A proper farewell to Kuznets' hypothesis

Luis Angeles*

July 20, 2007

Abstract

The aim of this paper is to offer a more appropriate test of Kuznets' "inverted-U" hypothesis than the one routinely used in the literature and implement it using panel and country-by-country regressions. We explore whether countries experiencing large shifts in population from the agricultural/rural sector to the urban one are characterized by an evolution of income inequality along the lines discussed by Simon Kuznets in its classical article. Our results show that there is no systematic relationship between income inequality and agricultural employment or rural population.

1 Introduction

In an enormously influential article, Kuznets (1955) speculated that income inequality should first rise and then fall as countries progress from low to high

*Department of Economics, University of Glasgow. Adam Smith Building, Glasgow G12 8RT, UK. Tel: +44 141 330 8517. Email: l.angeles@lbss.gla.ac.uk

development status. This “inverted U” hypothesis became one of the most widespread notions in development economics. During subsequent decades, no discussion of income inequality seemed to be complete without a mention of Kuznets’ hypothesis. At its zenith, researchers in the area of income inequality referred to it as a "stylized fact" (Adelman and Robinson 1989) or an "economic law" (Robinson 1976).

Kuznets’ thinking was clearly embedded in the dominant notions of economic development of the 1950s. At this time, development and industrialization were used largely as synonyms, and the process of economic development was universally seen as the process which transfers labor from a “traditional”, low productivity, rural sector (agriculture) to a "modern", high productivity, urban sector (industry). This line of thought guided much policy making in multilateral aid-agencies, international organizations and developing country governments[1]. Our aim here is not to evaluate the pros and cons of this way of understanding economic development but rather to make the obvious point that Kuznets’ hypothesis can only be understood within the intellectual framework that produced it.

Kuznets’ thesis was that as population shifts from the traditional to the modern sector income inequality would first rise and then fall, thus describing an "inverted-U" trajectory. In his own words:

[1] The main influence of this line of tought was Lewis’ (1954) model of a dual economy. Two good discussions of Kuznets’ hypothesis and how it relates to the intellectual environment of its time are Kanbur (2000) and Moran (2005).

> One might thus assume a long swing in the inequality characterizing the secular income structure: widening in the early phases of economic growth when the transition from the pre-industrial to the industrial civilization was most rapid; becoming stabilized for a while; and then narrowing in later phases. (Kuznets 1955, p.18)

Kuznets sustained his point with a simulation exercise where population shifted from a low income sector to a high income one and noted the changes as this happened[2]. Simulations and mathematical proofs aside, the best argument in favour of Kuznets' hypothesis might be its intuitive appeal. It is not difficult to visualize the initial and final stages of the developmental process as characterized by a low level of inequality: we start when everyone is a traditional farm labourer and we end when everyone is a modern urban worker. The intermediate stages would be characterized by higher inequality due to the difference between urban workers and farm labourers.

The shift of population from Agriculture to modern industries is central to Kuznets' hypothesis, as it is the mechanism that gives the inverted U a meaning and makes Kuznets' contribution a theory of the evolution of income inequality and not just a black-box guess. It follows that a test of Kuznets' hypothesis should try to measure whether population shifts from agriculture to other sectors are related in any systematic way to income inequality. This, however, is not

[2] More rigorous modellings of the process and of its effects on inequality were provided by Robinson (1976) and Anand and Kanbur (1993). A "Kuznets curve" is also a feature of more complex models of economic development with additional features such as Greenwood and Jovanovic (1990), Aghion and Bolton (1997), Banerjee and Newman (1998) or Glomm and Ravikumar (1998).

what the very large empirical literature on Kuznets' hypothesis has done over the last five decades.

As is well known, tests of Kuznets' hypothesis search for a relationship between income per capita and inequality. The econometric method, the measures of inequality and income per capita, and of course the spatial and time coverage of the analysis change; but all the empirical studies we are aware of take as granted that regressing measures of inequality on income per capita is the way to test for Kuznets' hypothesis. Here we propose an alternative to this widespread and largely unchallenged view.

One way to gain insight into the question is to step back and ask ourselves why do we think inequality and income per capita are related. The answer would be Kuznets' account of the development process. But in Kuznets' account the driving mechanism is the structural transformation of the economy from agriculture to modern sectors. This is what is causing both changes in inequality and changes in income per capita. It follows that to test this mechanism we should be direct and search for a relationship between the structure of the economy and inequality.

Another way to express the above is to say that since inequality and income per capita are driven by a common cause (i.e. the structural transformation of the economy), there is correlation but not causality between them. In econometric terms, the regression of inequality on income per capita suffers from an endogeneity biais since a regressor (income per capita) is correlated with an omitted variable (the structural transformation of the economy) which is the true driving factor of inequality.

Moreover, income per capita also presents the inconvenient of being affected by many other factors besides the structural change in the economy: think of capital accumulation, technological change or the discovery of natural resources. If income per capita is changing for any of these reasons, how will income inequality evolve? There is no easy answer to this last question, and Kuznets' hypothesis does not deal with it.

This paper distinguishes itself by taking the mechanism proposed by Simon Kuznets seriously and testing whether changes in the employment structure of an economy can be related to income inequality in any systematic way. The most appropriate measure for our study is the share of the population employed in agriculture, and falls in this share would be interpreted as advances along the developmental path. A close, though not perfect, substitute would be the share of the population living in rural areas. The empirical section of this paper will use these two measures of the employment structure of the economy as alternative explanatory factors of income inequality.

A second contribution of the present paper is the approach we take to deal with the important problems of data comparability in the area of income inequality. As we explain in section 3, measures of income inequality are typically not directly comparable across countries or, for a given country, across time. The solution we propose is to group the data in series in such a way that all observations within a series are mutually comparable. Within-series variability can then be used to estimate the effect of explanatory variables.

The result of our empirical analysis is that, both on a panel of countries and on a country by country framework, Kuznets' hypothesis is not supported by the data. Kuznets' hypothesis should therefore be regarded as an attractive idea that is simply not true.

2 Econometric tests of Kuznets' hypothesis

Researchers have looked for a relationship between income per capita and inequality using one of the following three econometric approaches: (i) Cross-country regressions (looking for a relationship across countries observed at a given moment in time), (ii) Panel regressions (looking for a relationship across countries and across time), (iii) Country by country regressions and case studies (looking for a relationship across time in a single country or in a group of countries taking each country separately).

Of these three, only the last one is fully acceptable, the second one is tolerable and the first one is completely inadequate. Due to data restrictions, the first approach dominated tests of the Kuznets hypothesis until the 1990s[3]. The problem with cross-country regressions in this context is that they implicitly assume not only that the relationship between income per capita and inequality

[3] Examples are Ahluwalia (1976), Anand and Kanbur (1993b), Dawson (1997), Paukert (1973)

is the same for all countries but also, and more problematically, that all other factors affecting inequality are either non-existent or constant across countries. Fundamentally, Kuznets' hypothesis describes what happens to a country over time. Testing it in a cross-section is considering that all countries in the world are images of each other at different developmental stages. One can try to remedy this problem by introducing some controls, like dummies for different continents, but results tend to be not robust to this type of exercises.

A panel regression is a superior approach since it allows to control for all time-invariant country characteristics by the inclusion of fixed effects. This approach has been used in the more recent empirical studies, specially since the publication of the extensive dataset of Deininger and Squire (1996)[4]. The downside of testing Kuznets' hypothesis with a panel regression is that we assume that income per capita affects inequality in the same way in all countries. It is not difficult to envisage problems with this assumption. Even if we limit our thinking to the mechanisms presented by Kuznets (1955), the particular inverted U pattern would differ across countries if parameters such as the average income in each sector and the within-sector inequality are not the same in all countries.

The last possibility, studying each individual country separately, is the only one that is fully consistent with the spirit of Kuznets' hypothesis. Only by following a given country along its development path can we say if it experienced

[4]Examples of panel regressions of inequality are Barro (2000), Deininger and Squire (1998), Frazer (2006), Higgins and Williamson (1999), Li et al. (1998) or Matyas et al. (1998).

an increase and subsequent fall in income inequality as a result of this process. This approach is also the most demanding in terms of data since it requires a number of observations over a relatively long period of time for each country we want to study. This is why most work in this category are case studies of developed countries, for example the analysis of the evolution of income inequality over the last three centuries in Britain and the United States by Lindert (2000) or the study of the evolution of inequality in European countries by Morrisson (2000). With the notable exception of Deininger and Squire (1998), we can think of no paper applying this methodology over a large number of developing countries.

We test whether the proportion of the population of a country living in urban areas is related to income inequality first with a panel regression and then with a country by country regression approach. While both methodologies point to the same conclusion, we consider that the country by country regressions provide the most convincing evidence.

3 Data selection and methodology

Our evaluation of Kuznets' hypothesis requires data on the structure of the economy and data on income inequality. Our two measures of the structure of the economy are the share of population employed in agriculture and the share of population living in rural areas. The source for these two measures is the World Bank's World Development Indicators (Edition 2006). The number of observations and country coverage for the population employed in agriculture is quite restricted for developing countries, and for all countries the earliest observation refers to 1980. This is not a severe problem when using a panel regression since the large number of countries provides enough variability, but it does make a country by country approach very difficult. For this reason we will also use the share of the population living in rural areas as a proxy for the employment in agriculture. For this second variable we have a balanced panel with 46 yearly observations (1960-2005) for each of the 226 countries and regions included in the World Bank's dataset. Clearly, not all of those living in rural areas are employed in agriculture and the definition of what constitutes a rural area might be problematic. These caveats notwithstanding, the two variables are highly correlated (0.72) and most regression results continue to hold when we change one variable for the other. The data on income inequality presents more delicate problems, to which we turn to next.

Our source for income inequality data is the World Income Inequality Database version 2.0, published by the World Institute for Development Economics Research at the United Nations University (UNU-WIDER). This is the largest secondary database on income inequality available and contains, among many other sources, the work of Deininger and Squire (1996), diverse estimates made by the World Bank and the data from numerous national statistical agencies. The full database contains over four thousand observations of Gini coefficients for most countries in the world, mainly over the last five decades.

Before turning to the empirical work we must address the important comparability problems that exist when using secondary datasets in general and datasets of income inequality in particular (see Atkinson and Brandolini 2001, Atkinson and Bourguignon 2000). For a given distribution of income, a measure of inequality can vary greatly according to the choices made in a large number of dimensions: (a) The reference unit (household, family, tax unit, person), (b) The equivalence scale used (if any), (c) The income definition (income, consumption, earnings, expenditure, monetary income and whether it is a gross or a net quantity. This without mentioning the many items that can be included or excluded in any definition of income), (d) The weighting of each observation (same weight to each observation or weighting by the number of household members), (e) The population covered (all population, income earners, taxpayers), (f) The age coverage (all ages, persons in age of working, persons above a certain age), (g) The treatment of very low incomes (reported as zero) and of very high incomes (reported as the lower limit of the top income band) and so on. Different researchers make different choices, so their estimates are not directly comparable.

The solution to this problem *is not* to control for each underlying characteristic of the observation by using dummy variables (that is, using a dummy for each type of income, for each reference unit, and so on); and this for at least two reasons. First, it is impossible to control for everything as some choices made in the construction of the inequality measure are not specified in the dataset. One can find several cases where two or more observations that refer to the same country, the same year and with identical *reported* characteristics still take different values. The explanation is simply that these observations have been calculated by different researchers making different choices that are not specified in the dataset.

The second problem is that using dummy variables in this way assumes that the difference between, say, Gini coefficients calculated on income and Gini coefficients calculated on consumption is equal across all countries. This assumption is clearly inadequate, as shown by Atkinson and Brandolini (2001) who compare Gini coefficients for OECD countries calculated from gross and net income, or with and without an equivalence scale. Even though the data they use is highly comparable (all estimates come from the same source: the Luxembourg Income Studies), they found that the differences between gross and net income observations or between observations that use an equivalent scale and those that do not are anything but constant across countries. The distance between observations range from 1 Gini point to 6 Gini points, according to the country. The potential for error is therefore considerable.

Our view is that an error of a couple of Gini points is tolerable if our aim is to explain inequality differences between countries. In this case we are concerned with differences of 20 Gini points or more, like the ones that exist between Latin American and European countries, and a couple of Gini points would not bias the results too much. The same is not true, however, if we are interested in explaining inequality changes for a given country over time. In most cases income inequality changes very slowly over time, maybe 4 or 5 Gini points over several decades, so that an error of a couple of Gini points is very large and can bias our results completely.

To deal with this issue we proceed as follows. First, we group observations in what we will call "series". A series is a set of observations over time referring to the same country, coming from the same source and where the same choices have been made in all relevant dimensions (reference unit, income definition and so on). Once this is done, we control for the level of each series and use the within-series variability to estimate the effect of explanatory factors. The idea is that we can confidently compare within-series observations. Within a given series we find not only that all reported characteristics are the same, but unreported characteristics should also be constant since all observations come from the same source. Moreover, we allow for the difference between (say) income-based Gini coefficients and consumption-based Gini coefficients to vary across countries.

To control for the level of each series our empirical work will include fixed effects for each individual series. This implies that any single-observation series will be dropped from the exercise: they have no within-series variation to be exploited; their single observation would be used to estimate their own fixed effect.

After this data selection procedure we have a total of 2752 observations of inequality grouped in 551 series and corresponding to 121 countries, for an average of 4.55 series and 22.74 observations per country[5].

4 Empirical results

4.1 Panel regressions

We start by grouping all countries together and using a panel data approach to investigate whether a relationship exists between the structure of the economy and income inequality. The empirical specification we use is the following:

$$I_{i,t} = \alpha_1 x_{i,t} + \alpha_2 x_{i,t}^2 + \sum_{i=1}^{N} \beta_i + \varepsilon_{i,t} \tag{1}$$

In equation (1) the subscript t indexes time and the subscript i indexes series. $I_{i,t}$ is our measure of income inequality, the Gini coefficient. The variable $x_{i,t}$ is one of our two measures of the structure of the economy. In order to obtain an inverted U relationship we need to use a variable that, according to Kuznets, would be positively related to inequality at the beginning of the development process and negatively related towards its end. This is why we use the "Share of population employed *outside* agriculture", which is simply one minus the share of population employed in agriculture, as our first choice of $x_{i,t}$. Similarly, our second choice is just the share of population living in urban areas and equals one minus the share of population in rural areas. The β_i are the series-specific fixed

[5] Besides the selection procedure described here, we have also excluded observations that refer only to urban or rural areas, observations whose quality rating is "unreliable" and observations from years before 1960 (the data for our explanatory variables start in 1960 so these observations would be dropped from the regressions anyway).

effects, N is the number of series and $\varepsilon_{i,t}$ is an error term uncorrelated with the regressors. All regressions include the series-specific fixed effects but their estimates are not reported in the tables of results. We note that the presence or absence of a "Kuznets' curve" will be given by the values taken by parameters α_1 and α_2. An inverted-U relationship corresponds to a positive α_1 and a negative α_2. A "non inverted U" relationship would correspond to a negative α_1 coupled with a positive α_2. Other combinations would yield monotonic relationships between x and I.

The results of estimating equation (1) in different panels of countries are shown in table 1 (using the Population employed outside Agriculture) and in table 2 (using the Population living in urban areas). The first column of these two tables present our baseline regression, when all countries are taken together. With the population employed in Agriculture as the regressor the signs of the coefficients correspond to a non inverted U relationship, and they are both statistically significant. When we use the urban population as regressor the signs of the two coefficients still denote a non inverted U but this time they are statistically not significant. Thus, when all countries are taken together the results do not support Kuznets' hypothesis.

A look at the R^2 coefficient of this regression reveals that most of the variation of the endogenous variable is being explained (95% in table 1, 94% in table 2). This high explanatory power is common for panel regressions of inequality and is due to the presence of the fixed effects[6]. If the fixed effects are not included and we estimate a single intercept for all countries the explanatory power of the regression falls to 8.8% when using the population employed outside Agriculture and to 4% when using urban population (not shown in the tables). We interpret this as additional evidence against Kuznets' hypothesis: x and x^2 are only marginally relevant to explain the variation in the data.

All other columns in tables 1 and 2 repeat the exercise in diverse sub-groups of countries. The second column investigates the possibility that Kuznets' hypothesis characterizes only developing countries by excluding high income countries from the sample. We use the World bank's definition of high income countries, which includes not just Western Europe, North America and Japan but also countries such as Korea, Hong Kong, Singapore and many small states. In table 1 we note that the relationship among developing countries is still a non inverted U, but this time the parameters are statistically not significant. The same regression in table 2 shows this time an inverted U relationship, but only one of the two parameters is significant. The evidence continues to be inconclusive.

The remaining columns of tables 1 and 2 group countries by geographical location. Once again we use the World Bank's definitions to create these groups. The limitations of the data on employment in agriculture start to show up here,

[6]Deininger and Squire (1998) obtain an R^2 coefficient of 0.9294 while Matyas et al. (1998) obtain 0.8425 and 0.932 according to the sample. In these two studies the effects are country-specific, not series-specific as here.

since we are not able to perform an estimate for sub-Saharan Africa and there are very few observations for the regions "Middle East and North Africa" and "South Asia". When we use the data on urban population we encounter no such problems. Table 1 shows that, when using employment in agriculture, we find an inverted U relationship in two regions (though none is statistically significant) and a non inverted U relationship in three regions (two of which are statistically significant). When using urban population our findings are two inverted U relationships (one of them significant), three non inverted U relationships (one significant) and one monotonously positive relationship (not significant). Support for Kuznets' hypothesis remains thus elusive. It is also noticeable that in most cases the signs of the coefficient are maintained when we change one measure of the structure of the economy for the other one.

Overall, the evidence from country groups seems inconclusive and in all regressions most of the explanatory power comes from the presence of fixed effects. As we explained earlier, panel regressions are not the ideal setting for testing Kuznets' hypothesis and we should not put too much confidence in the results of this methodology. It could be the case that each individual country satisfies Kuznets' hypothesis and evolves along its own inverted U path, but when we take all countries together no universal pattern emerges. Country-by-country regressions are thus preferable in this context and we turn to them next.

4.2 Country by country regressions

The specification we intend to estimate for each country is exactly as equation (1). Keep in mind that i indexes series, so we simply have to select all series from the country under study. Since data availability is a more pressing problem here we will use urban population as our standard regressor and present results with employment in agriculture for the few cases where it's possible.

Regressions should not be blindly applied to all countries where the number of observations allows for it. A problem we must be aware of is that testing Kuznets hypothesis in a given country requires that some significant degree of structural transformation takes place in that country over the period of observation. In other words, it would be erroneous to use a country where the urban population changes by, say, 5% over the period of analysis. The data availability problem is thus compounded by the fact that we need not just countries where urban population increases considerably; but countries where inequality observations exist over the period when urban population is increasing.

Table 3 provides a list of countries ordered by the change in urban population over the period of observation of inequality. The table shows the first and last year when inequality is observed, the share of urban population in those two points in time and the increase in the urban ratio between them. We include in this list all countries where the change in urban population is at least 12%. There are 32 countries that satisfy this condition and these are the ones we will use in our empirical exercise[7]. A change of 12% is a rather small value, the type of process described by Kuznets would probably require changes of two or three times that figure. We set the limit this low because very few countries experience large changes in their urban ratios and have measures of inequality during that time. A limit of 20% would reduce the number of countries to 12 and a limit of 30% would just leave three countries.

We estimate equation (1) separately for each country and group the countries in table 4 according to the results of these regressions. Most countries fit in one of the following four cases: countries with an inverted U relationship and statistically significant coefficients, countries with a non inverted U relationship and statistically significant coefficients, and the two corresponding cases with non significant coefficients. We have also marked with a star the countries whose change in urban population is above 20% over the period of observation.

The evidence in table 4 clearly refutes the existence of a Kuznets' curve as an empirical regularity. For 25 of the 32 countries the relationship between urban population and inequality is not statistically significant, showing that as a general rule these two variables are not clearly related. Of the remaining 7 countries, 4 present a statistically significant inverted U and 3 a statistically significant non inverted U relationship.

Moreover, the 4 countries with a statistically significant inverted U relationship are all high income countries (Greece, Spain, Netherlands and France) and

[7] One additional country, Botswana, also satisfies this condition but the number of observations available for it are not sufficient to carry a regression.

the change in their urban population is below 16% in all four cases. These are certainly not cases that we would use to defend Kuznets' hypothesis. In fact, the countries that one would think of as good candidates for Kuznets' hypothesis, i.e. large developing countries experiencing important rural-urban migration such as Brazil, Turkey, China or Mexico, do not show an inverted U relationship at all. Korea may be the closest case since it experienced much change over the period of observation and its estimates would be significant at the 10% level. But it is practically the sole case that one could use in support of Kuznets' hypothesis, against many other countries contradicting it.

If we concentrate only on those countries with a change in urban population greater than 20% the evidence against Kuznets' hypothesis becomes more marked. Out of these 12 countries only 3 present an inverted U relationship and this relationship is statistically non significant in all 3 cases. Thus, the countries offering the most appropriate conditions for testing Kuznets' hypothesis tend to reject it more strongly.

Finally, we have repeated this country by country exercise using employment in Agriculture as the explanatory variable and report the results in table 5. The number of countries is reduced to 20 and there are less observations per country but the general outcome remains the same. Half of the countries present inverted U relationships and half non inverted U relationships. The majority of them, 17 out of 20, are not statistically significant.

5 Conclusion

This paper provides a test of Kuznets hypothesis that is both different and superior to the numerous tests that can be found in the literature. As we have argued, the driving factor behind Kuznets' mechanism is the progressive shift of labour from agriculture to "modern" sectors as countries develop. An empirical assessment of Kuznets' hypothesis should then concentrate on the relationship between employment in agriculture and the level of income inequality. This is precisely the approach we have taken here. Analyzing the relationship between income per capita and inequality, as the rest of the literature does, might be an interesting question by itself but is an inferior approach to test Kuznets' hypothesis. A relationship between income per capita and inequality can be the outcome of some development process other than shifts in sectorial employment, but empirical papers typically do not make such links and cite only Kuznets (1955) as their theoretical sustain. Our approach is focused on this particular mechanism stressed by Simon Kuznets and not on some unidentified and systematic effect of the growth of income on its distribution.

Our results are conclusive: there is no systematic relationship between the share of labour employed in agriculture and the level of income inequality in a country. We obtain this result both when considering all countries together in a panel regression or when analyzing each country separately. Inequality has simply failed to change in any impressive way in the very numerous developing nations that have experienced large shifts in the employment of its population throughout Latin America, Africa and Asia.

The work of Simon Kuznets has been one of the most influential in Economics in general and in the study of income distribution in particular. He opened a whole new avenue of research by making us think of the way in which the distribution of income can change as countries develop. His insights into this area were numerous and are still very worth reading today. But the particular hypothesis relating shifting sectorial employment and inequality, the most cited part of Kuznets' work in income inequality, should now be definitely abandoned. This has arguably already been done by many in the profession (see Moran 2005), but the ghost of the inverted U continues to appear persistently in development textbooks and academic papers.

Our view is that Kuznets' curve should be dismissed, but not Kuznets' ideas and even less Kuznets' careful and insightful approach to the question. It would be, however, a disappointing fact if Kuznets' curve passed from glory to oblivion without ever having been tested with the appropriate variables that its underlying mechanism suggests. We believe we have remedied this in the present paper, and hope to have offered this ingenious and attractive hypothesis the proper farewell it deserves.

References

Aghion, P. and Bolton, P. 1997, A theory of trickle-down growth and development, Review of Economic Studies 64 (2), 151-172.

Ahluwalia, M. 1976, Inequality, poverty and development, Journal of Development Economics 3, 307-342.

Anand, S. and Kanbur, S.M.R. 1993a, The Kuznets process and the inequality-development relationship, Journal of Development Economics 40, 25-52.

Anand, S. and Kanbur, S.M.R. 1993b, Inequality and Development. A Critique, Journal of Development Economics 41, 19-43.

Atkinson, A.B. and Bourguignon, F. 2000, Income Distribution and Economics, in: Atkinson, A.B. and Bourguignon, F (eds.), Handbook of Income Distribution, Elsevier North Holland.

Atkinson, A.B. and Brandolini, A. 2001, Promise and Pitfalls in the use of "secondary" data-sets: Income Inequality in OECD countries as a case study, Journal of Economic Literature 39, 771-799.

Banerjee, A.V. and Newman, A.F. 1998, Information, the dual economy and development, Review of Economic Studies 65 (4), 631-653.

Barro, R.J. 2000, Inequality and Growth in a panel of countries, Journal of Economic Growth 5 (1), 5-32.

Dawson, P.J. 1997, On testing Kuznets' economic growth hypothesis, Applied Economic Letters 4, 409-410.

Deininger, K. and Squire, L. 1996, Measuring income inequality: a new data base, The World Bank Economic Review 10, 565-591.

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

Frazer, G. 2006, Inequality and Development Across and Within Countries, World Development 34 (9), 1459-1481.

Glomm, G. and Ravikumar, B. 1998, Increasing returns, human capital and the Kuznets curve, Journal of Development Economics, 55, 353-367.

Greenwood, J. and Jovanovic, B. 1990, Financing development, growth, and the distribution of income, Journal of Political Economy 98 (5), 1076-1107.

Higgins, M. and Williamson, J.G. 1999, Explaining inequality the world round: Cohort size, Kuznets curves and openness, NBER working paper n. 7224.

Kanbur, R. 2000, Income Distribution and Development, in: Atkinson, A.B. and Bourguignon, F (eds.), Handbook of Income Distribution, Elsevier North Holland.

Kuznets, S. 1955, Economic Growth and Income Inequality, American Economic Review 45 (1), 1-28.

Lewis, W.A. 1954, Economic development with unlimited supplies of labour, Manchester School 22, 139-191.

Li, H., Squire, L. and Zou, H. 1998, Explaining international and intertemporal variations in income inequality, Economic Journal 108, 26-43.

Lindert, P.H. 2000, Three Centuries of Inequality in Britain and America, in: Atkinson, A.B. and Bourguignon, F (eds.), Handbook of Income Distribution, Elsevier North Holland.

Matyas, L., Konya, L. and Macquarie, L. 1998, The Kuznets U-curve hypothesis: some panel data evidence, Applied Economic Letters 5, 693-697.

Moran, T.P. 2005, Kuznets' Inverted U-Curve Hypothesis: The Rise, Demise, and Continual Relevance of a Socioeconomic Law, Sociological Forum 20 (2), 209-244.

Morrisson, C. 2000, Historical Perspectives on Income Distribution: The Case of Europe, in: Atkinson, A.B. and Bourguignon, F (eds.), Handbook of Income Distribution, Elsevier North Holland.

Paukert, F. 1973, Income distribution at different levels of development: A survey of evidence, International Labour Review 108, 97-125.

Table 1
Panel regressions using employment outside agriculture as explanatory factor

	All countries	*Developing countries*	*East Asia and the Pacific*	*Europe and Central Asia*	*Latin America and the Caribbean*	*Middle East and North Africa*	*South Asia*
Employment outside Agriculture	-0.5304	-0.1295	0.4189	-2.9272	-0.6931	-61.0133	3.3428
p-value	0.0	0.4276	0.0241	0.0	0.0452	0.2675	0.3181
(Employment outside Agriculture)^2	0.0040	0.0005	-0.0029	0.0189	0.0050	0.3220	-0.029
p-value	0.0	0.6743	0.1044	0.0	0.0339	0.2604	0.3954
R^2	0.9586	0.9575	0.9396	0.8883	0.9412	0.9916	0.8389
Number of observations	1583	784	200	963	270	11	24
Number of series	374	228	57	201	79	4	10

Table 2
Panel regressions using urban population as explanatory factor

	All countries	*Developing countries*	*East Asia and the Pacific*	*Europe and Central Asia*	*Latin America and the Caribbean*	*Middle East and North Africa*	*South Asia*	*Sub-Saharan Africa*
Urban Population	-0.0487	0.2759	0.8766	-0.8556	-0.5364	-0.9016	0.1200	0.4830
p-value	0.5706	0.0282	0.0	0.0001	0.0803	0.1956	0.7932	0.6256
(Urban Population)^2	0.0011	-0.0018	-0.0093	0.0070	0.0044	0.0080	0.0015	-0.0055
p-value	0.1402	0.1297	0.0	0.0	0.0631	0.1307	0.9006	0.6934
R^2	0.9420	0.9429	0.9185	0.8634	0.9144	0.7521	0.7932	0.8355
Number of observations	2602	1442	336	1417	402	48	112	132
Number of series	534	357	71	261	103	16	24	45

Table 3
Countries with largest change in urban population

Country	*Year 0 (earliest observation of inequality)*	*Year 1 (last observation of inequality)*	*Urbanization rate, year 0*	*Urbanization rate, year 1*	*Change in urbanization rate*
Korea	1961	1998	28.64	79.04	50.4
Brazil	1960	2001	44.9	81.8	36.9
Bulgaria	1960	2002	37.1	69.34	32.24
Turkey	1968	2000	36.6	64.7	28.1
Philippines	1961	2000	30.56	58.5	27.94
Malaysia	1970	1999	33.5	60.56	27.06
Costa Rica	1961	2000	34.6	59	24.4
Puerto Rico	1963	1989	48.7	71.6	22.9
Mexico	1963	2002	53.26	75.22	21.96
China	1964	2003	17.28	38.56	21.28
Indonesia	1976	1999	19.86	40.72	20.86
Japan	1962	1998	44.82	64.96	20.14
Morocco	1960	1991	29.3	49.1	19.8
Nigeria	1971	1996	21.8	40.38	18.58
Botswana	1986	1994	29.74	47.58	17.84
Bangladesh	1963	2000	5.76	23.2	17.44
Greece	1960	2001	42.9	58.84	15.94
Finland	1966	2003	45.18	61.1	15.92
Bahamas	1973	1993	69.94	85.46	15.52
Spain	1965	2002	61.3	76.46	15.16
Colombia	1970	2000	56.6	71.2	14.6
Venezuela	1976	2000	76.52	91.1	14.58
Ukraine	1968	2002	53.12	67.38	14.26
Poland	1960	2002	47.9	61.86	13.96
Israel	1961	2001	77.62	91.44	13.82
Dominican Rep.	1976	1998	46.82	60.56	13.74
Netherlands	1977	2001	63.8	77.48	13.68
Tunisia	1975	2000	49.9	63.4	13.5
Hong Kong	1966	1996	86.66	100	13.34
Belarus	1981	2002	57.56	70.88	13.32
Panama	1989	2000	53.54	65.8	12.26
France	1962	2002	63.98	76.16	12.18
Chile	1968	2000	73.8	85.9	12.1

Table 4
Country by country regressions using urban population as explanatory factor.

Country	Urban population	p-value	(Urban population)^2	p-value	R^2	Number of observations	Number of series
Inverted U relationships, statistically significant							
Greece	4.867	0.009	-0.048	0.011	0.947	25	4
Spain	10.086	0.000	-0.073	0.000	0.922	55	14
Netherlands	5.778	0.001	-0.040	0.001	0.908	45	6
France	8.933	0.006	-0.071	0.003	0.986	27	5
Inverted U relationships, statistically non significant							
Korea*	0.786	0.096	-0.009	0.066	0.474	31	8
Puerto Rico*	0.069	0.985	-0.002	0.939	0.787	10	4
Japan*	1.639	0.278	-0.016	0.262	0.698	31	4
Morocco	3.406	0.390	-0.042	0.433	0.945	6	2
Nigeria	2.448	0.621	-0.035	0.634	0.875	16	6
Bahamas	29.271	0.052	-0.179	0.057	0.744	10	3
Ukraine	1.476	0.910	-0.013	0.902	0.586	36	8
Dominican Republic	1.940	0.529	-0.021	0.462	0.856	11	4
Belarus	4.768	0.099	-0.038	0.086	0.944	35	7
Non inverted U relationships, statistically non significant							
Brazil*	-0.808	0.720	0.006	0.676	0.234	37	8
Bulgaria*	-0.543	0.698	0.006	0.646	0.840	63	9
Turkey*	-1.425	0.346	0.010	0.441	0.949	11	5
Malaysia*	-0.279	0.717	0.002	0.797	0.740	18	6
China*	-0.028	0.982	0.020	0.346	0.808	32	6
Indonesia*	-1.089	0.184	0.017	0.214	0.809	19	4
Bangladesh	-1.418	0.133	0.067	0.048	0.655	29	7
Finland	-4.906	0.061	0.044	0.070	0.876	96	7
Venezuela	-9.878	0.075	0.060	0.072	0.933	51	9
Poland	-12.208	0.058	0.107	0.059	0.828	74	14
Israel	-12.585	0.237	0.080	0.207	0.795	14	3
Tunisia	-1.077	0.638	0.007	0.731	0.860	7	2
Hong Kong	-52.184	0.207	0.283	0.219	0.680	10	4
Panama	-0.906	0.743	0.008	0.722	0.631	15	4
Chile	-12.164	0.294	0.077	0.269	0.779	39	8
Non inverted U relationships, statistically significant							
Philippines*	-1.820	0.001	0.021	0.001	0.811	33	8
Costa Rica*	-2.168	0.007	0.021	0.009	0.614	28	6
Mexico*	-3.258	0.012	0.028	0.007	0.673	55	14
Monotonous relationship, statistically non significant							
Colombia	-0.061	0.996	-0.004	0.970	0.826	25	8

Table 5
Country by country regressions using employment outside agriculture as explanatory factor.

Country	*Urban population*	*p-value*	*(Urban population)^2*	*p-value*	R^2	*Number of observations*	*Number of series*
Inverted U relationships, statistically significant							
Netherlands	178.561	0.005	-0.931	0.005	0.909	43	6
Inverted U relationships, statistically non significant							
Korea	0.363	0.889	-0.005	0.777	0.717	15	6
Brazil	3.164	0.565	-0.018	0.622	0.694	30	7
Bulgaria	22.540	0.101	-0.149	0.092	0.893	36	7
Philippines	3.583	0.418	-0.029	0.459	0.773	23	7
Mexico	55.416	0.126	-0.344	0.127	0.821	14	6
Japan	34.287	0.321	-0.181	0.330	0.935	10	3
Greece	3.011	0.416	-0.022	0.366	0.749	11	3
Spain	0.247	0.915	-0.002	0.869	0.915	47	13
Poland	2.039	0.632	-0.009	0.744	0.898	66	13
Non inverted U relationships, statistically non significant							
Turkey	-0.358	0.950	0.001	0.988	0.988	6	3
Malaysia	-7.149	0.105	0.047	0.104	0.894	12	5
Costa Rica	-7.080	0.309	0.047	0.306	0.914	23	6
China	-0.249	0.874	0.009	0.644	0.697	15	4
Venezuela	-24.625	0.162	0.140	0.166	0.935	46	8
Israel	-61.013	0.268	0.322	0.260	0.992	10	3
Dominican Republic	-59.685	0.124	0.363	0.124	0.850	7	3
Panama	-3.939	0.354	0.026	0.346	0.661	15	4
Non inverted U relationships, statistically significant							
Finland	-23.376	0.000	0.134	0.000	0.976	85	7
Chile	-50.802	0.002	0.310	0.001	0.745	37	7