of a two equation recursive model, in which the agricultural productivity variables enter only in the other equation. This route is taken partly because the four variables are not significant if they are included in a single equation. However, the main reason is that the two should not be estimated independently. The Gini coefficient is an explanatory variable in the poverty equation and the dependent variable in the other equation, so the two should be estimated simultaneously. The other three models show the combinations of the four variables that are significant, but the results are not easy to interpret. For instance, the rural population variable changes sign when combined with the land labour ratio, although the other variables are reasonably well behaved. In all, attempts to the attempts to explain inequality were disappointing.

5.3 Single equations using residuals and recursive systems

Although the Gini regressions shed little light directly, they do suggest that a system of equations may be the best approach, so these models are developed below. First, model 1 in Table 19 shows that trade as a percentage of GDP can be added to the equation of model 3 in Table 17. It has a negative effect on poverty, confirming the correlation result of Table 1 that openness is poverty reducing. However, although the explanatory power is a little better, the F statistic is lower. It does seem that five variables is as many as the single equation model will accommodate and this is also the number of variables in the Hanmer and Naschold (2000) model. The choice is somewhat arbitrary: as model 2 shows, if trade is replaced by the GDP growth rate (not per capita) the model is about equally acceptable and GDP growth has a poverty elasticity of –0.16. In both these models, the yield elasticity stays very close to the value of –0.56, that was the estimate in model 3 of Table 17.

Table 19: Dependent Variable is % of Population with Less than $1 per Day

Explanatory Variables	*Expected Sign*				
		Model 1	**Model 2**	**Model 3**	**Model 4 Simultaneous**
VA/LAND	Negative	-0.55**	-0.54**	-0.25**	-0.49**
LAND/LABOUR	Negative	-0.67**	-0.65**	-0.36**	-0.64**
GINI	Positive	2.28**	1.74**	2.25**	2.33*

TRADE (%GDP)	Negative	-0.26*			-0.29*
GROSS FIX INV.	Negative	-0.80**	-0.77**	-0.84**	-0.78**
GDP GROWTH	Negative		-0.16*		
GDP PER CAPITA				-0.47**	
RURAL POPUPATION (%)	Negative				-0.15**
POPULATION GROWTH	Positive				0.24**
Constant		1.73	2.79*	2.44*	1.25
R square		0.59	0.52	0.59	0.51
F Test		18.20**	18.35**	19.3**	83.48**
Sample Size		109	83	107	102

** significant at the 5% level, two-tailed test. * significant at the 10% level, two-tailed test. *** The dependent Variable in the first part of model 4 is poverty and in the second the GINI Coefficient. The test statistic for the simultaneous model is Chi-Square rather than F.

This is not true if GDP per capita replaces GDP growth, as it does in model 3. Then, the poverty elasticity of the yield is reduced to –0.25, which is less than half what it was, and GDP per capita has a poverty elasticity of –0.47. We would argue that this does not indicate that the previous models were miss-specified, but rather that higher yields are an important component of higher GDP per capita and hence the two should not be explanatory variables in the same equation. For comparison, in Hanmer and Naschold (2000), the poverty elasticity of GDP per capita is –0.93 for the low inequality countries and –0.34 for those with high inequality. They included the ratio of value added per worker in the modern sector to value added per worker in agriculture as an explanatory variable. This seems to have a similar effect to including agricultural productivity as we have done. But, when productivity explicitly competes with GDP per capita, the yield elasticity was halved, relative to our earlier models.

Further work is required in modelling the relationship if both are to be included, but for now the problem is ignored and we progress to a recursive two-equation model. Thus, model 4 combines model 1 with model 2 of Table 18, to give equations (5a) and (5b), which are estimated simultaneously using Zellner's seemingly unrelated regression model or 3 stage least squares.[8] In this model and in all that follow, all of the variables are in logarithms, so that the coefficients can be interpreted as elasticities.

$$\$1 day\ poverty = \beta_0 + \beta_1\left[\frac{VA_{AG}}{Land_{AG}}\right] + \beta_2\left[\frac{Land_{AG}}{Labour_{AG}}\right] + \beta_3 Gini + \beta_4 Trade + \beta_5 Fix\ Invest + \varepsilon \quad \textbf{(5a}$$

$$Gini = \alpha_0 Rural\ Pop + \alpha_1 Pop\ Growth + \nu \quad \textbf{(5b)}$$

In this way the model can be stretched to seven explanatory variables and the results change what is probably the most theoretically correct model so far, since we conducted tests to ensure this. The classical assumptions of regression analysis suggest that the standard errors of the fitted coefficients will be increased if there are correlated independent variables in the model. This may cause the exclusion of variables that actually have a statistical relation with the dependent variable from the model of study. Ideally, all variables in a model should be independent; but in practice, it is too strict and may jeopardise the power of the model. So, the VIF (Variable Inflation Factor) is then used to test the seriousness of the multicolinearity problem. Many researchers rely on informal rules of thumb applied to the VIF. For example, Studenmund (1997, p.276) suggests that multicolinearity is severe if the VIF value for any variables is larger than 5. The land productivity (VA/LAND) and land over labour ratio (LAND/LABOUR) are correlated to many socio-economic variables. Therefore, in the models we selected, the VIF value is reported to show that the inclusion of these factors in a model is not violating the co linearity assumption. That is, the results inferred from these models appear to be reliable, since Table 20 shows that these variables are not highly correlated.

[8] Since the error terms are related, Zellner's model is required, but 3 stage least squares is not strictly necessary, as the model is not genuinely simultaneous. The two techniques give very similar results.

Table 20: Variable Inflationary Factors

Variable	VIF	1/VIF
LAND/LABOUR	2.09	0.477692
VA/LAND	1.85	0.54125
TRADE (%GDP)	1.33	0.749123
GINI	1.3	0.770887
GROSS FIX INV.	1.18	0.84705
Mean VIF	1.55	

Both the single equation approach and recursive models have been used in the growth and poverty alleviation literature. The World Bank poverty and growth studies have frequently applied the single equation approach, as did Hanmer and Naschold (2000), whereas IFPRI studies, such as Fan, Hazell and Haque (2000) opt for simultaneous methods. The advantage of the first is apparent simplicity, but as model 3 in Table 19 showed, there are clear limitations. GDP per capita depends on so many factors that it is difficult to avoid including the variables that determine it in the single equation. The effect is apparent when a major determinant of GDP per capita is included, as it was in this model. There are two ways of overcoming this difficulty and both are used in the next set of models. The alternative to the recursive equation approach is to resolve the conflict between GDP per capita and agricultural productivity prior to estimating the single equation.

Thus, model 1, below explicitly deals with the fact that agricultural labour productivity is a component of GDP per capita. The first equation (6a) has both as independent variables, but the second equation (6b) regresses GDP per capita on the labour productivity variables. The residual from this equation may be viewed as GDP per capita net of the contribution of agricultural labour productivity. Then, it is this agricultural productivity free GDP variable that is used in the first equation.

__Model 1:__

$$\$1poverty = \beta_0 + \beta_1\left[\frac{VA_{AG}}{Land_{AG}}\right] + \beta_2\left[\frac{Land_{AG}}{Labour_{AG}}\right] + \beta_3 Gini + \beta_4 Trade + \beta_5 Fixed\ Invest + \beta_6\varepsilon_{GDP\ PC} + \xi \qquad \textbf{(6a)}$$

$$where\ GDP\ per\ capita = \gamma_0 + \gamma_1\left[\frac{VA_{AG}}{Land_{AG}}\right] + \gamma_2\left[\frac{Land_{AG}}{Labour_{AG}}\right] + \varepsilon_{GDP\ PC} + \zeta \qquad \textbf{(6b)}$$

$$Gini = \alpha_0 + \alpha_1 Rural\ Population + \alpha_2 Population\ Growth + \varsigma \qquad \textbf{(6c)}$$

The effect of this transformation can be seen in the first column of Table 21. The poverty elasticity of agricultural productivity is now very close to its previous values, at –0.48 and the poverty elasticity of GDP per capita recovers to –0.68, which is considerable higher than the figure of –0.47 in model 3 of Table 19. Indeed, this is a little higher than the unweighted average of Hanmer and Naschold's (2000) results, which is –0.635. Thus, both agricultural productivity and GDP per capita can be included in same equation. The third equation (6c) again has the Gini coefficient as the dependent variable and this is fitted simultaneously with the first equation and produces much the same results as before. The problem with this model is that the Gini coefficient is insignificant in the first equation and the reason is clear in model 2.

Table 21: Dependent Variable is % of Population with Less than $1 per Day: Recursive Models

Explanatory Variables	**Expected Sign**	**Estimated Coefficients**			
		Model 1	**Model 2**	**Model 3**	**Model 4**
Poverty Reduction					

VA/LAND	Negative	-0.48**		-0.65**	
LAND/LABOUR	Negative	-0.52**		-0.48**	
GINI	Positive	1.00	1.65**		0.99*
TRADE (%GDP)	Negative	-0.36**	-0.35**	-0.34**	-0.34**
GROSS FIX INV.	Negative	-0.68**	-0.45**	-0.67**	-0.66**
GDP PER CAPITA	Negative		-0.88**		-0.76**
AGRICULTURE FREE	Negative	-0.65**		-0.59**	
GDP PER CAPITA					
AGRICULTURE FREE	Positive			2.21**	
GINI					
CONSTANT		6.01	6.72**	10.40**	8.33**
R-SQUARE		0.52	0.60	0.61	0.52
CHI-SQUARE OR F-STATISTIC		97.00**	105.61**	16.89**	90.16**
SAMPLE SIZE		100	107	107	100
Gini					
RURAL POPULATION	Negative	-0.18**			-0.18**
POPULATION GROWTH	Positive	0.23**			0.24**
VA/LAND	Negative		-0.05**		
LAND/LABOUR	Negative		0.06**		
CONSTANT		4.3**	3.92**		4.30**
R-SQUARE		0.21	0.20		0.21
CHI-SQUARE		27.58**	27.06**		29.61**
SAMPLE SIZE		100	107		100
GDP PER CAPITA					
VA/LAND	Positive		0.66**		0.65**
LAND/LABOUR	Positive		0.68**		0.67**
CONSTANT			3.13**		3.19**
R-SQUARE			0.80		0.78
CHI-SQUARE			423.20**		361.73**
SAMPLE SIZE			107		100

** significant at the 5% level, two-tailed test.   * significant at the 10% level, two-tailed test.

Model 2 shows the extent of the inter-relationships, by dropping the agricultural productivity variables from the first equation (7a). Then, the second column of Table 21 shows that the Gini coefficient is again significant, as are the percentage of GDP that is traded, gross fixed investment and GDP per capita. Sixty percent of the variance is explained by these variables. The second equation (7b) shows that the two labour productivity variables have as much power in explaining inequality as did the percentage of the population that is rural and the rate of population growth, but in both cases the R^2 remains low at about 0.2. Thus, as the models become more refined, it becomes clear that like GDP per capita, the Gini coefficient is a function of the labour productivity variables and should not be included in the same equation. The third equation (7c) uses just the agricultural productivity variables to explain GDP per capita and Table 21 shows that the explanatory power is amazingly high, at 80% and so is the Chi-Square test value for this regression.

<u>Model 2:</u>

$$\$1\,Poverty = \beta_0 + \beta_1 Gini + \beta_2 Trade + \beta_3 Fixed\ Invest + \beta_4 GDP\ per\ capita + \varepsilon \qquad \textbf{(7a)}$$

$$GDP\ per\ capita = \gamma_0 + \gamma_1\left[\frac{VA_{AG}}{Land_{AG}}\right] + \gamma_2\left[\frac{Land_{AG}}{Labour_{AG}}\right] + \xi \qquad \textbf{(7b)}$$

$$Gini = \alpha_0 + \alpha_1 \left[\frac{VA_{AG}}{Land_{AG}} \right] + \alpha_2 \left[\frac{Land_{AG}}{Labour_{AG}} \right] + \zeta \quad \textbf{(7c)}$$

Thus, model 3 is a single equation (8a) in which GDP per capita is the residual, with agricultural productivity effects removed as in (8b) and with the agricultural productivity effects similarly removed from the Gini coefficient by using the residual from (8c). The results reported in the third column of Table 21 show that the poverty elasticity of land productivity now reverts to the higher level it had in the earlier models (see Table 6), at –0.65. The land labour ratio, percentage of GDP traded and gross fixed investment are all significant, and so is agricultural productivity free GDP per capita, which retains a very reasonable elasticity of -0.59. However, this means that the poverty reducing effects of improving agricultural productivity are actually greater than the effect of increasing GDP per capita. This seems unlikely, as although agricultural productivity does seem to be the key variable, it is still only a component of GDP per capita, all be it the main one. Thus, the broader variable should have the greater impact. Other than this quibble, removing the double counting of effects caused by the impacts of the productivity variables on both GDP per capita and the Gini coefficient gives an equation that is entirely satisfactory, that explains 61% of the variance, with six independent variables. This does seem to be the limit for single equation models of this relationship, so we now return to recursive models to show how more variables can be incorporated and explanatory power increased.

Model 3:

$$\$1\,poverty = \beta_0 + \beta_1 \left[\frac{VA_{AG}}{Land_{AG}} \right] + \beta_2 \left[\frac{Land_{AG}}{Labour_{AG}} \right] + \beta_3 \varepsilon_{GINI} + \beta_4 Trade + \beta_5 Fixed\ Invest + \beta_6 \varepsilon_{GDP\ PC} + \xi \quad \textbf{(8a)}$$

$$where\ GDP\ per\ capita = \gamma_0 + \gamma_1 \left[\frac{VA_{AG}}{Land_{AG}} \right] + \gamma_2 \left[\frac{Land_{AG}}{Labour_{AG}} \right] + \varepsilon_{GDP\ PC} \quad \textbf{(8b)}$$

$$and\ Gini = \alpha_0 + \alpha_1 \left[\frac{VA_{AG}}{Land_{AG}} \right] + \alpha_2 \left[\frac{Land_{AG}}{Labour_{AG}} \right] + \varepsilon_{GINI} \quad \textbf{(8c)}$$

It is not really possible to identify one single formulation of the recursive model that dominates all others, but some of the best models can be briefly covered, to show the range of options. Model 4 is the same as model 2, except that in equation (9c), the percentage of the population that is rural and the population growth rate replace the agricultural productivity measures, increasing the number of variables to eight, instead of six. However, if model 4 is compared with model 2, the equations fit less well and this is shown by the R^2 values and the Chi Square test statistics, which fall considerably for equations (8a) and (8c).

Model 4:

$$\$1\ per\ day\ poverty = \beta_0 + \beta_1 Gini + \beta_2 Trade + \beta_3 Fixed\ \ Invest + \beta_4 GDP\ per\ capita + \varepsilon \quad \textbf{(9a)}$$

$$GDP\ per\ capita = \gamma_0 + \gamma_1 \left[\frac{VA_{AG}}{Land_{AG}} \right] + \gamma_2 \left[\frac{Land_{AG}}{Labour_{AG}} \right] + \xi \quad \textbf{(9b)}$$

$$Gini = \alpha_0 + \alpha_1 Rural\ \ Population + \alpha_2 Population\ \ Growth + \zeta \quad \textbf{(9c)}$$

Model 5 experiments with the structure of the model by reducing the explanatory variables in the poverty equation (10a) to just the Gini coefficient and GDP per capita. The, Gini coefficient is explained by the percentage of the population that is rural and the rate of population growth, in (10 c), while all the other variables explain GDP per capita growth in (10b). Together, the five exogenous variables explaining 84% of the variance in GDP per capita, which is reassuring and the model obviously has the highest value for the Chi-Square test for this equation. However, this formulation fails because the Gini coefficient in (10a) is not significant, as the first column in Table 22 shows.

Model 5:

$$\$1\, per\ day\ \ poverty = \beta_0 + \beta_1 Gini + \beta_2 GDP\ per\ capita + \varepsilon \qquad \textbf{(10a)}$$

$$GDP\ per\ capita = \gamma_0 + \gamma_1\left[\frac{VA_{AG}}{Land_{AG}}\right] + \gamma_2\left[\frac{Land_{AG}}{Labour_{AG}}\right] + \gamma_3 Trade + \gamma_4 Fixed\ Invest + \gamma_5 Illiteracy + \xi \qquad \textbf{(10b)}$$

$$Gini = \alpha_0 + \alpha_1 Rural\ Population + \alpha_2 Population\ \ Growth + \zeta \qquad \textbf{(10c)}$$

Thus, model 6 reverts to the previous distribution of the variables between equations and differs from model 4 only in including labour productivity in industry and services as explanatory variables in the GDP per capita equation (11b). The results, in the second column of Table 22 show that neither of these variables is significant, so model 2 remains preferred. Note though, that the test statistics, such as the Chi-Squares, are reduced not by the effect of having two insignificant variables, but by the reduction in sample size, due to the number of missing observation for these variables. If it were possible to construct a larger sample, it is likely that these productivity variables would have a significant impact, but far lower than for agricultural productivity.

Model 6:

$$\$1\, per\ day\ \ poverty = \beta_0 + \beta_1 Gini + \beta_2 Trade + \beta_3 Fixed\ Invest + \beta_4 GDP\ per\ capita + \varepsilon \qquad \textbf{(11a)}$$

$$GDP\ per\ capita = \gamma_0 + \gamma_1\left[\frac{VA_{AG}}{Land_{AG}}\right] + \gamma_2\left[\frac{Land_{AG}}{Labour_{AG}}\right] + \gamma_3\left[\frac{VA_{Services}}{Labour_{Services}}\right] + \gamma_4\left[\frac{VA_{Industry}}{Labour_{Industry}}\right] + \xi \qquad \textbf{(11b)}$$

$$Gini = \alpha_0 + \alpha_1 Rural\ Population + \alpha_2 Population\, Growth + \zeta \qquad \textbf{(11c)}$$

Model 7 differs from 6 only in dropping these two insignificant variables and adding illiteracy to equation (12b). The third column shows that this is significant and this is probably the best of the recursive models to date. It has the greatest number of significant explanatory variables (nine): the highest Chi-Square test statistics, and the best fit overall, in that 51% of the variance is explained in the poverty equation, 21% in the inequality equation and 82% in the GDP per capita equation. The values of the two key poverty elasticities are –0.72 for the poverty elasticity of GDP per capita and –0.40 for the poverty elasticity of land productivity. Since land productivity in the recursive model affects poverty through its effect on GDP per capita, its elasticity is the calculated as the product, -(0.72)*(0.55) = -0.40.

Table 22: Dependent Variable is % of Population with Less than $1 per Day: Recursive Models

Explanatory Variables	**Expected Sign**	**Estimated Coefficients**			
		Model 5	**Model 6**	**Model 7**	**Model 8**
Poverty Reduction					
VA/LAND	Negative				
LAND/LABOUR	Negative				
GINI	Positive	0.04	-0.216	0.94*	1.74**
TRADE (%GDP)	Negative		-0.17	-0.37**	-0.50**
GROSS FIX INV.	Negative		-1.16**	-0.64**	-0.75**
GDP PER CAPITA	Negative	-0.76**	-0.78**	-0.72**	-0.76**
AGRICULTURE FREE GDP PER CAPITA	Negative				
CONSTANT		8.66**	13.92**	8.32**	6.32**

R-SQUARE		0.31	0.34	0.51	0.65
CHI-SQUARE OR F-STATISTIC		45.15**	38.39**	85.24**	85.92**
SAMPLE SIZE		99	42	99	52
GINI					
RURAL POPULATION	Negative	-1.67**	-0.08	-0.18**	-0.2**
POPULATION GROWTH	Positive	0.21**	0.19**	0.23**	0.33**
VA/LAND	Negative				
LAND/LABOUR	Negative				
CONSTANT		4.27**	3.97**	4.31**	4.33**
R-SQUARE		0.21	0.21	0.21	0.42
CHI-SQUARE		24.86**	10.54**	28.50**	46.31**
SAMPLE SIZE		99	42	99	52
GDP PER CAPITA					
VA/LAND	Positive	0.55**	0.63**	0.55**	0.58**
LAND/LABOUR	Positive	0.62**	0.70**	0.59**	0.57**
VA/LABOUR(SERVICE)	Positive		0.12		
VA/LABOUR(INDUSTRY)	Positive		-0.07		
TRADE (%GDP)	Negative	-0.21**			
GROSS FIX INV.	Positive	0.19**			
ILLITERACY	Negative	-0.16**		-0.16**	
EXPENDITURE ON PRIMARY EDUCATION	Positive				-0.14*
CONSTANT		4.55**	2.76	4.31**	4.06
R-SQUARE		0.84	0.78	0.82	0.77
CHI-SQUARE		517.06**	145.53**	454.26**	177.48**
SAMPLE SIZE		99	42	99	52

** significant at the 5% level, two-tailed test. * significant at the 10% level, two-tailed test.

Model 7:

$$\$1\,per\,day\,poverty = \beta_0 + \beta_1 Gini + \beta_2 Trade + \beta_3 Fixed\ Invest + \beta_4 GDP\ per\ capita + \varepsilon \quad \textbf{(12a)}$$

$$GDP\ per\ capita = \gamma_0 + \gamma_1\left[\frac{VA_{AG}}{Land_{AG}}\right] + \gamma_2\left[\frac{Land_{AG}}{Labour_{AG}}\right] + \gamma_3 Illiteracy + \xi \quad \textbf{(12b)}$$

$$Gini = \alpha_0 + \alpha_1 Rural\ Population + \alpha_2 Population\ Growth + \zeta \quad \textbf{(12c)}$$

The only disappointing result in model 7 is the low explanatory power of the inequality equation. The last model shows that this is doubled to 42% if illiteracy is replaced by the expenditure on primary education in equation (13b) and there is the additional benefit that the explanatory power of the poverty equation increases to 65%, which is as high as it has been in any of the models. The costs are only a reduction in the R^2 for (13b), from 82% to 77% and a reduction in the Chi-Square test statistic for this regression, balanced by a big improvement in the Chi-Square for the Gini equation, and the average explanatory power of the three equations is over 60%. The only snag is that the effect of higher expenditures per student in primary education is negative. This is not so unreasonable, since the policies may well be donor funded and reactive, so that countries with the most severe problems are spending most. The literacy rate is in many ways a better measure since it records the success of past education expenditures. Current expenditures have an effect only with a considerable lag, which has not been allowed for here. Thus, this model is a viable alternative to model 7 and the single equation results of model 3, in Table 21. The key elasticities are almost identical to the model 7 results, except that the poverty elasticity of the yield rises by ten percent, to –0.44.

Model 8:

$$\$1\, per\ day\ poverty = \beta_0 + \beta_1 Gini + \beta_2 Trade + \beta_3 Fixed\ Invest + \beta_4 GDP\ per\ capita + \varepsilon \quad \textbf{(13a)}$$

$$GDP\ per\ capita = \gamma_0 + \gamma_1 \left[\frac{VA_{AG}}{Land_{AG}}\right] + \gamma_2 \left[\frac{Land_{AG}}{Labour_{AG}}\right] + \gamma_3 PEducation + \xi \quad \textbf{(13b)}$$

$$Gini = \alpha_0 + \alpha_1 Rural\ Population + \alpha_2 Population\, Growth + \zeta \quad \textbf{(13c)}$$

Summary

This section has shown that either single equation techniques of recursive equation methods can explain over 60% of the variance in the percentage of the population living on less than $1 per day. The two approaches give different answers because the constraints differ. If GDP per capita and the Gini coefficient are purged of the effects of agricultural productivity growth by prior regressions, so that all three can be included in the same equation, the poverty elasticity of yield growth is –0.65. If the yield can only affect poverty through its effects on GDP per capita, as in the recursive models, then the poverty elasticity is reduced to -0.40, because direct effects are precluded. Thus, since there are direct effects, we would argue that –0.65 is the more appropriate figure, and is this value is typical of the results we have produced, however the model has been formulated. Indeed, this elasticity does seem to be surprisingly robust.

The recursive models do make another point. The question we were asked to investigate was the effect of agricultural productivity growth on poverty, but it is clear that growth in yields not only affects poverty, but has an almost equally big impact on GDP per capita, where the elasticity is about 0.58. Thus, everybody gains from yield growth, not just the poor, so ignoring the GDP per capita gains would be foolish.

6) R&D COST OF YIELD INCREASES TO REDUCE POVERTY & INCREASE GDP PER CAPITA

There is a final stage in the development of the recursive model, which is adding a further equation in which land productivity is explained by agricultural R&D, fertiliser, land quality and illiteracy. The relationship, which has been extensively modelled (see Thirtle, 1999, for example), should also include extension expenditures and a weather index, but these data do not exist for most of the sample.[9] Thus, Model 9 adds an additional equation (14d), which completes the model by allowing estimation of the elasticity of yields with respect to R&D expenditures. The R&D data needs scaling, so it is expressed as per unit of land, to match up with the dependent variable in this yield equation and fertiliser is treated in the same way. The land quality index is from the USDA (2001) web site, but it is not included in equation (14d) because it was not significant.

[9] The rationale for this formulation is that R&D generates new technology, which usually requires more fertiliser. Extension takes the innovation from the trial plot to the farmers and literate farmers fare better at adapting the technology to their particular circumstances. The weather index accounts for a considerable proportion of the residual, so the fit would be improved if these data were available.

This allows calculation of the marginal internal rate of return to investment in agricultural R&D for the countries in this sample. Also, once the R&D expenditures for the sample are known, the model can determine with reasonable accuracy the expenditure necessary to increase GDP per capita by one percent or to reduce the percentage of the population living on $1 per day by one percent. This allows calculation of the extra income generated and the number of people who move out of the less than $1 per day poverty bracket. This final model is specified as Model 9, below, which is fitted to the same sample as Model 8, since for these countries R&D data is available from Pardey and Roseboom (1989) for all but four observations.

Model 9:

$$\$1\,per\ day\ poverty = \beta_0 + \beta_1 Gini + \beta_2 Trade + \beta_3 Fixed\ Invest + \beta_4 GDP\ per\ capita + \varepsilon \tag{14a}$$

$$GDP\ per\ capita = \gamma_0 + \gamma_1 \left[\frac{VA_{AG}}{Land_{AG}}\right] + \gamma_2 \left[\frac{Land_{AG}}{Labour_{AG}}\right] + \gamma_3 PEducation + \xi \tag{14b}$$

$$Gini = \alpha_0 + \alpha_1 Rural\ Population + \alpha_2 Population\ Growth + \zeta \tag{14c}$$

$$\frac{VA}{Land} = \phi_0 + \phi_1[R\&D/Land] + \phi_2[Fertiliser/Land] + \phi_3[Illiteracy] + \varsigma \tag{14d}$$

The results of model 8 are changed only slightly by the additional yield equation. The only cost is that the inequality equation reverts to explaining only 21% of the variance, but the yield equation is robust and explains 73% of the variance. R&D, fertiliser and illiteracy are highly significant and the elasticity of R&D is 0.62. This elasticity is the crucial piece of information needed to link R&D expenditures to the increase in the value of output that results from the higher yield, allowing the rate of return to be calculated.

Table 23: Dependent Variable is % of Population with Less than $1 per Day: Recursive Model

Dependent variable, followed by explanatory variables	***Expected Sign***	
Poverty Reduction Equation		**Model 13**
% of Population Living on Less than $1 per Day		
GINI	Positive	1.93**
TRADE (%GDP)	Negative	-0.57**
GROSS FIX INV.	Negative	-0.75**
GDP PER CAPITA	Negative	-0.73**
AGRICULTURE FREE GDP PER CAPITA	Negative	
CONSTANT		5.59**
R-SQUARE		0.64
CHI-SQUARE OR F-STATISTIC		68.05**
SAMPLE SIZE		48
Inequality Equation		
GINI		
RURAL POPULATION	Negative	-0.17**
POPULATION GROWTH	Positive	0.22**

CONSTANT		4.33**
R-SQUARE		0.21
CHI-SQUARE		13.25**
SAMPLE SIZE		48
GDP per Capita Equation		
GDP PER CAPITA		
VA/LAND	Positive	0.56**
LAND/LABOUR	Positive	0.57**
EXPENDITURE ON PRIMARY EDUCATION	Positive	-0.20**
CONSTANT		4.27**
R-SQUARE		0.80
CHI-SQUARE		167.94**
SAMPLE SIZE		48
Yield Equation		
VA/LAND		
R&D/LAND	Positive	0.62**
FERTILISER/LAND	Positive	0.86**
ILLITERACY	Negative	-0.35**
CONSTANT		14.02**
R-SQUARE		0.73
CHI-SQUARE		164.94**
SAMPLE SIZE		48

** significant at the 5% level, two-tailed test.

6.1 Calculating the Rate of return to Investment in R&D

The rate of return calculation is normally based on the estimated coefficients of R&D in explaining productivity (Lu, Cline and Quance, 1979, Davis, 1981, Thirtle and Bottomley, 1989, Alston, Norton and Pardey, 1995), or on the R&D coefficient in the estimation of the dual profit function (Jayne et al., 1994). The coefficient is an output elasticities relating R&D expenditures to the yield, but it can be converted to a marginal value products to allow calculation of the marginal internal rate of return (MIRR) to R&D. If the elasticity is • then

$$\alpha_i = [\frac{\partial LnYield_t}{\partial LnR\&D_{t-i}}] = [\frac{\partial Yield_t}{\partial R\&D_{t-i}}][\frac{\overline{R\&D_{t-i}}}{\overline{Yield_t}}] \qquad \textbf{(15)}$$

where $\overline{RD}$ *and* $\overline{Yield}$

can be approximated by the mean values, so that the marginal product of R&D in year i is

$$MP_{RD_{t-i}} = \alpha_i [\frac{\overline{Yield_t}}{\overline{RD_{t-i}}}] \qquad \textbf{(16)}$$

However, (16) is still in terms of the effect of R&D on yield, and for a rate of return to be calculated, the change in productivity must be converted into a value. Thus, both sides of equation (16) are multiplied by a conversion factor that is the change in the value of output (• V), that results from a change in the yield (• Y). This gives a value marginal product, with both R&D and value added being measured at constant prices

$$VMP_{RD_{t-i}} = \alpha_i [\frac{\overline{Yield_t}}{\overline{RD_{t-i}}}] [\frac{\Delta V}{\Delta Y}] \qquad \textbf{(17)}$$

The marginal internal rate of return (MIRR) can be calculated from equation (18)

$$\sum_{i=1}^{n} [\frac{VMP_{t-i}}{(1+r)^i}] -1 = 0 \qquad \textbf{(18)}$$

in which, i is the length of the lag, for each expenditure term, and the MIRR for a one unit change in R&D expenditure is calculated by solving for r. If the lag between R&D expenditures and yield increases is assumed to be five years, the MIRR is estimated at 52%, which is in line with the relatively high values normally found in the literature. If the lag were six years, the MIRR falls to 42% and if it were seven years, to 35%.

In this case it is possible to conduct a simple check on the MIRR calculation Data for the yields of all cereals, for the LDCs, from the FAO Agrostat database allow the rate of growth to be calculated for the period 1960-1998. The annual average growth rate of yields is 2.4% and Alston et al. (2000) report that in 1991, the agricultural research expenditures of the LDCs was $8 billion. Thus, the cost of increasing LDC yields by one percent is $3.3 billion 1991 dollars. This is the cost of reducing the percentage of the LDC populations living on less than $1 per day by between 0.4% (the lowest estimate, in model 7) and 0.65% (the single equation result in model 3) and at the same time increasing LDC GDP per capita by between 0.55% (model 7) and 0.58% Model 8).

Both these effects can be calculated from the estimated elasticities and the data. Alston et al (2000) do not specify the countries that are included in their group of LDCs, that are spent $8 billion on R&D in 1991. If we presume they are the same group the FAO puts in this category, their GDP in 1990 was about $2,853.9 billion (from World Bank, 2001). If a 1% increase in yields increases GDP per capita by 0.565% (the average of the two very similar figures), the gain in GDP for the LDC group is $16.13 billion and the cost was $3.3 billion.

Whilst this looks attractive, the lags between R&D expenditures and their impact on productivity are long. The lag is not known and discounting is powerful as Table 24 shows. If the output gains were in the next the rate of return would be a massive 389%, but if the lag is ten years the payoff is only 17%. This shows how important it is for LDC national agricultural R&D systems to concentrate much of their efforts on applied and adaptive research, which has a relatively rapid payoff. South Africa has been successful in this respect: Khatri and Thirtle (2000) found the peak lag to be at less than three years and as a result the minimum rate of return estimated was 77%. Most cross country studies of LDCs, such as Thirtle, Hadley and Townsend (1995) find that a peak lag of about five years fits the data. If this is about right, the rate of return to agricultural research in the LDCs should be about 37%, which is entirely satisfactory, and despite the different and somewhat crude method of calculation, is right in the range that has been found in a large number of studies.

Table 24: Marginal Internal Rate of Return to R&D, by Length of Lag from Expenditure to Benefit

Years of Lag	10	9	8	7	6	5	4	3	2	1
Rate of Return	17%	19%	22%	25%	30%	37%	49%	70%	121%	389%

The MIRR result, from Model 9, of 52% is thus corroborated and this can be also be used to calculate the gain in GDP per capita that results from a 1% increase in R&D expenditures. In Table 23, the elasticity of GDP per capita with respect to yield is 0.56 and the elasticity of yield with respect to R&D is 0.62. Thus, the increase in GDP per capita that results from a 1% increase in R&D expenditures is 0.56 x 0.62 = 0.34%. If this sample is taken to be representative of the LDCs, then the GDP per capita gain falls from $16.13 billion to 9.7 billion. The rate of return to R&D, calculated from the GDP gain falls to 24%, but this is still a decent return.

The payoff to agricultural R&D in terms of the number of people that it can shift out of the less than $1 per day poverty bracket can be similarly calculated. Table 25 first shows the percentages of the regional populations living on less than $1 per day, followed by the populations. The product of the two is the number of people living on less than $1 per day. Then if a 1% increase in yields reduces this count by 0.4%, the total number of people escaping from the less than $1 per day bracket is 4.8 million. If the alternative estimate of 0.65% is used, this total rises to nearly 7.8 million.

Table 25: Reduction in the Number of People on Less than $1 per Day from a 1% Increase in Yield

Region	% in $1 poverty	Population	# in $1 poverty	0.4% Reduction	0.65% Reduction	0.25% Reduction
East Asia	15.3	1836.9	281.0457	1.124	1.827	0.703
Latin America	15.6	509.2	79.4352	0.318	0.516	0.198
Middle East & N Africa	1.9	290.9	5.5271	0.022	0.036	0.014
SE Asia	40	1329.3	531.72	2.127	3.456	1.329
SSA	46.3	642.3	297.3849	1.189	1.933	0.698
Total Number of People Moving Out of $1 per Day Poverty, Millions				4.780	7.768	2.942

Thus, only this small proportion of the estimated (World Bank web site) 1, 200 million people living on less than $1 per day are removed from this category when $3.3 billion is spent on agricultural research. This does not sound nearly as promising as the gains in GDP, but we fear it is correct. To put it in perspective, the models suggest that a 1% increase in GDP per capita has hardly any more impact than a 1% increase in yields: the model 3 result was 0.59%, which is actually lower and models 7 and 8 gave 0.72 and 0.76.

The alternative to these simple estimates is to fully use the results of Model 9, which adds sophistication, but is les direct, as it accumulates the effects of three equations. In Table 23, the elasticity of $1 per day poverty with respect to GDP per capita is –0.73; the elasticity of GDP per capita with respect to yield is 0.56 and the elasticity of yield with respect to R&D is 0.62. Thus, the reduction in the percentage of the population living on less than $1 per day that results from a 1% increase in R&D expenditures is –0.73 x 0.56 x 0.62 = -0.25%. If this sample is taken b be representative of the LDCs, then the reduction in the numbers living on less than $1 per day falls to 2.94 million.

These are not spectacular reductions, but the evidence suggests these effects are in the right range. The data on the World Bank's Poverty Net web site show that from 1990 to 1998 the number of people living on less than $1 per day fell from 1,276.4 million to 1,174.9 million. This is a reduction of 8% over eight years, so the decline is only about 1% per year. Viewed in this way, agricultural research may well be a useful and cost effective instrument for reducing poverty, but it is a pretty blunt one as many have argued (Alston et al., 1995). Perhaps getting benefits to the poorest segment of the population is always the most difficult task.

REFERENCES

Alston J, Chan-Kang C, Marra M, Pardey P and Wyatt T (2000) *A meta-Analysis of Rates of Return to Agricultural R&D: Ex Pede Herculem?* IFPRI Research Report 113.

Alston, J.M., Norton, G.W. and Pardey, P.G. (1995). *Science Under Scarcity: Principles and Practice for Agricultural Research Evaluation and Priority Setting*, Cornell University Press, Ithica.

Bruno, Michael, Martin Ravallion and Lyn Squire (1996). *Equity and growth in developing countries: old and new perspectives on the policy issues*. World Bank working paper No 1563

Chen S, Datt G and Ravallion M (1994), Is Poverty Increasing in the Developing World?, *Review of Income and Wealth*, **40**,(4): 359-76.

Chen S and M Ravallion (2000), *How did the poorest fare in the 1990's*, World Bank, Washington, DC.

Davis, J. A. (1981). Comparison of Procedures for Estimating Returns to Research Using Production Functions, *Australian Journal of Agricultural Economics*, **25**, 60-72.

Fan, S, P Hazell and T Haque (2000), Targeting public investments by agro-ecological zone to achieve growth and poverty alleviation goals in rural India, *Food Policy*, **25**, 411-28.

Fan S, P Hazell and S Thorat S (1999), Linkages between Government spending, growth, and poverty in rural India, IFPRI research report 110

Hanmer L and Nashchold F (2000), *Attaining the International Development Targets: Will Growth be Enough?*, working paper

Hayami Y and Ruttan V (1986), *Agricultural Development: An International Perspective*, Johns Hopkins University Press, Baltimore

Jayne, T.S., Khatri, Y., Thirtle, C. and Reardon, T. (1994). A Profit Function Approach to Productivity Change in Zimbabwean Agriculture: Implications for Research and Policy, *American Journal of Agricultural Economics*, **76**, 613-18.

Khatri Y and Thirtle C, (2000). Cointegration and modelling the length and shape of the research lag Chapter 10 in Colin Thirtle, Johan Van Zyl and Nick Vink (Eds), *South African Agriculture at the Crossroads*, Macmillan, Basingstoke.

Lu, Y., Cline, P. and Quance, L. (1979). *Prospects for Productivity Growth in US Agriculture*, Agriculture Economic Report, No.435, ESCS, USDA, Washington, DC.

Maxwell S (2001) Innovative and Important, yes, but also Instrumental and Incomplete: the treatment of redistribution in the new poverty agenda, ODI, London

Pardey, P and Roseboom, H (1989) ISNAR Agricultural Research Indicator Series: A Global Database on National Agricultural Research Systems, Cambridge University Press

Ravallion M (1997), Can High Inequality Developing Countries Escape Absolute Poverty? *Economics letters* **56:** 51-57

Ravallion M and G Datt (1999), When is growth pro-poor? Evidence from the diverse experience of India's states World Bank Policy Research Working Paper Series, 2263, Studenmund, A. H. (1997), 'Using Econometrics: A Practical Guide', 3rd Edition, New York: Addison-Wesley.

Thirtle, C., Producer Funding of R&D: Productivity and the Returns to R&D in British Sugar, 1954-93, *Journal of Agricultural Economics*, Vol.50, No.3, September, 1999, pp.450-67.

Thirtle, C. and Bottomley, P. (1989). The Rate of Return to Public Sector Agricultural R&D in the UK, 1965-80, *Applied Economics*, **21**, 1063-86.

Thirtle C, Hadley D and Townsend.R (1995). A Multilateral Malmquist Productivity Index Approach to explaining Agricultural Growth in Sub-Saharan Africa, *Development Policy Review*, **13**: 323-348.

USDA (2001), http://www.ers.usda.gov/Briefing/GlobalResources/Questions/slide4.htm

World Bank (1999), *World Development Report 1998/99*, World Bank/Oxford University Press

World Bank (2000) World Development Indicators, CD, World Bank, Washington DC.

World Bank (2001), *World Development Report 2000/01 Attacking Poverty*, World Bank/ Oxford University Press