Journal of Development Economics 40 (1993) 25–52. North-Holland

The Kuznets process and the inequality–development relationship*

Sudhir Anand

St. Catherine's College, Oxford, UK

S.M.R. Kanbur

The World Bank, Washington DC, USA

Received December 1985, final version received January 1992

In his classic paper on economic growth and income inequality, Kuznets discussed the process of population shift from traditional to modern activities as the basis for a theory of distributional change during the course of development. In this paper, we present a formalization of the Kuznets process, conduct a general analysis of distributional change under this process, and derive the functional forms of, and conditions for a turning point in, the inequality–development relationship for six commonly used indices of inequality. The functional form appropriate to each index is then estimated using cross-section data on 60 developing and developed countries. Finally, some extensions to the initial formalization of the Kuznets process are considered.

1. Introduction

In a classic paper, Kuznets (1955) introduced the idea of a link between inequality and development. Based on evidence from time-series data on England, Germany and the United States, he hypothesized the now famous 'inverse-U' relationship between inequality and development. The *mechanisms* underlying this relationship were also discussed by Kuznets. He took a somewhat eclectic view, stressing economic, political and social factors as explanations of the statistical regularities he had observed. But the foremost of these factors, one which provided the focus of Kuznets's analysis and has

Correspondence to: Sudhir Anand, St. Catherine's College, Oxford OX1 3UJ, UK.

*We are grateful to the Economic and Social Research Council of the UK for financial assistance under grant no. B 0023 0001, and to Afsaneh Farzin and George Zanias for research assistance. This is a revised version of our paper 'Intersectoral Shifts Theory and the Inequality–Development Relationship', first presented at the EADI Symposium on Income Distribution at the University of Paderborn, West Germany in April 1981. It has also been presented at seminars at the Universities of Oxford, Cambridge, London (University College and LSE), Princeton and Essex, and was published as University of Essex, Department of Economics, Discussion Paper no. 249, August 1984.

become important in the recent literature, is the shift of population from traditional to modern activities. This process of population shift, together with a formalization of what he regarded as the 'stylized facts' of economic development, allowed Kuznets to derive predictions of the behaviour of inequality during the course of development:

> 'An invariable accompaniment of growth in developed countries is the shift away from agriculture, a process usually referred to as industrialization and urbanization. The income distribution of the total population, in the simplest model, may therefore be viewed as a combination of the income distributions of the rural and of the urban populations. What little we know of the structures of these two component income distributions reveals that: (a) the average per capita income of the rural population is usually lower than that of the urban; (b) inequality in the percentage shares within the distribution for the rural population is somewhat narrower than in that for the urban population. ... Operating with this simple model, what conclusions do we reach? First, all other conditions being equal, the increasing weight of urban population means an increasing share for the more unequal of the two component distributions. Second, the relative difference in per capita income between the rural and urban populations does not necessarily drift downward in the process of economic growth: indeed, there is some evidence to suggest that it is stable at best, and tends to widen because per capita productivity in urban pursuits increases more rapidly than in agriculture. If this is so, inequality in the total income distribution should increase' [Kuznets (1955, pp. 7–8)].

This underlying model of population shift may be termed the Kuznets process. It appears to have inspired much of the recent empirical literature on inequality and development [see Paukert (1973), Adelman and Morris (1973), Ahluwalia (1974, 1976a, b), Ahluwalia et al. (1979), Lydall (1977), Loehr (1981), etc.]. The focus of this literature has been the cross-country estimation of a relationship between inequality and per capita national income, but it has appealed to the Kuznets process to justify: (i) a turning point in the relationship between inequality and development, with inequality first increasing and then decreasing; and (ii) the specific functional form of the inequality–development relationship [see, for example, the influential papers of Ahluwalia (1976b) and Ahluwalia et al. (1979)].

Kuznets's original analysis was based solely on a numerical illustration. Fields (1979) attempts to analyze the Kuznets process in the absence of within-sector inequality, but Robinson (1976) is the only paper we know of which tries to investigate the properties of the Kuznets process in the presence of within-sector inequality. His object is to:

> '... demonstrate that the U-hypothesis can be derived from a very simple model with a minimum of economic assumptions. One need only assume that the economy can be divided into two sectors with different sectoral income distributions and that there is a monotonic increase in the relative population of one of the sectors over time. ... The U-hypothesis will be seen to be a necessary implication of the workings of such models' [Robinson (1976, p. 437)].

However, Robinson's analysis is restricted to one specific index of inequality – the variance of the logarithm of income. There is no indication how inequality in general, and alternative indices of inequality in particular, might be expected to change during the course of development, and under what conditions the U-hypothesis for alternative indices would be 'a necessary implication' of the workings of the Kuznets process. Furthermore, he does not test the model by estimating the derived relationship, but rather illustrates it by simulation.

Our object in the present paper is to: investigate further the formal properties of the Kuznets process; use the model to generate estimating equations for the inequality–development relationship for alternative inequality indices; and provide estimates of these equations using the cross-section data set of Ahluwalia (1976b), which forms the basis for the projections of inequality and poverty by Ahluwalia et al. (1979).

The plan of this paper is as follows. In section 2 we present a formal statement of Kuznets's (1955) discussion of population shift in the course of development, and derive *general* properties of the relation between inequality and development. Section 3 specializes the discussion to six alternative indices of inequality; for each index we derive the specific functional form of and condition for a turning point in its inequality–development relationship. Section 4 provides estimates of these functional forms for cross-section data on 60 developing and developed countries. Section 5 presents some implications of generalizing the definition of the Kuznets process. Section 6 concludes the paper.

2. General properties of the Kuznets process

Let the economy be divided into two sectors indexed by $i=1,2$. Denote by $x(t)$ the fraction of the total population in Sector 1 at time t; $[1-x(t)]$ is thus the fraction of the total population in Sector 2. We assume that Sector 1 is the modern (urban/industrial/advanced) sector of the economy, while Sector 2 is the traditional (rural/agricultural/backward) sector. In keeping with Kuznets's analysis, an increase in x over time is taken to be synonymous with development.

Letting y denote income, the cumulative income distribution in Sector i

can be written quite generally as $F_i(y, x(t), t)$ and its associated income density function as $f_i(y, x(t), t)$. The cumulative national income distribution $F(y, x(t), t)$ is then given by

$$F(y, x(t), t) = x(t)F_1(y, x(t), t) + [1 - x(t)]F_2(y, x(t), t). \tag{2.1}$$

As stated, expression (2.1) is simply an identity. What turns it into a theory, from which predictions can be derived, is the imposition of restrictions on F_1 and F_2. In the simplest version of the Kuznets process we restrict F_1 and F_2 to be independent of x and t. Hence F and its associated density function f may be written as

$$F(y, x(t)) = x(t)F_1(y) + [1 - x(t)]F_2(y), \tag{2.2}$$

$$f(y, x(t)) = x(t)f_1(y) + [1 - x(t)]f_2(y). \tag{2.3}$$

This process is, of course, somewhat special. The sectoral distributions are assumed to remain unchanged over time, which in turn implies that the sectoral mean incomes and inequality levels are constant. In fact, we would expect the mean income of the urban sector to increase relative to that of the rural sector, at least in the early stages of development; this case is analyzed in section 5. The assumption that F_1 and F_2 are independent of x implies a migration process in which a representative subsample of people from the rural sector moves to the urban sector and relocates itself representatively there.

The general statement of distributional change in $F(y, x)$ as x changes is given by

$$\partial F(y, x)/\partial x = F_1(y) - F_2(y). \tag{2.4}$$

Some general results can be derived immediately from (2.4). Hence,

$$F_1(y) \leqq F_2(y) \quad \text{for} \quad \text{all } y$$

implies

$$\partial F(y, x)/\partial x \leqq 0 \quad \text{for} \quad \text{all } y. \tag{2.5}$$

Thus if the urban sector distribution first-order dominates the rural sector distribution, then an increase in x leads to a first-order dominating shift in the national distribution. Such a change would be preferred by all social welfare functions which are increasing in individual incomes [see Hadar and Russell (1969)].

Similarly, a result on second-order dominance can be derived by integrating (2.4) to give

$$\partial\left\{\int_0^y F(w,x)\,\mathrm{d}w\right\}/\partial x=\int_0^y [F_1(w)-F_2(w)]\,\mathrm{d}w. \tag{2.6}$$

Thus if the urban sector distribution second-order dominates the rural sector distribution, i.e.

$$\int_0^y [F_1(w)-F_2(w)]\,\mathrm{d}w \leqq 0 \quad \text{for} \quad \text{all } y,$$

then

$$\partial\left\{\int_0^y F(w,x)\,\mathrm{d}w\right\}/\partial x \leqq 0 \quad \text{for} \quad \text{all } y, \tag{2.7}$$

so that an increase in x leads to a second-order dominating shift in the national distribution. As Hadar and Russell (1969) have shown, these shifts are preferred by all social welfare functions which are increasing, symmetric and quasi-concave in individual incomes.

It should be clear that the above results on first- and second-order dominance can be extended to third- and higher-order dominance. However, such results are all that can be derived at this level of generality. If the two sectoral distributions do not satisfy a dominance relation, then we cannot make general statements about the behaviour of the national distribution as development proceeds. But if we make more specific assumptions about the sectoral distributions, the behaviour of the national distribution can be analyzed in more detail. Thus, let μ_i denote the mean income, and I_i the value of the inequality index, in Sector i. The stylized facts of economic development (see the Kuznets quote in the introduction) may then be represented as the two Kuznets conditions:

$$\mu_1>\mu_2, \quad \text{and} \quad I_1>I_2. \tag{2.8}$$

The overall mean income μ is simply

$$\mu=x\mu_1+(1-x)\mu_2 \tag{2.9}$$

and under condition (2.8) national income per capita increases with x,

$$\partial\mu/\partial x=\mu_1-\mu_2>0. \tag{2.10}$$

What about the behaviour of overall inequality? With the above formalization of the Kuznets process we can ask, and answer, a number of precise questions about how inequality behaves as development proceeds. Does I

increase with μ at the start of the development process? Is there a turning point in the relationship between I and μ? What is the exact functional form of the relationship between I and μ? We begin by considering the behaviour of the Lorenz curve.

2.1. *Behaviour of the Lorenz curve*

One of the most influential estimations of the inequality–development relationship is that by Ahluwalia (1976b). This work has formed the basis, for example, of the poverty projections in Ahluwalia et al. (1979) and in the World Bank's *World Development Reports* (1978, 1979, 1980). Ahluwalia uses as his (in-)equality index the income share of the lowest 40 percent of the population. This index is obviously very special and is completely determined by a *single* point on the Lorenz curve of the income distribution. Here we investigate the behaviour of the *entire* Lorenz curve under the Kuznets process.

The Lorenz curve of an income distribution shows the percentage of total income received by the lowest ($100p$) percent of the population as p varies from 0 to 1. Define z as the income level which cuts off the lowest ($100p$) percent of the *national* population. Then z is clearly a function of p but it is also a function of x, since a change in x will generally change the national income distribution and hence the income level corresponding to p. The function $z(p,x)$ is defined as the solution to the equation

$$p=F(z,x)=xF_1(z)+(1-x)F_2(z). \tag{2.11}$$

The *income share* of the lowest ($100p$) percent of the population is a function of z, and thus of p and x. Denote this function as $L(p,x)$. The graph of $L(p,x)$ as a function of p is the Lorenz curve of the national income distribution. We can write $L(p,x)$ in terms of the sectoral Lorenz curves $L_1(.)$ and $L_2(.)$ using (2.3):

$$\begin{aligned} L(p,x)&=\frac{1}{\mu}\int_0^{z(p,x)} yf(y)\,\mathrm{d}y \\ &=\frac{1}{\mu}\left\{x\int_0^{z(p,x)} yf_1(y)\,\mathrm{d}y+(1-x)\int_0^{z(p,x)} yf_2(y)\,\mathrm{d}y\right\} \\ &=\frac{x\mu_1}{\mu}L_1(F_1(z))+\frac{(1-x)\mu_2}{\mu}L_2(F_2(z)), \end{aligned} \tag{2.12}$$

where $L_i(F_i(z))$ is the income share of the lowest ($100F_i(z)$) percent of the

Sector i population. Thus the Lorenz curve of the national distribution evaluated at percentile p is a weighted average of the Lorenz curves of the sectoral distributions evaluated *not* at p but at $F_1(z)$ and $F_2(z)$, respectively – the weights being the income shares of the two sectors in total income. Note that $F_i(z)$ will be identically equal to p if and only if $F_1(y) \equiv F_2(y)$, i.e. the two sectoral distributions are *identical.*

In general the lowest $(100p)$ percent of the national population will *not* be composed of the lowest $(100p)$ percent of Sector 1's population and the lowest $(100p)$ percent of Sector 2's population. In fact, from (2.11), either

$$F_1(z) < p < F_2(z)$$

or (2.13)

$$F_1(z) > p > F_2(z),$$

so that the lowest $(100p)$ percent of the national population will *either* contain less than the lowest $(100p)$ percent of the Sector 1 population and correspondingly more of the Sector 2 population, *or* vice versa. If we wish to express $L(p,x)$ in terms of $L_1(p)$ and $L_2(p)$, adjustment factors will have to be introduced to eq. (2.12). Thus $L(p,x)$ may be written as

$$\begin{aligned} L(p,x) = {} & \frac{x\mu_1}{\mu} L_1(p) + \frac{x\mu_1}{\mu}[L_1(F_1(z)) - L_1(p)] \\ & + \frac{(1-x)\mu_2}{\mu} L_2(p) + \frac{(1-x)\mu_2}{\mu}[L_2(F_2(z)) - L_2(p)], \end{aligned} \tag{2.14}$$

where the adjustment factors represented by the second and fourth terms are of opposite sign because of the inequalities (2.13) and the fact that the Lorenz curve is monotonic-increasing. Since their relative magnitudes are not known in general, there seems to be no obvious relationship between $L(p,x)$ and $[(x\mu_1/\mu)L_1(p) + ((1-x)\mu_2/\mu)L_2(p)]$.

To investigate the behaviour of $L(p,x)$ as a function of x, we fix p and examine the partial derivative of $L(p,x)$ with respect to x. Thus, for $p=0.4$ we will obtain the behaviour of the income share of the lowest 40 percent as a function of x. The expression (2.14) for $L(p,x)$ can be rewritten with the adjustment factors expressed as integral terms:

$$\begin{aligned} L(p,x) = {} & \frac{x\mu_1}{\mu} L_1(p) + \frac{(1-x)\mu_2}{\mu} L_2(p) \\ & + \frac{x}{\mu} \int_{F_1^{-1}(p)}^{z} y f_1(y)\,\mathrm{d}y + \frac{(1-x)}{\mu} \int_{F_2^{-1}(p)}^{z} y f_2(y)\,\mathrm{d}y. \end{aligned} \tag{2.15}$$

Differentiating this partially with respect to x, we get

$$\frac{\partial L(p,x)}{\partial x}=\frac{\mu_1\mu_2}{\mu^2}[L_1(p)-L_2(p)]$$

$$+\frac{\mu_1}{\mu^2}\left\{\int_{F_2^{-1}(p)}^{z} F_2(y)\,\mathrm{d}y+pF_2^{-1}(p)-zp\right\}$$

$$-\frac{\mu_2}{\mu^2}\left\{\int_{F_1^{-1}(p)}^{z} F_1(y)\,\mathrm{d}y+pF_1^{-1}(p)-zp\right\}. \tag{2.16}$$

In order to sign this expression it is obvious that some restrictions must be imposed on the underlying distributions $F_1(y)$ and $F_2(y)$. The second Kuznets condition in (2.8) requires inequality in Sector 1 to be greater than in Sector 2. Since inequality here is represented in terms of Lorenz dominance, the second Kuznets condition can be written as

$$L_2(p)\geqq L_1(p) \tag{2.17}$$

for all p between 0 and 1 (with strict inequality for some p). The characterization in terms of Lorenz dominance corresponds to Kuznets's 'inequality in the percentage shares within the distribution for the rural population (being) somewhat narrower than in that for the urban population'.

We are now in a position to investigate the implications of Kuznets's assumptions for the behaviour of inequality as x increases. In the case where $\mu_1=\mu_2$, the shift in population to the more unequal sector simply leads to a mean-preserving spread in the overall distribution [see Atkinson (1970) and Rothschild and Stiglitz (1970)], and hence leads to a distribution which is more unequal in the Lorenz sense.[1]

In the case where $\mu_1>\mu_2$, the overall mean μ increases with a shift of population to Sector 1. Looking at the expression for $\partial L(p,x)/\partial x$ it is clear that we cannot sign it unambiguously for *general* distributions F_1 and F_2 which satisfy the Kuznets conditions. However, we can investigate the nature of inequality change at the *start* of the development process, $x=0$. At $x=0$, we have $\mu=\mu_2$, $p=F_2(z)$, and hence $z=F_2^{-1}(p)$. Substituting in (2.16), it immediately follows that

[1]This can be confirmed by putting $\mu_1=\mu_2$ in (2.16), rearranging and using the condition that $L_2(p)\geqq L_1(p)$ *for all* p between 0 and 1, not simply the p at which the derivative $\partial L(p,x)/\partial x$ is being evaluated.

$$\frac{\partial L(p,0)}{\partial x}=\frac{\mu_1}{\mu_2}[L_1(p)-L_2(p)]$$

$$-\frac{1}{\mu_2}\left\{\int_{F_1^{-1}(p)}^{F_2^{-1}(p)} F_1(y)\,\mathrm{d}y-p[F_2^{-1}(p)-F_1^{-1}(p)]\right\}, \tag{2.18}$$

which is non-positive by the second Kuznets assumption in (2.8) and the fact that $F_1(y)$ is a monotonic increasing function of y. Thus the entire Lorenz curve for the national income distribution moves *down* as x increases from 0 up, i.e. there is an unambiguous increase in inequality. Hence the share of the lowest $(100p)$ percent decreases *for all* p in the interval $(0, 1)$ at the start of the development process. In particular, the share of the lowest 40 percent decreases at the start of the development process.

We can also attempt to answer the specific question about the direction of inequality change at the *end* of the development process, $x=1$. At $x=1$, we have $\mu=\mu_1$, $p=F_1(z)$, and hence $z=F_1^{-1}(p)$. The expression (2.16) therefore reduces to

$$\frac{\partial L(p,1)}{\partial x}=\frac{\mu_2}{\mu_1}[L_1(p)-L_2(p)]$$

$$+\frac{1}{\mu_1}\left\{\int_{F_2^{-1}(p)}^{F_1^{-1}(p)} F_2(y)\,\mathrm{d}y-p[F_1^{-1}(p)-F_2^{-1}(p)]\right\}. \tag{2.19}$$

The first term on the right-hand side is non-positive by the second Kuznets assumption, while the second term is non-negative since $F_2(y)$ is a monotonic increasing function of y. Thus, in general, the sign of (2.19) is ambiguous.

While at the start of the Kuznets process inequality increases unambiguously, in the sense that the new national distribution is Lorenz dominated by the old one, there is an ambiguity at the end of the process. These results allow us to characterize the behaviour of inequality at $x=0$ and $x=1$ for the entire Lorenz class of inequality indices, since Lorenz dominance is a sufficient condition for inequality as measured by *any* of these indices to be lower.[2] This class of indices includes almost all of the commonly used measures of inequality, and all but one of the indices studied in the next section (the exception is the variance of log-income). For all these indices therefore, assuming (2.17), there will be an ambiguity in the behaviour of inequality at the end of the development process. One purpose of the next section is to state the precise conditions that resolve this ambiguity.

[2]See Anand (1983, appendix D). An inequality index belongs to the Lorenz class if it satisfies symmetry, mean independence, population-size independence, and the Pigou–Dalton condition.

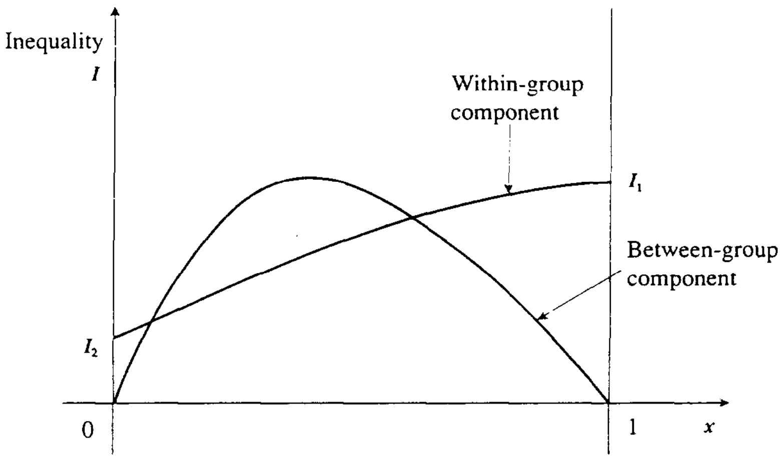

Fig. 1

An alternative insight into the behaviour of inequality at the beginning and the end of the Kuznets process can also be obtained by considering a decomposition of inequality into within- and between-sector components. Define between-sector inequality as the value of the index when everyone in a sector receives the mean income of that sector. In other words, it is the inequality in the hypothetical distribution where a fraction x of the total population receive income μ_1 and a fraction $(1-x)$ receive income μ_2. It is clear from its definition that the between-group component of inequality must be *zero* both at $x=0$ and at $x=1$, and it must be *positive* elsewhere (by definition of inequality). It must therefore have a *negative* slope at $x=1$, a positive slope at $x=0$, and a turning point as a function of x (see fig. 1). Define the within-sector component of inequality as the difference between overall inequality and the between-sector component as defined above. If the within-sector component of inequality can be written as a weighted sum of the sectoral inequality indices, we refer to the index as being *weakly decomposable*. If the weights are sectoral population shares or income shares, which sum to unity, we call the index *decomposable*. If the weights turn out to be simply the *population shares* of the two sectors, we refer to the index as *strictly decomposable*.[3]

For indices which are *decomposable* the within-sector component increases

[3]For a discussion of decomposability as defined here, see Anand (1983). Decomposability is, of course, neither necessary nor sufficient for an index to belong to the Lorenz class. The Atkinson index belongs to the Lorenz class but is not decomposable; the same holds true for the Gini coefficient. On the other hand, the variance of income is strictly decomposable, while the variance of log-income is decomposable around sectoral *geometric* means, but neither belongs to the Lorenz class: the first does not satisfy mean independence, the second does not satisfy the Pigou–Dalton condition.

with x since the weight on the more unequal distribution increases (and that on the less unequal distribution decreases) whether the weights are population shares *or* income shares.[4] Combining this with the argument that the between-sector component also increases with x at $x=0$, we arrive at the conclusion that for decomposable indices inequality must increase at the start of the Kuznets process. Further, since the within-sector component is (still) increasing at $x=1$, while the between-sector component must be decreasing at $x=1$, the overall behaviour of these indices at the end of the Kuznets process is in general *ambiguous*. We must derive particular conditions at $x=1$ to resolve the ambiguity. For indices that are only *weakly* decomposable, so that the weights do not sum to unity or to even a constant number, it needs to be checked by direct calculation whether or not the behaviour of the within-sector component follows the same pattern as that for decomposable indices. Thus, for example, it can be shown that for the squared coefficient of variation the within-sector component is indeed increasing at both $x=0$ and $x=1$.[5]

To end this section, we note one implication of our analysis for Ahluwalia's (1976b) use of the Kuznets process to justify a turning point in the relationship between the share of the lowest 40 percent and per capita national income. From our analysis of $\partial L(p,x)/\partial x$, letting $p=0.4$, we have shown that the income share of the lowest 40 percent does *not* necessarily have a turning point as a function of x, and hence as a function of μ. Still less can be asserted about the functional form relating this index to the level of per capita income – see eq. (2.15). Substituting for x from (2.9) into (2.15), it is clear that the share of the lowest 40 percent will *not* for general F_1, F_2 have the *log-quadratic* form in μ used by Ahluwalia (1976b) in his empirical work.[6] The next section provides, for general F_1, F_2, the functional forms generated by the Kuznets process for six inequality indices.

3. The inequality–development relationship for six inequality indices

The previous section attempted a general characterization of distributional change under the Kuznets process. We showed that under the Kuznets

[4]Let the income share of Sector 1 be denoted s so that the income share of Sector 2 is $(1-s)$. Hence, $s=(x\mu_1)/[x\mu_1+(1-x)\mu_2]$, therefore $(\partial s/\partial x)=(\mu_1\mu_2)/\mu^2>0$. Notice that this does not depend upon $\mu_1>\mu_2$ (but merely on μ_1 and μ_2 being positive). When it is a population-weighted average of sectoral inequality coefficients, the within-sector component increases linearly with x from I_2 to I_1. But when the within-sector component is an income-weighted average of the sectoral inequality coefficients, it increases non-linearly with x from I_2 to I_1. Since $\partial^2 s/\partial x^2<0$, the within-sector component in this case is a concave function of x.

[5]It can be demonstrated that $2\mu_2>\mu_1$ is a sufficient condition for the within-sector component of the squared coefficient of variation to be increasing in x for *all* values of x between 0 and 1.

[6]It is an open question as to what restrictions are implied on F_1 and F_2 if $L(0.4,x)$ *were* to have a log-quadratic form in μ.

assumptions the entire Lorenz class of inequality indices increases at the start of the process, but the behaviour of inequality at the end of the process – and hence the existence of a turning point in the inequality–development relation – is ambiguous. We also showed that the Kuznets process does *not* imply a closed functional form between the share of the lowest $(100p)$ percent of the population and overall mean income, μ. In order to derive specific functional forms which can be estimated, we have to specify the inequality index under analysis more closely. The object of this section is to derive the functional form of the inequality–development relation for six commonly used inequality indices, and the conditions for a turning point in this relation. The next section estimates these functional forms on data from 60 developing and developed countries.

The six inequality indices we consider are: (i) Theil's entropy index T, which was developed by Theil (1967) using concepts from information theory; (ii) Theil's second measure L, which is the logarithm of the ratio of the arithmetic to the geometric mean income of the distribution [see Theil (1967)]; (iii) the squared coefficient of variation S^2, which was used, for example, by Swamy (1967) in his analysis of inequality change in India; (iv) a decomposable transform of the Atkinson (1970) inequality index $I(\varepsilon)$, where ε is the inequality aversion parameter – the transform being $[1-I(\varepsilon)]^{1-\varepsilon}$; (v) the Gini coefficient G, for the case in which the sectoral distributions are non-overlapping; and (vi) the variance of log-income σ^2, which was analyzed by Robinson (1976). Table 1 presents the functional form of the inequality–development relationship, and the condition for a turning point in this relationship, for each of these six indices.

The first five indices in table 1 all belong to the Lorenz class of indices. Moreover, they are *aggregable*, in the sense that overall inequality can be written as a function of the sectoral means, the sectoral inequalities, and the sectoral population share(s):

$$I=I(\mu_1,\mu_2,I_1,I_2,x).$$

From our general analysis of the behaviour of the Lorenz class of indices (see previous section) we know that $[\partial I/\partial x]_{x=0}>0$, i.e. that inequality will increase at the start of the development process. A sufficient condition for the existence of a turning point is therefore that $[\partial I/\partial x]_{x=1}<0$. For aggregable indices the above condition can be stated explicitly as a condition on μ_1, μ_2, I_1, I_2. This is precisely the condition reported in the last column of table 1, where we have expressed the condition in terms of the *ratio* of the sectoral means, $\theta=(\mu_1/\mu_2)$. Note that $\theta>1$ by the first Kuznets condition in (2.8).

Since the overall mean is given by $\mu=[x\mu_1+(1-x)\mu_2]$, we can substitute $x=(\mu-\mu_2)/(\mu_1-\mu_2)$ in $I=I(\mu_1,\mu_2,I_1,I_2,x)$ to get an explicit relationship between I and μ, viz. the inequality–development relationship. For *aggreg-*

Table 1

Six inequality indices: Functional form of and condition for turning point in inequality–development relationship.

Inequality index	Functional form of the inequality–development relationship	Condition for turning point
Theil T	$T=A+B(1/\mu)+C\log\mu$ where $A=[\mu_1(T_1+\log\mu_1)-\mu_2(T_2+\log\mu_2)]/(\mu_1-\mu_2)$ $B=\mu_1\mu_2(T_2-T_1+\log\mu_2-\log\mu_1)/(\mu_1-\mu_2)$ $C=-1$	$(T_1-T_2)<(\theta-1-\log\theta)$
Theil L	$L=A+B\mu+C\log\mu$ where $A=(\mu_1L_2-\mu_2L_1+\mu_2\log\mu_1-\mu_1\log\mu_2)/(\mu_1-\mu_2)$ $B=(L_1-L_2-\log\mu_1+\log\mu_2)/(\mu_1-\mu_2)$ $C=1$	$(L_1-L_2)<[(1/\theta)-1+\log\theta]$
Squared coefficient of variation, S^2	$S^2=A+B(1/\mu)+C(1/\mu)^2$ where $A=-1$ $B=[\mu_1^2(S_1^2+1)-\mu_2^2(S_2^2+1)]/(\mu_1-\mu_2)$ $C=\mu_1\mu_2[\mu_2(S_2^2+1)-\mu_1(S_1^2+1)]/(\mu_1-\mu_2)$	$S_1^2[(2/\theta)-1]-(1/\theta)^2S_2^2$ $<[(1/\theta)-1]^2$
Decomposable transform of Atkinson Index, $[1-I(\varepsilon)]^{1-\varepsilon}$	$[1-I(\varepsilon)]^{1-\varepsilon}=A+B\mu^{\varepsilon}+C\mu^{\varepsilon-1}$ where $A=0$ $B=[\mu_1^{1-\varepsilon}(1-I_1)^{1-\varepsilon}-\mu_2^{1-\varepsilon}(1-I_2)^{1-\varepsilon}]/(\mu_1-\mu_2)$ $C=[\mu_1\mu_2^{1-\varepsilon}(1-I_2)^{1-\varepsilon}-\mu_2\mu_1^{1-\varepsilon}(1-I_1)^{1-\varepsilon}]/(\mu_1-\mu_2)$	<u>Case $0<\varepsilon<1$</u>: $[(1-I_2)/(1-I_1)]^{1-\varepsilon}$ $<[\varepsilon\theta+(1-\varepsilon)]/\theta^{\varepsilon}$ <u>Case $\varepsilon>1$</u>: $[(1-I_2)/(1-I_1)]^{1-\varepsilon}$ $>[\varepsilon\theta+(1-\varepsilon)]/\theta^{\varepsilon}$
Gini coefficient, G with non-overlapping sectoral distributions	$G=A+B\mu+C(1/\mu)$ where $A=[(\mu_1^2-\mu_2^2)-2\mu_1\mu_2(G_1+G_2)]/(\mu_1-\mu_2)^2$ $B=[\mu_1(1-G_1)-\mu_2(1+G_2)]/(\mu_1-\mu_2)^2$ $C=\mu_1\mu_2[\mu_1(1-G_2)-\mu_2(1+G_1)]/(\mu_1-\mu_2)^2$	$(1+G_1)/(1-G_1)<\theta$
Variance of log-income, σ^2	$\sigma^2=A+B\mu+C\mu^2$ where $A=[(\mu_1-\mu_2)(\mu_1\sigma_2^2-\mu_2\sigma_1^2)-\mu_1\mu_2(m_1-m_2)^2]/(\mu_1-\mu_2)^2$ $B=[(\mu_1-\mu_2)(\sigma_1^2-\sigma_2^2)+(\mu_1+\mu_2)(m_1-m_2)^2]/(\mu_1-\mu_2)^2$ $C=-(m_1-m_2)^2/(\mu_1-\mu_2)^2$	$(\sigma_1^2-\sigma_2^2)/(m_1-m_2)^2<1$

able indices this can always be written in closed form, and the exact relationship is stated in the second column of table 1. It should be emphasized that the Gini coefficient is aggregable only when the sectoral distributions are non-overlapping [Anand (1983, pp. 321–322)], and both its functional form and its turning point condition are valid only under this restriction.

The sixth index, the variance of log-income σ^2, is not a member of the Lorenz class of indices, nor is it aggregable with respect to *arithmetic* mean incomes. However, it is aggregable with respect to the means of log-income, m_1 and m_2. Inequality can be shown to increase at the start of the process, and a condition derived for the turning point which involves geometric (rather than arithmetic) sectoral means. This was done by Robinson (1976) and is included in table 1 for comparison. Also, an explicit functional form between σ^2 and μ has been derived, involving coefficients which depend on sectoral arithmetic and geometric means as well as sectoral inequalities.

The detailed analysis underlying table 1 is illustrated below for the case of the Theil T index. This index can be decomposed into the sum of within- and between-sector components T_W and T_B as follows [see Anand (1983)]:

$$\begin{aligned} T &= [T_W] + [T_B] \\ &= [(x\mu_1/\mu)T_1 + ((1-x)\mu_2/\mu)T_2] \\ &\quad + [(x\mu_1/\mu)\log\mu_1 + ((1-x)\mu_2/\mu)\log\mu_2 - \log\mu]. \end{aligned}$$

Thus the within-sector component is simply the income share weighted sum of the sectoral Theil T indices while the between-sector component is the value of T for the hypothetical distribution in which everyone in Sector 1 receives income μ_1 and everyone in Sector 2 receives income μ_2. Now

$$\partial T_W/\partial x = \theta(T_1 - T_2)/[x\theta + (1-x)]^2 > 0 \quad \text{for} \quad T_1 > T_2,$$

so that the within-sector component increases continually with x, from T_2 at $x=0$ to T_1 at $x=1$. On the other hand, T_B is equal to 0 at both $x=0$ and $x=1$, and

$$\partial T_B/\partial x = \theta\log\theta/[x\theta + (1-x)]^2 - (\theta-1)/[x\theta + (1-x)].$$

Hence,

$$[\partial T_B/\partial x]_{x=0} = \theta\log\theta - (\theta - 1) > 0$$

and

$$[\partial T_B/\partial x]_{x=1} = (\log\theta - \theta + 1)/\theta < 0 \quad \text{for} \quad \theta > 1.$$

Thus the between-sector component first increases and then decreases with x. Since both T_W and T_B are increasing functions of x at $x=0$, it follows that the same must be true of $T=T_W+T_B$. However, at $x=1$, T_W is increasing in x while T_B is decreasing in x. A turning point in T is thus guaranteed if

$$[\partial T/\partial x]_{x=1}=[\partial T_W/\partial x]_{x=1}+[\partial T_B/\partial x]_{x=1}<0,$$

i.e.

$$(T_1-T_2)<(\theta-1-\log\theta),$$

which is the turning point condition reported in table 1 for the Theil T index. Turning now to the question of functional form, by substituting $x=(\mu-\mu_2)/(\mu_1-\mu_2)$ in the expression for T we can derive a relation of the form

$$T=A+B(1/\mu)+C\log\mu,$$

where the coefficients A and B depend on μ_1, μ_2, T_1, T_2, and the coefficient C is restricted to equal (-1). This relationship can be estimated, and a test of $C=(-1)$ conducted as a test of the validity of the model.

A similar analysis applies to each of the five aggregable indices and, with some modifications, to the variance of log-income. The important point to note is that each index has its own functional form and its own turning point condition. If the Kuznets process is being invoked as the theoretical underpinning of the inequality–development relationship, the right index must be used with the right functional form for estimation purposes.[7]

4. Cross-section estimation of the inequality–development relationship

Because of the lack of adequate time-series data for individual developing countries, the recent literature has attempted to estimate and test inequality–development relationships using cross-section data. The best known and most influential of these estimations is that by Ahluwalia (1976b) on data from 60 developing and developed countries, based on sources reported in Jain (1975). While being fully aware of the limitations of these sources [see Anand and Kanbur (1986, 1991)], our object here is to estimate inequality–development relationships for each of the six indices of inequality analyzed in

[7]For example, Lydall (1977) uses the functional form $Y=a_0+a_1\log X+a_2X^{-1}$ where Y is a measure of inequality and X is per capita GNP. Lydall uses various well-known indices of inequality for Y (e.g. the Gini coefficient, various quintile shares, etc.) but he does *not* use the Theil T index for which this is actually the appropriate estimating form (see table 1).

the previous section, using the same data sources as Ahluwalia (1976b). These regression estimates form the basis for testing the implications of the Kuznets process.

Table A.1 in the appendix sets out the data used in these regressions. The per capita GNP figures for the countries are taken from Ahluwalia (1976b), but the values of the six inequality indices are calculated from the decile share figures reported in Jain (1975).[8] Ahluwalia (1976b) presents ordinary least squares estimates of the inequality–development relationship for the full sample of 60 developing and developed countries, and for a subsample of 40 developing countries only. We have followed the same procedure. Like Ahluwalia (1976b), we also introduce a dummy variable for the six socialist countries in the sample; these countries show consistently lower values of inequality.[9]

Tables 2 and 3 present our regression results respectively for the full sample of 60 countries and the restricted sample of 40 developing countries only. The notation is the same as in section 3. D denotes the socialist dummy, which is 1 for the six socialist countries in the sample and zero otherwise. An obvious feature of the estimates is the wide disparity among them in the fitted inequality–development relationship. They differ vastly in goodness-of-fit, in turning point, and in the predicted behaviour of inequality in the long run. For example, in the case of the full sample, the turning point for inequality measured by the T index occurs at U.S.$284.6, while for the L index it occurs at U.S.$616.7.[10] Even more dramatic comparisons can be drawn from other estimates in tables 2 and 3.

These estimates may also be compared with those obtained by Ahluwalia (1976b, p. 311) for his (in-)equality index, the income share of the lowest 40 percent. His estimated functional form for this index, a quadratic in the logarithm of per capita GNP, yields a turning point of U.S.$468.0 for the full sample and of U.S.$371.1 for the subsample. Given the large differences in the shape and turning point of the inequality–development relationship for different indices it seems clear that the choice of index will be crucial in any

[8]Since our six indices are all mean-independent the decile share data are sufficient to calculate their values. Ahluwalia (1976b), table 8, provides figures on quintile shares, which, however, were estimated by a *freehand fit* of the Lorenz curve to data contained in the survey sources. Jain's (1975) decile shares are estimated by fitting a Lorenz curve to the survey data following the econometric methodology suggested by Kakwani and Podder (1976). While this methodology is itself not free from criticism, it does at least have the merit of being *replicable*. Notice, however, that our use of decile share data to calculate inequality indices assumes, in effect, that there is perfect equality within each decile: overall inequality is thus understated. This is consistently so for all indices other than the Gini coefficient (see note b to table A.1 in the appendix).

[9]The nature and significance of the split between the developing and developed countries, and the role of the socialist dummy, are discussed further in Anand and Kanbur (1986).

[10]Thus, if a country's per capita GNP were growing at 3% per annum, its turning point for inequality as measured by the L index would occur a quarter of a century later than its turning point for inequality as measured by the T index.

Table 2

Estimates of the inequality–development relationship for the full sample of 60 developing and developed countries.[a]

Dependent variable	Independent variables							Statistics			Turning point (1970 U.S. dollars)
	Constant	$1/\mu$	$\log\mu$	$(1/\mu)^2$	μ	μ^2	D	$\bar{R}^2$	F	SEE	
T	1.644	−49.117	−0.173				−0.320	0.492	20.07	0.130	284.6
	(7.29)	(−4.60)	(−5.60)				(−5.61)				
L	0.092		6.905×10^{-2}		-0.112×10^{-3}		−0.343	0.345	11.34	0.152	616.7
	(0.44)		(1.78)		(−2.97)		(−5.03)				
S^2	0.646	2.920×10^{2}		-24.493×10^{3}			−0.784	0.379	13.01	0.470	167.8
	(5.20)	(4.52)		(−4.43)			(−3.80)				
$[1-I(2)]^{-1}$	2.491				-0.712×10^{-4}	-0.206×10^{-7}	−1.144	0.093	3.01	0.976	−1,727.1
	(11.83)				(−0.22)	(−0.27)	(−2.68)				
G	0.565	−9.500			-0.537×10^{-4}		−0.252	0.538	23.86	0.077	420.6
	(23.57)	(−2.44)			(−4.90)		(−7.33)				
σ^2	0.860				-0.203×10^{-4}	-0.101×10^{-7}	−0.604	0.202	5.97	0.349	1,003.1
	(11.43)				(−0.17)	(−0.36)	(−3.96)				

[a] t-ratios are given in parentheses below the coefficient estimates.

Table 3

Estimates of the inequality–development relationship for the subsample of 40 developing countries.[a,b]

Dependent variable	Independent variables						Statistics			Turning point (1970 U.S. dollars)
	Constant	$1/\mu$	$\log\mu$	$(1/\mu)^2$	μ	μ^2	$\bar{R}^2$	F	SEE	
T	1.664 (2.65)	−51.050 (−2.43)	−0.174 (−1.83)				0.129	3.88	0.148	294.1
L	−1.001 (−2.34)		0.304 (3.37)		-0.728×10^{-3} (−2.96)		0.198	5.80	0.155	418.4
S^2	1.161 (4.66)	1.023×10^{2} (0.96)		-11.415×10^{3} (−1.39)			0.065	2.35	0.521	223.2
$[1-I(2)]^{-1}$	1.727 (3.72)				0.412×10^{-2} (1.94)	-0.375×10^{-5} (−2.01)	0.050	2.02	1.075	549.0
G	0.636 (12.31)	−16.917 (−2.88)			-0.134×10^{-3} (−1.86)		0.144	4.29	0.080	355.0
σ^2	0.504 (3.25)				0.188×10^{-2} (2.65)	-0.164×10^{-5} (−2.63)	0.116	3.57	0.360	571.0

[a]Table A.1 in the appendix in fact lists 41 developing countries. The 40 countries' subsample used for the regressions reported above excludes Spain. This is in keeping with Ahluwalia's (1976b) developing countries only regressions, which excludes Spain from his developing countries subsample. A fuller discussion of the Ahluwalia (1976b) data set is provided in Anand and Kanbur (1986, 1991).

[b]t-ratios are given in parentheses below the coefficient estimates.

analysis, especially policy analysis, of the inequality–development relationship.

What light can such cross-sectional estimation shed on the Kuznets process? From table 1, it is seen that for four of our six indices the inequality–development relationship derived from the Kuznets process implies a restriction on one of the coefficients. For the Theil T index the coefficient of $\log \mu$ is restricted to equal (-1), while for the Theil L measure the corresponding coefficient is restricted to 1. For the squared coefficient of variation S^2 the constant is restricted to (-1), while for the transform of the Atkinson index the constant is restricted to 0.[11] As can be seen from tables 2 and 3, in every case except one (viz. the coefficient of $\log \mu$ in the L regression in table 2), the restriction implied in table 1 is rejected at the 5% level of significance.

Of course it would be quite wrong to conclude from these tests that the Kuznets process is rejected for any *individual* country. For a cross-section of countries the hypothesis being tested is that all countries are following the *same* Kuznets process. The rejection of this hypothesis still leaves open the possibility that the Kuznets process operates in each country but that each country is on a *different* inequality–development relationship (because of differences in sectoral means *or* inequality levels), or that the process does not operate in some countries. Without adequate time-series data we cannot hope to resolve these questions. But the evidence presented here does at least suggest that it would be prudent to consider some extensions of the simple version of the Kuznets process we have modelled in sections 2 and 3. That is the aim of the next section.

5. Extensions of the Kuznets process

In this section we investigate the implications for the inequality–development relationship of extending the model of the simple Kuznets process in sections 2 and 3 by allowing μ_1 and μ_2 to change over time. Our interest is in examining the consequences of such an extension for: (i) the possibility of a turning point; and (ii) the appropriate functional form of the inequality–development relationship.

5.1. Turning point

If all incomes grow at the same proportionate rate, then since all of the six indices considered in section 3 are mean-independent, the overall inequality

[11]For the other two indices, the Gini coefficient G and the variance of log-income σ^2, no such restrictions are implied. But, as was noted in section 3, the latter is not an aggregable index and the former is aggregable only under the implausible assumption of non-overlapping sectoral distributions.

and turning point conditions will not be affected. In order for the turning point conditions to be influenced, we have to look at the case where all incomes do *not* grow at the same rate. Consider, therefore, the case where the ratio of sectoral mean incomes $\theta=(\mu_1/\mu_2)$ changes over time, while the sectoral inequality indices remain constant. The relationship between inequality and the population share x of Sector 1 will now be influenced by the changing value of θ. We illustrate the new considerations that arise for the case of the decomposable Theil T index. A similar analysis applies for the other indices.

Suppose we take the stylized assumption of Kuznets (1955) that θ increases with time (see quote in introduction). Is a turning point in the inequality–development relationship more, or less, likely under this assumption? There are two cases to consider – where θ increases to a finite upper bound θ_1, and where it increases without limit – as x increases from 0 to 1. The analysis of the first case is straightforward. Writing the mean-independent Theil T index as $T=T(x,\theta)$, at the end of the process of θ increasing to θ_1, T as a function of x will follow $T=T(x,\theta_1)$. Hence a sufficient condition for a turning point is that given in table 1 with $\theta=\theta_1$, i.e. $(T_1-T_2)<(\theta_1-1-\log\theta_1)$. Since the right-hand side of the sufficient condition is an increasing function of θ, the inequality is more likely to be satisfied when θ increases to θ_1 than when it does not.

The case where θ increases without limit as x increases from 0 to 1 is more complicated, and opens up the possibility that there may be no turning point at all in inequality over time. The key to understanding this proposition lies in the decomposition of the Theil T index into its within- and between-sector components:

$$T(x,\theta)=T_{\mathrm{W}}(x,\theta)+T_{\mathrm{B}}(x,\theta),$$

where

$$T_{\mathrm{W}}(x,\theta)=[x\theta/(x\theta+(1-x))]T_1+[(1-x)/(x\theta+(1-x))]T_2$$

and

$$T_{\mathrm{B}}(x,\theta)=[x\theta/(x\theta+(1-x))]\log\theta-\log(x\theta+(1-x)).$$

It can be checked that T_{W} is an increasing function of θ bounded above by T_1. T_{B} has a turning point as a function of x at $x^*=(\theta\log\theta-\theta+1)/(\theta-1)^2$. The whole function shifts upward and becomes more 'peaked' as θ increases, with its turning point moving to the left and the slopes at the endpoints becoming steeper. If θ increases without limit, so that the function T_{B} shifts upward without limit, it becomes possible for the decrease in T_{B} with x (beyond x^*) to be more than offset by the shift upward of the function T_{B} with θ. To this we add the increase in the component T_{W} with x, and thus T

can increase continually over time. The appropriate necessary and sufficient conditions for T to increase or decrease over time are:

$$\dot{T} \gtreqless 0 \quad \text{according as} \quad \frac{\dot{\theta}/\theta}{\dot{x}/x} \gtreqless \frac{-(x/T)(\partial T/\partial x)}{(\theta/T)(\partial T/\partial \theta)}.$$

Following Kuznets (1955), we have so far considered the case of $\dot{\theta}(t) \geqq 0$ for all t. However, Ahluwalia (1976b) suggests that a turning point in inequality as a function of time is more likely when θ itself first increases and then decreases with development:

> 'The assumption(s) of ... equal growth rates in sectoral incomes (is) obviously unrealistic. In fact we would expect ... mean income differences to change systematically with development. Interestingly, there are plausible reasons for supposing that these changes may reinforce the U-shaped pattern in overall inequality. ... The ratio of mean incomes between sectors may also follow a U-shaped pattern with intersectoral differences widening in the early stages. ... These differentials can be expected to narrow in the later stages of development ...' [Ahluwalia (1976b, pp. 317–318)].

We can make the above heuristic argument precise by modelling the process as one in which $\dot{\theta}(t) > 0$ for $t < t^*$, say, and $\dot{\theta}(t) < 0$ for $t > t^*$. The behaviour of inequality over time for the T index can then be seen directly from

$$\dot{T} = \dot{x}(\partial T/\partial x) + \dot{\theta}(\partial T/\partial \theta).$$

Since $(\partial T/\partial \theta) > 0$, the sign of the second term will be the same as the sign of $\dot{\theta}$. From the arguments presented earlier, it is easy to check that at $t=0$, $\dot{T} > 0$. Now if $\theta(t)$ increases from θ_0 to its maximum $\theta(t^*)$ and then decreases to θ_1, with $\theta_1 \geqq \theta_0$, then at the *end* of this process the first term is the same as it would be if θ increased continually from θ_0 to θ_1 but the second term is negative, thus making a turning point more likely. These are the precise conditions under which the Ahluwalia claim is valid.[12]

5.2. Functional form

How is the functional form of the inequality–development relationship changed when we allow μ_1 and μ_2 to vary with time? As in the previous subsection, we will restrict the analysis to the case of the Theil T index.

[12]It should be noted, however, that if θ increases from θ_0 to $\theta(t^*)$ and then decreases to $\theta_2 < \theta_1$, then it is not clear that a turning point is more likely than when θ increases continually from θ_0 to θ_1. This is because $[\partial T/\partial x]_{x=1}$ is more negative for higher values of θ, and this may counteract the influence of the second term.

Consider first of all the case where μ_1 and μ_2 change at the same proportionate rate, i.e. $\theta=\bar{\theta}$, a constant. Then, substituting for $\mu_2=(1/\bar{\theta})\mu_1$ in the functional form for T in table 1, it can be shown that

$$T=A+B[1/(\mu/\mu_2)]+C\log(\mu/\mu_2),$$

where

$$A=(\bar{\theta}T_1-T_2+\bar{\theta}\log\bar{\theta})/(\bar{\theta}-1),$$

$$B=(T_2-T_1-\log\bar{\theta})/(\bar{\theta}-1),$$

$$C=-1.$$

The independent variables of the regression are thus $[1/(\mu/\mu_2)]$ and $\log(\mu/\mu_2)$, and the restriction to be tested is that the coefficient on the latter is (-1). As before, with cross-section data a rejection of $C=(-1)$ is quite consistent with each country following the Kuznets process with its own values of $\bar{\theta}$, T_1 and T_2, where these values differ across countries.

Finally, consider the case where $\dot{\theta}\neq 0$. Then it can be shown that the appropriate functional form is

$$\begin{aligned}T=A&+B[\mu_1/(\mu_1-\mu_2)-\mu_1\mu_2/\mu(\mu_1-\mu_2)]\\&+C[-\mu_2/(\mu_1-\mu_2)+\mu_1\mu_2/\mu(\mu_1-\mu_2)]\\&+D[(\mu_1\log\mu_1-\mu_2\log\mu_2)/(\mu_1-\mu_2)]\\&+E\log\mu,\end{aligned}$$

where the restrictions are that $A=0$, $D=1$ and $E=(-1)$, and the coefficients B and C depend on T_1 and T_2. A rejection of these restrictions in a cross-section regression implies a rejection of the hypothesis that T_1 and T_2 are the same across countries. A fortiori the hypothesis is rejected that T_1 and T_2 are the same across countries and constant over time, i.e. that all countries mutatis mutandis (with μ_1's and μ_2's different) are following the same Kuznets process. The estimation of these modified functional forms of course requires data on sectoral as well as overall mean incomes.

6. Conclusion

The Kuznets process plays a central role in the literature on inequality and development. Kuznets's (1955) classic discourse is often invoked to provide theoretical support for the estimation of specific functional forms of the inequality–development relationship, and for a turning point in this relationship such that inequality first increases and then decreases with development. Yet Kuznets (1955) only illustrated his discussion with a numerical example, and the subsequent literature has provided an analysis of the process for just one index [Robinson (1976)]. In this paper we have presented a formalization of the Kuznets process, conducted a general analysis of distributional change under this process, and derived functional forms of the inequality–development relationship for six commonly used indices of inequality.

Our formalization of the process has followed the spirit of Kuznets's (1955) numerical example. During the course of development, the population is seen as shifting from a low-mean low-inequality sector to a high-mean high-inequality sector, the sectoral means and inequality levels remaining unchanged over time. The general characterization of distributional change and the results for specific indices have been derived under these assumptions. When we estimate the functional forms on cross-section data, the restrictions implied by the above formalization of the process are rejected. Such results suggest that our formalization should allow sectoral means and sectoral inequalities to change over time. Assuming constancy in sectoral inequalities, we have provided an analysis of how the functional form and turning point condition for the Theil T index are altered when the ratio of the sectoral means has an exogenous time trend. If, however, the change in sectoral means (or sectoral inequalities) is not exogenous but depends on the population share, then a different type of analysis will be required. It raises the issue of the economic rationale underlying the migration process in the Kuznets model. Can the Kuznets process of population shift be represented as the result of individual migration decisions? If not, what can be said about the inequality–development relationship in such a model? These questions must be left for further research.[13]

[13]A start has been made in Anand and Kanbur (1985).

Appendix

Table A.1[a, b]

Country	Per capita GNP in 1970 U.S.$	Theil entropy index T	Theil second measure L	Squared coefficient of variation S^2	Atkinson index $I(2)$	Gini co-efficient G	Variance of log-income σ^2
Developing countries							
1. Chad 58 (1)	79.5	0.2241	0.2060	0.5715	0.3074	0.3687	0.3676
2. Malawi 69 (1)	80.0	0.3705	0.3358	1.0385	0.4379	0.4696	0.5755
3. Dahomey 59 (1)	91.3	0.3699	0.3345	1.0399	0.4532	0.4675	0.5939
4. Pakistan 63/64 (1)	93.7	0.2406	0.2394	0.5830	0.3702	0.3865	0.4676
5. Tanzania 67 (1)	103.8	0.4282	0.3897	1.2289	0.4830	0.5033	0.6614
6. Sri Lanka 69–70 (3)	108.6	0.2270	0.2225	0.5434	0.3406	0.3771	0.4238
7. India 63–65 (4)	110.3	0.3641	0.3342	1.0238	0.4499	0.4668	0.5922
8. Malagasy 60 (1)	138.7	0.5410	0.4561	1.7275	0.5080	0.5618	0.6853
9. Thailand 62 (1)	142.8	0.4519	0.4031	1.3123	0.4791	0.5103	0.6540
10. Uganda 70 (1)	144.3	0.2573	0.2552	0.6244	0.3822	0.4007	0.4910
11. Kenya 69 (1)	153.2	0.7073	0.6223	2.3322	0.6141	0.6368	0.9355
12. Botswana 71–72 (1)	216.6	0.5545	0.7383	1.4106	0.8579	0.5740	1.9925
13. Philippines 65 (1)	224.4	0.4294	0.4439	1.1342	0.5750	0.5139	0.8802
14. Egypt 64–65 (1)	232.8	0.3004	0.3273	0.6978	0.4899	0.4337	0.6934
15. Iraq 56 (1)	235.5	0.6896	0.7386	1.9886	0.7270	0.6288	1.4012
16. El Salvador 61 (1)	267.4	0.5239	0.4720	1.5583	0.5299	0.5460	0.7597
17. Korea 70 (5)	269.2	0.2188	0.2158	0.5130	0.3304	0.3719	0.4108
18. Senegal 60 (1)	281.8	0.5884	0.5743	1.7414	0.6307	0.5874	1.0241
19. Honduras 67–68 (1)	301.0	0.6586	0.7625	1.8211	0.7776	0.6188	1.6279
20. Tunisia 70 (2)	306.1	0.4157	0.4247	1.0593	0.5358	0.5019	0.8061

Developing countries							
21. Zambia 59 (1)	308.2	0.4730	0.4182	1.4030	0.4897	0.5226	0.6707
22. Ecuador 70 (1)	313.6	0.8363	0.8862	2.6009	0.7734	0.6826	1.6012
23. Turkey 68 (1)	322.2	0.5435	0.5606	1.5229	0.6459	0.5679	1.0803
24. Ivory Coast 70 (2)	328.7	0.4801	0.4768	1.3044	0.5661	0.5342	0.8708
25. Guyana 55–56 (1)	350.8	0.2823	0.3357	0.6291	0.5501	0.4192	0.7996
26. Taiwan 64 (3)	366.1	0.1716	0.1703	0.3962	0.2816	0.3290	0.3336
27. Colombia 70 (7)	388.2	0.5225	0.5167	1.4718	0.5966	0.5557	0.9408
28. Malaysia 70 (3)	401.4	0.4341	0.4765	1.1270	0.6372	0.5131	1.0304
29. Brazil 70 (4)	456.5	0.5595	0.5655	1.5867	0.6341	0.5744	1.0496
30. Jamaica 58 (1)	515.6	0.5610	0.6352	1.5034	0.7256	0.5766	1.3637
31. Peru 70/71 (4)	546.1	0.5984	0.7101	1.6109	0.7844	0.5941	1.6169
32. Lebanon 55–60 (1)	588.3	0.5000	0.4457	1.4926	0.5120	0.5370	0.7174
33. Gabon 68 (2)	608.1	0.7214	0.6631	2.3332	0.6494	0.6439	1.0531
34. Costa Rica 71 (2)	617.1	0.3217	0.3171	0.8140	0.4426	0.4445	0.5987
35. Mexico 69 (5)	696.9	0.5925	0.5411	1.8094	0.5794	0.5827	0.8745
36. Uruguay 67 (1)	720.8	0.2918	0.3290	0.6639	0.5096	0.4279	0.7290
37. Panama 69 (2)	773.4	0.5191	0.5436	1.4320	0.6480	0.5567	1.0800
38. Spain 64–65 (1)	852.1	0.2453	0.2522	0.5738	0.3661	0.3893	0.5055
39. Chile 68 (1)	903.5	0.4315	0.4018	1.2204	0.5005	0.5065	0.7008
40. Argentina 61 (1)	1,004.6	0.3263	0.2869	0.9089	0.3799	0.4375	0.4749
41. Puerto Rico 63 (1)	1,217.4	0.3300	0.3450	0.8067	0.4917	0.4526	0.6970
Developed countries							
42. Japan 68 (2)	1,712.8	0.2488	0.3042	0.5384	0.5343	0.3932	0.7529
43. Finland 62 (1)	1,839.8	0.3625	0.4728	0.8178	0.7100	0.4729	1.2421
44. Netherlands 67 (2)	2,297.0	0.3261	0.3705	0.7733	0.5651	0.4493	0.8398
45. France 62 (1)	2,303.1	0.4379	0.5427	1.0602	0.7268	0.5176	1.3343
46. Norway 63 (1)	2,361.9	0.2150	0.2929	0.4276	0.5708	0.3622	0.8097
47. United Kingdom 68 (2)	2,414.3	0.1781	0.1953	0.3761	0.3390	0.3385	0.4215
48. New Zealand 70–71 (3)	2,501.5	0.2201	0.2754	0.4529	0.5043	0.3708	0.6920
49. Denmark 63 (1)	2,563.9	0.2368	0.2687	0.5197	0.4505	0.3863	0.6050
50. Australia 67–68 (1)	2,632.4	0.1580	0.1725	0.3346	0.3098	0.3185	0.3738
51. West Germany 70 (2)	3,208.6	0.2458	0.2539	0.5715	0.3926	0.3939	0.5107
52. Canada 65 (1)	3,509.6	0.1731	0.1894	0.3674	0.3323	0.3333	0.4095
53. Sweden 70 (2)	4,452.2	0.2375	0.2697	0.5199	0.4512	0.3872	0.6069
54. United States 70 (3)	5,244.1	0.2638	0.2975	0.5916	0.4794	0.4074	0.6632

Table A.1 (continued)

Country	Per capita GNP in 1970 U.S.$	Theil entropy index T	Theil second measure L	Squared coefficient of variation S^2	Atkinson index $I(2)$	Gini co-efficient G	Variance of log-income σ^2
Socialist countries							
55. Bulgaria 60 (1)	406.9	0.0937	0.0960	0.1983	0.1767	0.2459	0.1956
56. Yugoslavia 68 (2)	602.3	0.1890	0.2027	0.4120	0.3439	0.3474	0.4283
57. Poland 64 (1)	660.8	0.1070	0.1063	0.2325	0.1857	0.2635	0.2080
58. Hungary 67 (1)	872.7	0.0999	0.1140	0.1955	0.2300	0.2508	0.2599
59. Czechoslovakia 64 (1)	887.7	0.0567	0.0567	0.1181	0.1059	0.1938	0.1127
60. East Germany 70 (1)	2,046.3	0.0650	0.0709	0.1269	0.1426	0.2044	0.1541

[a]The sample of 60 countries is the same as that in Ahluwalia (1976b). For each country Ahluwalia provides a year and a source number (in parentheses) which are supposed to identify the income distribution from Jain (1975). However, in some cases (Tanzania, Honduras, Malaysia, Brazil, Costa Rica, Uruguay, Spain, Argentina and Yugoslavia) Ahluwalia's source references do not identify a unique distribution in Jain (1975). In these cases we have chosen the distribution whose income share for the lowest 40 percent is closest to the figures reported in Ahluwalia's table 8.

For India, the distribution 63–64 (4) referred to in Ahluwalia (1976b) is non-existent in Jain (1975). In this case we have chosen the distribution identified in Jain (1975) as 63–65 (4), IR-NL, on the assumption that there is an error in reporting the year to which the distribution refers rather than in the source number. For Peru, the Ahluwalia source reference is 70 (4), but there is no such distribution in Jain (1975). Again, we assume the source number to be correct and choose the 70/71 (4), EAP-NL, distribution from Jain (1975). For Bulgaria, Ahluwalia's table 8 does not give a year at all. In this case there are three income distributions – for the years 1957, 1960, 1962; we choose 60 (1), WRK-NL, which has an income share for the lowest 40 percent (24.5%) closest to the figure (25.0%) in Ahluwalia's table 8. For Taiwan, Ahluwalia does not make use of any source in Jain (1975) – rather he uses Kuo (1975). However, this source does not provide data on decile shares. Hence we have chosen Taiwan 64 (3) from Jain (1975), which has a share for the lowest 40 percent closest to that reported by Ahluwalia.

For all these countries we continue to use Ahluwalia's per capita GNP figures; the correct procedure would have been to use per capita GNP figures for the year of the distribution chosen. This is unlikely to make a significant difference to the conclusions. For a discussion of this problem, as well as a review and critique of the data sources used by Ahluwalia (1976b), see Anand and Kanbur (1986, 1991).

[b]For each country the inequality indices (with the exception of the Gini coefficient) are calculated from the decile share information provided for the source in Jain (1975). Estimates for the Gini coefficient were taken directly from Jain (1975). The following formulae, where s_i $(i=1,2,\dots,10)$ are the decile shares, were used to estimate the other five indices:

$$T=\sum_{i=1}^{10} s_i \log_e s_i + \log_e 10$$

$$L=\log_e\left[\sum_{i=1}^{10} s_i/10\right]-\frac{1}{10}\sum_{i=1}^{10}(\log_e s_i)$$

$$S^2=\left[10\sum_{i=1}^{10} s_i^2\right]-1$$

$$I(\varepsilon)=1-\left[10^{-\varepsilon}\sum_{i=1}^{10} s_i^{1-\varepsilon}\right]^{1/(1-\varepsilon)}$$

$$\sigma^2=\frac{1}{10}\sum_{i=1}^{10}(\log_e s_i)^2-\frac{1}{100}\left[\sum_{i=1}^{10}\log_e s_i\right]^2.$$

References

Adelman, Irma and Cynthia T. Morris, 1973, Economic growth and social equity in developing countries (Stanford University Press, Stanford, CA).
Ahluwalia, Montek S., 1974, Income inequality: Some dimensions of the problem, in: H.B. Chenery et al., eds., Redistribution with growth (Oxford University Press, London).
Ahluwalia, Montek S., 1976a, Income distribution and development: Some stylized facts, American Economic Review 66, May, 128–135.
Ahluwalia, Montek S., 1976b, Inequality, poverty and development, Journal of Development Economics 3, 307–342.
Ahluwalia, Montek S., Nicholas G. Carter and Hollis B. Chenery, 1979, Growth and poverty in developing countries, Journal of Development Economics 6, 299–341.
Anand, Sudhir, 1983, Inequality and poverty in Malaysia: Measurement and decomposition (Oxford University Press, New York).
Anand, Sudhir and S.M.R. Kanbur, 1984, Inequality and development: A reconsideration, in: H.-P. Nissen, ed., Towards income distribution policies, EADI Book Series 3 (Tilburg, Netherlands).
Anand, Sudhir and S.M.R. Kanbur, 1985, Poverty under the Kuznets process, Economic Journal 95, Supplement, Jan., 42–50.
Anand, Sudhir and S.M.R. Kanbur, 1986, Inequality and development: A critique, Paper presented at the 25th Anniversary Symposium of the Economic Growth Center, Yale University, New Haven, CT, April 11–13, Forthcoming in Journal of Development Economics, June 1993.
Anand, Sudhir and S.M.R. Kanbur, 1991, International poverty projections, Ld'A/QEH Working Paper no. 32 (Queen Elizabeth House, Oxford), Feb.; and PRE Working Paper WPS 617 (World Bank, Washington DC) March.
Atkinson, Anthony B., 1970, On the measurement of inequality, Journal of Economic Theory 2, no. 3, Sept., 244–263.
Chenery, Hollis B. and Moises Syrquin, 1975, Patterns of development, 1950–1970 (Oxford University Press, New York).
Fields, Gary S., 1979, A welfare economic approach to growth and distribution in the dual economy, Quarterly Journal of Economics XCIII, no. 3, Aug., 325–353.
Hadar, Josef and William R. Russell, 1969, Rules for ordering uncertain prospects, American Economic Review 59, no. 1, March, 25–34.
Jain, Shail, 1975, Size distribution of income: A compilation of data (Johns Hopkins University Press (for the World Bank), Baltimore, MD).
Kakwani, N.C. and N. Podder, 1976, Efficient estimation of the Lorenz curve and associated inequality measures from grouped observations, Econometrica 44, no. 1, Jan., 137–148.
Kuo, Wan-yong, 1975, Income distribution by size in Taiwan area – changes and causes, in: Income distribution, employment and economic development in Southeast and East Asia, Volume 1, Papers and proceedings of the seminar sponsored jointly by the Japan Economic Research Center and the Council for Asian Manpower Studies, Dec. 16–20, 1974.
Kuznets, Simon, 1955, Economic growth and income inequality, American Economic Review 45, no. 1, March, 1–28.
Loehr, William, 1981, Economic growth, distribution and incomes of the poor, Journal of Economic Studies 7, no. 3, 127–139.
Lydall, Harold F., 1977, Income distribution during the process of development, World Employment Programme Research Working Paper no. WEP2-23/WP52 (International Labour Office, Geneva) Feb.
Paukert, Felix, 1973, Income distribution at different levels of development: A survey of evidence, International Labour Review 108, nos. 2–3, Aug.-Sept., 97–125.
Robinson, Sherman, 1976, A note on the U-hypothesis relating income inequality and economic development, American Economic Review 66, no. 3, June, 437–440.
Rothschild, Michael and Joseph E. Stiglitz, 1970, Increasing risk I: A definition, Journal of Economic Theory 2, no. 3, Sept., 225–243.
Swamy, Subramaniam, 1967, Structural changes and the distribution of income by size: The case of India, Review of Income and Wealth, series 13, 155–174.
Theil, Henri, 1967, Economics and information theory (North-Holland, Amsterdam).
World Bank, 1978, 1979, 1980, World development report (Oxford University Press (for the World Bank) New York).