

GROWTH AND THE POOR: A COMMENT ON DOLLAR AND KRAAY

MALTE LÜBKER, GRAHAM SMITH and JOHN WEEKS*

SOAS, University of London, London, UK

Abstract: In a recent paper Dollar and Kraay come to sweeping conclusions about economic growth and the poor. On the basis of empirical work they assert that standard World Bank and IMF policy packages are good for the poor. This paper 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, nor are they tested in a consistent manner. These problems imply that the policy conclusions of the authors are unsafe. Copyright © 2002 John Wiley & Sons, Ltd.

1 INTRODUCTION

A recent paper by Dollar and Kraay (2000) might have gone largely unnoticed had it not apparently been associated with the well-publicized controversy over the *World Development Report 2000*, including the resignation of one of the major contributors.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. More precisely stated, their results showed that the mean income of the lowest quintile had an elasticity of unity with respect to overall mean income. In other words, this quintile's income tended to remain stable in the process of growth. The additional hypothesis tests, involving various policy indicators which were statistical tests of association in the absence of a theoretical framework,² proved to be non-significant in most cases. Thus, the most important result of the paper may be a 'finding', but not a discovery, since the stability of income distributions is both well established and well known (see Weisbrot *et al.*, 2000). Further, how 'good' a unit elasticity is for the poor is a matter of opinion.

Nonetheless, it is instructive to deconstruct the Dollar–Kraay exercises because: first, the fervour of the paper ('growth *is* good for the poor', emphasis in original) and the wide

*Correspondence to: John Weeks, SOAS, University of London, Thornhaugh Street, London WC1H 0XG, UK.
E-mail: cdpr@soas.ac.uk

¹In mid-2000, Ravi Kanbur resigned as co-ordinator of WDR 2000. For journalistic and 'insider' reports, see *The Guardian* (London) 15 June; 22 June 2000, and Chang (2001, p. 5). For the views of Kanbur on the distribution of income and growth, see Kanbur (1998, 2001), and Kanbur and Lustig (1999).

²See Dagdeviren *et al.* (2001), where distribution neutral growth is considered for its poverty reducing potential.

debate over it calls for consideration of the anti-distributional policy perspective it implicitly incorporates; second, a review of the empirical procedure provides a tutorial in how statistical work should and should not be done; and, third, the authors' interpretation of results demonstrates that the absence of theoretical modelling provides fertile ground for ideological inferences.

The Dollar and Kraay paper can be summarized 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. They seek to establish the first point through a dubious statistical framework, though the finding is hardly controversial. The second point is a *non sequitur*; what is true on average is not necessarily true for the components of an average. The third point they establish from a biased reading of the growth literature, including work from the World Bank itself. Since they do not establish points two and three, the sweeping policy conclusion (point iv) does not follow.

2 NATURE OF THIS CRITIQUE

In the context of ideological predilections, be they personal or institutional, we should make clear the purpose of the critique that follows. Economics is a social science that since its origins has addressed issues of public policy. Public policy is never decided by technical considerations alone, but must arise from a political process. In turn, all politics reflect or are influenced by an underlying ideology. Thus, it is to be expected that policy recommendations from economists will provoke political, even ideological debate. This has been particularly the case for the Dollar and Kraay paper. We disagree with the political implications of Dollar and Kraay, but that disagreement is secondary to the issue of methodological soundness.

While economists may disagree among themselves over politics, they share a common commitment to analytical rigour ('science'). This rigour involves the procedure presented to first year students: theoretical foundations, model construction, empirical testing (including evaluation of the quality of data) and statistical inference. Considerably more important than the politics of the Dollar and Kraay paper is whether their policy conclusions emerge from this rigorously disciplined process that is the *ration d'être* of the profession.

Therefore, this paper focuses upon how the Dollar and Kraay hypotheses were formulated, empirically tested and the validity of their statistical inferences. While our critique can be correctly read as carrying a pro-redistribution message, it is essentially technical, and we would equally apply the critique to studies whose outcomes support our political predilections. Superficially the Dollar and Kraay argument that growth reduces poverty appears robust, in part because it would be quite surprising if growth did not have this effect. However, the authors go well beyond this broad generalization to argue that growth stimulated by World Bank type policies is good for the poor. This more specific conclusion is not robust. It is not robust because the statistical modelling used to sustain it is flawed, the data used to test the model are unsafe, and the statistical inferences drawn from that testing do not follow best professional practice.

To carry out our critique as described above, it is obviously necessary to discuss in some detail hypothesis formulation, the characteristics of the data set and statistical testing. For

the statistically trained readers our technical points will be familiar and their interest will be in implications for the credibility of the Dollar and Kraay policy conclusions. For those less familiar with modelling and statistics, we avoid jargon and excessively technical points in order to demonstrate that while economists may have political opinions, they also have a methodological rigour to which they must subject those opinions.

Our critique is elaborated as follows. First, we consider the relationship between growth and distribution, to indicate that basic theory should be made explicit 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.

3 HYPOTHESIS FORMULATION

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, one must be especially vigilant to adhere to the rules of rigour established in the profession. Rigour requires clarification of concepts and basic causality. The basic Dollar and Kraay estimating equation would seem less than clear and transparent:

$$\ln[Y_{1ct}] = \alpha_0 + \alpha_1 \ln[Y_{ct}] + \alpha_2 [X_{ct}] + \mu_c + \varepsilon_{ct} \quad (1)$$

$\ln[Y_{1ct}]$ 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, c and t indicate countries and years, $[X_{ct}]$ is a vector of additional explanatory variables, $\mu_c + \varepsilon_{ct}$ are the composite error term, and quintiles are numbered 1, 2...5, with only the first playing a role.

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 all the information that modelling procedure requires. It is obvious that equation (1) is a reduced form; i.e. it is the result of algebraic substitutions for unspecified behavioural relationships. Average income does not act directly on the income of a component of itself, except tautologically (an average is derived from its components). Since we are inspecting 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. Specifically, there is no presentation of the theory that predicts that α_1 should be unity (this being the basis of the author’s conclusion, ‘growth *is* good for the poor’). In effect, Dollar and Kraay treat this outcome as serendipitous rather than analytically anticipated.

Notwithstanding the ambiguousness of their estimating equation, Dollar and Kraay focus on the elasticity between the two per capita income variables. Prior to considering their argument, we point out that if this elasticity is positive and statistically significant,

one can conclude that the 'poor' benefit from growth, even if the value is near zero. Dollar and Kraay put forward the specific null hypothesis that the coefficient is 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, *ad hoc* reasoning would predict growth to be distribution-neutral. Income distributions arise from political, social and economic factors that over the short to medium term are stable. While the past decades have witnessed major shifts in income distribution even in industrial countries (see Atkinson 1998), in the absence of major policy changes or social upheaval such as war, one would expect the determinants of distributions to change slowly, though in some countries the changes might be dramatic.

Further, there is a purely statistical reason to expect the coefficient to be close to unity. The lower limit of the share of the income of the lowest quintile is zero by definition, and its upper limit is restricted by the condition that the poorest quintile cannot have a share equal to or greater than the second quintile. Thus, the possible range for variation is quite small. 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. The fact that the dependent variable (income share of the bottom quintile) is a component of the explanatory variable (average income) creates the possibility that the equation is a tautology. By definition,

$$Y_{ct} = p_1 Y_{1ct} + p_2 Y_{2ct} + \dots + p_5 Y_{5ct} \quad (2a)$$

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 0.2, one can write,

$$Y_{ct} = 0.2[Y_{1ct} + Y_{2ct} + \dots + Y_{5ct}] \quad (2b)$$

If Y_{1ct} 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. Again by definition, the income of the bottom quintile (Y_{1ct}) is:

$$0.2Y_{1ct} = Y_{ct} - 0.2[Y_{2ct} + \dots + Y_{5ct}] \quad (2c)$$

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

$$0.063Y_{ct} = Y_{ct} - 0.2[Y_{2ct} + \dots + Y_{5ct}]. \quad (2d)$$

It follows that:

$$Y_{1ct} = 0.315Y_{ct}. \quad (2e)$$

The logarithmic form of this definition is:

$$\ln[Y_{1ct}] = -1.155 + 1.000 \ln[Y_{ct}]. \quad (2f)$$

This is suspiciously similar to the Dollar–Kraay result when they estimate the simple regression between $\ln[Y_{1ct}]$ and $\ln[Y_{ct}]$, using three methods, ordinary least squares,

instrumental variables, and a system estimator. With the last two methods their estimated coefficients are not significantly different from those predicted by the manipulation of the definition of per capita income (equation (2f)). Our own OLS estimates on the data set, which do not include some of the Dollar and Kraay observations, confirms the prediction of the definition (standard errors in parenthesis),

$$\ln[Y_{1ct}] = -1.364 + 1.016 \ln[Y_{ct}]$$

(0.366) (0.044)

R – square = 0.83.

For all regression analysis, the coefficients must be such that the linear estimation passes through the average value of the variables. Equations (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 -square 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, an arithmetic determination should run the other way.

The absence of a behavioural (causal) mechanism in the Dollar and Kraay basic equation is surprising, since in this intensely controversial field there is general agreement on the basic determinants of income distribution. It is generally accepted 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. One should first specify the determinants of the income of the bottom quintile in a model that incorporates a theory of income generation. 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.

The importance of an explicit theory of income determination can be demonstrated in another way. Dollar and Kraay regress the average income of the lowest quintile against overall mean income. The income share of the bottom quintile determines the former. This

share has a lower limit of zero, and an upper limit of 0.2 if income is equally distributed across quintiles, and a practical upper limit of about 0.12 (see World Bank, 2000). If the long run elasticity of the average income of the bottom twenty percent with respect to per capita income were greater than or less than one, then the income share of the poorest twenty percent would asymptotically approach either the lower or the upper limit. As these values were approached, the change in the share would become less and less. Thus, by statistical definition (dividing the population by quintiles), the elasticity must be unity in the long run, and statistical proof of it is redundant.

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.

4 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. The purpose of this discussion is not to labour detail, but to demonstrate that the technical basis upon which the authors' sweeping generalizations are made is not sound.

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* of 1999 and 2000. Since some of the data from the last source were unpublished at the time this paper was written, we had access to a slightly smaller sample with 759 observations from just 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), individuals (365), and income receiving units whose definition is unknown in 22 cases. Some observations derive from income (591) and others from expenditure (168). Of the income-based observations, some were reported after of taxes (174), some before taxes (370), and forty-seven observations are not identified. 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.³

³These 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* (US) 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 (Atkinson and Brandolini, 1999, p. 7).

Table 1. Adjustments to Gini coefficients and income shares

Category	Gini Coefficient		Income Share of Bottom Quintile	
	Coefficient	<i>t</i> statistic	Coefficient	<i>t</i> statistic
Expenditure	-0.041	1.7	0.045	1.0
Net Income	-0.087	3.6	0.016	0.3
Unknown Income	0.044	2.3	-0.088	2.3
Household Unit	0.069	4.9	-0.047	1.6
Unknown Unit	0.033	1.8	0.005	0.2

Note: The omitted categories are income before tax and individual incomes.

Given these data inconsistencies, the appropriate procedure would be to sort the observations into consistent sub-sets. However, Dollar and Kraay choose to attempt adjustments to 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 definitions of coverage, with household-based pre-tax (gross) 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).⁴ 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'. In other words, one does not know whether these observations are pre- or post-tax, though one presumes the category includes both. Since the omitted ('benchmark') category is income before taxes, simple deduction implies that no adjustment should be made. If an 'unknown' observation is in fact pre-tax, making the adjustment with any coefficient would incorrectly alter the original statistic. If the unknown observation is post-tax, no adjustment is called for, either, because the coefficient on net income is not significantly different from zero. Thus, it is either a logical mistake or a statistical error to use the coefficient on the 'unknown' category to adjust any statistic in the database. 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 also adjust Gini coefficients. The same inconsistencies described above apply for the inequality statistics. And for these there are additional problems. It is puzzling that the coefficients for the 'net income' and 'unknown income' categories are both significant at the 0.05 level. In words, the regression statistics state with confidence that the Gini must be corrected downwards when dealing with net income and taxed income, but upwards when it is unknown whether the original statistic was before tax or after tax. Again, simple deduction tells us that only one of the two adjustment coefficients can be correct. Any 'unknown income' that in fact is net income

⁴After 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.

should be corrected *downwards*, not *upwards*, if one was to believe the coefficient for this category, and any gross income should remain unchanged. One can conclude from these contradictory results that either the method used to establish them is invalid, the data are unreliable, or both. Since the adjustments do not stand the test of simple logic, one's strong suspicion is that the entire adjustment exercise is hopelessly flawed. Even if the logical inconsistencies were not present, it would be invalid to attempt to adjust for net and gross income. Across countries, the rate of taxation and progressivity of rates vary 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 distributive impact.⁵

In summary, the data used by Dollar and Kraay are problematical and unusable in their raw form. Adjustments the authors make are not justified by statistical procedure. Since logic tells one that the adjustments do not improve data reliability, and one cannot rule out the possibility that they made the data worse. The authors might object that our criticism in this section involves nit picking, and point out that all empirical work on developing countries requires use of problematical data. That general defence, with which we have some sympathy, does not apply in the Dollar and Kraay study. The authors themselves stress the importance of the accuracy of their data, as manifested in their reference to the 'high quality' data set of Deininger and Squire (1996), the quality of which, they maintain, lends their results particular credibility. Finally, were one to accept the general defence that all data are to some degree problematical, the Dollar and Kraay adjustments to their data set render it unsafe for econometric use.

5 STATISTICAL ESTIMATES OF GROWTH AND THE 'POOR'

In this section we consider the more 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 ordinary least squares (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. They consider this an inaccurate estimate and attribute it to the unreliability of a simple OLS regression (Dollar and Kraay 2000, p. 19). No theoretical explanation is given as to why this result is unacceptable.

Using a system estimator they obtain a point elasticity of 1.046, which 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 and low income countries, high and low growth, positive and negative growth, and observations segregated by decades (Dollar and Kraay 2000, p. 36). On the basis of these variations, they conclude that unit elasticity is a near-universal outcome among and within countries.

⁵Atkinson and Brandolini (1999, p. 21 ff) pointed 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.

Table 2. Estimations of the basic Dollar and Kraay equation:
 [average income of bottom quintile] = $a_0 + a_1$ [overall mean income] + [error term]

	\hat{a}_0	\hat{a}_1	H_0 : $a = 1$	R^2
Dollar and Kraay, OLS	-1.728 (7.48)	1.060 (39.26)	2.22	0.9
Dollar and Kraay, system estimator	-1.613 (1.90)	1.046 (10.25)	0.45	na
Lübker, Smith and Weeks replications:				
OLS total sample	-1.364 (3.73)	1.016 (23.25)	0.62	0.8
Developing countries only:				
Full sample	-0.677 (1.27)	0.923 (12.95)	-1.08	0.7
For which $Y_{1cr} < \text{US\$500}$	2.299 (1.72)	0.477 (2.705)	-2.97	0.3
For which $\text{US\$500} < Y_{1cr} < \text{\$1500}$	3.839 (6.33)	0.355 (4.476)	-8.14	0.3
For which $Y_{1cr} > \text{\$1500}$	4.712 (3.26)	0.350 (2.033)	-3.77	0.2

Notes: \hat{a}_0 is the intercept and \hat{a}_1 the estimated coefficient on mean income. The value of the t -statistic is in parenthesis under the coefficients in columns one and two. Column three reports the t -statistic for the hypothesis that the coefficient a is unity. OLS estimation with levels of variables and t statistics are calculated from Newey-West adjusted standard errors. The data used differ slightly from those of Dollar and Kraay, as explained in the text.

The composition of the data set casts doubt upon the allegation of generality. First, and for reasons previously discussed, it is counter-productive to adjust for differences in measurement definitions. Second, it is unsound to estimate missing quintile shares.⁶ Third, the sample includes observations from 22 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 it involves fundamental social and economic transitions.⁷ For reasons of analytical caution alone, the data from the developed and developing countries should be separated.

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, this is not significantly different from one. Unit elasticity appears to be confirmed, *at this the most general level*. The important question is whether analysis of the components sustains the general result, as Dollar and Kraay seek to establish with their various decompositions of the data set.

The most analytically interesting of the Dollar and Kraay divisions of the data set are by level of per capita income, though they are not explicit as to what this division tests.

⁶If 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. The Gini, which is a measure rather than a functional form, is consistent with considerable variation in income shares at the extreme of the distribution.

⁷These 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.

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 the division of a sample can bias results.

If the unit elasticity hypothesis is general and robust, any analytically justifiable division of the data set should sustain it, not only the numerous ones chosen by the authors. In other words, because Dollar and Kraay provide no theoretical justification of their various regression experiments, other no less arbitrary divisions of the data set should be no less acceptable to them as valid tests of the hypothesis.

Dollar and Kraay divided the sample on the basis of the overall average per capita income of countries. In a variation on this *ad hoc* division, we ordered the developing country observations according to *increasing income of the bottom 20 per cent*. 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. In the absence of theoretical guidelines for division of the data set, their article suggests no reason why they should not accept ours to be as valid as the ones they choose.

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 0.48 (low income), 0.36 (middle income), and 0.35 (high income), all significantly less than one. With the collapse of the elasticity hypothesis, the *R*-squares also collapse, into the 0.2 to 0.3 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

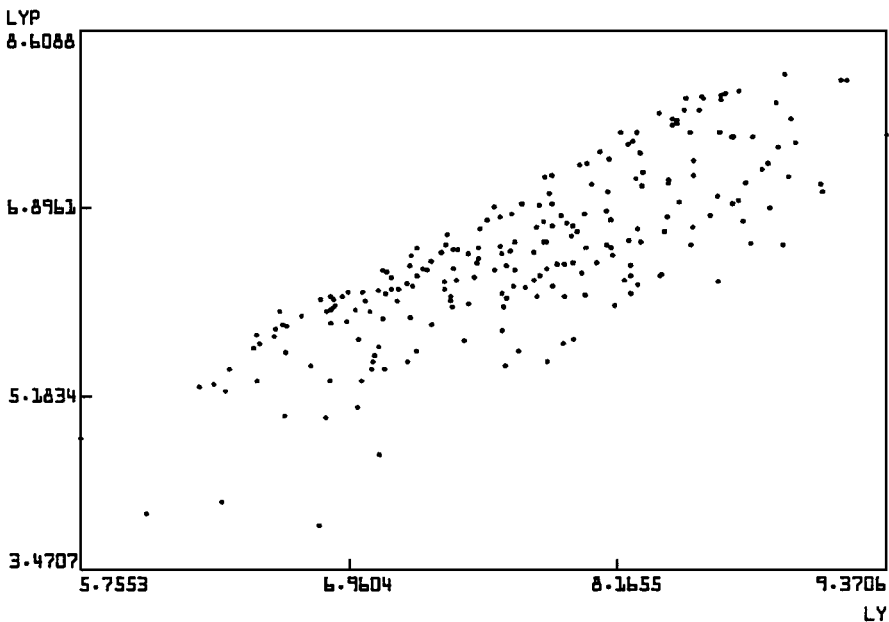


Figure 1. Scatter plot of the logarithms of mean income of the lowest quintile (LYP) on overall mean income (LY)

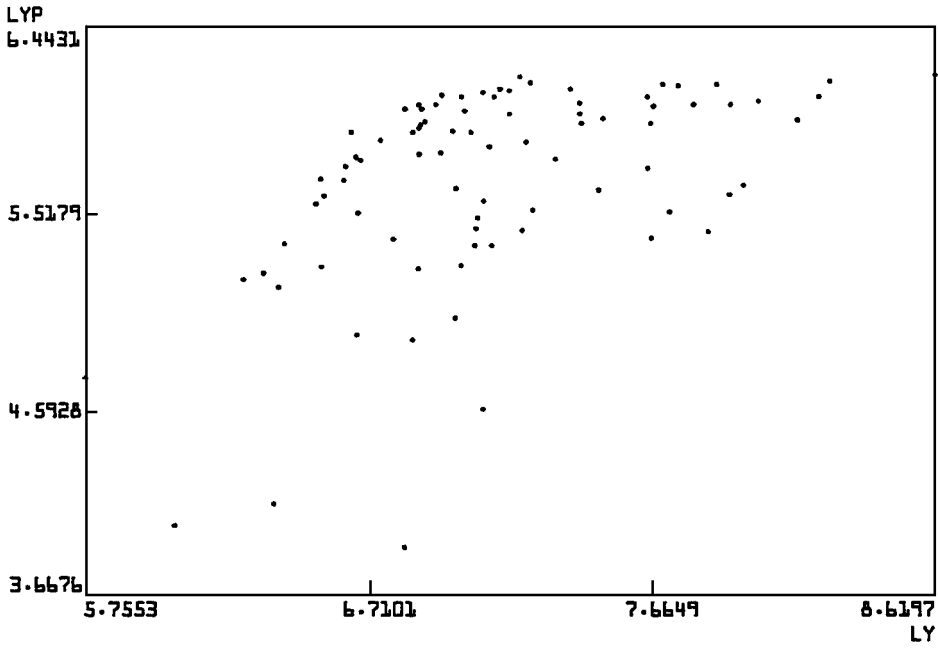


Figure 2. Scatter plot of the logarithms of mean income of the lowest quintile (LYP) on overall mean income (LY), low income poor observations

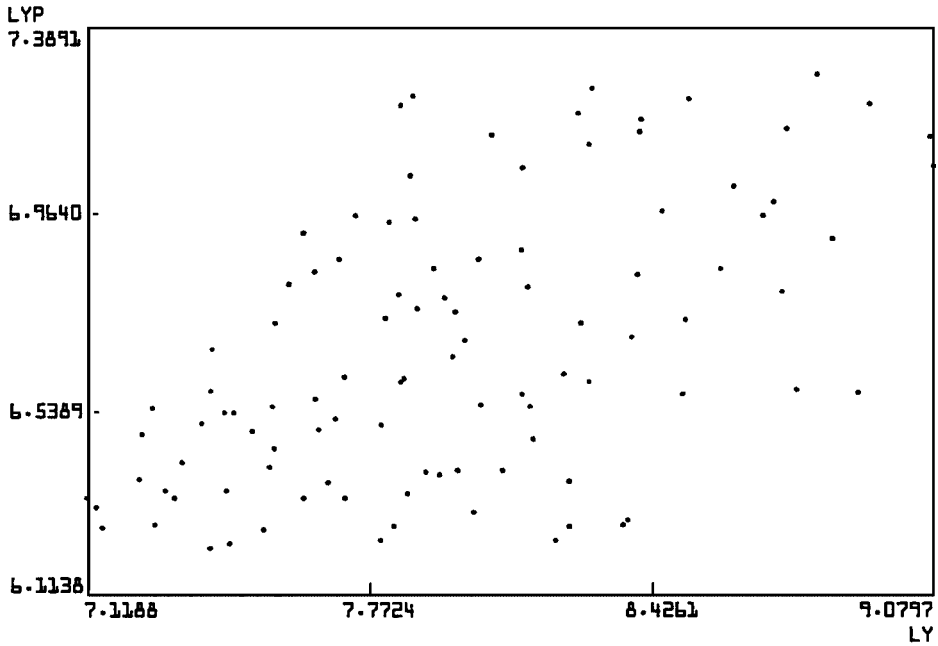


Figure 3. Scatter plot of the logarithms of mean income of the lowest quintile (LYP) on overall mean income (LY), middle income poor observations

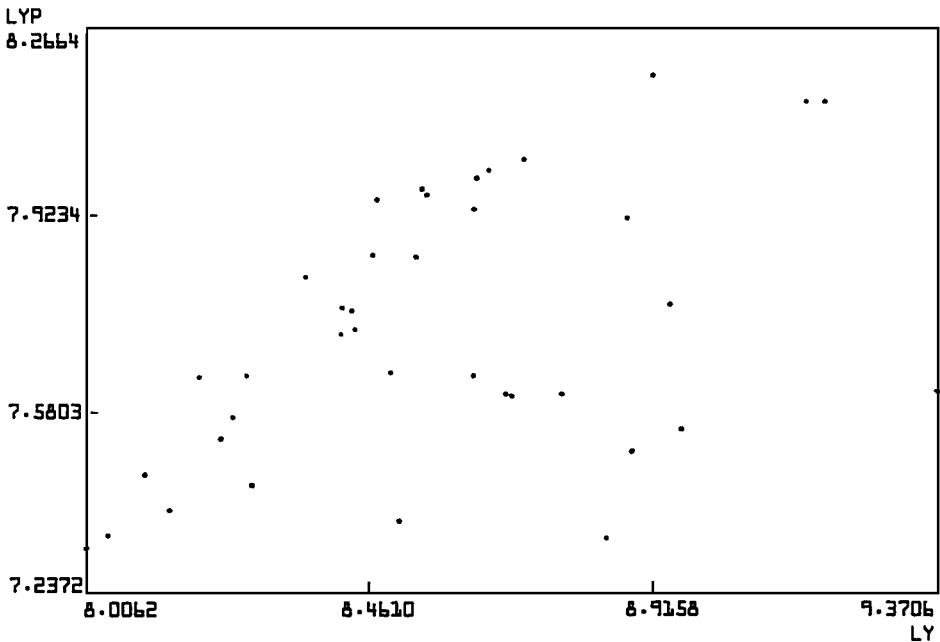


Figure 4. Scatter plot of the logarithms of mean income of the lowest quintile (LYP) on overall mean income (LY), middle income poor observations

elasticity. 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 a pattern, appears non-linear, with several extreme values at low incomes (Figure 2). For the middle-income poor the scatter is quite diffuse (Figure 3). In the case of the high-income poor, 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_c^* ; our criterion draws a horizontal line (at a point such as Y_{c1}^*). 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.

Our division may be no more justified than Dollar and Kraay's, since neither is based on theoretically justified *ex ante* criteria; *equally, it is no less justified*. 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. We would argue that the division by decades is more suspect for at least two reasons. First, it is arbitrary; and, second, that arbitrariness is compounded because the decade samples do not cover the same set of countries.

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 itself and as a vehicle for

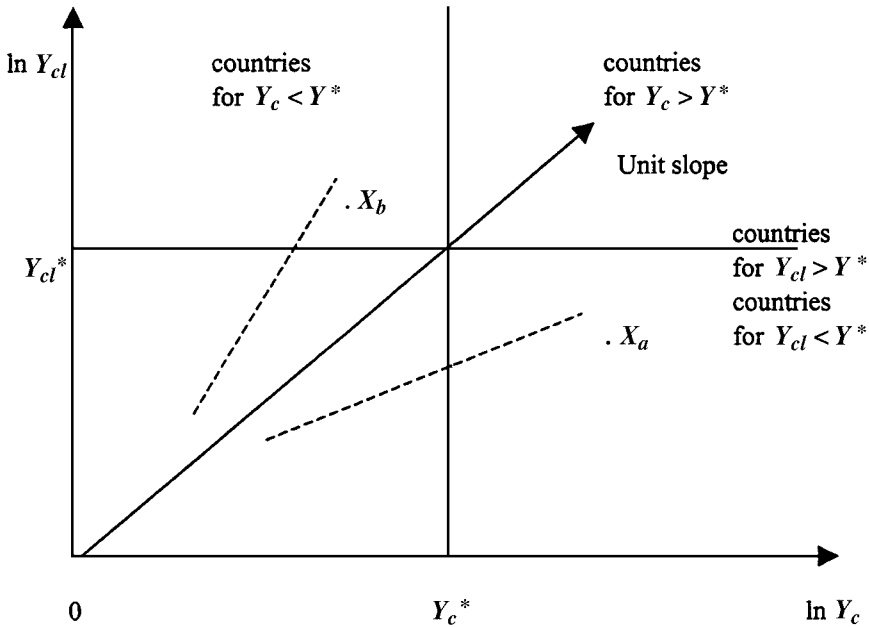


Figure 5. Likely estimation bias resulting for division of the Dollar-Kraay sample

poverty reduction, which we consider analytically justified (see Dagdeviren *et al.*, 2001; Dagdeviren and Weeks, 2001). 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 per cent 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 unreliable 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.

6 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 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 is inflation, with a *t* statistic of 1.75, significant at the 0.05 level on a one-tailed test. If the three non-significant variables are omitted, the coefficient on

inflation ceases to be significantly different from zero. Even the 0.05 outcome for the inflation variable is suspect. There are no *a priori* reasons to predict the sign of the coefficient on inflation with regard to distribution. A one-tailed test is appropriate if the sign of a coefficient is theoretically predicted. If the sign is not theoretically specified *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.

These so-called policy indicators, common in empirical work by World Bank professionals,⁸ suffer from a number of serious problems. First, inflation and trade as a portion of GDP are not policy variables; they are 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 may 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: a change in the terms of trade, the impact of whether on agriculture, and other 'shocks'. No conclusion about the impact of policy on any variable can be drawn from a cross-country regression using inflation as an indicator. For example, a 20 per cent inflation rate might be experienced in a country that pursued 'good' macro policy (but was unlucky), and in one whose government 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 outcome indicators, the relevant policy variables should be, for example, nominal (not real) interest rates and monetary emission (for inflation), and the nominal exchange rate (for trade).

If one were to suspend objections and accept the appropriateness of the variables, there remains a serious theoretical inconsistency in the way they enter the model. Their impact 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 simple logic, 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 (see Mosley *et al.*, 1995, p. 1463).

⁸See, 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 *et al.*, 1995; Weeks 1997).

Yet another problem with the ‘policy’ variables would seem to undermine the entire model. 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: ‘...the “growth effect” shows direct effects of the indicated variable on incomes of the poor that operates through its effect on overall incomes’ (Dollar and Kraay, 2000, p. 5). A second equation, with the income of the bottom quintile as the dependent variable, employs the same policy variables as arguments along with overall per capita income as an independent variable. This the authors call the ‘distribution effect’: ‘The... “distribution effect” captures the indirect effect of [a] variable on the incomes of the poor through its effects on the distribution of income’ (Dollar and Kraay, 2000). 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 and 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 are treated as a composite variable (X_{ct}). None of these affect our point.

$$\ln[Y_{1ct}] = \alpha_0 + \alpha_1 \ln[Y_{ct}] + \alpha_2[X_{ct}] \quad (3)$$

and

$$\ln[Y_{ct}] = \beta_0 + \beta_1 \ln[X_{ct}] \quad (4)$$

It is immediately obvious that the two equations are not independent of each other and should not be estimated separately. Equation (3) has as one of its arguments the dependent variable in equation (4). Straightforward substitution yields,

$$\ln[Y_{1ct}] = (\alpha_0 + \alpha_1\beta_0) + (\alpha_1\beta_1 + \alpha_2)[X_{ct}] \quad (5)$$

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 (despite the absence of standard arguments such as growth of the capital stock), 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 useful 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; 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 accuse them of holding that position, though it is accurate to describe their paper as ignoring policies of redistribution as a way of improving the condition of the poor and reducing poverty. If ‘growth is all that is needed’ were their position, it would be relatively benign, for that opinion has a long and distinguished pedigree. Our objection is to their conclusion *via* statistical inference they show 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 increases the income of the poor to the same extent that it increases the income of the other households in society.

... [W]e find... 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 either 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. Indeed, William Easterly of the World Bank produced a paper in 2001 with econometrics that suggest structural adjustment policies tended to generate a growth pattern in which income distribution worsened (Easterly, 2001). 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 complex and politically contentious such 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.

ACKNOWLEDGEMENTS

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).

REFERENCES

- Atkinson AB. 1998. The distribution of income in industrialized countries. In *Income Inequality, Issues and Policy Options*. Federal Reserve Bank of Kansas City (ed.). Federal Reserve Bank of Kansas City: Kansas City; 11–32.
- Atkinson AB, Brandolini A. 1999. Promise and pitfalls in the use of ‘Secondary’ data-sets: income inequality in OECD Countries. Nuffield College, University of Oxford, and Banca d’Italia, Research Department: Oxford and Rome.
- Chang HJ (ed.). 2001. *Joseph Stiglitz and the World Bank: The Rebel Within*. Anthem Press: London.
- Dagdeviren H, van der Hoeven R, Weeks J. 2001. Redistribution does matter: growth and redistribution for poverty reduction. *International Labour Office Employment Paper 2001/10* ILO: Geneva.
- Dagdeviren H, Weeks J. 2001. How much poverty could the HIPC reduce? *WIDER Development Conference: Debt Relief*, 17–18 August 2001 (Helsinki: WIDER). [available at www.wider.unu.edu]
- Deining K, Squire L. 1996. A new data set measuring income inequality. *The World Bank Economic Review* 10(3): 565–591.
- Demery L, Squire L. 1996. Macroeconomic adjustment and poverty in Africa: an emerging picture. *The World Bank Research Observer* 11(1).
- Dollar D, Kraay A. 2000. *Growth is Good for the Poor*. World Bank: Washington, DC.
- Easterly W. 2001. The effect of IMF and World Bank programmes on poverty. *WIDER Discussion Paper* No. 2001/102 (WIDER Helsinki) [available at www.wider.unu.edu].
- The Guardian* 2000. Poverty row author quits World Bank. *The Guardian* (London) 15 June.
- The Guardian* 2000. Victims of ‘criminal complacency’, *The Guardian* (London) 22 June.
- Kanbur R. 1998. *Income Distribution and Development*. WP 98–13. World Bank: Washington, DC.
- Kanbur R. 2001. Economic policy, distribution and poverty: the nature of disagreements. *World Development* 29(6): 1083–1094.
- Kanbur R, Lustig N. 1999. Why is inequality back on the agenda? Paper prepared for the Annual Bank Conference on Development Economics. World Bank: Washington, DC.
- Lübker M. 2000. Growth, income distribution and the poor. Department of Development Studies MSc Dissertation. SOAS: London.
- Lundberg M, Squire L. 1999. Growth and inequality: extracting the lessons for policymakers. Updated draft version as of 11 August 2000. World Bank: Washington, DC.
- Mosley P, Subasat T, 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, Neta G. 2000. 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 University Press: Oxford.
- World Bank. 2000. Distribution of income or consumption. World Bank data base. [<http://www.worldbank.org/data/wdi2001/pdfs/tab2-8.pdf>]