Journal of Economics and Business Research,
ISSN: 2068 - 3537, E – ISSN (online) 2069 – 9476, ISSN – L = 2068 – 3537
Year XX, No. 1, 2014, pp. 161-172

Estimating the liaison between Unemployment and GDP beyond the Mean of the Distribution

M. Aamir Khan, M. Rasheed, M. Hanif

Muhammad Aamir Khan
Department of Economics/Agri.Economis PMAS Arid Agriculture University Rawalpindi.
Mamoona Rasheed, Muhammad Hanif
Department of Statistics and Mathematics, PMAS Arid Agriculture University Rawalpindi.

Abstract
Okun law postulates a negative relationship between the movements of unemployment rate and the real gross domestic product. This study applied Quantile regression analysis to estimate the relationship between unemployment and its predictor GDP, and compared the results to parameter estimates using OLS regression. Conditional mean has been used as a solution to minimize the error variance. Time series annual data has been used from period 1973 to 2010. Predictor of unemployment used was GDP. Empirical Results of quantile regression analysis showed that for GDP parameter estimates were significant only for certain quantiles. Parameters for GDP were significant only for 5%, 10%, 95%. The GDP parameter had non- significant on OLS but not on all quantile from quantile regression.
Keywords: Okun law, Quantile Regression, OLS regression, GDP parameter

Introduction

Conventional regression analysis usually depends on the mean to summarize the relationship between the response variable and predictor variables by describing the mean of the response for each fixed value of the predictors, using a function we refer to as the conditional mean of the response. The thought of modeling and fitting the conditional mean function is at the heart of a extensive family of regression-modeling approaches, including the proverbial simple linear regression model, multiple regression and models with heteroscedastic errors using weighted least squares.

In situations in which the model assumptions were not met or there were outliers, the conditional mean cannot accurately reflect the conditional distribution of the data. If the effect of predictors was different across varying percentiles of conditional distributions, then the effect of the predictors on the upper tail of the distribution may be cancelled out by the effect or predictors on the lower tail of the distribution, which in turn make the effects seem to be zero.

For the results of the study to be accurate, what is needed is a statistical technique that can provide more information about the relationships between variables at varying locations of the distributions of the data. Another limitation in parameter estimation using OLS that was not investigated in the previous research is that of parameters estimated by OLS procedure being influenced by outliers. The existence of outliers violates one key model assumption: only one regression line is needed to represent the relationships for the whole distribution.

It was suggested that the outliers can be excluded from the analysis if they were, from thorough investigations, proven to be non-valid observations. But when the outliers are valid, it can give new insights about the nature of the data. It means that a statistical technique is needed, that will capture the outliers in the analysis and yet is less influenced by their presence.

An alternative technique that has more capability to solve some issues mentioned earlier is called quantile regression. Quantiles are values that give us information about location of a case in a distribution related to proportion of cases having smaller values. It was developed from a conditional median regression introduced by Boscovich in the 18th century, even before the idea of least squares regression estimators emerged (Koenker, 2005; Koenker and Bassett, 1978). Quantile regression was developed by applying estimation and minimization methods for the conditional median, which is quantile.5, and to other quantiles, rather than the conditional mean, as is done in OLS regression.

Quantile regression has some advantages over OLS regression. It provides information of location shift not only in terms of central tendency location but also other quantile locations. This means that we may have more than one regression line can be modeled, covering the whole conditional distribution including the outliers. For this reason, quantile regression may give us more information about relationship between variables, not only the relationship in term of location shift but also distributional shift including scale shift and skewness shift.

OLS-R is claimed to produce parameters with desirable characteristics - best, linear, unbiased estimators (BLUE). This means that parameters estimated using OLS-R have the smallest variance, model a linear relationship between response and outcome variables, and resemble value of parameters in population. However, these characteristics only hold if there are not serious violations of the model assumptions or presence of influential. Heteroscedasticity will make parameter estimates no longer BLUE, while the presence of influential outliers will cause the regression line to be leveraged in the outliers.

OLS-R still has inherent disadvantages even when procedures to overcome effects of violation assumptions and outliers are applied. Models suggested by OLS-R cannot be immediately extended to other locations in the distribution that may be more interesting to be investigated in other studies. For example, the study of unemployment focuses on over-achieving or under-achieving students.

OLS-R also assumes that response variables only affect the location shift of the conditional distribution, while response variables may affect other parameters of the distribution in some instances. This means that OLS-R provides limited information about the relationship between variables. OLS-R may give inaccurate information about the nature of the relationship between variables. When heteroscedasticity occurs and the slope of the regression line on the conditional mean is zero, OLS-R or related approaches to overcome heteroscedasticity will suggest no relationship between variables, although there are relationships between variables on non-central locations or on other distributional parameters (e.g. scale, skewness).

Another advantage of using quantile regression is its monotone equivariance property. Hao and Naiman (2007) explain that if we apply a monotone transformation to the outcome variable and then conduct a quantile regression analysis, the predicted values from this procedure will be approximately the same with predicted values from a procedure in which we conduct quantile regression first and then apply monotone equivariance to its prediction. But in OLS the mean does not have transformation equivariance since

$$Eh(Y) \neq h\ (E(Y)$$

Applications of Quantile Regression are still limited to economics or environmental studies, but currently there are more and more studies using Quantile Regression as a data analysis tool. The current paper will apply quantile regression on data of Unemployment and its predictor GDP. The term relationship is meant to be used in a broader sense: not only relationships in term of conditional locations but also conditional distributions. This study will also compare information provided by this method to those provided by OLS regression to get the sense of how both methods provide different information about the data. This study gauges what is the relationship between educational attainment and its predictors either using OLS regression or quantile regression methods. Last but not the least is there any differences in information given by OLS regression and quantile regression.

Methodology

A Quantile is a value that gives us information about the location of a case or a score in a group corresponds to a specified proportion of the sample or population. A person's score on a test is said to be in the p-th quantile in his/her group if his/her score in the test is bigger than a proportion of p of his/her group and smaller than a proportion (1-p) of his/her group. The median is at the 0.5 quantile, because there is half of the group that have values bigger than the median, and half of the group that have values smaller than the median. The lower quartile is at .25 quantile and the higher quartile is at .75 quantile. A function that gives us the value of a certain quantile is called a quantile function (QF) denoted as Q (p). For example, if a median of a group has a value of 50, it can also be said that Q (.5) is 50. The quantile function is an inverse of cumulative distribution function (CDF) denoted as F (x). The CDF can show us a proportion of a group that has a value equal to or smaller than a certain value of x. It can be formulated as:

$$F(x) = P(X \leq x) \quad (1)$$

The relationship between quantile function and CDF is denoted as

$$Q^{(p)} = F^{-1}(p)$$

OLS-R uses conditional mean E (yi | xi) on the solution of minimization problem while line, which is the regression lines on the conditional mean

$$E(y_i \mid x_i) = ß_o + \Sigma\, ß_i^{(p)}\, x$$

While Quantile regression can produce more than one regression line, one for any quantile of interest

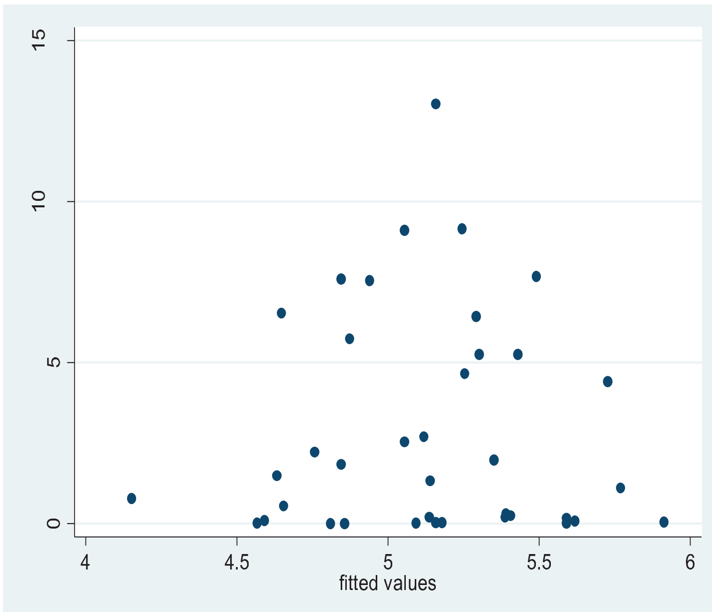
$$Q^{(p)}(y_i \mid x_i) = ß_o + \Sigma\, ß_i^{(p)}\, x$$

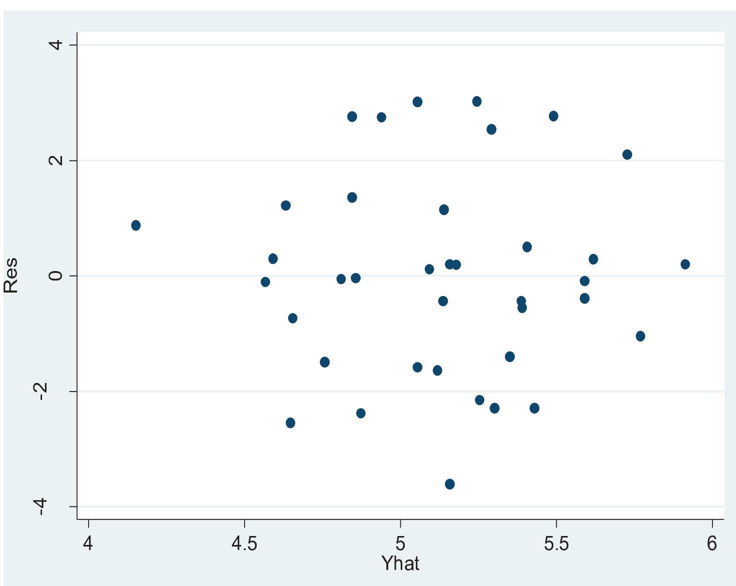
Results and Discussion

Quantile Regression can provide a regression line with non-central locations because of its ability to examine the relationship between variables on any quantiles in a conditional distribution. This will enable researchers to conduct inequality studies involving non-central area of the conditional distribution.

Assumption Check and Diagnostic for OLS regression

Homogeneity Assumption





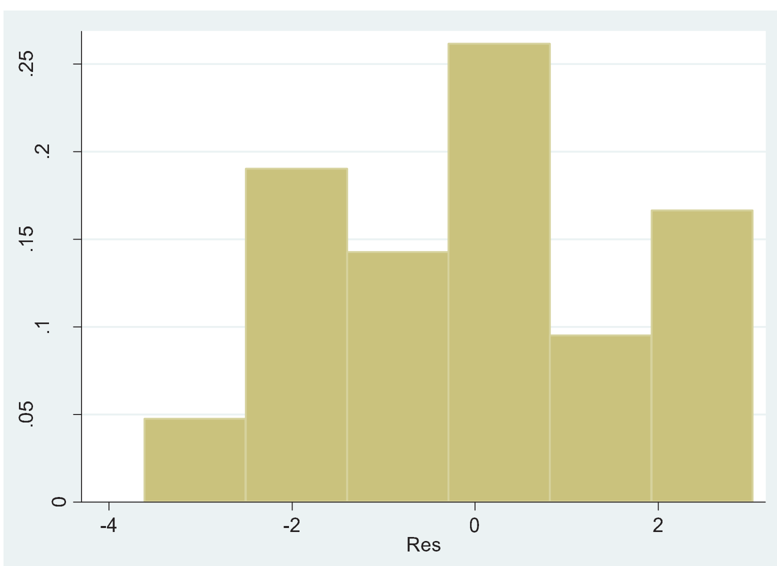
Homogeneity assumption was checked by plotting residuals and predicted values. If the plots form a megaphone-like pattern, we concluded that the heterogeneity assumption was violated. The plots are shown in the figures above. Based on the plots, it can be concluded that there is no violation of homogeneity assumptions.

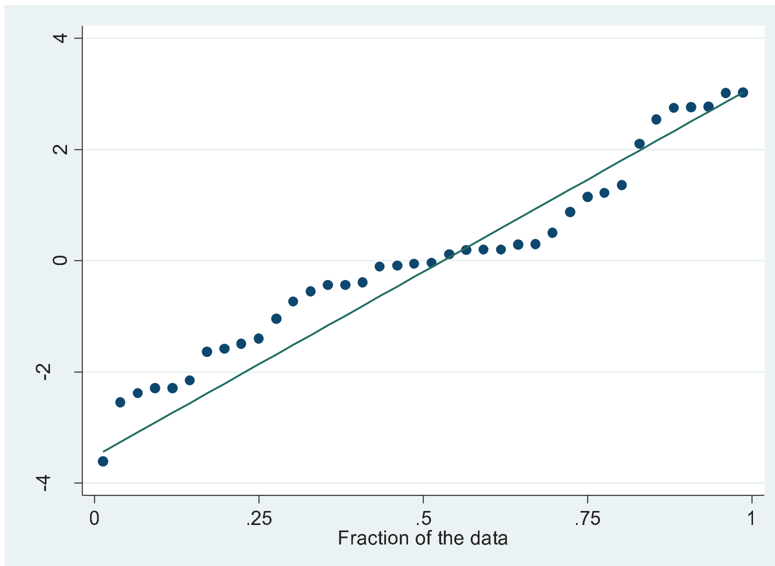
Linear Relationship between Variables

The linear relationship assumption was investigated using the same plots to check the homogeneity assumption. It can be seen that there are no patterns suggested that the relationship between variables is not linear.

Errors term are distributed normally

Normal distributions of error term can be confirmed from the Q-Q plot of the residuals, as it can be seen in Figure below. From the plot, it can be seen that there is no violation of the normality of error distribution. It can be confirmed also that there are some potential outliers to be checked.





Descriptive Statistics

Descriptive statistics for each variable can be seen in the table below:

Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
Y	38	5.132632	1.768097	1.55	8.27
X	38	5.095	2.011632	1.01	10.2

Summary Tables for OLS Regression and Quantile Regression on each Quantile

Parameters, Standard Error, t-Value and p-Value for OLS Regression

UN	Coef.	Std. Err	t	P>|t|	[95% Conf. Interval]	
constant	| 6.1069	.78182	7.81	0.000	4.521296	7.692505
GDP	-.1912206	.1429806	-1.34	0.189	-.4811986	.0987574

Small parameter differences between median regression and OLS regression were due because the assumptions of OLS regression were not violated and there were no influential outliers. The conditional distribution was also normal and symmetrical which makes the mean and median have relatively similar values.

Comparison to other Quantiles

Regression parameters for other quantiles were also estimated. The parameters of several important quantiles can be seen below

Parameters, standard error, t-value and p-value for Quantile Regression Q.05 (5%)

UN	Coef.	Std. Err	t	P>|t|	[95% Conf. Interval]	
constant	4.048802	.4900027	8.26	0.000	3.055031	5.042574
GDP	-.2567238	.0543921	-4.72	0.000	-.3670361	-.1464114

Parameters, standard error, t-value and p-value for Quantile Regression Q.10 (10%)

UN	Coef.	Std. Err	t	P>|t|	[95% Conf. Interval]	
constant	4.142515	.9575896	4.33	0.000	2.200433	6.084596
GDP	-.2690059	.133747	-2.01	0.052	-.5402573	.0022456

Parameters, standard error, t-value and p-value for Quantile Regression Q.25 (25%)

UN	Coef.	Std. Err	t	P>|t|	[95% Conf. Interval]	
constant	5.21483	1.034103	5.04	0.000	3.117573	7.312088
GDP	-.2754717	.1809402	-1.52	0.137	-.6424354	.091492

Parameters, standard error, t-value and p-value for Median Regression Q.50 (50%)

UN	Coef.	Std. Err	t	P>|t|	[95% Conf. Interval]	
constant	5.978125	.6845197	8.73	0.000	4.589854	7.366395
GDP	-.1770833	.125756	-1.41	0.168	-.4321283	.0779617

Parameters, standard error, t-value and p-value for Quantile Regression Q.75 (75%)

UN	Coef.	Std. Err	t	P>|t|	[95% Conf. Interval]	
constant	7.130151	1.564695	4.56	0.000	3.956802	10.3035
GDP	-.1660377	.2701254	-0.61	0.543	-.7138774	.3818019

Parameters, standard error, t-value and p-value for Quantile Regression Q.90 (90%)

UN	Coef.	Std. Err	t	P>|t|	[95% Conf. Interval]	
constant	8.897292	.5828414	15.27	0.000	7.715235	10.07935
GDP	-.1979167	.1142385	-1.73	0.092	-.4296032	.0337697

Parameters, standard error, t-value and p-value for Quantile Regression Q.95 (95%)

UN	Coef.	Std. Err	t	P>|t|	[95% Conf. Interval]	
constant	9.705716	.2396031	40.51	0.000	9.219778	10.19165
GDP	-.3190479	.045639	-6.99	0.000	-.411608	-.2264878

Summary of tables

UN	constant	GDP
OLS	6.1069**	-.191221
5%	4.048802**	-.256724**
10%	4.142515**	-.269006*
25%	5.21483**	-.275472
50%	5.978125**	-.177084
75%	7.130151**	-.166038
90%	8.897292**	-.197917
95%	9.705716**	-.319048**

Interpretation of a constant

In the median regression the constant is the median of the sample while in the .75 quantile regression the constant is the 75th percentile for the sample.

The parameters for Quantile Regression on quantile .05 and .1 were significant (b.05=-.256724, $p<.05$; b.1=-.269006, $p<.05$) but those from OLS regression (b=-.191221, $p>.05$) and Quantile Regression on all other quantiles ($-.275472 \leq b \leq -.166038$, $p>.05$) were not significant. This means that GDP was related to unemployment only in lower quantiles.

Conclusion

These results suggested that the relationships between unemployment and its predictor might differ at different location across conditional distributions. We apply a monotone transformation to the outcome variable and then conduct a quantile regression analysis, and then apply monotone equivariance to its prediction. But in OLS the mean does not have transformation equivariance. This information could not be obtained if researchers only used the OLS regression method to analyze the data. Suggesting relationships as they were presented only by OLS regression could neglect important issues, for example using OLS regression it would be suggested that there was no relationship between unemployment and GDP, while actually there was significant relationship on only the very low quantiles.

REFERENCES

Barro, R., Xavier, S. M (1995). *Economic growth,* MIT Press Cambridge, London

Cahill, M. B., (2006). "Estimating key macroeconomic relationships at the undergraduate level: Taylor rule and Okun's Law examples", Meeting of the Allied Social Science Associations Boston, MA

Dornbusch, R., Fischer, S. (2007). *Macroeconomics* 9th Edition

Hao, L., Naiman, D.Q. (2007). *Quantile Regression,* Sage Publications, Thousands Oaks

Knotek, E. S., II. (2007). How useful is Okun's law? *Economic Review.* Federal Reserve Bank of Kansas City, Fourth Quarter, p. 73-103

Koenker, R. (2005). "Quantile Regression Software" available at ^http://www. econ.uiuc.edu/;roger/research/rq/rq.html&.

Koenker, R., Bassett, G. (1978). "Regression Quantiles." *Econometrica.* January, 46:1, p. 33–50

Lang, D., Peretti, C. (2009). "A strong hysteretic model of Okun's Law: theory and a preliminary investigation", *International Review of Applied Economics* Vol. 23, No. 4, p. 445-462

Lee, J. (2000). The Robustness of Okun's Law: Evidence from OECD countries. *Journal of Macroeconomics,* 22, p. 331-56

Malley, J., Molana, H. (2008). "Output, unemployment and Okun's law: Some evidence from the G7", Elsevier, Economics Letters 101, p. 113-115

Mankiw, G. (2004). *Macroeconomics,* Worth Publishers, New York

Okun, A.M. (1962). Potential GNP: Its Measurement and Significance, *Proceedings of the Business and Economic Statistics,* p. 98-103

Okun, A. (1983). *Economics for Policy making,* MIT Press Cambridge, London

Pierdzioch, C., Rulke, J. K., Stadtmann, G. (2009). "Do professional economists' forecasts reflect Okun's law? Some evidence for the G7 countries", *Applied Economics,* p. 1-9

Soègner, L., Stiassny, A. (2002). "An analysis on the structural stability of Okun's law-a cross-country study", *Applied Economics*, 14, p. 1775-1787

Thirlwall, A.P. (1969). Okun's Law and the natural rate of growth, Southern Economics Journal, 36(1), p. 87-89

Villaverde, J., Maza, A. (2008). "The robustness of Okun's law in Spain, 1980-2004 Regional evidence", *Journal of Policy Modeling 31*, p. 289-297

Zaleha, M. N., Nor, M., Ghani, N., Ghani, A. (2007). The relationship between output and unemployment in Malaysia: Does Okun's Law exist? *International Journal of Economics and Management,* 1(3), p. 337-344.