GROWTH AND THE POOR: A COMMENT ON DOLLAR AND KRAAY

Malte Lübker
Graham Smith
John Weeks*

October 2000

Centre for Development Policy and Research
& The Department of Economics
School of Oriental & African Studies
University of London
Thornhaugh Street
London
WC1H 0XG

cdpr@soas.ac.uk

The authors are, respectively, Researcher, Senior Lecturer and Professor at the School of Oriental and African Studies.

* Corresponding author. Tel.: +44 20 7898 4487; fax: +44 20 7898 4559; Email: jw10@soas.ac.uk.

Growth and the Poor: A Comment on Dollar and Kraay*

Abstract: In a recent paper Dollar and Kraay come to sweeping conclusions about economic growth and the poor. On the basis of the empirical work they assert that standard World Bank and IMF policy packages are good for the poor. This critique demonstrates that (i) the empirical work is based on theoretically unsound equations; (ii) the data are seriously flawed; and (iii) the policy variables are not defined appropriately and are tested in an inconsistent manner. These problems imply that the policy conclusions of the authors are unsafe.

1 INTRODUCTION

The recent exercise by Dollar and Kraay (2000) that implicitly excludes asset and income redistribution from the policy agenda for poverty reduction might have gone largely unnoticed had it not apparently been associated with the well-publicised controversy over the *World Development Report 2000*.[1] Their major finding of statistical significance was that across countries the income of the lowest quintile varies as overall average income varies with an elasticity of unity;[2] *i.e.*, that this quintile's income tended to remain stable in the process of growth. The additional statistical exercises involving various policy indicators, which were merely hypothesis tests,[3] proved to be non-significant in most cases. Thus, the most important result of the paper could hardly be called a 'finding' at all, since the stability of income distributions is both well established and well known.[4] How 'good' a unit elasticity is for the poor is a matter of opinion.[5]

None-the-less, it is instructive to deconstruct the Dollar-Kraay exercises: first, because the fervour of the paper ('growth *is* good for the people', emphasis in original) alone calls for consideration of the anti-distributional policy perspective it advocates; second, because a review of the empirical procedure provides a tutorial in how statistical work should and should not be done; and, third, because the authors' interpretation of the results demonstrates that the absence of theoretical modelling provides fertile ground for ideological inferences.

The Dollar and Kraay paper can be summarised as follows: (i) on average across countries and over time, growth is distribution neutral; thus (ii) any factor which increases the growth rate is good for the poor; (iii) World Bank and IMF policy packages increase the growth rate; therefore, (iv) these policy packages should be the core of poverty reduction strategies. Though it is hardly controversial, they seek to establish the first point through a dubious statistical framework. The second point is a *non-sequitur*; what is true on average is not necessarily true for the components of an average, a point we establish below. The third point they establish from a biased reading of the growth literature. Since they do not establish points two and three, the sweeping neoliberal policy conclusion does not follow.

Our critique is elaborated as follows. First, we consider the relationship between growth and distribution, to indicate that a bit of common sense can be enlightening prior to launching into complex empirical exercises. Second, we treat the limitations of the data set used by Dollar and Kraay. Third, we analyse the Dollar-Kraay statistical results, and cast doubt on the robustness of their basic empirical finding, that growth is distribution neutral. We also consider statistical inference as practiced by Dollar and Kraay.

2 INCOME DISTRIBUTION AND COMMONSENSE

In what follows, we adhere to the approach of Dollar and Kay and restrict ourselves to the size distribution of income, across households or individuals. In all fields of economics, there is a danger that ideology will disguise itself as objective analysis, and treatments of asset and income distribution suffer perhaps most of all from this malady. This tendency results in part because asset and income distributions arise from extremely complex economic and social interactions. This complexity fosters a temptation to treat these interactions in a partial rather than a general framework. Therefore, in this field one must be especially vigilant to adhere to professional rules of rigour. At the outset rigour requires clarification of concepts and basic causality. The need for rigour and clarity is demonstrated by the ambiguities in the basic Dollar-Kraay estimating equation:

$$\ln Y_{ct}^{p} = \alpha_0 + \alpha_1 \ln Y_{ct} + \alpha_2' X_{ct} + \mu_c + \epsilon_{ct} \qquad (1)$$

where $\ln Y_{ct}^{p}$ is the logarithm of per capita income of the bottom or first quintile (the 'poor'), $\ln Y_{ct}$ is the logarithm of per capita average income, X_{ct} is a vector of additional explanatory variables, $\mu_c + \epsilon_{ct}$ is the composite error term, *c* refers to countries and *t* to years.

While they have a lengthy discussion of various aspects and potential problems arising from this estimating equation (see Section 3, 'Empirical Strategy'), the authors do not provide the information that modelling procedure requires as a basic minimum. It is obvious that equation (1) is a reduced form. Average income does not act directly on the income of a component of itself, if one accepts the general rule that an average is made up of its parts. Since it must be a reduced form, the following questions require answering in order to interpret the statistical results: (i) from what model of

growth is it generated? (ii) is this an equilibrium model, or are the observations treated as partially adjusted to equilibrium values? and (iii) what assumption is made about the rate of return on assets and payment for services across quintiles? Without answers to these questions the coefficients cannot be interpreted, because their predicted values are not specified.

Not withstanding the ambiguousness of their estimating equation, Dollar and Kraay focus on the elasticity between the two per capita income variables. If this is positive, the 'poor' benefit by definition. Dollar and Kraay put forward the specific null hypothesis that the coefficient is not significantly different from one; *i.e.*, that the income of the bottom quintile grows at the same rate as mean income. This is not a very interesting hypothesis to test. When averaging across countries, commonsense would predict growth to be distribution-neutral. Income distributions arise from political, social and economic factors that, relatively to income flows, have the status of parameters or 'initial conditions'. In the absence of major policy changes or social upheaval such as war, one would expect the determinants of distributions to change slowly. The interesting exercise would be, by using some objective criteria, to separate out those countries that have undergone major policy change or conflicts, and test these for the distributional neutrality of growth. Combining a large number of countries with stable policies, institutions, and politics with a small number racked by instability has the effect of averaging away the outcomes in the latter.

It is tempting to formulate the hypothesis in the simple Dollar-Kraay manner, but the estimating equation suffers from a number of analytical weaknesses, which arise from the lack of an explicit theoretical model. First, the equation is not specified correctly, because the dependent variable is a component of the explanatory variable. By definition,

$$Y_{ct} \equiv p_1 Y_{ct}^p + p_2 Y_{2ct} + \ldots + p_5 Y_{5ct} \quad (2a)$$

where p_i is the share of households or individuals in each successive fifth of the distribution. Since by definition these are all equal to .2, one can write,

$$Y_{ct} \equiv .2[Y_{ct}^p + Y_{2ct} + \ldots + Y_{5ct}]. \quad (2b)$$

If Y_{ct}^p is a function of X_{ct} then Y_{ct} is also a function of X_{ct}. We take up this problem below. Here, we drop X_{ct} and consider only the two income variables. By definition, the income of the bottom quintile, Y_{ct}^p, is:

$$.2 Y_{ct}^p \equiv Y_{ct} - .2[Y_{2ct} + \ldots + Y_{5ct}]. \quad (2c)$$

From the data set supplied by Dollar and Kraay, we can calculate averages to obtain,

$$Y_{ct} - .2[Y_{2ct} + \ldots + Y_{5ct}] = .063\, Y_{ct}. \quad (2d)$$

It follows that

$$Y_{ct}^p = .315\, Y_{ct}. \quad (2e)$$

The logarithmic form is

$$y_{ct}^p = -1.155 + 1.000 y_{ct} \quad (2f)$$

where lowercase variables represent logarithms.

This is suspiciously similar to the Dollar-Kraay result when they estimate a simple regression between y_{ct}^p and y_{ct} using three methods: ordinary least squares, instrumental variables, and a system estimator.[6] With the last two methods their estimated coefficients are not significantly different from those predicted by the elaboration of the definition of per capita income (expression 2f). Our own OLS estimates on the data set supplied, which does not include some of the Dollar and Kraay observations, confirms the prediction of the definition (standard errors in parenthesis),

$$y_{ct}^{p} = \underset{(.366)}{-1.364} + \underset{(.044)}{1.016}\,y_{ct} \qquad R^2 = .83$$

For all regression analysis, the coefficients must be such that the linear estimation passes through the average value of the variables. Expressions (2a) through (2f) demonstrate more than that: namely, from a definition and the particular data set to be used for regression analysis, they predict a unique pair of coefficients, and the regression-estimated coefficients prove to be not statistically different. This might account for the rather high R^2 the authors obtain.

That the basic Dollar-Kraay statistical outcome is very close to what one would deduce from manipulating the definition of average per capita income highlights a second fundamental problem of their estimating equation: there is no obvious causal mechanism by which average income would transmit its increase to one of its components. On the contrary, since averages are by definition made up of their components, numerical determination should run the other way.

In this intensely controversial field, few generalisations are accepted by all. One of these few is that the most important determinant of the primary distribution of income at any point in time is the distribution of assets; and changes in the primary distribution of income result from changes in the distribution of those assets. In the abstract, one can conceive of a technically determined primary distribution resulting from a matrix of assets by households and an associated vector of returns on these assets. Public sector interventions via fiscal instruments produce the secondary distribution of income, which in most countries is more equitable than the primary distribution, but need not be. The observed or actual distribution of income differs from this abstract secondary distribution due to conjunctural factors (*e.g.* cyclical unemployment), relative bargaining power among social classes, and a range of social and political influences (such as ethnic discrimination).

The question Dollar and Kraay ask, does the income of the 'poor' rise proportionately, less than proportionately, or more than proportionately to average income, cannot be answered in a meaningful manner by regressing a component of average per capita income on that average itself. There is no obvious causality in this regression. One should first specify the determinants of the income of the bottom quintile in a model that incorporates a theory of income growth. One possible approach would be to specify the income of the bottom quintile as a function of the assets of that quintile, some measure of the redistributive effect of fiscal measures, and mediating social and political factors. The next step would be to model a credible mechanism by which average per capita income is a determinant of some or all of these variables.

It is quite likely that ordinary least squares, or even the more complicated techniques used by Dollar and Kraay are not appropriate to the task. By definition, all income quintiles vary simultaneously. Once average per capita income is introduced as an argument for the income of one quintile, it is implicitly impacting on all other quintiles. Consider, for example, a reduced form equation in which a theoretically consistent relationship between average per capita income and the income of the lowest quintile were modelled. Since the elasticity of average income with respect to itself must be unity, the weighted elasticities across all fifths of the distribution must also be unity (the relative loss or gain of one quintile must be the relative gain or loss of the others). Thus, one would need to constrain the five coefficients on average per capita income across quintiles to be equal to one.[7] Estimating only one quintile generates a biased coefficient.

The point of our discussion is that economists have established a standard method for carrying out empirical work. First, one specifies a model of causation on

the basis of certain a priori principles, and formulates the model in a consistent manner that clearly differentiates between dependent and independent variables. The resultant model typically provides predictions as to the signs and reasonable range of values for the coefficients. Second, that model is tested with relevant data, using a statistical procedure consistent with the model. Third, the results are rigorously evaluated in light of statistical theory. In this section we have demonstrated that the Dollar-Kraay regressions do not conform to the first two rules of procedure: there is no theoretical framework, and the statistical procedure is inappropriate.

3 THE DATA AND ADJUSTMENTS TO THE DATA

In any econometric investigation the data must be adequate for the purpose. The Dollar-Kraay data set does not pass this test. Further, the authors' attempts to correct and adjust these data for inconsistencies are either inconsistent themselves or unlikely to improve matters.

Dollar and Kraay rely on 682 so-called high-quality observations from 108 countries, drawn from the Deininger and Squire (1996) data set. To this they add 143 observations from Lundberg and Squire (1999) and the *World Development Reports* for 1999 and 2000. Since some of the data from the last source were unpublished at the time this paper was written, we had access only to a slightly smaller sample with 759 observations from a bit over 100 countries. All of the data measure 'inequality', but the definitions vary among and within countries. The Dollar and Kraay compilation includes observations for households (in 372 cases), for individuals (365), and the unit is unknown in 22 cases. Some observations refer to income (591) and others to expenditure (168). Of the income-based observations, some report net of taxes (174), some gross (370), and forty-seven observations are not identified with

respect to taxes. These inconsistencies and unknowns compromise comparability and bias any statistical results in an unpredictable manner. Household data hide intra-household inequality, income saved is not included as expenditure, and progressive taxes mean that inequality before taxes is greater than after taxes.[8]

Facing these inconsistencies, Dollar and Kraay adjust the data to correct for differences in definition. In doing so, they provide a guide to how adjustment should not be done. Their procedure is to assign dummy variables to different combinations of coverage, with household-based pre-tax income as the omitted category ('benchmark'). After some intermediate steps, this yields a set of 'adjustment coefficients' for the five deviations from the benchmark category (see Table 1).[9] Consider the adjustments for the income share of the bottom quintile. Four of the five estimated coefficients are not significantly different from zero, which, according to rules of statistics calls for no adjustment. The only significant adjuster is for 'unknown income'. The omitted category is income before taxes. If an 'unknown' observation is in fact pre-tax, then no adjustment should be made; if the unknown observation is post-tax, then, again, no adjustment is called for, because the coefficient on net income is not significantly different from zero. Thus, commonsense tells us that there is no point in using the coefficient on 'unknown' to adjust data (it is redundant). The same point applies to 'household unit'. If unknown observations fall into two mutually exclusive categories, it is a logical contradiction to adjust all of them with the same coefficient.

As part of their exercise, Dollar and Kraay must also adjust Gini coefficients. The same inconsistencies arise. The puzzling thing is that the coefficients for the 'net income' and 'unknown income' categories are both significant at the .05 level. Hence, the regression states with confidence that the Gini must be corrected downwards when

there is a clear deviation from the benchmark, but upwards when it is unknown whether there is such a deviation or not. Only one of the two can be correct. One can conclude from these contradictory results that either the method used to establish them is invalid, the data are unreliable, or both.[10] Commonsense suggests that the entire adjustment exercise is hopelessly flawed, most obviously so for the attempt to adjust for net and gross income. Across countries, the rate of taxation and progressivity of rates varies greatly. The proposition that one could produce a meaningful set of adjustment coefficients, applicable to all countries, implicitly assumes that all tax systems have the same impact.[11]

To state the problem simply, the adjustments by use of dummy variables violate the most basic rules of logic. In both cases, the dummies for unknown observations refer to binary categories. An observation is or is not category A; if it is not, it is category B. It is logically impossible for a coefficient to both adjust a category A observation into category B, and at the same time leave B as a category B variable.

In summary, the raw data used by Dollar and Kraay are far from perfect, which they concede. Adjustments the authors make are either not justified by statistical procedure or are implausible. Logic tells one that the adjustments do not improve data reliability, and one cannot rule out the possibility that they made the data worse.

4 STATISTICAL ESTIMATES OF GROWTH AND THE 'POOR'

In this section we consider the controversial conclusions Dollar and Kraay draw from their empirical work. The most fundamental of these derive from their estimation of the so-called elasticity of the average income of the bottom quintile with respect to

overall average income. They make a number of assertions about this statistic that do not bear close scrutiny.

Table 2 reports their OLS regression using pooled country-year observations. The point estimate of the elasticity is 1.06 and significantly greater than one, implying growth is better for the bottom twenty percent than for the eighty percent above. For reasons that are unclear, they consider this an inaccurate estimate and attribute it to the unreliability of a simple OLS regression (Dollar and Kraay 2000, p.19). Using a system estimator they obtain a point elasticity of 1.046, that is not significantly different from one. The unit elasticity hypothesis is not rejected in several variants of the basic specification, that include regional dummies, observations split between high or low income countries, high or low growth, positive and negative growth, and observations segregated by decades (Dollar & Kraay 2000, p 36). On the basis of these variations, they conclude that unit elasticity is a near-universal outcome among and within countries.

The composition of the data set does not seem appropriate to us. First, and for reasons previously discussed, we consider it counter-productive to adjust for differences in measurement concepts. Second, we consider it unsound to estimate missing quintile shares.[12] Third, the sample includes observations from twenty-two developed economies including the United States, Japan, and Germany. As discussed above, income distributions result in part from the institutions of society and structure of the economy. Development is more than the growth of income per capita, for involves fundamental social and economic transitions.[13] For reasons of analytical caution alone, the data from the developed and developing countries should be separated. As dubious as use of the entire sample is, we shall show that statistical

exercises become even more problematical when the sample is divided, as we now demonstrate.[14]

We begin with the entire set of observations, and obtain a point estimate of the growth elasticity of 1.016, which is not significantly different from one (Table 2). Next, we exclude the developed country observations (leaving 223), and OLS estimation results in a growth elasticity of 0.923, again, not significantly different from one. Unit elasticity appears to be confirmed (though it would be reckless to conclude that the distribution dynamics are the same for developing and developed countries).

Dollar and Kraay do not restrict themselves to the entire data set, but divide it according to several criteria. The most analytically interesting of these is the division by level of per capita income, though they are not explicit as to what this division tests. One presumes that it seeks to reveal if there are different distributional dynamics associated with level of development. In practice what it demonstrates is how results can be biased by the division of a sample.

While Dollar and Kraay divided the sample on the basis of the overall average per capita income of countries, we ordered the developing country observations according to *increasing income of the bottom twenty percent.* These are divided into three sub-sets: the low-income poor (bottom quintiles with average per capita income less than US$ 500), the middle-income poor (quintiles in the range $500-1,500), and the high income poor (greater than $1,500), with sample sizes of 84, 101 and 38, respectively. We do not make this division because we believe that the data set will produce sensible statistical results. We do it to demonstrate that apparently sensible, but ad hoc, divisions can result in unanticipated statistical bias.

When OLS regressions are carried out on our three subsets, the unit elasticity hypothesis collapses (see Table 2, last three rows). The regressions produce estimated elasticities of .48 (low income), .36 (middle income), and .35 (high income), all significantly less than one. With the collapse of the elasticity hypothesis, the R^2s also collapse, into the .2 to .3 range.

We can now consider the reason for the radical difference between the results (i) for whole sample and for the parts, and (ii) for the Dollar and Kraay division by per capita incomes and ours by the income of the poor division. The observations are shown in Figures 1-4. If one inspects the full sample (Figure 1), there is a scatter along a line of unit elasticity, with the compactness of the scatter exaggerated by the scale of the graph. For the low-income poor, those observations for per capita incomes of US$ 500 or less for the bottom quintile, the relationship, if there is one, appears non-linear, with several extreme values at low incomes (Figure 2). For the middle-income poor ($\text{US\$ } 500 < Y_{ct}^{p} < \text{US\$ } 1{,}500$) the scatter is diffuse, indeed (Figure 3). In the case of the high-income poor ($Y_{ct}^{p} > \text{US\$ } 1{,}500$), about two-thirds of the observations stretch along a line of roughly unit elasticity, but the remaining third seem randomly scattered to the right of that line (Figure 4).

Figure 5 demonstrates why the outcomes differ. Assume, as is the case (see Figure 1), that the all observations lie loosely around a line of unity elasticity (the solid, 45 degree line. The Dollar and Kraay criterion for dividing the sample draws a vertical line at some Y*; our criterion involves a horizontal line. The former division (Dollar and Kraay's) eliminates extreme values such as point X_a, which will tend to rotate the estimated line counter-clockwise. It is obvious that this favours the unity elasticity hypothesis. On the other hand, our division, by the income of the lowest

quintile, eliminates extreme values such as X_b, and tends to generate a lower elasticity.

We do not consider our division any more justified than Dollar and Kraay's, since neither is based on theoretically justified *ex ante* criteria. In the absence of theory, all divisions of the sample must be treated as arbitrary, even apparently innocuous ones such as the authors' division by decades.[15]

The unit elasticity hypothesis does not appear robust across different divisions and combinations of the sample. Further, it would seem that unit elasticity is an artefact of a large data set seriously flawed by measurement errors. It collapses under close inspection. At this point we confess to an ideological bias in favour of redistribution for greater equality. As a result of this predilection, it is an obvious temptation to make much of the elasticities generated by our sub-division of the developing country sample of observations, for they could be interpreted to indicate that the 'poor' would gain far less than the households in the upper eighty percent of the distribution. With some reluctance, but considerable conviction, we resist this temptation. Our regression exercise does not demonstrate that 'growth is bad for the poor,' any more than the Dollar-Kraay ones demonstrate that 'growth is good for the poor'. Rather, our regressions and the scatter diagrams indicate that the Dollar-Kraay model and the data set used to test it produce meaningless results.

The income distribution of any country and the changes in that distribution are extremely complex phenomena, affected by history, political power, level of development, and many other factors. An attempt to provide a statistical summary of these phenomena for a single country with a few variables is profoundly unsound and non-scientific. To attempt to provide a universal summary across scores of countries is pointless.

5 POLICY VARIABLES AND THE POOR

After their travails with the elasticity between average income and the income of the poor, Dollar and Kraay turn to the impact of four main policy indicators, inflation, government consumption, exports and imports relative to GDP, and rule of law. To do this they expand their basic equation using a system estimator. When all four variables are included, only one is significant. This exception is inflation, with a *t* statistic of 1.75, significant at the .05 level on a one-tailed test. If the three insignificant variables are omitted, the coefficient on inflation ceases to be significantly different from zero. In any case, as Easterly and Fischer (2000, p.2) argue, there are no *a priori* reasons to predict the sign of the coefficient on inflation, and so one should insist on a two-tailed test.[16]

These so-called policy indicators, typical of empirical work by World Bank professionals,[17] suffer from a number of serious problems. First, inflation and trade as a portion of GDP are not policy variables, but policy outcomes. The distinction is fundamental to any consideration of the effects of public sector actions. A policy is something that governments formulate to achieve a desired outcome, and implement this policy by use of instruments. A policy many be sound (i.e., reduction of inflation), the instruments to achieve it appropriate (an increase in nominal interest rates and tax rates), but the outcome a failure due to unforeseen factors, such as a change in the terms of trade. No conclusion about the impact of policy on any variable can be drawn from a cross-country regression using inflation as an indicator, for it would include countries that pursued 'good' macro policy (but were unlucky), and ones that fecklessly allowed the money supply to grow out of control.

The same criticism applies, even stronger, to trade as a portion of GDP. This is obviously an outcome indicator, which is only partially determined by policy. It is

strongly influenced by changes in the external terms of trade and the elasticity of non-tradable production with respect to tradable production, among other non-policy factors. Even more important, it is well documented that trade shares are inversely correlated with measures of size of country. The governments of two countries could have the same broad trade policy, yet quite different outcomes in terms of the share of trade in national income, due to size of the economy, natural resource endowments, and level of development. Using their openness index, Dollar and Kraay might conclude that small, mineral-rich countries have a remarkable record for pursuing 'good' trade policy. In place of these outcomes, the relevant policy variables would be, for example, real interest rates and monetary emission (for inflation), and the nominal exchange rate (for trade).[18]

Whatever the appropriateness of the variables, there is a serious theoretical inconsistency with how they enter the model. Their impacts on both average per capita income and the income of the poor are treated as monotonic, which implies the more they decrease (inflation) and increase (trade), the higher will be the level or growth of incomes. This defies both theory and commonsense, for there must be some value of each that is optimal, such that incomes decrease each side of that value. Given the wide range of values for each variable in the data set, one cannot make the pragmatic argument that a monotonic relationship holds over the observed range, though not for all values. To state the matter simply, it is not credible that for every country in the sample lower inflation and more trade would raise incomes (Mosley et al. 1995, p.1463).

A further problem with the inclusion of the 'policy' variables would seem to undermine the entire enterprise. The impact of the policy variables is estimated with two equations. In one, overall per capita income is the dependent variable and the

policy indicators are the arguments. The authors claim that this equation measured the 'growth effect' of the 'policy' indicators.[19] A second equation, with the income of the bottom quintile as the dependent variable, employs the same policy variables as arguments. This they call the 'distribution effect'.[20] On the basis of these two equations, the authors reach quite strong conclusions, such as the following,

> Reducing government consumption and stabilizing inflation are examples of policies that are 'super-pro-poor'. Not only do both of these raise overall incomes, but they appear to have an additional positive effect on the distribution of income, further increasing incomes of the poor (Dollar& Kraay 2000, pp. 5-6).

It is worth inspecting the two equations in detail. To simplify the algebra, lags are ignored, the error terms have been omitted, and the 'policy' indicators, X_{ct}, are treated as a composite variable. None of these affect our point.

$$y_{ct}^{p} = \alpha_0 + \alpha_1 y_{ct} + \alpha_2 X_{ct} \tag{3}$$

and

$$y_{ct} = \beta_0 + \beta_1 X_{ct} \tag{4}$$

It is immediately obvious that the two equations are not independent of each other, and should not be estimated separately. Straightforward substitution yields,

$$y_{ct} = (\alpha_0 + \alpha_1 \beta_0) + (\alpha_1 \beta_1 + \alpha_2) X_{ct} \tag{5}$$

In this, the logically consistent form to estimate the relationship between the 'policy' variables and the income of the 'poor', overall per capita income drops out, and the income of the poor is a function of the policy variables only. The elasticity between the two incomes α_1 is found in both the constant term and in the coefficient on X_{ct} and it cannot be extracted. If equation 2 is the appropriate estimator, then equation 3

should not be estimated; or, if equation 3 is the appropriate growth model,[21] then equation 2 should not be estimated.

From this discussion of 'policy' variables, we reach the following conclusions: (i) the data set is too corrupted by measurement errors to be used for multivariate analysis; (ii) the policy variables are inappropriate, either because they measure outcomes not policies, or because they suffer from the ambiguities of ordinality;[22] and (iii) were the data satisfactory and the variables true indicators of policy, the model is estimated in a logically inconsistent form. Any one of these problems would undermine the conclusions of Dollar and Kraay; taken together, they render the conclusions useless at best and misleading at worst.

6 FINAL COMMENTS

Early in their paper, the authors entreat the reader not to misinterpret them 'as arguing that growth is all that is needed to improve the lives of the poor'. We do not want to be misinterpreted as accusing them of holding that position. Quite the contrary, if this were their position, it would be relatively benign, for that opinion has a long and distinguished pedigree. Our objection is to their conclusion allegedly via statistical inference that the World Bank definition of 'good policy' is good for the poor:

> From [our statistics] we conclude that the basic policy package of private property rights, fiscal discipline, macro stability, and openness to trade increase the income of the poor to the same extent that it increases the income of the other households in society.

> . . .[W]e find that contrary to popular myths, standard pro-growth macroeconomic policies are good for the poor as they raise mean incomes with no significant effect on the distribution of income. [p. 6]

These assertions are perhaps the most troublesome aspect of the Dollar and Kraay paper. There exists a massive literature on the effect of World Bank and IMF policy packages on economic growth and social welfare. Not even the most fervent partisan on each side of the debate would claim that the empirical evidence is conclusive with regard to growth, exports, or inflation, much less with respect to the distribution of income. The debate over purely economic aspects of stabilisation and adjustment is fiercely contested; theoretical and empirical analysis of the impact of non-economic factors such as 'rule of law' has hardly begun. To state without regard to a country's economic characteristics, social structure, or political power relations that one specific package of economic and political policies will prove in practice to be universally good for the poor is to move well beyond the boundaries of normative economics into ideology.

When dealing with matters as complex and as politically contentious as poverty reduction, economists must be true to the rigour of their discipline. This involves careful theoretical modelling, with causality clearly specified; a sceptical inspection of data; close adherence to the rules of statistical inference; and caution in making generalisations. Slippage on any of these principles does not move our understanding forward.

[*] The authors thank Aart Kraay, who provided us with a data set that included the vast majority of observations used in the Dollar-Kraay statistical exercises. This paper represents an elaboration of points made in chapter 2 of Lübker (2000).

[1] In mid-2000, Ravi Kanbur resigned as coordinator of WDR 2000. For journalistic and 'insider' reports, see *The Guardian* (London) 15 June and 22 June 2000. For the views of Kanbur on the distribution of income and growth, see Kanbur (1998) and Kanbur and Lustig (1999).

[2] More precisely, the mean income of the lowest quintile has an elasticity of unity with respect to overall mean income, which is strictly equivalent to the relationship between the quintile income share and overall mean income.

[3] That is, they represented tests for correlation without a theoretical framework.

[4] This point is made in detail in Weisbrot*, et. al*. (2000).

[5] See Dagdeviren, van der Hoeven and Weeks (2000), where distribution neutral growth is considered for its poverty reducing potential.

[6] Dollar and Kraay estimate the equation with and without the policy variables; and their approach to those variables is to include all in one specification, then each alone in others. Econometric procedure frowns upon the practice of taking variables in and out of an estimating equation with no theoretical justification for doing so. A researcher's formal model should be the guide to what goes in and what is omitted.

[7] If one divided the distribution into two equal parts in terms of population, the elasticity of mean income for the lower fifty percent with respect to average income would be the mirror image of the elasticity for the upper fifty percent. This is strictly analogous to estimations of production functions with only two factors: the elasticity of substitution between capital and labour must equal the elasticity of substitution between labour and capital. Dollar and Kraay recognize the necessity of constraining coefficients when using their 'system estimator' (Dollar & Kraay 2000, pp. 15-16), but do not comment on the similar need across quintiles. This oversight is despite their comment that estimating equations by quintiles 'has gained some popularity' (*ibid*., p. 17).

[8] These inconsistencies can have serious implications for the robustness of econometric modelling. Atkinson and Brandolini (1999) assessed the quality of the Deininger-Squire dataset for OECD countries, using estimates derived from the *Luxembourg Income Study* (LIS) as a benchmark. They found that the Gini-ranking of the latest available observations for 16 countries differs considerably. The simple correlation between the two sets is only 0.48 (*ibid*., p.7).

[9] It is unnecessary for our purposes to go into full detail of the estimation of the 'adjusters'. However, we note that after obtaining the coefficients in Table 1, Dollar and Kraay regress the logarithms of the Gini coefficient on these and a full set of country dummies. This is repeated for the logarithm of the income share of the bottom quintile. Under the assumptions that the differences are caused by the different concepts of measurement and that the underlying Gini coefficients and quintile shares stay stable, all of the non-household non-gross income data are adjusted by the

estimated coefficients. This assumption may be reasonable for most purposes. However, the question Dollar and Kraay aim to answer in their paper is precisely *whether* Ginis and quintile shares stay stable over time. Hence, adjusting data based on the assumption that they are stable will bias the data in the direction of stability.

[10] The third such contradiction is between 'household unit' and 'unknown unit' for the income share of the bottom quintile.

[11] Atkinson and Brandolini (1999, p.21ff) point out that the difference between gross and net (disposable) income is largely a function of the tax system, but this varies considerably across countries in its progressive impact. Deininger and Squire (1996, p.580) reported that the impact on the Gini ranges from 0.0187 to 0.0566 points in countries included in the LIS database (for a range zero to one), when both net and gross figures are available for the same year. Hence, adjusting the gross income data for Sweden and the United Kingdom with a *common* coefficient is not a sensible thing to do, especially when this coefficient is in turn calculated from within-country changes in Peru, Romania and Guyana, among others.

[12] If the share of the lowest quintile was not known, Dollar and Kraay (2000, p. 9f) used a lognormal approximation to estimate the income share of the lowest quintile from the Gini coefficient. This functional form is notoriously inaccurate at the tails of distributions.

[13] These are the famous three transitions: in population distribution, from a rural to an urban society; in demography from high death rates and birth rates to low birth rates

and death rates; and in production, from primary sectors to secondary and tertiary. Each of these has implications for income distribution.

[14] It was not possible to employ exactly the same set of data as Dollar and Kraay. They had access to unpublished data from the *World Development Report 2000* that could not be included here. We have followed the rule of Dollar and Kraay that any two observations for a single country should be separated by at least five years but in implementing it we start with the most recent observation and work backwards through time. As a result, our regressions are based upon a slightly smaller sample of 320 observations for 111 countries.

[15] The division by decades is especially suspect for at least two reasons. First, such a chronological division is arbitrary; and, second, the decade samples would not cover the same set of countries.

[16] A one tailed test is appropriate when the sign of a coefficient is predicted. If the sign is not predicted *ex ante*, a two tailed test should be used. Since Dollar and Kraay seem to profess *ex ante* agnosticism for the sign for inflation, a two tailed test would be the correct one. This would render the t-statistic non-significant.

[17] See, for example, the use of various macro indicators in the 1994 World Bank report on adjustment in Africa (World Bank 1994), again in a paper by Demery and Squire (1996), and critiques of these (Mosley, Subasat & Weeks 1995, and Weeks 1997).

[18] Since inflation is an outcome, not a policy, the real exchange rate is also an outcome since it is the nominal exchange rate adjusted for inflation.

[19] '…[T]he 'growth effect' shows direct effects of the indicated variable on incomes of the poor that operates through its effect on overall incomes' (Dollar & Kraay 2000, p. 5).

[20] 'The… "distribution effect" captures the indirect effect of [a] variable on the incomes of the poor through its effects on the distribution of income' (*ibid*.).

[21] It is highly unlikely that equation 3 is a reliable growth estimator, for it excludes major determinants of growth, perhaps the most important being the capital stock.

[22] It is beyond the scope of this paper to discuss the problems with the ordinal variables, of which 'rule of law' is perhaps the most problematical.

REFERENCES

Atkinson, A. B. and Brandolini A. (1999). Promise and Pitfalls in the Use of 'Secondary' Data-Sets: Income Inequality in OECD Countries. Oxford and Rome: Nuffield College, University of Oxford, and Banca d'Italia, Research Department.

Dagdeviren, H., van der Hoven, R. and Weeks J. (forthcoming). Redistribution Does Matter: Growth and Redistribution for Poverty Reduction. Draft. London and Geneva: SOAS and ILO.

Deininger, K. and Squire L. (1996). 'A New Data Set Measuring Income Inequality', *The World Bank Economic Review*, **10**(3), 565-91.

Demery, L. and Squire L. (1996). 'Macroeconomic Adjustment and Poverty in Africa: an Emerging Picture', *The World Bank Research Observer*, **11**(1).

Dollar, D. and Kraay A. (2000). *Growth is Good for the Poor*. Washington, D. C.: World Bank.

The Guardian (15 June 2000). 'Poverty row author quits World Bank', *The Guardian* (London).

The Guardian (22 June 2000). 'Victims of 'criminal complacency'', *The Guardian* (London).

Kanbur, R. (1998). *Income Distribution and Development*. WP 98-13. Washington, D.C.: World Bank.

Kanbur, R. and Lustig N. (1999). *Why is Inequality Back on the Agenda?* Paper prepared for the Annual Bank Conference on Development Economics. Washington, D. C.: World Bank.

Lübker, M. (2000). *Growth, Income Distribution and the Poor*. London: SOAS Department of Development Studies MSc Dissertation.

Lundberg, M. and Squire L. (1999). Growth and Inequality: Extracting the Lessons for Policymakers. Updated draft version as of 11 August 2000. Washington, D. C.: World Bank.

Mosley, P., Subasat, T., and Weeks, J. (1995). 'Assessing Adjustment in Africa', *World Development*, **23**(9), 1459-1473.

Weeks, J. (1997). 'Analysis of the Demery and Squire 'Adjustment and Poverty' Evidence', *Journal of International Development*, **9**(6).

Weisbrot, M., Baker, D., Naiman, R., and Neta, G. 'Growth May be Good for the Poor But are IMF and World Bank Policies Good for Growth?' Mimeo. http://www.cepr.net

World Bank (1994). *Adjustment in Africa: Reforms, Results and the Road Ahead* Oxford: Oxford University Press.

Table 1 Adjustments to Gini Coefficients and Income Shares

	Gini Coefficient		Income Share of Bottom Quintile	
	Coefficient	*t* statistic	Coefficient	*t* statistic
Expenditure	–.041	1.7	0.045	1
Net Income	–.087	3.6	0.016	0.3
Unknown Income	0.044	2.3	–.088	2.3
Household Unit	0.069	4.9	–.047	1.6
Unknown Unit	0.033	1.8	0.005	0.2

Table 2 $y_{ct}^{p} = \alpha_0 + \alpha_1 y_{ct} + \epsilon_{ct}$

	$\hat{\alpha}_0$	$\hat{\alpha}_1$	H_0: $\alpha_1 = 1$	R^2
Dollar and Kraay, OLS	–1.728	1.060	2.22	0.9
	(-7.48)	(-39.26)		
Dollar and Kraay, system	–1.613	1.046	0.45	na
	(-1.90)	(-10.25)		
Replication of OLS	–1.364	1.016	0.62	0.8
	(-3.73)	(-23.25)		
Developing Countries: full sample	–0.677	0.923	–1.08	0.7
	(-1.27)	(-12.95)		
Developing Countries: $Y_{ct}^{p} < \$\ 500$	2.299	0.477	–2.97	0.3
	(-1.72)	(-2.705)		
Developing Countries: $\$\ 500 < Y_{ct}^{p} < \$\ 1{,}500$	3.839	0.355	–8.14	0.3
	(-6.33)	(-4.476)		
Developing Countries: $Y_{ct}^{p} > \$\ 1{,}500$	4.712	0.350	–3.77	0.2
	(-3.26)	(-2.033)		

($|t|$ statistic)

Unless otherwise stated, OLS estimation with levels of variables and *t* statistics calculated from Newey-West adjusted standard errors.

Replication based upon Dollar-Kraay data with differences as discussed.

Figure 1 Scatter Plot of y_{ct}^{p} on y_{ct}, Full Sample

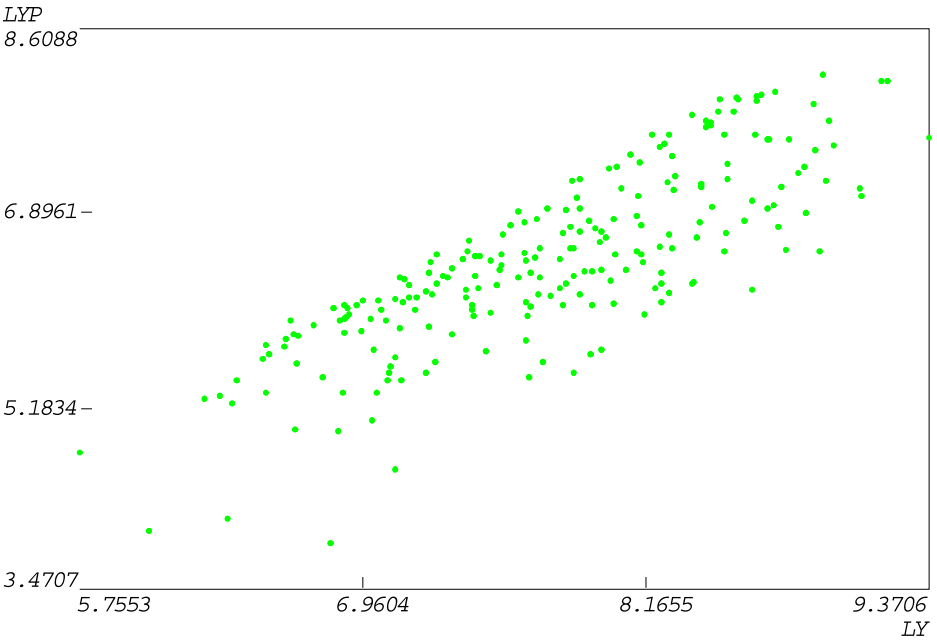

Figure 2 Low-Income Poor

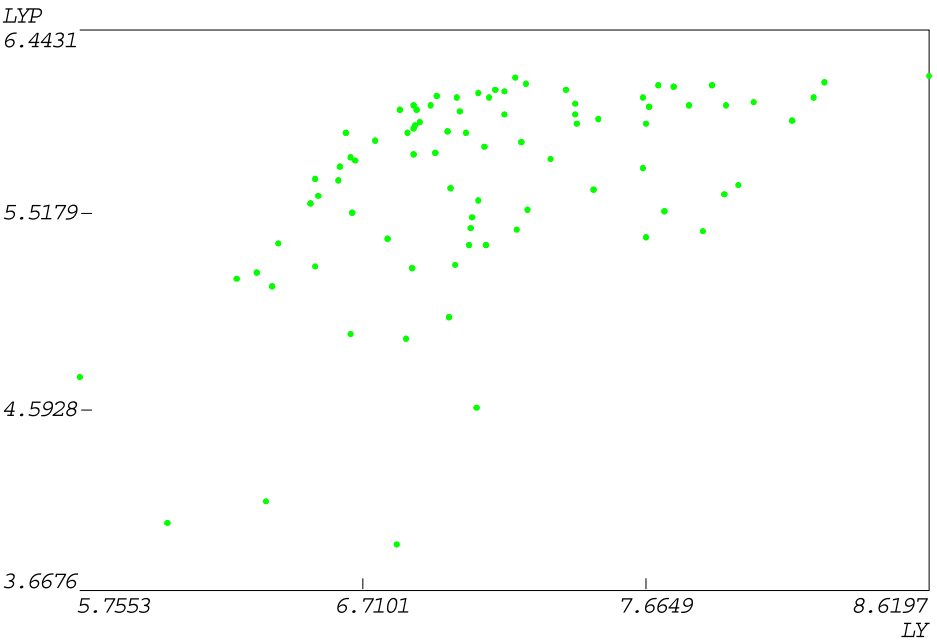

Figure 3 Middle-Income Poor

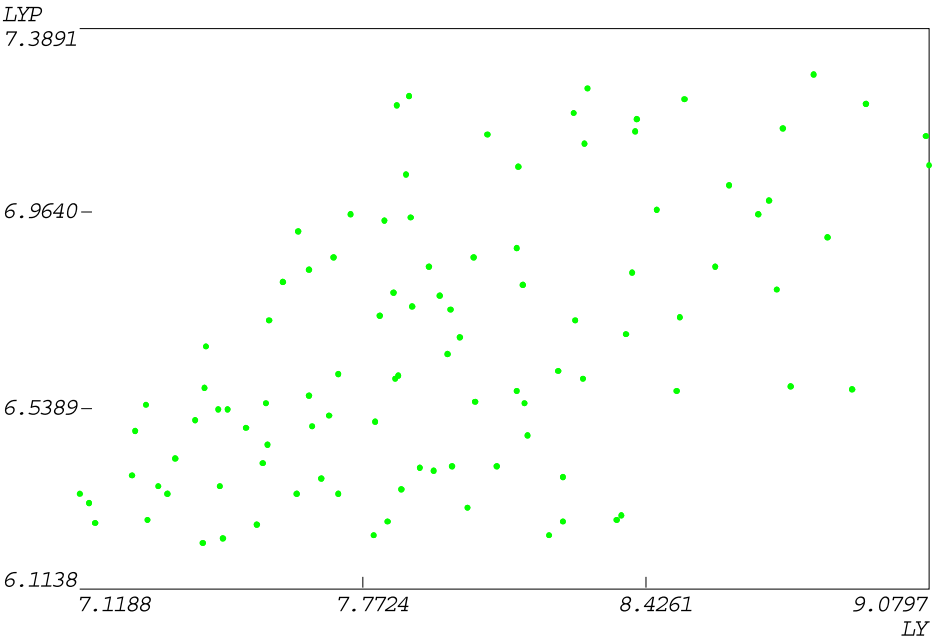

Figure 4 High-Income Poor