KUZNETS CURVE: THE CASE OF THE INDIAN ECONOMY

Arjun Teotia*

*Department of Economics,
Delhi School of Economics.

ABSTRACT

This paper examines the functional relationships between income inequality, economic factors, institutions, and Kuznets’ inverted-U hypothesis. A model that incorporates interactive as well as direct effects of several factors to capture their combined effect on inequality is developed. The model is estimated using the Gini coefficient using a cross-sectional data set for 25 Indian states in the years 1983, 1993 and 2000. The results do not support Kuznets’ hypothesis; however, the relationship between growth and inequality is conditioned by a host of economic and institutional factors.

INTRODUCTION

The relationship between income inequality and growth was first given by Simon Kuznets and is known as the Kuznets curve. It shows that there is an inverted U-shaped relationship between growth and income inequality i.e. the inequality first increases in the early stages of development, reaches a certain maximum and then declines as the country reaches a high level of development. As a poor economy develops there is a shift of labour force from agricultural to higher paying non-agricultural sector of the economy. This leads to an initial increase in the inequality between the sectors and as a whole. This happens to the point where factor movement equalizes returns across the factors. The above hypothesis derives a great deal of interest due to the concern that development hurts the poor.

Empirical tests of the Kuznets hypothesis have shown the effect to be true for some cases and not true for some. Early tests of the hypothesis for England, Germany and the United States seemed to support Kuznets’ hypothesis, but in a large number of later studies, a wide variety of results emerged with some challenging the hypothesis and others supporting it. Most of these studies used cross-sectional data while a few used time-series data. These earlier studies often suffered from serious problems with the quality and availability of the underlying data.
BACKGROUND

Estimates of the distribution of income or consumption provided by a World Bank project (Deininger and Squire, 1996) seemed to provide some general support for Kuznets’ hypothesis when cross-section data were examined. For instance, Jha (1996) reports that despite problems with data comparability, the Kuznets hypothesis holds. Huang et al. (2007) also verify the Kuznets prediction for countries with mild income inequality but do not find such support for countries where inequality is either too high or too low. Quite a few empirical studies refute the Kuznets hypothesis, however. For example, Fields (1991) finds that there is no tendency for income inequality in poor countries to increase (rather than decrease) and no tendency for income inequality in rich countries to decrease (rather than increase). Similarly, Ravallion and Chen (1997) in their study of 67 developing and transitional economies covering the 1981-94 period find that income distribution improved with economic growth as often as it worsened. And as the first users of the new dataset, Deininger and Squire (1998) found no support for the Kuznets hypothesis in the cross-country data on income and asset distribution. Lundberg and Squire (2003) claim that growth and inequality move together, determined by a simultaneous process.

METHODOLOGY AND DATA

There is a long run relation between inequality and growth and is influenced by many factors including income. The effect of income on inequality can be dampened by other factors such as education, financial development, social and political freedom etc.

This paper tests the effect of income and some other factors on inequality of the states in India. We use the following model to estimate the inequality in states at a particular time.

\[ G_{it} = F(INC_{it}, INC_{it}^2, EDU_{it}, POP_{it}, DUM) \]

The subscript i refers to state and t refers to the time period.

\[ G_{in} = \text{Gini Coefficient of the state} \]
\[ INC = \text{Per Capita Income of the state} \]
\[ INC^2 = \text{Square of Per Capita Income of the state} \]
\[ EDU = \text{Level of Education in the state} \]
\[ POP = \text{Population density of the state} \]
\[ DUM = \text{Dummy variable for the years} \]

This paper uses the following regression equation for state \( i \) at period \( t \) :

\[ G_{it} = b_1 + b_2 INC_{it} + INC_{it}^2(b_3 + b_4 EDU_{it} + b_5 POP_{it}) + b_6 DUM_{83} + b_7 DUM_{93} + b_8 EDU_{it} + b_9 POP_{it} + e_{it} \]
The above equation includes interactive as well as direct effect of several factors to capture their combined effect on inequality.

The paper estimates the model using the Gini Coefficient. In addition to rural and urban gini coefficient we use a combined gini coefficient. To combine the rural and urban gini to form a single index we use the geometric mean. The use of arithmetic mean is avoided to give appropriate weights to both the coefficients. The state-wise per capita income is used as Income. EDU is the state-wise literacy rate in India. Due to lack of data the states of Arunachal Pradesh, Meghalaya, Nagaland, Sikkim, Tripura, Chandigarh, Dadra & Nagar Haveli and Lakshwadeep have been dropped from the analysis.

ESTIMATION AND RESULTS

If we run the regression with Rural gini coefficient as the dependent variable and the above independent variables we get the following results:

\[
RURGIN = 0.1934431 + 3.84 \times 10^{-6} INC_{it} - 3.47 \times 10^{-11} INC_{it}^2 - 2.56 \times 10^{-8} [INC_{it}^2 \times EDU_{it}]
\]

\[
-1.17 \times 10^{-14} [INC_{it}^2 \times POP_{it}] + 0.0879655 DUM_1 + 0.0295793 DUM_2 - 0.0000176 EDU_{it}
\]

\[
R^2 = 0.3737 \quad \text{No. of Observations} = 73
\]

t-statistics are written in parenthesis

The above results show that the coefficients of all independent variables except the dummy variables are insignificant and thus have not affected rural inequality significantly. The significant coefficients of the dummy variables show that rural inequality has fallen over time from 1983 to 2000.

If we run the regression with Urban gini coefficient as the dependent variable and the above independent variables we get the following results:

\[
URBGIN = 0.1954142 + 3.51 \times 10^{-6} INC_{it} - 1.44 \times 10^{-10} INC_{it}^2 - 1.59 \times 10^{-6} [INC_{it}^2 \times EDU_{it}]
\]

\[
-1.09 \times 10^{-14} [INC_{it}^2 \times POP_{it}] + 0.0389706 DUM_1 + 0.0172478 DUM_2 - 0.0009961 EDU_{it}
\]
The above results show that the coefficients of all independent variables except the dummy variables are insignificant and thus have not affected urban inequality significantly. The significant coefficients of the dummy variables show that urban inequality has fallen over time from 1983 to 2000.

If we run the regression with Combined gini coefficient as the dependent variable and the above independent variables we get the following results:

\[
\text{CMBGIN} = 0.1905684 + 3.70e^{-06}\text{INC}_it - 8.48e^{-11}\text{INC}_it^2 - 8.74e^{-07} \left[ \text{INC}_it^2 \times \text{EDU}_it \right] \\
-7.32e^{-16} \left[ \text{INC}_it^2 \times \text{POP}_it \right] + 0.0654369 \text{DUM}_1 + 0.0239259 \text{DUM}_2 - 0.0005074 \text{EDU}_it \\
\]

\[
(4.23) \quad (0.72) \quad (-0.44) \quad (-0.97) \\
(0.03) \quad (2.87) \quad (2.04) \quad (0.82) \\
\]

\[R^2 = 0.2351 \quad \text{No. of Observations} = 73\]

The above results show that the coefficients of all independent variables except the dummy variables are insignificant and thus have not affected urban inequality significantly. The significant coefficients of the dummy variables show that urban inequality has fallen over time from 1983 to 2000.

The above three regressions give similar results and the coefficients of the income term is positive whereas the coefficient of square of income term is negative giving evidence for inverted-U shaped Kuznets hypothesis but as these coefficients as insignificant in all the above cases we can conclude that the inverted-U shaped hypothesis does not hold true for the above empirics.

Now we look at the effect of dropping the square of income terms for all the regression and how it alters the result:
RURGIN = 0.191724 + 2.28e^{-06}\text{INC}_{it} + 0.001884\text{EDU}_{it} - 6.14e^{-06}\text{POP}_{it}

(7.13) (1.60) (0.43) (-1.84)*

+ 0.0869433\text{DUM}_1 + 0.0324231 \text{DUM}_2

(5.40)* (3.05)*

R^2 = 0.3621 No. of Observations = 73

The above results show that the coefficient of population density and the dummy variables are significant. The negative sign of the coefficient of the population density implies that as the population density in rural areas rises, it leads to an increase in inequality. A possible explanation could be that the increase in population density in the rural areas is due to an increase in the dependent population and thus has not reached a stage where its demographic dividend for most people can be realized, leading to an increase in inequality. The significant coefficient of the dummy variable shows that rural inequality has fallen over time from 1983 to 2000.

URBGIN = 0.2425145-1.18e^{-06}\text{INC}_{it} + 0.000778\text{EDU}_{it} + 9.60e^{-06}\text{POP}_{it}

(6.49) (-0.60) (1.29) (2.07)*

+ 0.0279526\text{DUM}_1 + 0.019821 \text{DUM}_2

(1.25)(1.34)

R^2 = 0.3621 No. of Observations = 73

t-statistics are written in parenthesis

In the above regression only the coefficient of population density is significant and is of positive sign. The difference between the urban and rural cases could be because the urban area is at a higher stage of demographic dividend and thus an increase in population density means that the earning population increases (due to migration etc.) and thus leading to a fall in inequality. As the levels of health, education, social participation and economic as well as social awareness are still lower in the rural areas of the country as compared to the urban areas the population in the rural areas is not able to exploit the resources available to them efficiently and thus the difference between the developed and underdeveloped increases leading to an increase in inequality. For urban areas the difference between underdeveloped and developed is lower due to the global facilities available to all and thus leading to a fall in inequality.

CMBGIN = 0.2147254 + 7.01e^{-07}\text{INC}_{it} + 0.0004698\text{EDU}_{it} + 4.86e^{-06}\text{POP}_{it}

(7.52) (0.46) (1.01) (0.14)
.0595875DUM₁ + .0264706 DUM₂

(3.49)* (2.34)*

R² = 0.1970  No. of Observations= 73

t-statistics are written in parenthesis

In the above regression only the coefficients of the dummy variables are significant implying that the overall inequality in Indian states has fallen as a whole.

We removed all the square of income terms from the regression and then regressed the rural, urban and combined gini coefficients on income, education, population density and the dummies. The coefficient of the population density term in the case of rural and urban regressions becomes significant at 10% and 5% levels respectively. In the case of combined gini it still remains insignificant. But removing the squares of incomes from the regressions would not be an ideal assumption as it would make the model weaker.

CONCLUSION

From the above regressions we conclude that if the square of per capita income is included in our regression model, there seems to be no significant relation between income and inequality. Furthermore, there is no significant evidence that supports the inverted U-shaped Kuznets hypothesis leading to the conclusion that it does not hold true in the Indian context. If square of per capita income term is removed from the regression a relation between population density and inequality arises which is different for rural and urban areas. In rural areas inequality rises with population density whereas it falls in urban areas.

The model has been tested for heteroskedasticity as well as autocorrelation for all the above regressions. There is no biasness in the model and the model is robust.

The model is not free of all limitations. The no. of observations is 73 and it uses a common dataset of the independent variables for three different dependent variables i.e. rural gini coefficient, urban gini coefficient and the combined gini coefficient.

REFERENCES


degradation: the environmental Kuznets curve and sustainable development. World development, 24(7), 1151-1160.


