Non-parametric investigation of the Kuznets hypothesis for transitional countries

Demidova Olga[1]
State University – Higher School of Economics

Abstract

The Kuznets hypothesis states that the relation between income distribution and economic development is characterized by an inverted - U curve. In other words, inequality in the income distribution first rises but then falls. Many researchers have verified this hypothesis, using the data for developed countries excluding transitional countries.

While testing for the Kuznets inverted-U relationship, most studies follow the parametric quadratic specification by regressing the Gini index on GDP per capita and its squared term. But goodness of fitting for such models is usually quite low.

The purpose of this paper is to test the Kuznets hypothesis using the data for 29 transitional countries. We estimate the unknown relationship between the Gini index and GDP per capita estimating the Nadaraya-Watson nonparametric regression with Gaussian kernel. The graphical outcomes show clear evidence of an inverted-U relationship between the Gini index and per capita GDP. We use an alternative measure of income distribution, namely, the ratio of incomes for 20% richest and 20% poorest people; the ratio of incomes for 10% richest and 10% poorest people reproduces that estimation and results in the same inverted-U shape of the curve. The graphical evidence confirms the validity of the Kuznets' hypothesis for transitional countries.

All the estimated functions $Y_i = f(X)$, i = 1,2,3 (where Y_1 is the Gini index, Y_2 is the ratio of incomes for 20% richest and poorest people, Y_3 is the ratio of incomes for 10% richest and poorest people, X is per capita GDP) decrease for GDPs per capita which are greater than ca. 11000 (PPP USD). The GDP per capita for Russia equals 9902 (PPP USD) and we can expect reduction of the inequality level in this country with increasing GDP per capita.
.

JEL Classification: O11, O15 , C14

Keywords: Kuznets' hypothesis; Inverted -U curve; Transitional countries

1. Introduction

The central topic of this paper is the relation between the income distribution inequality and the mean income increase. When does the inequality in the distribution of income decrease as the mean income increases? In 1955 S.Kusnets (Kusnets S., 1955) suggested that the value of income inequality rises initially, and then drops after reaching a turning point. This hypothesis is referred to as the Kuznets' hypothesis of the inverted - U curve. The Gini index is the most commonly used measure of income inequality. That is why many investigators search how the Gini index or the low income share depends on per capita GDP using cross – countries survey. Barro R. (Barro R., 2000), Papanek G. and Kyn O. (Papanek G.1986) regress the Gini index on the per capita GDP and its squared term, D.Mushinski (D. Mushinski, 2001) used fourth-degree polinom. Auluwalia M. (Auluwalia M., 1976) observed a U-shaped relationship between the low income share and per capita GDP.

[1] E-mail:demidova@hse.ru

Ho-Chuan River Huang (Ho-Chuan River Huang, 2004) used nonparametric approach to test the validity of the Kuznets hypothesis' inverted-U shape.

Some authors note that the nature of the relationship differs according to a country's level of economic development and divide countries into two groups (developed and less developed) as a prerequisite to testing the Kuznets hypothesis. Savvides A. (Savvides A., 2000) used the threshold regression model.

Sukiassyan G. (Sukiassyan G., 2007) remarks that the existing literature on the inequality and economic development "has virtually ignored transition economies" and his "paper fills an important gap on the theme". The author indicates that the effect of inequality on growth is negative for the transition economies of Central and Eastern Europe and the Commonwealth of Independent States.

This paper pursues the following two objectives:

1) To test the Kuznets hypothesis on the theoretical level, to determine the conditions on which the inverted-U dependence of the Gini coefficient on the mean income might take place, and to give an economic interpretation of the obtained mathematic al results;

2) To test whether the Kuznets hypothesis is valid for countries with transition economy.

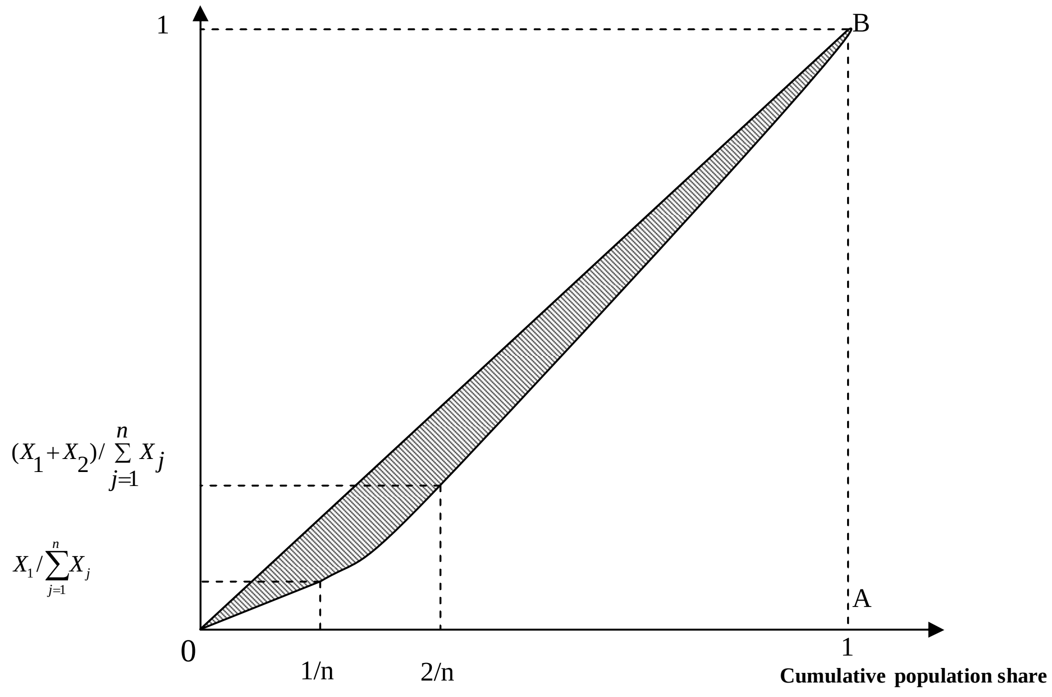
2. Theoretical approach

Suppose the population of a country is divided into n equal groups and the group's income increases with its number.

Let X_1 be the income of the poorest group, X_n be the income of the richest group, $Z = \frac{1}{n}\sum_{i=1}^{n} X_i$ is the mean income, $p_i = \frac{X_i}{\sum_{j=1}^{n} X_j}$ is the income share of i-th group, i=1,…,n.

The ratio of the crosshatched region area to the area of the triangle OAB multiplied by 100% is called the Gini index G (Fig.1)

Figure 1. The Gini index



One can show that

$$G = \left(1 - \frac{1}{n} - 2\frac{n-1}{n^2}\cdot\frac{X_1}{Z} - 2\frac{n-2}{n^2}\cdot\frac{X_2}{Z} - ... - \frac{2}{n^2}\cdot\frac{X_{n-1}}{Z}\right)\cdot 100\% \quad (1)$$

or $G = \left(1 - \frac{1}{n} - 2\frac{n-1}{n}\cdot p_1 - 2\frac{n-2}{n}\cdot p_2 - ... - \frac{2}{n}\cdot p_{n-1}\right)\cdot 100\%$ (2)

Thus G linearly depends on X1,…,Xn-1 and inversely on Z. The Gini index G also linearly depends on the income shares p1,…,pn-1. The coefficients of the income shares p1,…,pn-1 are negative and their absolute values decrease as the number of the income share (and the corresponding income) increases.

Remark 1. For the case of quantile groups (n = 5) from formula (2) it follows that

$$G = 100\% \cdot (0.8 - 1.6 p_1 - 1.2 p_2 - 0.8 p_3 - 0.4 p_4)$$

Generally, the Gini index G is a function of n variables: X_1, X_2,…,X_{n-1}, Z.

Note that

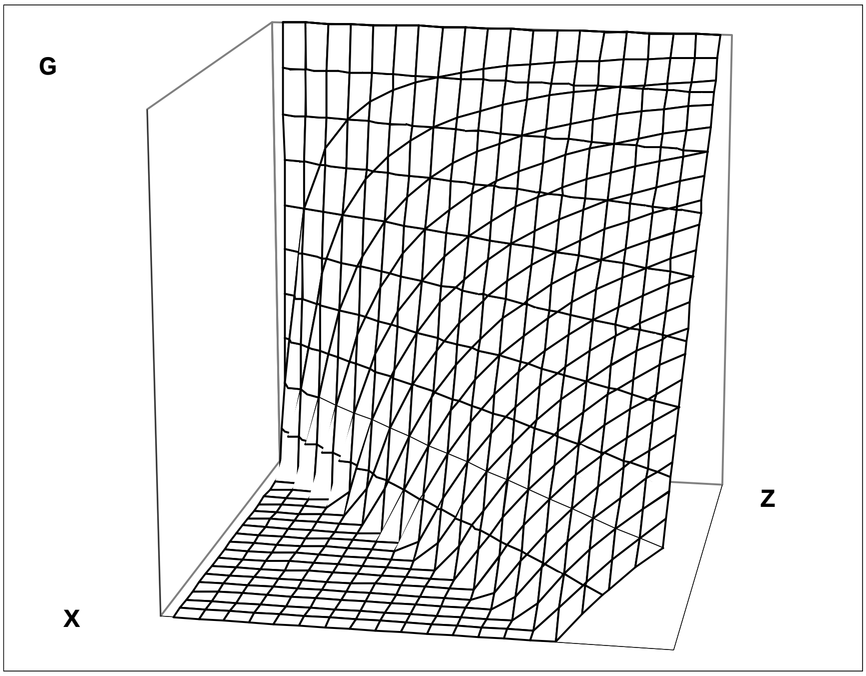
$$\frac{\partial G}{\partial Z} = 2 \cdot \left(\frac{n-1}{n^2} \cdot X_1 + \frac{n-2}{n^2} \cdot X_2 + \ldots + \frac{1}{n^2} \cdot X_{n-1} \right) \cdot \frac{1}{Z^2} \cdot 100\% > 0, \quad (3)$$

$$\frac{\partial G}{\partial X_i} = -2 \cdot \frac{n-1}{n^2} \cdot \frac{1}{Z} \cdot 100\% < 0, \quad i = 1, \ldots, n-1 \quad (4)$$

Hence, the Gini index increases as the mean income increases (the incomes of n-1 poorest groups being constant) and decreases as the income of any income group with number 1,…,n-1 increases (other factors being constant).

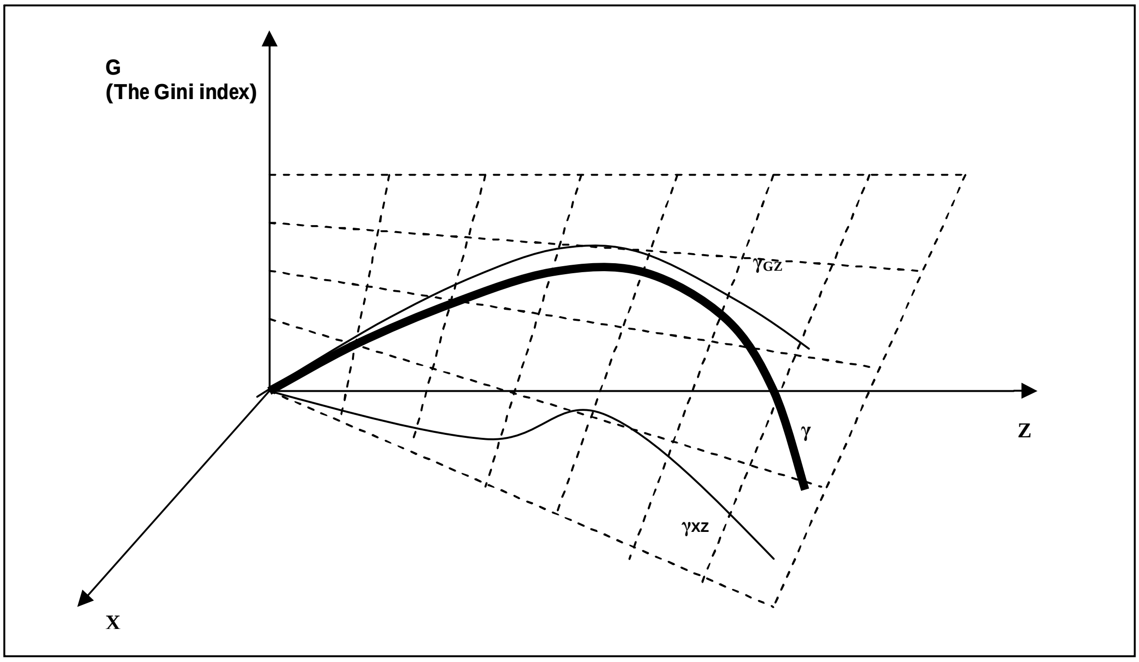
The graph of the function G is presented in Fig. 2. Hereafter we keep only variables Z and X_1 for simplicity. In general case, the graph of the function G coincides with n – dimensional manifold $\tilde{G}_n$.

Figure 2: Manifold



Suppose γ is a smooth curve on the manifold $\tilde{G}_n$, γ_{GZ} is the projection of the curve γ onto the plane GOZ, γ_{XZ} is the projection of the curve γ onto the plane X_1OZ (Fig.3). Let γ_{GZ} be inverted-U curve. It is quite interesting to determine the form of the curve γ_{XZ} in this case.

Figure 3: Projections

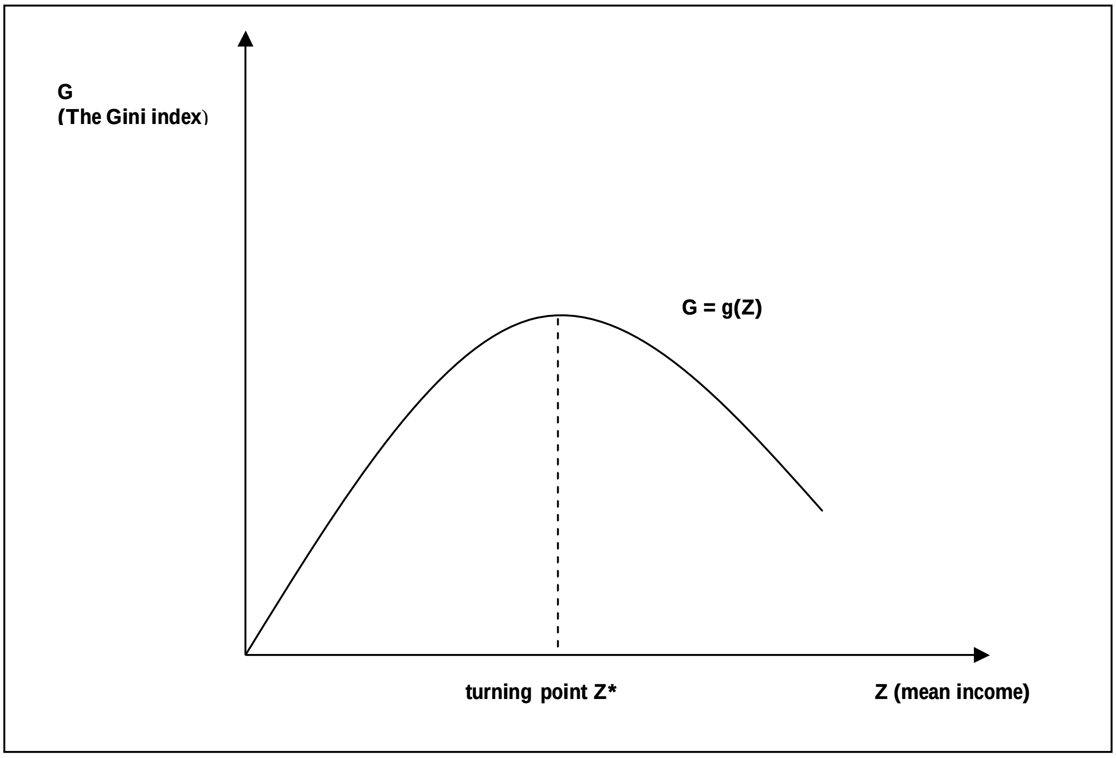


Let the curve γ_{GZ} be presented as G = g(Z). Suppose that

$$\begin{cases} g'(Z) > 0 \;\; for \; 0 \le Z < Z^*, \\ g'(Z^*) = 0, \\ g'(Z) < 0 \;\; for \; Z > Z^* \\ and \;\; g''(Z) < 0 \end{cases} \qquad (5)$$

The graph of the function g(Z) is presented in Fig. 4.

Figure 4: Inverted – U form



Substituting g(Z) for G and zero for $X_2, \ldots, X_{n-1}$ in (1), we obtain expression for the curve γ_{XZ}: $X_1 = \varphi(Z)$, where

$$\varphi(Z) = \frac{1}{2} Z(n - \frac{n^2}{n-1} g(Z)) \quad (6)$$

Differentiating both sides (6) two times, we obtain

$$\varphi'(Z) = \frac{1}{2}\left(n - \frac{n^2}{n-1} g(Z)\right) - \frac{n^2}{2(n-1)} \cdot Z \cdot g'(Z), \quad (7)$$

$$\varphi''(Z) = -\frac{n^2}{2(n-1)}(2g'(Z) + Zg''(Z)) \quad (8)$$

Substituting Z^* for Z in (7) and (8) and note that $g'(Z^*) = 0$, we get

$$\varphi'(Z^*) = \frac{1}{2}\left(n - \frac{n^2}{n-1} g(Z^*)\right) \quad (9)$$

$$\varphi''(Z^*) = -\frac{n^2}{2(n-1)} Z^* g''(Z^*) \quad (10)$$

Using (6) for Z = Z^*, we get $\varphi(Z^*) = \frac{1}{2} Z^* \cdot \left(n - \frac{n^2}{n-1} g(Z^*)\right)$, hence

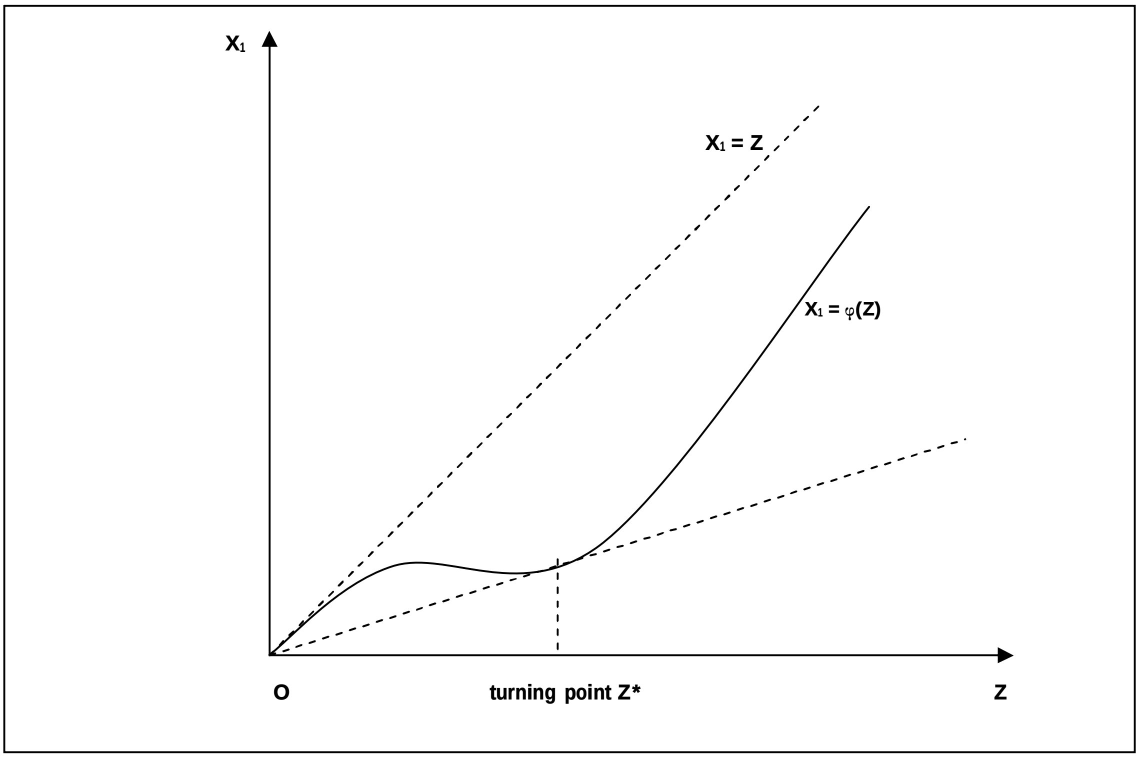
$$\varphi'(Z^*) = \frac{\varphi(Z^*)}{Z^*} \quad (11)$$

Taking into account (7), (8) and (5), we obtain

$$\varphi'(Z) > 0 \textit{ and } \varphi''(Z) > 0 \textit{ for } Z \geq Z^* \quad (12)$$

From (10), (11), (12), we get the following graph for the function $\varphi(Z)$ (Fig.5). Function $\varphi(Z)$ is convex for $Z \geq Z^*$.

Figure 5: Projection on the plane X_1OZ



Remark 2. If g(Z) is a polynomial of degree k then φ(Z) is also a polynomial with degree k+1. E.g. if g(Z) is a quadratic function then φ(Z) is a cubical function.

Remark 3. If the projection of the curve γ onto the plane GOZ has inverted-U form, then the projection of this curve onto any plane XjOZ (j = 2,…,n-1) has the same form as function $\varphi(Z)$ (Fig.5).

Remark 4. Suppose G = g(Z) where g(Z) has inverted-U form. Then the relationship between p_i, i = 1,…,n - 1 and mean income Z is U-shaped. It follows from (2).

Let us state the main result of this section. In order for the Gini index to begin dropping starting from a certain level of the mean income Z^*, it is essential for the income of low-income groups to increase with the mean income growth. In particular, for the Gini coefficient to decrease quadratically, the income of the most low-income group X_1 should increase cubically.

3. Empirical approach

Suppose we have a sample for m countries. Let us denote the set of observations for i – th country by $A_i = (X_1^i, X_2^i, ..., X_{n-1}^i, Z^i, G^i)$, i = 1,…,m, where X_j^i is an income of the j – th group in the i – th country, j = 1,…,n-1, Z^i is the mean income in the i – th country, G^i is the Gini index for the i – th country. Then points A1,…,Am belong to the manifold $\tilde{G}_n$. This set of points is a proxy for the curve γ.

Using the projections of the points A1,…,Am onto the planes GOZ and X1OZ we estimate the functions g(Z) and $\varphi(Z)$. First, we try to estimate the parameters of the functions g(Z) and $\varphi(Z)$ using quadratic and cubical specification correspondingly. Then we use nonparametric specification.

4. Data

The data set used in this study is taken from the Human development report (2006) of World Bank (Appendix). 29 transitional countries were chosen. "One attractive feature of this group of countries is that their starting points were remarkably similar. Yet, they subsequently have experienced substantial divergence in growth rates and income inequality" (Sukiassyan G., 2007).

For each of the countries, three variables are considered. Those include the Gini index (denoted by GINI, a measure of inequality), the GDP per capita (PPP USD, denoted by GDP, a proxy for the mean income), and the 20% low income share (denoted by $P20_-$). We also create the new variable X20 – the income of the 20 % low income share, where $X20_- = 0.05 \cdot P20_- \cdot GDP$.

5. Empirical results

5.1. Parametric specification

The traditional regressions have been estimated in the following specification:

$$GINI = \beta_0 + \beta_1 GDP + \beta_2 GDP^2 + \varepsilon \quad (13)$$

$$P20_- = \beta_0 + \beta_1 GDP + \beta_2 GDP^2 + \varepsilon \quad (14)$$

$$X20_- = \beta_0 + \beta_1 GDP + \beta_2 GDP^2 + \beta_3 GDP^3 + \varepsilon \quad (15)$$

The results of running are presented in Tables 1, 2 and 3.

Table 1. Regression of the Gini coefficient on per capita GDP and its squared term

Source	SS	df	MS	Number of obs = 29	F(2, 26) = 1.25
Model	81.6520979	2	40.8260489	Prob > F = 0.3041	
Residual	851.578306	26	32.7530118	R-squared = 0.0875	Adj R-squared = 0.0173
Total	933.230404	28	33.3296573	Root MSE = 5.723	

GINI	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
GDP	.00059	.0007261	0.81	0.424	-.0009024	.0020825
GDP^2	-4.00e-08	3.49e-08	-1.15	0.262	-1.12e-07	3.17e-08
_cons	31.37331	3.020073	10.39	0.000	25.16546	37.58116

Table 2. Regression of the 20 % low income share on per capita GDP and its squared term

Source	SS	df	MS	Number of obs = 29	F(2, 26) = 1.53
Model	7.23092625	2	3.61546312	Prob > F = 0.2358	
Residual	61.5070073	26	2.36565413	R-squared = 0.1052	Adj R-squared = 0.0364
Total	68.7379336	28	2.4549262	Root MSE = 1.5381	

P20-	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
GDP	-.0002719	.0001951	-1.39	0.175	-.000673	.0001292
GDP^2	1.52e-08	9.38e-09	1.63	0.116	-4.03e-09	3.45e-08
_cons	8.777689	.8116476	10.81	0.000	7.109324	10.44605

Table 3. Regression of the 20 % low income on per capita GDP, its squared and cubic terms

Source	SS	df	MS	Number of obs = 29	F(3, 25) = 175.86
Model	174959631	3	58319877.1	Prob > F = 0.0000	
Residual	8290576.31	25	331623.052	R-squared = 0.9548	Adj R-squared = 0.9493
Total	183250208	28	6544650.27	Root MSE = 575.87	

X20-	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
GDP	.4405587	.1918244	2.30	0.030	.0454889	.8356285
GDP^2	-.0000155	.0000211	-0.73	0.469	-.0000589	.0000279
GDP^3	8.61e-10	6.54e-10	1.32	0.200	-4.86e-10	2.21e-09
_cons	31.50217	463.1688	0.07	0.946	-922.4119	985.4163

The output indicates that regressions (13) and (14) are insignificant although all coefficients have expected signs. Increasing the polynomial degree doesn't change the situation. The coefficients of the nonlinear powers of per capita GDP in regression (15) are insignificant.

5.2. Non - Parametric specification

The low goodness of fitting and insignificance of the polynomial regressions coefficients was a reason why we estimated the unknown relationship between the Gini index and per capita GDP (and two other relationships) using the following relationship:

$$GINI = m(GDP) + \varepsilon \quad (16)$$
$$P20\_ = m(GDP) + \varepsilon \quad (17)$$
$$X20\_ = m(GDP) + \varepsilon \quad (18),$$

The conditional expectation function, *m(…)* was estimated using the Nadaraya - Watson estimator and the Gaussian kernel.
Figures 6 - 8 contain kernel regressions (16) – (18) results.

Fig.6 The estimated conditional mean of the Gini coefficient on per capita GDP

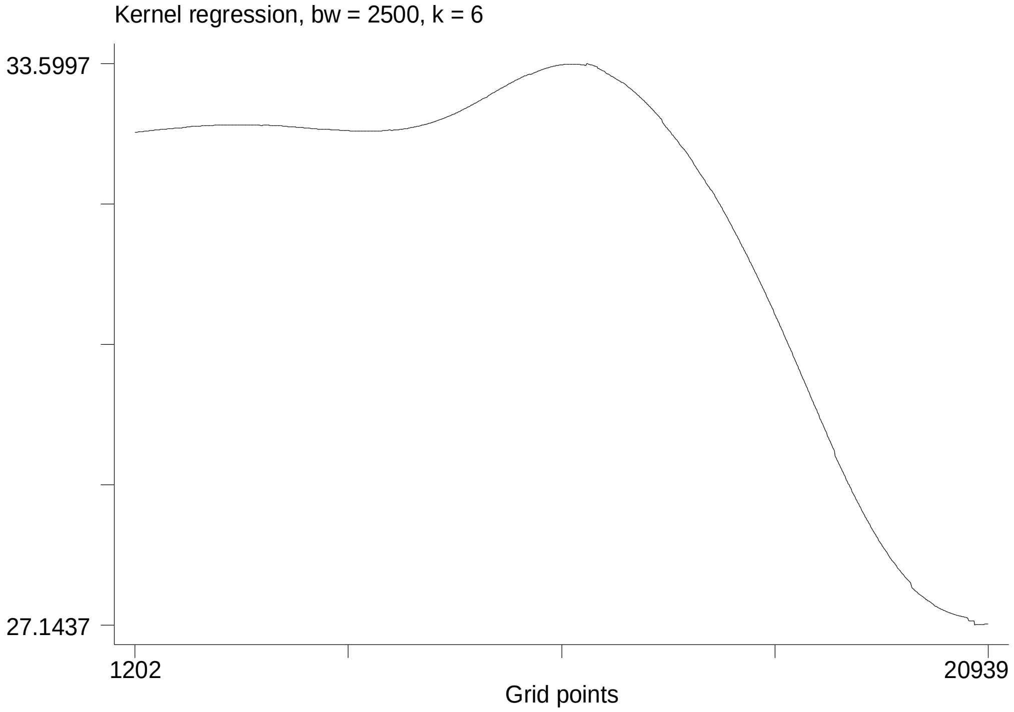


Fig.7 The estimated conditional mean of the 20 % low income share on per capita GDP

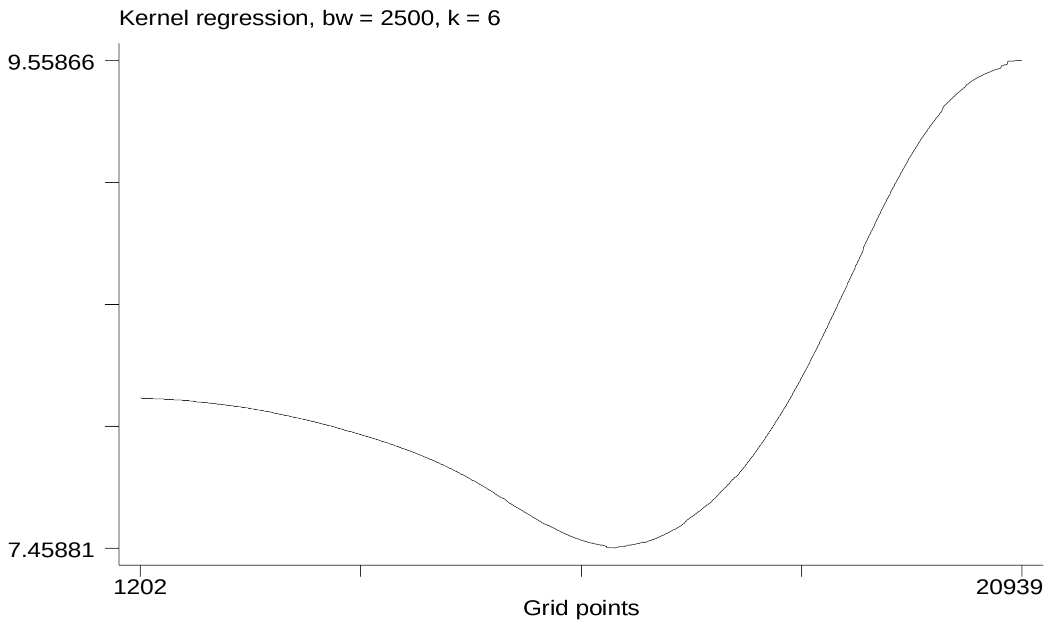
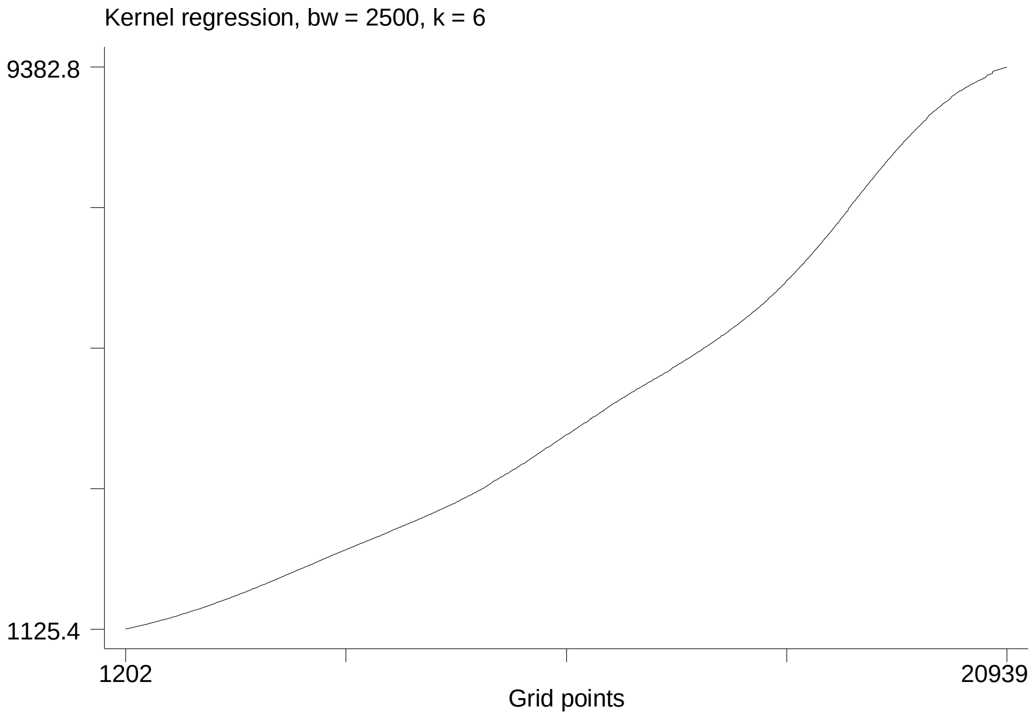


Fig.8 The estimated conditional mean of the 20 % low income on per capita GDP



As seen from Fig. 6 – 8, the Kuznets hypothesis is supported for the transition countries. Indeed, the dependence of the Gini index on per capita GDP takes an inverted-U form, the dependence of the 20 % low income share on per capita GDP takes a U- form, the dependence of the 20 % low income on per capita GDP looks alike the graph of the function φ in Fig. 5.

According to Fig. 6, the turning point is ca. 11000 (PPP USD). The GDP per capita for Russia equals 9902 (PPP USD). All countries with greater GDP per capita have the Gini index smaller than Russia. That is why we can expect reduction of the inequality level in this country with increasing GDP per capita.

Some deviation of the practical results from the theoretical ones is observed only at the edges that is typical for kernel regression. This problem can be solved, say, by using splines.

Remark 5. We use an alternative measure of income distribution, namely, the ratio of incomes for 20% richest and 20% poorest people; the ratio of incomes for 10% richest and 10% poorest people with the same result.

Remark 6. The dependence of the 10 % low income share on per capita GDP also has a U-form. The dependence of the 10% low income on per capita GDP has the same form as the one for 20% low income.

6. Conclusion

This article is devoted to the verification of the Kuznets hypothesis about the relation between income distribution and the mean income. It has been shown that the Gini index is a function of the mean income and the incomes of all income groups except the richest group.

The suggestion about an inverted-U shape of the dependence of the Gini index on the mean income was formalized using the conditions for the first and second derivatives of certain functions. When meeting these conditions, the form of the low-income groups' income dependence on the mean income was determined. The drop in the Gini index after reaching the turning point is possible only with the low-income groups' income growth being ahead. The shape of the low-income groups' income dependence on the mean income after reaching the turning point must be convex.
The data on 29 countries confirm the validity of the Kuznets hypothesis for transition countries. For this group of countries the turning point of ca. 11000 PPP USD was found. Among the countries with lower GDP per capita, Russia is the closest one to the turning point. We can expect reduction of the inequality level in this country with increasing GDP per capita.

References

Ahluwalia M. (1976), "Income distribution and development", *American Economic Review* 66, pp. 128–135

Barro R. (2000), "Inequality and growth in a panel of countries", *Journal of Economic Growth* 5, pp. 5–32.

Ho-Chuan River Huang (2004), "A flexible nonlinear inference to the Kuznets hypothesis", Economics Letters, vol. 84, issue 2, pp. 289-296

Kusnets, S. (1955), "Economic growth and income inequality", *American Economic Review* 45, pp.1–28

Mushinski D.(2001), "Using non-parametrics to inform parametric tests of Kuznets' hypothesis", *Applied Economics Letters*, pp. 77–79

Papanek G. and Kyn O. (1986), "The effect on income distribution of development, the growth rate and economic strategy", *Journal of Development Economics* 23, pp. 55–65

Savvides A. and Stengos T. (2000), "Income inequality and economic development: evidence from the threshold regression model", *Economics Letters,* vol. 69, issue 2, pp. 207-212

Sukiassyan G. (2007), "Inequality and growth: What does the transition economy data say?", *Journal of comparative economics,* vol.35, issue 1, pp. 35-56

Appendix

	Gini index	GDP per capita. PPP ($)	Share of income or consumption. poorest 10% (%)	Share of income or consumption. poorest 20% (%)
27 Slovenia	28.4	20939	3.6	9.1
30 Czech Republic	25.4	19408	4.3	10.3
35 Hungary	26.9	16814	4	9.5
37 Poland	34.5	12974	3.1	7.5
40 Estonia	35.8	14555	2.5	6.7
41 Lithuania	36	13107	2.7	6.8
42 Slovakia	25.8	14623	3.1	8.8
44 Croatia	29	12191	3.4	8.3
45 Latvia	37.7	11653	2.5	6.6
54 Bulgaria	29.2	8078	3.4	8.7
60 Romania	31	8480	3.3	8.1
62 Bosnia and Herzegovina	26.2	7032	3.9	9.5
65 Russian Federation	39.9	9902	2.4	6.1
66 Macedonia, TFYR	39	6610	2.4	6.1
67 Belarus	29.7	6970	3.4	8.5
73 Albania	28.2	4978	3.8	9.1
77 Ukraine	28.1	6394	3.9	9.2
79 Kazakhstan	33.9	7440	3	7.4
80 Armenia	33.8	4101	3.6	8.5
81 China	44.7	5896	1.8	4.7
97 Georgia	40.4	2844	2	5.6
99 Azerbaijan	36	4153	5.4	12.2
105 Turkmenistan	40.8	4584	2.6	6.1
109 Viet Nam	37	2745	3.2	7.5
110 Kyrgyzstan	30.3	1935	3.8	8.9
113 Uzbekistan	26.8	1869	3.6	9.2
114 Moldova	33.2	1729	3.2	7.8
122 Tajikistan	32.6	1202	3.3	7.9
133 Lao People’s Dem. Rep.	34.6	1954	3.4	8.1