3. Nonlinearities in the Inequality-Development Relationship: Examining the Kuznets Hypothesis with the Threshold Regression Model

Andreas Savvides and Thanasis Stengos

1. Introduction

Following the publication of Kuznets's (1955) presidential address to the American Economic Association, considerable research effort has been devoted to what has become known as the Kuznets hypothesis. According to the Kuznets inverted-U hypothesis, income inequality increases during the early stages of economic development and after reaching a turning point declines. Early empirical tests of the inverted-U hypothesis relied exclusively on cross-sectional data and found evidence on the whole supportive of the hypothesis.[1] The major weakness of these tests is that Kuznets was concerned with the intertemporal pattern of income inequality[2] an attribute that cross section empirical tests are unable to capture. In order to remedy this deficiency, panel data sets on income inequality have recently become available, allowing researchers to investigate the temporal aspects of income inequality. The panel results are mixed. Jha (1996) and Fields and Jakubson (1994) develop limited panel data sets and report contrasting results (the former claiming support for the inverted-U hypothesis). The most comprehensive panel data set on income inequality is that of Deininger and Squire (1996). Using a subset of their data set they deem to be of high quality, Deininger and Squire (1997) claim that, once country-specific effects are allowed for, no discernible relationship between income inequality and per capita income (the inequality-development relationship) can be detected . On the other hand, Ram (1997) claims that if the sample is confined to the developed countries in the Deininger/Squire high quality set, not only is an inverted-U absent but a significant uninverted-U pattern is evident (after allowing for country-specific effects)[3] In a recent paper, Savvides and Stengos (1998) have pointed out that most tests of the Kuznets hypothesis divide countries arbitrarily into two groups: less developed (LDCs) and developed.[4] The criterion, however, according to which countries are divided into the two groups is arbitrary: separation is based on a researcher's subjective notion of the category (LDC or otherwise) to which each country belongs or on a notional level of per capita income according to which a country is classified into each category. Instead, Savvides and Stengos (1998) employ the threshold regression (TR) model[5] that lets the level of economic development (per capita income) determine the existence and significance of a threshold level in the Kuznets relationship rather than imposing *a priori* an arbitrary classification scheme. If indeed there exists a well-defined relationship between income inequality and per capita income (inverted- or uninverted-U) that depends on the level of economic development, the TR model can identify the threshold level endogenously and test for such a relationship above and below the threshold.

The purpose of this chapter is to reexamine the nonlinearity in the inequality-development relationship implied by the Kuznets hypothesis with the TR model. In accordance with previous research, we begin with a functional form that encompasses the possibility of an inverted-U curve. Our point of departure is to examine an alternative functional form that previous researchers have not considered. This form allows the possibility of a more complex relationship between income per capita and inequality. While Kuznets suggested the possibility of an inverted-U, recent research by Fields and Jakubson (1994), Ram (1997) and Savvides and Stengos (1998) has reported that an uninverted-U curve may characterize the inequality-development relationship. Consequently, we explore an alternative functional form that allows for an inverted-U pattern over a certain range of income per capita and an uninverted-U over another. To the best of our knowledge, this is the first attempt to allow for differences in the shape of the inequality-development relationship (inverted- or uninverted-U) according to the range of income per capita. The following section applies the TR model to the high quality data in the Deininger/Squire panel data set in order to test nonlinearities in the inequality-development relationship. The final section concludes the chapter.

2. Testing the Kuznets hypothesis with the TR Model

We test the Kuznets hypothesis via the high-quality panel data set of Deininger and Squire (1996). As already mentioned, this data set is the most comprehensive compilation of income inequality statistics. This data set is designed to ensure intertemporal and international comparability of income inequality in terms of the recipient unit (household or personal), variable being measured (income or expenditure), and income concept (gross or net) used by income surveys across various countries. The Deininger-Squire estimates of the Gini coefficient serve as our measure of income inequality. As in all previous tests of the Kuznets hypothesis the level of economic development is measured by per capita income. We use the Summers-Heston estimates of per capita income in international dollars at 1985 prices. Our sample consists of the 618 observations in the Deininger-Squire data set for which Summers-Heston provide estimates of per capita income. Given that some of the 95 countries in this sample are represented by one or a very small number of observations, we follow some previous researchers (e.g. Bruno, Ravallion and Squire 1998; Deininger and Squire 1997) and consider an alternative sample that contains only countries for which at least four observations are available. This alternative sample contains 547 observations from 52 countries. In what follows we test the sensitivity of our results to the alternative sample and note any differences in results between the two samples.

In testing the inverted-U hypothesis we consider two alternative specifications of the TR model as follows:

$$INEQ_{it} = a_0 + a_1 INC_{it} + a_2 (INC)^2 + e_{it} \quad (1)$$

$$INEQ_{it} = b_0 + b_1 (INC_{it}) + b_2 (INC_{it})^2 + b_3 (INC_{it})^3 + e_{it} \quad (2)$$

where $INEQ_{it}$ is the Deininger-Squire estimate of the Gini coefficient and INC_{it} is the Summers-Heston estimate of per capita income. The specification in (1) contains per capita income and its square. This is one of the most popular specifications adopted by previous researchers when testing the inverted-U hypothesis. It has been used, *inter alia* by Fields and Jakubson (1994), Ram (1995) and Sundrum (1990). Specification (2) includes a cubic term to allow for a more rich modeling of any nonlinearities present in the inequality-development relationship. In particular, it allows us to detect the presence (if any) of an inverted-U or an uninverted-U relationship over different ranges of per capita income. For both specifications (1) and (2), the threshold variable is per capita income.

The first test we conduct is whether the TR model is statistically significant relative to a linear specification, i.e. we test the null hypothesis that there is no threshold. The values of the Lagrange-multiplier (LM)\ test are 42.6 and 38.9 for specifications (1) and (2), respectively. Based on 1000 bootstrap replications, the null hypothesis can be decisively rejected regardless of specification (the p-value for this test is less than 0.01 in both cases). There is strong evidence in favor of the presence of a threshold (per capita income) level in the Kuznets relationship. Having uncovered a threshold level, we proceed to fix the income level which subdivides the sample to test for the possibility of additional threshold levels. For the sub-sample of observations that exceed the threshold level, the LM tests for specifications (1) and (2) yield p-values of 0.3572 and 0.4431 respectively. Thus, we conclude that no additional threshold levels are present. Our results suggest that the inequality-development relationship is described by a two-regime split of the sample based on per-capita income.

Given that our LM test indicates the presence of only one threshold level in the inequality-development relationship, it is necessary to estimate this threshold level and split the sample accordingly. The estimate of the threshold is $PPP 3297 for specification (1) and $PPP 3731 for (2).[6] Tables 1 and 2 contain the empirical results for models (1) and (2), respectively. The first two columns in each table display estimates from a linear specification that ignores threshold effects: column (1) contains the OLS\ results (with a common intercept across countries) and column (2) the results from a panel model that includes country-specific intercepts[7].

The estimates in columns (1) and (2) of Table 1 reveal the presence of a significant uninverted-U curve: the coefficients of both INC and its square are highly significant. This result holds regardless of whether country-specific effects are included or excluded. Our estimate of the turning point (the value of per capita income at which inequality achieves its minimum level) is $PPP 15209 for the specification in column (1) and\$PPP 10251 for the model in column (2)[8]. The estimates in columns (1) and (2) of Table 2 reveal a different shape for the relationship between inequality and per capita income that previous research has not looked into. The estimates reveal that as income per capita increases income inequality first follows an inverted-U pattern and at higher levels of income it is succeeded by an uninverted-U pattern. According to the estimates of column (1), income inequality first increases and reaches a maximum at $PPP 2202. Subsequently, it declines reaching a minimum at $PPP 12603, after which it begins to increase again. The same pattern is observed for the model of column (2) that includes country-specific effects; the coefficient estimates, however, are no longer significant. The uninverted-U pattern at higher levels of income has been confirmed by Ram (1997) for a group of 19 developed economies and Savvides and Stengos (1998) for observations above the threshold level identified by the TR model. Our results indicate that the relationship between income inequality and per capita income, though nonlinear, may be more complex than that previously hypothesized by researchers. This finding points to the need of additional research on this issue, a sentiment shared by Ram (1995).

The remaining columns of Tables 1 and 2 provide estimates of the TR model: columns (3) and (4) correspond to observations above the threshold and columns (5) and (6) below the threshold[9]. The results in columns (5) and (6) of Table 1 indicate that no relation between per capita income and inequality exists for observations below the threshold. The same cannot be said for the model that includes the cubic term. It is clear from the estimates of column (5) in Table 2 that for observations below the endogenous threshold level there is a significant relationship between per capita income and income inequality. Moreover, it is described by an uninverted-U pattern at lower levels of income and an inverted-U at higher levels of per capita income. Income inequality first decreases reaching a minimum at $PPP 876; subsequently it rises and reaches a maximum at $PPP 2652. A\ similar pattern is observed for the model in column (6) that includes country-specific effects, though the coefficient estimates are no longer significant. This pattern is the opposite to that observed for the complete sample. Once more it points to the complex nature of the inequality-development relationship.

When it comes to observations above the threshold, an uninverted-U pattern is evident for specification (1) in Table 1, regardless of inclusion or exclusion of country-specific effects. This confirms Ram's (1997) result for a group of 19 'developed' economies in the Deininger/Squire sample. Except for stating that these "... are the fairly standard industrialized market economies," Ram does not provide any clear guidelines for his choice of countries to include as developed[10]. We conclude that the uninverted-U pattern reported previously for a sample of developed economies extends to the TR model for observations above the endogenous threshold level. The estimate of the turning point (the minimum income level) is $PPP 11490 for the specification in column (3) and \$PPP\ 10713 for the specification in column (4) that includes country specific effects.[11] These estimates are broadly similar though slightly higher than Ram's estimate of $PPP 9970. This may be due to the fact that the TR model reports estimates for all observations above the threshold level (regardless of country) while all previous studies report estimates for specific countries classified arbitrarily as developed or less developed. When we estimate the more general specification in (8) with observations above the threshold, no significant relationship between inequality and per capita income can be discerned (columns (3) and (4) of Table 2). The signs of the coefficients indicate a pattern similar to that reported for the observations below the threshold (columns (5) and (6) of Table 2), but they are not significant regardless of the inclusion or exclusion of country-specific effects. The same result holds true for the smaller sample of 547 observations, though in this case the coefficient estimates are marginally significant (10 percent level) for the specification that excludes country-specific effects (results available on request).

3. Conclusion

This chapter investigates the well-known and extensively-researched Kuznets hypothesis according to which the relationship between income inequality and per capita income is described by an inverted-U pattern. Our study uses the most comprehensive panel data set on income inequality. Our aim is to explore in greater detail the nonlinear relationship between inequality and per capita income. Rather than dividing the sample arbitrarily, we employ the threshold regression (TR) model. Initially we test the significance of the TR model relative to a linear specification. We find conclusive evidence in favor of the TR model. Subsequently we obtain endogenous estimates of the threshold level for two alternative functional forms. One of the functional forms has been extensively researched. The other allows for a more complex relationship by including a cubic per capita income term. Estimation of the TR model with both specifications allows the regression coefficients to differ between observations above/below the threshold in order to shed light on the existence of a Kuznets curve for the high/low income group. Several conclusions emerge from the estimates of the TR model. In the first place, there is no evidence of an inverted-U curve. This conclusion holds for observations both above and below the threshold and regardless of the functional form used to test the Kuznets hypothesis. This is an important but, perhaps, unsurprising finding. As Williamson (1997) suggests, one could hardly expect an 'unconditional' Kuznets curve to emerge given the numerous factors that condition the income inequality experience[12]. Thus any previous claims that the inverted-U hypothesis describes the inequality-development relationship, either at relatively high or low income levels, should be laid to rest. Second, the specification that includes the cubic term indicates that the relationship between income inequality and per capita income may be more complex than previously thought. Ignoring the existence of a threshold effect we find that income inequality follows an inverted-U pattern at lower levels of per capita income followed by an uninverted-U pattern at higher levels of income. When we take into account the threshold effect, the opposite pattern is observed for observations below the threshold: an uninverted-U pattern at lower levels of income followed by an inverted-U at higher levels. These results hold only for the model that excludes country-specific effects. For observations above the threshold level, no significant relationship between per capita income and inequality can be discerned. As previously discussed, our findings point to the nonlinearities in the inequality-development relationship as a fruitful avenue for future research.

Table 1. OLS and Threshold Regression Estimates of Equation (7)

Explanatory Variables	**OLS Estimates**		**TR Model Estimates: > Threshold**		**TR Model Estimates: < Threshold**	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Constant	41.32 (51.45)		44.77 (15.13)		33.14 (10.53)	
INC	-0.112×10^{-2} (4.565)	-0.492×10^{-3} (2.287)	-0.214×10^{-2} (3.697)	-0.791×10^{-3} (3.049)	0.465×10^{-2} (1.162)	-0.101×10^{-2} (0.454)
$(INC)^2$	0.368×10^{-7} (2.442)	0.240×10^{-7} (2.470)	0.929×10^{-7} (3.496)	0.369×10^{-7} (3.290)	-0.110×10^{-6} (0.102)	0.262×10^{-6} (0.469)
$\overline{R}^2$	0.084	0.919	0.050	0.927	0.121	0.888
LM-heteroscedasticity [significance level]	36.51 [0.00]	21.39 [0.00]	93.12 [0.00]	12.50 [0.00]	2.78 [0.10]	1.87 [0.17]
Number of Observations	618	618	393	393	225	225

Note: Columns (1), (3), and (5) report estimates without country-specific effects and columns (2), (4), and (6) with country-specific effects. Numbers in parentheses below coefficient estimates are t-statistics. LM-heteroscedasticity is the Breusch-Pagan Lagrange multiplier statistic for heteroscedasticity (p-values are shown in brackets). Where the hypothesis of homoscedasticity is rejected, t-statistics are based on heteroscedasticity-consistent standard errors.

Table 2. OLS and Threshold Regression Estimates of Equation (8)

Explanatory Variables	**OLS Estimates**		**TR Model Estimates: > Threshold**		**TR Model Estimates: < Threshold**	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Constant	38.28 (34.73)		50.13 (5.513)		45.77 (8.28)	
INC	0.118×10^{-2} (1.678)	0.195×10^{-3} (0.397)	-0.380×10^{-2} (1.314)	-0.456×10^{-3} (0.498)	-0.023 (2.329)	-0.505×10^{-2} (0.995)
$(INC)^2$	-0.314×10^{-6} (3.323)	-0.660×10^{-7} (1.145)	0.251×10^{-6} (0.898)	-0.335×10^{-8} (0.035)	0.177×10^{-4} (3.337)	0.269×10^{-5} (0.927)
$(INC)^3$	0.141×10^{-10} (4.056)	0.327×10^{-11} (1.634)	-0.470×10^{-11} (0.558)	0.140×10^{-11} (0.465)	-0.334×10^{-8} (3.950)	-0.421×10^{-9} (0.877)
$\overline{R}^2$	0.103	0.919	0.051	0.923	0.135	0.895
LM-heteroscedasticity [significance level]	72.06 [0.00]	19.68 [0.00]	97.36 [0.00]	9.80 [0.00]	1.14 [0.29]	4.31 [0.04]
Number of Observations	618	618	369	369	249	249

Note: See note to Table 1.

Notes

1 The literature has been surveyed by Bruno\textit{, }Ravallion and Squire (1998) and Sundrum (1990). For recent tests with cross section data see Eusufzai (1997), Fishlow (1996) and Ram (1995).

2 Kuznets (1955, p. 4) explored long-term trends in income inequality for England, Germany and the United States, concluding that for these countries the "... relative distribution of income ... has been moving toward equality - with these trends particularly noticeable since the 1920's but beginning perhaps in the period before the first world war."

3 Using observations from both developed and less developed countries in the Deininger/Squire high quality data set, Sarel (1997) claims to have detected an inverted-U pattern. This study, however, omits country-specific effects (it includes

only a binary classification variable that divides countries into socialist/nonsocialist and a year-specific effect). Deininger and Squire\ (1997) and Fields and Jakubson (1994) have shown that inclusion of country-specific effects invalidates the inverted-U hypothesis

4 For instance, as already mentioned, Ram (1997) confines his sample to developed economies, while Fields and Jakubson (1994) and Ram (1995) consider less developed economies (LDCs)\ alone. Ahluwalia (1976), Anand and Kanbur\ (1993), Fishlow (1996) and Jha (1996) present results for a sample that combines developed and less developed economies as well as for an LDC-only sample. Deininger and Squire (1997) divide countries into low and high income (based on various arbitrary cutoff levels of per capita income that range from \$1000 to \$10,000) and examine the possibility of a linear trend between income inequality and per capita income for countries above and below the cutoff (negative and positive, respectively).

5 For a description of the TR model, see Hansen (1996a) and 1996b).

6 Significant threshold effects are also found for the limited sample of 547 observations. The values of the LM statistic are 55.2 and 52.1 for specification (1) and (2) respectively. Both of these are highly significant (0.01 level). The estimates of the threshold level are $PPP 4642 for specification (1) and $PPP 3731 for specification (2).

7 Where the hypothesis of homoscedasticity is rejected, significance tests are based on heteroscedasticity-consistent standard errors.

8 Similar results are obtained for the more limited sample. A significant uninverted-U curve is obtained and the estimates of the turning point are $PPP 13937 for the model in column (1) and $PPP 9995 for column (2).

9 Once more, we present OLS estimates without country-specific intercepts (columns 3 and 5) and panel estimates with country-specific intercepts (columns 4 and 6).

10 Ram (1997, fn. 6) comments on one `marginal' case included in the sample (Greece) and argues that inclusion or exclusion of this country does not change the results appreciably. He does not, however, comment on other `marginal' cases (e.g. Portugal and Turkey) excluded from his sample

11 A significant uninverted-U pattern is also observed for the smaller sample of 547 observations. The estimates of the turning point are $PPP 11054 corresponding to column (3) and $PPP 11144 for column (4).

12 Williamson (1997) discusses four of these (globalization, education supply, demography and demand) focusing on the impact of globalization, and mass migrations in particular.

References

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

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

Bruno, M., M. Ravallion and L. Squire, 1998, Equity and growth in developing countries: Old and new perspectives on the policy issues, in V. Tanzi and K-y Chu (eds.), Income Distribution and High-Quality Growth, Cambridge MA: MIT Press.

Deininger K., and L. Squire, 1996, A new data set measuring income inequality, The World Bank Economic Review, 10:565-591.

Deininger K., and L. Squire, 1997, New ways of looking at old issues: Inequality and growth, manuscript, World Bank.

Eusufzai, Z., 1997, The Kuznets hypothesis: An indirect test, Economics Letters, 54:81-85.

Fields, G.S., and G.H. Jakubson, 1994, New evidence on the Kuznets curve, manuscript, Cornell University.

Fishlow A., 1996, Inequality, poverty, and growth: Where do we stand?, in M. Bruno and B.Pleskovic (eds.), Annual World Bank Conference on Development Economics 1995, Washington D.C.: The World Bank.

Hansen, B.E., 1996a, Inference when a nuisance parameter is not identified under the null hypothesis, Econometrica, 64:413-430.

Hansen, B.E., 1996b, Sample splitting and threshold estimation, manuscript, Department of Economics, Boston College.

Jha, S.K., 1996, The Kuznets curve: A reassessment, World Development, 24:773-780.

Kuznets, S., 1955, Economic growth and income inequality, American Economic Review, 45:1-28.

Ram, R., 1995, Economic development and income inequality: An overlooked regression constraint, Economic Development and Cultural Change, 43:425-434.

Ram. R., 1997, Level of economic development and income inequality: Evidence from the postwar developed world, Southern Economic Journal, 64: 576-583

Sarel, M., 1997, How macroeconomic factors affect income distribution: The cross-country evidence, International Monetary Fund Working Paper No. WP/97/152.

Savvides, A., and T. Stengos, 1998, Income inequality and economic development: Evidence from the threshold regression model, working paper.

Sundrum R.M., 1990, Income Distribution in Less Developed Countries, New York: Routledge.

Williamson, J.G., 1997, Growth, distribution and demography: Some lessons from history, National Bureau of Economic Research, Working Paper No. 6244.