Quantile Regression Analysis on Convergence of China's Regional Economic Growth

Kun HE [*]

School of Economics, Ocean University of China, Qingdao 266100, China

Abstract　Using quantile regression method, this paper made an empirical analysis on convergence of China's regional economic growth since the reform and opening-up. It firstly introduced principle of quantile regression method and related theories of convergence of economic growth. Through discussing interprovincial variation coefficient of GDP per capita, it carried out σ convergence analysis on economic growth and divided 3 decades since the reform and opening-up into 3 stages. Then, it made a comparative analysis of absolute β convergence on 3 stages using least-squares estimation and quantile regression method, and also stressed the advantage of quantile regression method. On this basis, it made an in-depth study on conditional β convergence at 3 stages. Empirical results indicate that there is absolute and conditional convergence at the first stage, no convergence at the second stage, and weak convergence at the third stage. Finally, it discussed weak points in this study and came up with recommendations for future studies.

Key words　Quantile regression, Convergence of economic growth, σ convergence, β convergence

1　Introduction

Since the introduction of neoclassical economic growth model (Solow, 1956), a lot of economists studied characteristics of convergence of economic growth on the basis of this model. Compared with theoretical studies on convergence, empirical studies started from the middle and later period of the 1980s, but the development momentum is considerable. In foreign countries, Baumol firstly made empirical study on convergence in 1986. He made a regression using Maddicon method and reached the conclusion: there is a high correlation between the growth rate and initial output level. At early stage, countries with higher productivity have slower economic growth rate. In other words, backward countries get convergent to advanced countries and the convergence speed is gradually accelerating. From then on, more and more economists started studying convergence of economic growth. For example, in the classical book, Economic Growth, Barro, R&Xavier Sala-i-Matin took neoclassical economic growth model as basic framework, divided convergence into σ convergence, absolute β convergence and conditional β convergence. Their studies play a great role in development of empircal studies of convergence, no matter in expansion of connotation and extension of convergence concept. Besides, some scholars put forward "club convergence" concept, economic regions with the same human capital and market opening degree have certain growth convergence trend. In China, the studies on convergence started from the middle and later period of the 1990s. Most scholars made empirical analysis on related theories of neoclassical economic growth model. For example, Wei Houkai (1997) reached the conclusion that GDP per capita of provinces in China was convergent at an annual speed of 2% in 1978 – 1995. Xu Xianxiang et al. (2004) studies convergence of 216 cit-ies above prefecture level in 1989 – 1999, and proved that there was σ convergence and absolute β convergence, and there was neoclassical convergence mechanism and new growth convergence mechanism, etc. From the above introduction, we can see that most scholars use general linear regression analysis method. Such method can only analyze the average level of study object influenced by other factors, and it is difficult to solve the problem of degree of different levels of study objects influenced by various factors.

In this study, quantile regression method is firstly used to analyze convergence of China's economic growth. The quantile regression method can select different quantile and can fully describe panorama of the study object, rather than the average value analysis. Besides, parameter estimation in different quantile is also different, indicating that the same influence factor has different function to the study object at different level. Especially, the distribution in study object takes on heterogeneity, such as asymmetric, fat-tailed distribution and truncation, such method generally provides more detailed information and has prominent advantages.

2　Introduction to quantile regression method

Quantile regression is a type of regression analysis used in statistics and econometrics. Whereas the method of least squares results in estimates that approximate the conditional mean of the response variable given certain values of the predictor variables, quantile regression aims at estimating either the conditional median or other quantiles of the response variable. The median regression is a special case of quantile regression. It uses symmetric weight to solve the problem of minimal residual error, while other conditional quantile regression methods use asymmetric weight to solve the problem of minimal residual error.

General linear regression model can be set as follows:

$$\rho_x(t) = t(\tau - i(t<0)), \ \tau \in (0, 1)$$

On the precondition of Gauss − Markov assumption, it can be expressed as follows:

$$E(y|x) = a_0 + a_1x_1 +_a 2x_2 + a_kx_k$$

where u is the random perturbation tem, and a_0, a_1, a_2, $\cdots$, a_k is coefficient of variable to be estimated. This is the expression of mean value regression OLS model, similar to mean value regression model and the quantile regression model can be defined as follows:

$$Q_y(\tau|x) = a_0 + a_1x_1 + a_2x_2 + \cdots + a_kx_k + Q_u(\tau)$$

For quantile regression model, it is feasible to adopt linear programming (LP) to estimate the weighted least absolute deviation, so as to obtain regression coefficient of explanatory variable. It can be expressed as follows:

$$\min E\rho_x(y - a_0 - a_1x_1 - a_2x_2 - \cdots - a_kx_k)$$

Through calculation, we get $\hat{Q}_y(\tau|x) = \hat{a}_0 + \hat{a}_1x_1 + \hat{a}_2x_1 + \cdots + \hat{a}_kx_k$

where, $\log(y_{i,T}|y_{i,0}) = \beta_0 + \beta_1\ln(y_{i,0}) + \beta_2\ln(I/GDP) + \beta_3\ln(n + g + \partial) + \beta_4\ln(h) + \varepsilon_{i,0,T}$.

From estimation method of parameters, the principle of general linear regression model is to make the square sum of difference between explanatory variable y and fitted value (residual error) being the least. The quantile regression is just an expression making absolute value of this residual error become the least. This expression is not differential. Thus, traditional derivation method is not suitable, we have to adopt linear programming or simplex algorithm. This is also one of the largest differences from general linear regression. With constant breakthrough in computer technology, the above algorithm can be realized with the aid of suitable software. In this study, qreg command of STATA10.0 is used to process data.

3 Model and data description

3.1 Principle of convergence of economic growth
Convergence of economic growth refers to negative correlation between static indicators (output per capita and income per capita) and the growth rate of countries or regions of different economic entities in an effective scope in closed economic condition. In other words, backward economic entity has higher growth rate than developed economic entity. As a result, the difference in static indicators of economic entities at initial stage gradually disappears. In 1992, Barro and Sala-i-Martin introduced a model for determining convergence of per capita income:

$$(1/T)\log y_{i,T}/y_{i,0} = \alpha - [(1 - e^{-\beta T})/T]\log(y_{i,0} + \mu_{i,0,T})$$
$$(1)$$

where $y_{i,T}$ denotes final per capita income, $y_{i,0}$ denotes base per capita income, disturbance term $\mu_{i,0,T}$ means average original error term, β is the coefficient of speed from convergence of control income to equilibrium status; constant $a = g[(1 - E^{-\beta T})/T]\log(\hat{y}^*)$, in which g refers to technological growth rate and $\hat{y}^*$ refers to per capita income at equilibrium level. If estimation result of parameter β is positive, the income is convergent, otherwise it is dispersed. This mode reflects that countries at the same technological level are convergent to the same equilibrium status. There-

fore, we should pay attention to whether there is negative correlation between initial income level and the income growth rate.

Linear expression of Equation 1 introduced by Baumol (1986) and Barro (1997) is as follows:

$$(1/t)\log(y_{i,T}/y_{i,0}) = a + b\log y_{i,0} + \mu_{i,0,T}$$
$$(2)$$

where negative correlation coefficient $b = -[(1 - e^{-\beta T})/T]$. Generally, using this expression, we can calculate convergence rate, and T refers to interval span of samples. The half-life period is calculated according to the formula $\ln(0.5)/\beta$.

Based on the above theoretical models (1) and (2), we adopt the following calculation equation:

$$\ln(\frac{Y_t}{Y_0}) = \beta_0 + \beta_1\ln Y_0 + \varepsilon_i$$
$$(3)$$

Mankiw, Romer and Weil (1992) (their method is the well-known MRW analysis framework) realized the limitation of neoclassical theories. They mainly corrected the assumption that all countries have the same consumption preference and technology, and re-interpreted the neoclassical economic growth theory. In empirical analysis, they took C-d production function $Y = K^\alpha H^\beta (AL)^{1-\alpha-\beta}$ as a tool. Through estimating approximate stable status, they built MRW model and obtained the convergence rate of a country getting towards stable status, i.e. the convergence rate $\lambda = (n + g + \delta)(1 - \alpha - \beta)$. Through introduction of human capital, this model set up an expanded econometric model using material capital investment, human capital investment and population growth rate, and proved effectiveness of neoclassical growth model. They believed that human capital slows down decreasing speed of marginal benefit of material capital. Despite the convergence rate is slower than predicted by traditional Solow model, it proved the existence of convergence. It should be noted that what the MRW model implies is only the conditional convergence, rather than absolute convergence. All countries have the trend of self equilibrium status in long-term development, rather than convergence towards the same equilibrium status. In this study, we selected two types of equation of this model. The first equation:

$$\log(y_{i,T}/y_{i,0}) = \beta_0 + \beta_1\ln(y_{i,0}) + \beta_2\ln(I/GDP) + \beta_3\ln(n + g + \partial) + \varepsilon_{i,0,T}$$
$$(4)$$

The second equation is expanded from the first equation plus human capital (h):

$$\log(y_{i,T}/y_{i,0}) = \beta_0 + \beta_1\ln(y_{i,0}) + \beta_2\ln(I/GDP) + \beta_3\ln(n + g + \partial) + \beta_4\ln(h) + \varepsilon_{i,0,T}$$
$$(5)$$

where $y_{i,T}$ denotes final GDP per capita; $y_{i,0}$ denotes base GDP per capita; I/GDP refers to the average investment rate from base period to final period; n stands for average growth rate of employees; $g + \delta$ denotes average technological growth rate and depreciation rate; in MRW model, it assumes reasonable value of $g + \delta$ is 0.05 or 0.1; h stands for human capital.

3.2 Data description
On the basis of theoretical model, we made quantile regression for annual data of related indicators of China's 29 provinces (except Tibet, Hainan and Chongqing) in 1978 − 2007. Sample data were collected from *China Compendium of Statistics* 1949 − 2004 and *China Statistical Yearbook* 2006 −

2008. denotes actual per capita GDP of provinces in the t-th year, corrected on the basis of chain price index in the same year; I/GDP refers to average investment rate of provinces in the sampling period, the original data were calculated by simple arithmetical average method; we took number of students of provinces in sampling base period as substitution variable of human capital, expressed with h; n stands for average growth rate of employees in provinces in sampling period, calculated by logarithm difference method; technological growth rate and depreciation rate are expressed with $g +$; east stands for dummy variable of eastern areas. For provinces in eastern areas (including Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, Guangxi, and Hainan), east takes 1; for provinces in other areas, it takes 0.

4   Quantile regression analysis on convergence of China's regional economic growth

According to assumptions of convergence theory, from the perspective of β convergence (conditional convergence), we made an in-depth analysis on convergence of China's regional economic growth using quantile regression method. Based on theoretical models (4) and (5), we studies 3 stages of economic growth in all provinces since the reform and opening-up.

(1) The first stage (1978 – 1991): At the first stage, we established Equation 1 according to MRW model. We adjusted this model, used dummy variables to replace insignificant variables, because eastern areas include state policies, marketization, and opening degree. Then we established the Equation 2, as listed in Table 1.

Table 1   Quantile regression results of conditional convergence for China's economic growth in 1978 – 1991

Variables	Quantile	Equation 1	Equation 2
Constant	0.1	−0.1890 (−0.39)	2.9029 (9.32 * * *)
	0.25	2.1301 (0.75)	3.0932 (7.16 * * *)
	0.5	2.2856 (0.93)	3.7570 (8.06 * * *)
	0.75	0.7617 (0.17)	3.9649 (5.53 * * *)
	0.9	−1.1393 (−0.48)	5.3440 (21.7 * * *)
$x_1 (\ln y t_0)$	0.1	−0.4113 (−11.01 * * *)	−0.2813 (−4.88 * * *)
	0.25	−0.3137 (−2.19 * *)	−0.3085 (−4.48 * * *)
	0.5	−0.3458 (−2.47 * *)	−0.3897 (−5.10 * * *)
	0.75	−0.2393 (−1.03)	−0.3988 (−3.41 * * *)
	0.9	−0.5884 (−3.53 * * *)	−0.5884 (−14.38 * * *)
$x_2 (\ln(I/GDP))$	0.1	0.1023 (6.09 * * *)	−0.1100 (−1.84 *)
	0.25	−0.1218 (−3.15 * * *)	−0.1040 (−2.35 * *)
	0.5	−0.0780291 (−1.46)	−0.0421 (−0.91)
	0.75	−0.0339676 (−0.35)	0.0414 (0.54)
	0.9	0.121177 (1.62)	0.1317 (2.8 * *)
$x_3 (\ln(h))$	0.1	0.0082 (0.38)	
	0.25	0.0378 (0.59)	
	0.5	0.0886 (1.08)	
	0.75	0.0051 (0.03)	
	0.9	0.1699 (1.6)	
$x_4 (\ln(n + g + \partial))$	0.1	−2.0164 (−6.1 * * *)	
	0.25	−0.4831 (−0.32)	
	0.5	−0.5471 (−0.39)	
	0.75	−1.1217 (−0.45)	
	0.9	−3.1899 (−2.02 *)	
east	0.1		0.1028 (0.84)
	0.25		0.1320 (1.25)
	0.5		0.1788 (1.63)
	0.75		0.3504 (1.79 *)
	0.9		0.4180 (4.95 * * *)

Through analysis on Equation 1 and Equation 2 in Table 1, we also obtained that there is a negative correlation between base GDP per capita and economic growth at this stage. In the Equation 1, when the investment rate in original theoretical model is in the decile and ninth-decile, the influence on economic growth takes on significant positive correlation. In other words, at this stage, increasing investment in low-income areas and high-income areas, economic growth rate will be higher. However, the human capital fails to pass the test in Equation 1, indicating that it has little con-

tribution to convergence of economic growth at this stage. For factors like technological progress, β coefficient in high-income areas is −3.1899, while it is −2.0164 in low-income areas, showing that the influence of technological progress on convergence of economic growth in high-income areas is more significant than that in low-income areas. In Equation 2, the dummy variable takes place of human capital and technological progress factors. The east variable passes the significance level test in high-income areas, indicating that there is club convergence in high-income areas. Be-

sides, the coefficient of per capita GDP in Equation 2 is more significant than Equation 1. With the increase in income level, β drops from -0.2813571 to -0.5884487, indicating that high-income areas have higher convergence rate, which is favorable for narrowing the gap between rich and poor, also indicating that there is conditional convergence of economic growth at this stage.

(2) The second stage (1992 – 2003): At the second stage, according to theoretical model of convergence of economic growth,

we have attempted to substitute all variables to the regression equation (1). Although the goodness of fit is high, per capita GDP, as major factor in the study on convergence of economic growth is not significant. To improve representativeness of the model, we made regression analysis through separately introducing variables. Through adjusting numbers of variables and introducing different variables, we selected Equation 2, Equation 3 and Equation 4. Detailed results are listed in Table 2.

Table 2　Quantile regression results of conditional convergence for China's economic growth in 1992 – 2003

Variables	Quantile	Equation 1	Equation 2	Equation 3	Equation 4
Constant	0.1	1.681 0(2.9***)	0.213 5(0.68)	0.531 3(1.97*)	0.241 0(0.61)
	0.25	0.433 9(0.26)	−0.286 6(−0.33)	0.629 0(1.47)	0.014 8(0.02)
	0.5	0.346 8(0.37)	−0.130 2(−0.16)	0.389 4(0.51)	−0.36 97(−0.52)
	0.75	0.918 2(0.94)	−1.003 9(−1.67)	0.283 3(0.27)	0.112 0(0.12)
	0.9	0.539 1(0.43)	−0.997 1(−2.38**)	−0.69 4 1(−2.32**)	−0.316 0(−0.69)
$\ln(y_{t0})$	0.1	−0.060 7(−0.72)	0.099 4(2.72**)	0.091 5(2.64***)	0.128 3(2.63**)
	0.25	0.101 9(0.55)	0.201 7(2.17**)	0.079 6(1.30)	0.172 3(2.04*)
	0.5	0.109 3(−0.46)	0.187 5(2.06**)	0.135 6(1.32)	0.204 6(2.31**)
	0.75	0.052 7(0.46)	0.297 8(4.11***)	0.15 52(1.09)	0.207 4(1.81*)
	0.9	0.103 6(0.69)	0.281 5(5.11***)	0.289 4(6.92***)	0.299 0(5.6***)
$\ln(i/GDP)$	0.1	0.342 4(3.00***)	−0.361 4(−3.5***)		
	0.25	−0.153 5(−0.35)	−0.170 1(−0.68)		
	0.5	−0.114 7(−0.46)	−0.235 0(−1.02)		
	0.75	−0.344 5(−1.61)	−0.344 6(−2.23**)		
	0.9	−0.359 0(−1.76*)	−0.552 6(−7.11***)		
$\ln(h)$	0.1	0.116 8(4.9***)		0.069 3(4.12***)	
	0.25	0.098 5(1.06)		0.095 6(1.92*)	
	0.5	0.000 5(0.01)		0.051 4(0.83)	
	0.75	−0.023 6(−0.42)		0.072 9(1.30)	
	0.9	−0.022 4(−0.28)		0.124 4(6.4***)	
$\ln(n+g+\partial)$	0.1	−0.081 0(−1.28)			−0.060 7(−0.82)
	0.25	0.046 9(0.58)			−0.032 5(−0.41)
	0.5	−0.075 5(−0.82)			−0.123 5(−1.14)
	0.75	0.046 9(0.36)			0.025 3(0.19)
	0.9	0.050 3(0.5)			0.125 8(2.65**)
east	0.1	0.184 3(1.61)			
	0.25	−0.017 8(−0.15)			
	0.5	0.1533 372(1.39)			
	0.75	0.237 995 8(2.17**)			
	0.9	0.281 080 7(2.67**)			

In Equation 2, apart from major factor per capita GDP, we also introduced the investment rate. Through analysis on Equation 2, we studied the influence of per capita GDP on convergence of economic growth. Coefficient value of β increased from decile 0.099 4 to ninth decile 0.281 5, indicating that there is no conditional β convergence at this stage, and the dispersion speed in high-income areas is higher than that in low-income areas. In Equation 3, we introduced the human capital; the coefficient value of per capita GDP is increasing and the coefficient of human capital increases from 0.069 3 in low-income areas to 0.124 4 in high-income areas, indicating that the contribution of human capital to economic growth in high-income areas is higher at this

stage. In Equation 4, we introduced the technological progress factor. Through analysis, we found that the coefficient value of per capita GDP is also increasing, and there is a significant positive correlation between technological progress and economic growth in high-income areas, indicating that increasing technological investment in high-income areas will promote rapid economic development. However, in backward areas, increasing technological investment will waste fund and not promote economic growth. Generally speaking, there is positive correlation between per capita GDP and economic growth at this stage. From low-income areas to high-income areas, the economic growth rate is increasing. In other words, the gap between rich and poor is gradually widening.

（3）The third stage （2004 – 2007）：At the third stage, based on the above model （5）, we made quantile regression analysis on convergence of China's economic growth. The results are listed in Table 3.

Table 3　Quantile regression results of conditional convergence for China's economic growth in 2004 – 2007

Variables	Quantile				
	0.1	0.25	0.5	0.75	
$\ln y_{t_0}$	−0.115 4(−7.66)***	−0.054 2(−0.79)	−0.057 1(−0.73)	−0.082 6(−0.65)	−0.137 7(−3.07)***
$\ln(i/GDP)$	0.219 3(4.8)***	0.283 5(1.45)	0.206 7(1.01)	0.395 2(1.3)	0.512 1(8.37)***
$\ln(h)$	0.005 6(0.6)	0.017 1(0.56)	0.033 3(0.78)	0.024 0(0.58)	0.059 9(4.07)***
$\ln(n+g+\partial)$	−0.155 1(−1.89)*	0.049 1(0.23)	−0.010 7(−0.03)	0.088 9(0.16)	0.512 7(2.33)**
Constant	1.177 628(6.27)***	1.238 6(1.48)	1.037 5(1.10)	1.774 6(1.56)	3.526 1(9.67)***

In models at the third stage, only the decile and ninth decile have higher significance level and have good fitting to models. Therefore, we carried out analysis mainly on high-income areas and low-income areas at this stage. From the coefficient value of per capita GDP, it is known that there is conditional β convergence in both low-income and high-income areas, and the convergence rate of high-income areas is slightly higher than that of low-income areas; the investment rate indicates that the effect of investment on economic growth is positive in areas with different income level, in other words, increasing investment in all areas has positive effect on economy. However, the coefficient value of investment rate in high-income areas is 0.512 1, much higher than that in low-income areas, showing that economic effect brought by investment in high-income areas is much higher than in low-income areas. From coefficient value of human capital, increasing human capital in high-income areas also can promote economy, and the effect is higher than that on low-income areas. From the perspective of technological progress, increasing technological investment in high-income areas can promote economic growth, but increasing technological investment in low-income areas exerts little effect on economic growth.

5　Conclusions and further discussions

This study made analysis on convergence of China's economic growth using quantile regression method. Through analysis on variation coefficient of provinces in China, we divided convergence of China's economic growth into 3 stages. Through study on the first stage （1978 – 1991）, it has convergence, absolute β convergence and conditional β convergence, indicating that the gap between rich and poor is firstly narrowed, then widened and finally narrowed, which is basically consistent with conclusions of most domestic scholars. The difference is that we reached richer information and conclusion using quantile regression and comparison with OLS estimation：the control over the gap between rich and poor in low-income areas is excellent; the dispersion speed in middle income areas is gradually quickening, and the gap between rich and poor is gradually widening; in high-income areas, the convergence rate at the first stage is highest; at the second stage, the dispersion speed is the highest, and it exerts greatest influence on the overall gap between rich and poor; at the third stage, this situation is preliminarily controlled and there appears absolute β convergence of economic growth. From the perspective of factors influencing convergence of economic growth, the influence degree of investment rate, human capital, and technological progress on economic growth rate is higher in areas with higher economic growth rate than that with lower economic growth rate at all three stages, so these three factors can bring more economic effect. Although we applied advanced quantile regression method to study convergence of China's economic growth and reached rich information and conclusion, our study is preliminary, and there will be weak points. For example, in data processing, we did not consider extracting more information through analysis of panel data, and making in-depth studies. In conditional convergence, we did not consider increasing more control variables. These questions need further studies in the future.

References

[1] CAI F, DU Y. Convergence and divergence of regional economic growth in China[J]. Economic Research Journal, 2000(10):30 −37. (in Chinese).

[2] CHEN AG, LI GP. Convergence of regional economic growth：An time series analysis of China[J]. The Journal of Quantitative & Technical Economics, 2004, 21(11)：31 −35. (in Chinese).

[3] CHEN J, LIN L, YE AZ. A quantile regression analysis on Chinese resident's consumption[J]. The Journal of Quantitative & Technical Economics, 2008, 25(2):16 −27. (in Chinese).

[4] CHEN JB, DING JJ. A review of technologies on quantile regression[J]. Statistics & Information Tribune, 2008, 23(3):89 −96. (in Chinese).

[5] DENG L, ZHENG Z. A statistical approach of depicting different research objects −quantile regression[J]. Statistics and Decision, 2009(4)：154 −155. (in Chinese).

[6] JIN XY. Analytical methods on regional economic growth convergence[J]. The Journal of Quantitative & Technical Economics, 2006, 23(3)：102 −110. (in Chinese).

◦•◦○•◦◦

[16] WU JP, XIE K, WANG MM. Information economics[M]. Beijing：Higher Education Press, 2002. (in Chinese).

[17] NIU S. Study on the countermeasures of enhancing county rural informatization construction of Harbin City[J]. Economic Research Guide, 2011 (2)：135 −136. (in Chinese).

[18] WANG X, HU XM. A brief analysis on the difficulties and acting point of a new socialist countryside construction −Bijie[EB/OL]. http：// www. stats −bj. gov. cn/tjfx/20061117/110432 shtml. 2012 −02 −27. (in Chinese).

[19] WU JP. Economic information[M]. Dalian：Dongbei University of Finance & Economics Press, 1996. (in Chinese).