ANALYSIS OF GROSS DOMESTIC PRODUCT DATA USING QUANTILE REGRESSION MODEL

BY

YUNUSA, Zainab Bello

DEPARTMENT OF STATISTICS

AHMADU BELLO UNIVERSITY, ZARIA, NIGERIA.

APRIL, 2021

ANALYSIS OF GROSS DOMESTIC PRODUCT DATA USING QUANTILE REGRESSION MODEL

BY

YUNUSA, ZAINAB BELLO
(P16PSST8018)

A THESIS SUBMITTED TO THE SCHOOL OF POSTGRADUATE STUDIES
AHMADU BELLO UNIVERSITY, ZARIA, NIGERIA.

IN PARTIAL FULFILLMENT FOR THE AWARD OF MASTER OF SCIENCE (M.SC) DEGREE IN STATISTICS

DEPARTMENT OF STATISTICS
AHMADU BELLO UNIVERSITY, ZARIA, NIGERIA.

APRIL, 2021

DECLARATION

I declare this thesis titled “ANALYSIS OF GROSS DOMESTIC PRODUCT DATA USING QUANTILE REGRESSION MODEL” was carried out by me in the department of statistics. The information derived from literature has been duly acknowledged in the text and a list of references provided. No part of this thesis was previously presented for another degree or diploma in this or any other institution.

YUNUSA Zainab Bello	Signature	Date

CERTIFICATION

This thesis titled "ANALYSIS OF GROSS DOMESTIC PRODUCT DATA USING QUANTILE REGRESSION MODEL" by Zainab Bello Yunusa (P16PSST8018) meets the regulations governing the award of the degree of Master of Science of the Ahmadu Bello University, Zaria and is approved for its contribution to knowledge and literary presentation.

____________________	____________________	____________________
Dr. S.I.S Doguwa Chairman, Supervisory Committee	Signature	Date
____________________	____________________	____________________
Dr. Aisha Umar Member, Supervisory Committee	Signature	Date
____________________	____________________	____________________
External Examiner	Signature	Date
____________________	____________________	____________________
Dr. Abubakar Yahaya Head of Department	Signature	Date
____________________	____________________	____________________
Prof Sani A. Abdullahi Dean, School of Postgraduate Studies	Signature	Date

DEDICATION

First of all, I thank Allah for keeping me alive. This work is dedicated to my parent.

ACKNOWLEDGEMENT

First and foremost, I express my happiness and glorification towards Allah, the eternally besought of all who in His infinite Mercy bestowed on me intellect and sound reasoning faculties and above all showered me with His blessings of which my research is an integral aspect.

My sincere thanks and appreciation to my supervisors Dr. S.I.S Doguwa and Dr. A. Umar and my Head of Department Dr. Abubakar Yahaya whose constructive criticism has been a source of inspiration to a dedicated research thesis. You have not only taken the pain guiding me through this project work but also gave me crucial facts which lead to the underlying foundation on which this research is based.

I shall always remain indebted to my parent Prof Mohammed-Bello Yunusa, Haj Maryam Yunusa for their role in my upbringing, education, encouraging and believing in me, your support in its entire ramification cannot be quantified, may Allah (S.W.A) reward and bless you with Jannah.

My whole hearted gratitude goes to Dr (Mrs.) A. Umar of Department of Mathematics I say thank you for the motherly advice, support and also believing in me.

I am immensely grateful to members of my family, my siblings and nephews Haj Zainab, Barr M.A Yunusa, Bld M.H Yunusa, Jamila, Ummul-Khair, Baba Isah, Shehu, Imam, Mama, Mujittaba, Aqib, Mahe, Nana, khalil, Farid, Khalid and Abdul-Rahman for their moral support.

I also gratefully acknowledge the excellent togetherness of my entire class members, am grateful to the post graduate coordinator Mal. Tasi’u Musa all students and staff of Department of Statistics.

In the process of this work I had to disturb the likes of Yahaya Jibril, Khadija Abdullahi, John. And I say a big thank you to National Bureau of statistics for making the data available. The contribution and encouragement of my friend Habiba Abdulkadir.

ABSTRACT

Quantile Regression (QR) models have been increasingly employed in many applied areas of statistics. The study used quarterly data on real GDP percentage contribution to Growth, of Agriculture, Industry, Building & construction, Wholesale & retail and Services. The data was analyzed using QR model at 0.25 and 0.5 (median) quantile for a good understanding of the relationships between variables beyond the mean of the data and also investigated the differences in effect of individual indicator on GDP. Kernel residual estimator was used for sparsity and the Huber Sandwich and bootstrap standard errors are also employed and the asymptotic covariance matrix is estimated from sample variance of the bootstrap result. We Further went on to run quantile slope equality test using Wald test and symmetric quantile test using chi-square test. The results provide further evidence that QR is a dependable method for modeling Gross Domestic Product data. Thus, the work recommends that the involvement of the quantile regression model in the control process of economic dataset intended for policy purposes help to boost the quality of their output for better inference.

TABLE OF CONTENTS

List of Abbreviations

G.D.P = Gross Domestic Product

L.A.D = Least Absolute Deviation

L.R.M = Linear Regression Model

M.L.R = Multiple Linear Regression

N.B.S = National Bureau of Statistics

O.L.S = Ordinary Least Square

Q.R = Quantile Regression

Q.E.P = Quantile Estimation Procedure

Q.R.M = Quantile Regression Model

R.A.S.W = Residual Absolute Sum of Weighted differences

S.E = Standard Error

T.A.S.W = Total Absolute Sum of Weighted Differences

CHAPTER ONE

1.1 Background of the Study

Regression analysis is a statistical methodology that utilizes the relation between two or more quantitative variables so that one variable can be predicted from the other, or others. It serves three major purposes: (1) description (2) control and (3) prediction.

If Y denotes the dependent variable and X the independent variable, a functional relation between the two variables is expressed by a mathematical formula:

$Y = F(X)$

Given a particular value of X, the function F indicates the corresponding value of Y. Depending on the nature of the relationships between Y and X, regression approach may be classified into two broad categories viz., linear regression models and nonlinear regression models. The response variable is generally related to other causal variables through some parameters. The models that are linear in these parameters are known as linear models, whereas in nonlinear models' parameters are nonlinear. Linear models are generally satisfactory approximations for most regression applications. There are occasions, however, when nonlinear models are more appropriate.

The purpose of regression analysis is to expose the relationship between a response variable and predictor variables. In real applications, the response variable cannot be predicted exactly from the predictor variables. Instead, the response for a fixed value of each predictor variable is a random variable. For this reason, we often summarize the behavior of the response for fixed values of the predictors using measures of central tendency.

Traditional regression analysis is focused on the mean; that is, we 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 idea of modeling and fitting the conditional-mean function is at the core of a broad family of regression-modeling approaches, including the familiar simple linear-regression model, multiple regression, models with "heteroscedastic" errors using weighted least squares, and nonlinear regression models.
Conditional-mean models have certain attractive properties. Under ideal conditions, they are capable of providing a complete and parsimonious description of the relationship between the covariates and the response distribution. In addition, using conditional-mean models leads to estimators (least squares and maximum likelihood) that possess attractive statistical properties that are easy to calculate, and are straightforward to interpret. Such models have been generalized in various ways to allow for "heteroscedastic" errors so that given the predictors, modeling of the conditional mean and conditional scale of the response can be carried out simultaneously. Conditional-mean modeling has been applied widely in the social sciences, particularly in the past half century, and regression modeling of the relationship between a continuous response and covariates via least squares and its generalization is now seen as an essential tool. More recently, non-linear models for binary response data, such as Logistic, Probit and Poisson regression models for count data have become increasingly popular in social science research. These approaches fit naturally within the conditional mean modeling framework. While quantitative social-science researchers have applied advanced methods to relax some basic modeling assumptions under the conditional-mean framework, this framework itself is seldom questioned.

Ordinary least-squares regression models the relationship between one or more covariates X and the conditional mean of the response variable Y given $X = x$. Least-squares regression assumes that the covariates affect only the location of the conditional distribution of the response, and not its scale or any other aspect of its distributional shape. Quantile regression, which was introduced by Koenker and Bassett (1978), extends the classical least squares methods of estimating conditional mean by offering a variety of methods for estimating conditional quantiles of the response variable, such as the 90th percentile thereby enabling the researcher to explore heterogeneous covariate effects (Koenker, 2005). Quantile regression is particularly useful when the rate of change in the conditional quantile, expressed by the regression coefficients, depends on the quantile. The main advantage of quantile regression over least-squares regression is its flexibility for modeling data with heterogeneous conditional distributions. Data of this type occur in many fields, including econometrics, survival analysis, and ecology (Koenker and Hallock 2001).

Time Series is a study of event over a period of time which could be monthly, yearly, weekly, hourly etc. Time series comprises of methods that attempt to understand the underlying observations which asks questions like; where is the origin? What lead to them? Or to forecast i.e. predicting the future. In the last two decades time series has been one of the most important and widely used branch of mathematical statistics. Its field application covers areas like economic forecasting, study of biological data (e.g health cases, population). Examples of time series are the monthly sales of businessman, the daily traded share of stock market, the hourly temperature announced by a city. A forecasting exercise is usually carried out to provide an aid in decision making and planning the future.

1.2 Motivation of the Study

One of the most popular and commonly used data mining techniques is Linear Regression Model (LRM). LRM can provide insightful information in cases when the rigid assumptions associated with it are met. It summarizes the average relationship between a set of regressors and the dependent variable.

1.3 Statement of Research Problem

Linear Regression Model (LRM) is one of the most commonly used data mining techniques, and can provide insightful information in cases where the rigid assumptions associated with linear regression are met. The assumptions include (1) linearity of the coefficients (2) normal or Gaussian distribution for the response errors (ε) and (3) the errors ε are independent. This is a robust in application for various kind of research, especially when provisions are made to control for problems dealing with heteroskedasticity, due to the violation of assumptions. What is the implication of heteroskedasticity? However, the ordinary least squares (OLS) estimators and regression predictions based on them remain unbiased and consistent. The OLS estimators are no longer the BLUE because they are no longer efficient. In contrast to LRM, Quantile Regression (QR) does not impose any strict parametric assumptions. It is intended to offer a comprehensive strategy for completing the regression picture. As Mosteller and Tukey (1977) noted in their influential paper as cited by Koenker (2005), the regression curve gives a grand summary for the averages of the distributions corresponding to the set of *Xs* and so regression often gives a rather incomplete picture. Just as the mean gives an incomplete picture of a single distribution, so the regression curve gives a correspondingly incomplete picture for a set of distributions.

The purpose of this study is to analyze Gross Domestic Product GDP data and propose the use of Huber sandwich quantile regression as a good model that will confirm quantile regression suitable for modelling economic data set.

1.4 Scope of Study

The study seeks to model quantile regression using and economic data on Gross Domestic Product keeping in mind that the quantile regression does not take into account the failure or otherwise of the existing assumptions. Discussing quantile regression, various tests of the importance of independent variables, model reliability of the models and result interpretation in the model will be considered.

1.5 Aim and Objectives

The aim of the research is to model GDP growth rate using quantile regression through the following objectives:

1. To evaluate whether the location-shift model is appropriate in the modelled GDP growth rate using quantile regression.
2. To construct a test of the less restrictive hypothesis of symmetry.
3. To estimate asymptotic covariance matrices of the estimates using bootstrapping technique.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

The theoretical developments in time series analysis started early with Stochastic process. However, the first actual application of autoregressive models to data can be brought back to work of G. U Yule and Walker in the 1920s and 1930s as cited by Leila Zoubir (2017).

Quantile regression (QR) was introduced by Koenker and Bassett (1978) and it is intended to offer comprehensive strategy for completing the regression picture, Koenker (2008). Quantile regression does not impose any strict parametric assumptions, Koenker (2005). Quantile Regression seeks to estimate conditional quantile functions ie the varying values of covariates are estimated based on the quantiles asymmetrically weighted absolute residual of the mean distribution, Buhai (2004). QR allows us to examine the behavior of the target variable (Y) beyond its average of the 80^{th} percentile, 95^{th} percentile and so on. By examining these quantiles, we get greater insights into the underlying process

2.2 Linear Regression Model

In statistics, regression analysis is a statistical process for estimating the relationships among variables. It includes many techniques for modeling and analyzing several variables, when the focus is on the relationship between a dependent variable and one or more independent variables (or ‘predictors’). More specifically, regression analysis helps one understand how the typical value of the dependent variable changes when any one of the independent variables is varied, while the other independent variables are held fixed Scott (2012). The Quantile Regression Model (QRM) facilitates analysis of the full conditional distributional properties of the response variable. The QRM model deals with a continuous response variable that is linear in unknown parameters. While extending to more than one covariate necessarily introduces additional complexity, the ideas remain essentially the same.

The continuous response variable depends on the independent variables and considering data consisting of pairs. By LRM, we mean the standard linear regression model

Where sample values of the response and exploratory variables are random disturbance terms assumed to be normally distributed with mean and variance. The intercept parameter is the value predicted for the response variable when the explanatory variable takes the value zero. The slope parameter is the change in the response variable predicted when the explanatory variable is increased by one unit. The parameters, also known as regression coefficients, can be estimated by least squares.

2.3 Quantile Regression

Quantile Regression (QR) was introduced by Koenker and Basssett (1978) and it is intended to offer a comprehensive strategy for completing the regression picture (Koenker 2008).Quantile regression does not impose any strict parametric assumptions (Koenker 2005), quantile regression seeks to estimate conditional quantile functions i.e. the varying values of covariates are estimated based on the quantiles asymmetrically weighted absolute residuals of the mean distribution Buhai (2004). Quantile regression allows us to examine behavior of the target variable(Y) beyond its average of the 80^{th} percentile, 95^{th} percentile and so on. By examining these quantiles, we get greater insight into the underlying process we are studying. Given the nature of the median statistics, this leads to a more accurate and robust representation of the relationship between the covariates and response variable. Buhai (2004) states “instead of assuming that covariates shift only the location or the scale of the conditional distribution”. Quantile regression looks at the potential effects on shape of the distribution as well.

Quantile regression was also used in ecological studies by Cade and Noon (2003). Dunham, Cade and Terrell (2002) applied quantile regression to analyze fish – habitat relationship, they studied the relationship between the conditional quantiles of trout density and the width – to depth ratio. Ibrahim (2015) presented quantile regression as an alternative to ordinary least squares based on analytical solution in statgrahics software. He illustrated goodness of fit statistic called QR coefficient of determination using a data set on fuel consumption in miles per gallon in highway driving, “sometimes Ordinary Least Square estimates can be misleading, what the true relationship between response variable and covariate as the effect can be different for different subsection. Therefore, quantile regression gives a better and more complete view of the relationship among random variables. Quantile regression model was considered with non-stationary and nearly non-stationary time series, Yini Wang (2012). Where inference in quantile regression with co-integrated variables allowing for multi-structural changes “the conditional quantile estimator has a non-standard limit distribution”, fully modified estimator is proposed to remove the second-order bias and nuisance parameters and the resulting limit distribution is mixed normal. Therefore, the study showed that the fully modified Quantile estimator has good finite sample properties.

Quantile regression models have been increasingly employed in many applied areas economic research. At the early stage, application took place usually using the cross-section data, but recent development has been a surge of the use of quantile regression in both time series and panel data sets. However how to test for possible autocorrelation especially in the context of time series has been paid little attention. One might attempt the usual Breusch-Godfrey LM-test to the residual from the baseline quantile regression. Application of the LM test can result in potentially large size distortion especially in either low or high quantile (we then propose two correct tests named F-test and the quantile regression LM-test) for autocorrelation in quantile models which do not suffer from any size distortion demonstrate two tests perform fairly well in finite samples, across either different quantiles or different underlying error distributions.

2.4 Conditional Median and Quantile Regression Models

With a skewed distribution, the median may become the more appropriate measure of central tendency; therefore, conditional-median regression, rather than conditional-mean regression, should be considered for the purpose of modeling location shifts. Conditional-median regression was proposed by Boscovich in the mid-18th century and was subsequently investigated by Laplace and Edgeworth. The median-regression model addresses the problematic conditional-mean estimates of the LRM. Median regression estimates the effect of a covariate on the conditional median, so it represents the central location even when the distribution is skewed. To model both location shifts and shape shifts, Koenker and Bassett (1978) proposed a more general form than the median-regression model, the Quantile Regression Model (QRM). The QRM estimates the potential differential effect of a covariate on various quantiles in the conditional distribution, for example, a sequence of 19 equally distanced quantiles from the 0.05th quantile to the 0.95th quantile. With the median and the off-median quantiles, these 19 fitted regression lines capture the location shift (the line for the median), as well as scale and more complex shape shifts (the lines for off-median quantiles). In this way, the QRM estimates the differential effect of a covariate on the full distribution and accommodates heteroscedasticity. Buhai(2004) eloquently states, “Instead of assuming that covariates shift only the location or the scale of the conditional distribution, Quantile Regression looks at the potential effects on the shape of the distribution as well.

Following Koenker and Bassett (1978), the QRM corresponding to the LRM

where $0 < p < 1$ indicates the proportion of the population having scores below the quantile at p. Recall that for LRM, the conditional mean of Y_i given X_i is $E(Y_i|X_i)$, and this is equivalent to requiring that the error term ε_i have zero expectation. In contrast, for the corresponding QRM, we specify that the pth conditional quantile given x_i is $Q^{(p)}\ (Y_i|X_i)$. Thus, the conditional pth quantile is determined by the quantile-specific parameters, and, and a specific value of the covariate X_i. As for the LRM, the QRM can be formulated equivalently with a statement about the error terms ε_i. Since the term is a constant, we have $Q^{(p)}\ (Y_i|X_i) = + Q^{(p)}(\varepsilon_i)$, so an equivalent formulation of QRM requires that the pth quantile of the error term be zero.

Dictates that the QRM can have numerous conditional quantiles. For example, if the QRM specifies 19 quantiles, the 19 equations yield 19 coefficients for X_i, one at each of the 19 conditional quantiles. The quantiles do not have to be equidistant, but in practice, having them at equal intervals makes them easier to interpret.

QRM has many applications, and was originally developed for statistical use, as the first Quantile Regression publication was in Econometrical (Koenker and Bassett 1978) insightfully envisioned a more robust regression approach capable of modeling conditional quantile functions beyond the classic linear regression least squares approach to model building. Koenker and Bassett (1978) note, “estimators are suggested, which have comparable efficiency to least squares for Gaussian linear models while substantially outperforming the least-squares estimator over a wide class of non-Gaussian error distributions”. Researches began applying these new techniques to the Management Science field of study and introduce an adaptive filtering procedure for exploring regression quantiles. These models are used as part of their Quantile Estimation Procedure (QEP) and are utilized to signal preventative actions and therefore avoid undesirable system states (Gorr and Hsu 1985).

Yu and Stander (2003) studied applications of quantile regression typical applications of quantile regression to medical reference charts, survival analysis, financial economics, environmental modelling and the detection of heteroscedasticity.

CHAPTER THREE

RESEARCH DESIGN AND METHODOLOGY

3.1 Introduction

This chapter presents the methodology that will be used to carry out the study. In this chapter, we discuss the research design that would be used. We also discuss how collected data is analyzed giving details of the model used in the analysis with reasons as to why this particular model is applied.

Research methodology is the process or methods used to carry out a research or study. This refers to the method used to collect data or information to be used for the purpose of research. The Statistical model employed in this study is quantile regression.

- The data was obtained from Central Bank of Nigeria official website at http://www.cbn.gov.ng

3.2 Specification of The Research Model

The independent variables in the model include:

i. Agriculture (quarterly percentage "%" contribution to the GDP)

ii. Industry (quarterly percentage "%" contribution to the GDP)

iii. Building and Construction (quarterly percentage "%" contribution to the GDP)

iv. Wholesale and Retail (quarterly percentage "%" contribution to the GDP)

v. Services (quarterly percentage "%" contribution to the GDP)

The dependent variable of the model is quarterly percentage "%" Gross Domestic Growth (GDP) Rate

3.3 Quantile Regression Model

The θth (0<θ<1) quantile of F_Y is $q_Y(\theta)$ is an order statistic and it can also be obtained by minimizing the asymmetric (linear) loss function.

$$\theta\int_{Y>q}|Y-q|\,dF_Y(y)+(1-\theta)\int_{y<q}|y-q|\,dF_Y(y)$$

The first order condition of this minimization problem is

$$0=-\theta\int_{Y>q}dF_Y(y)+(1-\theta)\int_{Y<q}dF_Y(y) \tag{3.0}$$

$$0=-\theta[1-F_Y(q)]+(1-\theta)F_Y(q) \tag{3.2}$$

$$0=-\theta+F_Y(q) \tag{3.3}$$

The sample counterpart of the asymmetric linear loss function is

$$\frac{1}{T}\sum_{t=1}^{T}\rho_\theta(y_t-q)=\frac{1}{T}\left[\theta\sum_{t:y_t\geq q}|y_t-q|+(1-\theta)\sum_{t:y_t<q}|y_t-q|\right] \tag{3.4}$$

Where

$$\rho_\theta(u)=(\theta-1_{\{u<0\}})u \tag{3.5}$$

Is the check function given θ, minimizing this function yields θth sample quantile of y.

Let Y be discrete random variable that takes value from 1,2……9 with equal probabilities. The task is to find median 0f Y hence the value of τ = 0.5 is chosen. The expected loss function L(u) is

$$L(u)=\frac{(\tau-1)}{9}\sum_{y_i<u}(y_i-u)+\frac{\tau}{9}\sum_{y_i\geq u}(y_i-u) \tag{3.6}$$

$$L(u)=\frac{0.5}{9}\left(-\sum_{y_i<u}(y_i-u)+\sum_{y_i\geq u}(y_i-u)\right) \tag{3.7}$$

Since $\frac{0.5}{9}$ is a constant it can be taken out of the expected loss function if τ=0.5 then U=3

$$L(3)=\sum_{i=1}^{2}(i-3)+\sum_{i=1}^{3}(i-3) \tag{3.8}$$

$$\left[(2+1)+(0+1+2+.......+6)\right]=24 \tag{3.9}$$

Supposed it is increased by 1 unit then the expected loss will be changed by -3 if U = 5 the expected loss is

$$L(5)=\sum_{i=1}^{4}i+\sum_{i=1}^{4}i=20 \tag{3.10}$$

Define q as the quantile to be estimated; the median is q = 0.5. For each observation i, let γ_i be the residual

$$\gamma_i = y_i - \sum_j \beta_j x_{ij} \tag{3.11}$$

The basic quantile regression model specifies the linear dependence of the conditional quantiles of *Y* on **x**. In other words, we assume that the relationship between the 100% quantile of *Y* and covariates **x** is given by $q_p(X)=x^{\tau}\beta$

Given a data set **{xi, yi}** i=1, we can estimate the parameters $\boldsymbol{\beta}$ by minimizing the sum of absolute residual.

$$\beta_n(\tau) = \arg\min_{\beta(\tau)} \sum_{i=1}^{n} \rho_p(y_i - x_i\beta^\tau) \tag{3.12}$$

$$\beta_n(\tau) = \arg\min_{\beta(\tau)} \left\{ \sum_i \rho_\tau(y_i - x_i\,\beta^\tau) \right\}$$
$$\beta_n(\tau) = \sum_{i:yi \geq xi\beta_\tau}^{n} \rho_\tau \left| y_i - x_i\,\beta^\tau \right| + \sum_{i:yi < xi\beta_\tau}^{n} (1-\rho_\tau)\left| y_i - x_i\,\beta^\tau \right| \tag{3.13}$$

Where where ρ_τ (.) is the loss function defined as ρ_τ (u) = u (τ −I) (u < 0)) or equivalently, in the form of a simple optimization problem

$y_i - x_i\beta_\tau$ yields the residuals $\varepsilon_{i\tau}$, correctly acknowledge what that means the absolute values of positive residuals are weighted by τ, while the absolute values of negative residuals are weighted by 1-τ. Suppose that we have a random variable Y with probability distribution function.

$F(y) = \Pr ob(Y \leq y)$ and F(x) the probability density function of a random variable X. The τ-th quantile, where τ= 0.50 quantile is defined as

$$L(q) = -0.5\int_{-\infty}^{q} (y-q)dF_Y(y) + 0.5\int_{q}^{\infty} \tag{3.14}$$

$$L(q) = \int_{-\infty}^{q} 1dF_Y(y) - \int_{q}^{\infty} 1dF_Y(y) \tag{3.15}$$

In equation (3.14) $F_Y(q)$ and in equation (3.15) 1- $F_Y(q)$. Therefore, the change in of expected loss function is negative if and only if $F_Y(q) < 0.5$ ie if q is less than the median.

So that for $0 < \tau < 1$, the T-th quantile of Y may be defined as the smallest y satisfying $F(y) \geq T$:

$$Q_Y(\tau) = F_Y^{-1} \tag{3.16}$$

$$Q_Y(\tau) = \inf\{y : F_Y(y) \geq \tau\} \quad (3.17)$$

Where τ∈(0,1). Define the loss function $\rho_\tau(y) = y(\tau - \mathbb{I}_{(y<0)})$ where ‖ is an indicator function. A specific quantile can be found by minimizing the expected loss of Y-u with respect to U

$$\min_u E(\rho_\tau(Y-u) = \min_u \left\{(\tau-1)\int_{-\infty}^{u} y-u)dF_Y(y) + \tau\int_{u}^{\infty}(y-u)dF_Y(y)\right\} \quad (3.18)$$

This can be shown by setting the derivative of the expected loss function to 0 and letting q be the solution

$$0 = (1-\tau)\int_{-\infty}^{q_\tau} dF_Y(y) - \tau\int_{q_\tau}^{\infty} dF_Y(y) \quad (3.19)$$

$$0 = F_Y(q_\tau) - \tau \quad (3.20)$$

$$F_Y(q_\tau) = \tau \quad (3.21)$$

Hence q_τ is the τ th quantile of the random variable Y

Quantile regression extends this simple formulation to allow for regressors X, we assume a linear specification for the conditional quantile of the response variable Y given values for the P-vector of explanatory variables

3.4 Huber Sandwich Estimator

The "Huber Sandwich Estimator" can be used to estimate the variance of the MLE when the underlying model is incorrect. If the model is nearly correct, so are the usual standard errors, and robustification. On the other hand, if the model is seriously in error, the sandwich may help on the variance side.

To begin the mathematical exposition, let i index observations whose values are y_i let $\cap\ \theta \epsilon R^p$be a $p\ x\ 1$ parameter vector. Let $y \to f_i(y/\theta)$ be a positive density. If $y_i$takes only the values 0 or 1, which is the case of interest here, then $f_i\,(0/\theta) > 0, f_i\,(1/\theta) > 0,$ and $f_i\,(0/\theta) + f_i\,(1/\theta) = 1.$ Some examples involve real or vector-valued--, and the notation is set up in terms of integrals rather than sums. We assume $\to f_i\,(y/\theta)$ is smooth. (Other regularity conditions are elided.) let Y_i be independent with density $f_i\,(y/\theta)$. Notice that the Y_i are not identically distributed: f_i depends on the subscript i. in typical applications, cannot be identically distributed, as will be explained below.

The data are modeled as observed values of Y_i for i = 1,…,n. the likelihood function is viewed $\prod_{i=1}^{n} f_i(Y_i/\theta)$, as a function of θ. The log likelihood function is therefore.

$$L(\theta) = \sum_{i=1}^{n} \log f_1(Y_i/\theta) \tag{3.22}$$

The first and second partial derivates of L with respect to θ are given by

$$L^{'}(\theta) = \sum_{i=1}^{n} gi\,(Y_i/\theta),\ \ L^{''}(\theta) = \sum_{1=1}^{n} h_i\,(Y_i/\theta) \tag{3.23}$$

To unpack the notation in (2), let $\phi^{'}$ denote the derivative of the function ϕ: differentiation is with respect to the parameter vector θ. Then

$$gi\,(y/\theta) = [\log f_i\,(y/\theta)]^{'} = \frac{\partial}{\partial\theta}\log f_i\,(y/\theta), \tag{3.24}$$

a 1 × p-vector. Similarly,

$$h_i(y/\theta) = [\log f_i(y/\theta)]^{''} = \frac{\partial^2}{\partial\theta^2}\log f_i(y/\theta), \qquad (3.25)$$

a symmetric p × p matrix. The quantity - $E_\theta h(Y_i/\theta)$ is called the "Fisher information matrix." It may help to note that - $E_\theta h_i(Y_i/\theta) =$ $E_\theta\left((gi(Y_i/\theta)^T gi(Y_i/\theta)\right)$ > 0, where T stands for transportation.

Assume for the moment that the model is correct, and θ_0 is the true value of θ. So the Y_i are independent and the density of Y_i is $f_i(y/\theta_0)$. The log likelihood function can be expanded in a Taylor series around θ_0:

$$L = L(\theta_0) + L^{'}(\theta_0)(\theta - \theta_0) + \frac{1}{2}(\theta - \theta_0)^T L^{''}(\theta_0)(\theta - \theta_0) + \ldots \qquad (3.26)$$

If we ignore higher terms and write □ for: "nearly equal" this is an informal exposition the log likelihood function is essentially a quadratic, whose maximum can be found by solving the likelihood equation $L^{'}(\theta) = 0.$ Essentially, the equation is

So $$L^{'}(\theta_0) + (\theta - \theta_0)^T L^{''}(\theta_0) = 0.$$

$$\hat{\theta} - \theta_0 \,\square\, [-L^{''}(\theta_0)]^{-1} L^{'}(\theta_0)^T$$

Then

$$\mathrm{cov}_{\theta_0}\hat{\theta} \,\square\, [-L^{''}(\theta_0)]^{-1}[\mathrm{cov}_{\theta_0} L^{'}(\theta_0)][-L^{''}(\theta_0)]^{-1} \qquad (3.27)$$

The covariance being a symmetric p × p matrix.

In the conventional textbook development, $L^{''}(\theta_0)$ and $\text{cov}_{\theta_0} L^{'}(\theta_0)$ are computed-approximately or exactly-using Fisher information. Thus, $-L^{''}(\theta_0) = \sum_{1-1}^{n} E_{\theta_0} h_i(Y_i)$. Furthermore, \The sandwich idea is to estimate $L^{''}(\theta_0)$ directly from the sample data, as $L^{''}(\hat{\theta})$. Similarly, $\text{cov}_{\theta_0} L^{'}(\theta_0)$ is estimated as

$$\text{cov}_{\theta_0} L^{'}(\theta_0) = -\sum_{1-1}^{n} E_{\theta_0} h_i(Y_i) \tag{3.28}$$

$$\text{cov}_{\theta_0} L^{'}(\theta_0) = \sum_{i=1}^{n} gi(Y_i / \theta)^T gi(Y_i / \theta_0). \tag{3.29}$$

Where

$$A = L^{''}(\hat{\theta}) \, and \; B = \sum_{i=1}^{n} gi(Y_i / \hat{\theta})^T gi(Y_i / \hat{\theta}) \tag{3.30}$$

$$E_{\theta}[gi(Y_i / \theta)] = 0, \tag{3.31}$$

$$\text{cov}_{\theta}[gi(Y_i / \theta) = E_{\theta}(Y_i / \theta)^T gi(Y_i / \theta)]. \tag{3.32}$$

Indeed,

$$E_{\theta}[gi(Y_i | \theta) = \int gi(y|\theta) f_i(y|\theta) dy \tag{3.33}$$

$$= \int \frac{\partial}{\partial \theta} f_i(y|\theta) dy$$

$$= \frac{\partial}{\partial \theta} \int f_i(y|\theta) dy$$

$$= \frac{\partial}{\partial \theta} 1$$

$$= 0. \qquad (3.34)$$

A derivative was passed through the integral sign. Regularity conditions are needed to justify such maneuvers, but we finesse these mathematical issues.

If the motivation for the middle factor in (3.29) I still obscure, try recipe. Let U_i be independent 1 × p-vector, with E(Ui) = 0. Now $\text{cov}(\sum U_i) = \sum \text{cov}(U_i) = \sum E(U_i^T U_i)$ Estimate $E(U_i^T U_i) by U_i^T U_i$. take $U_i = gi(Y_i \mid \theta_0)$. Finally, substitute $\hat{\theta}$ for θ_0.

$$\sum_{i=1}^{n} gi(Y_i \mid \hat{\theta}) = 0. \qquad (3.35)$$

Remember, $\hat{\theta}$ was chosen to solve the likelihood equation $L^{'}(\theta) \ = \sum_{i=1}^{n} gi(Y_i \mid \theta) = 0,.$

In textbook examples, the middle factor will be of order n, being the sum of n terms. Similarly, $-L^{''}(\theta_0) \ = \sum_{i=1}^{n} gi(Y_i \mid \theta)$ will be of order n. Thus, $L^{'}(\theta)$ will be of order 1/n. under suitable regularly conditions, the strong law of large numbers will apply to $-L^{''}(\theta_0)$ so $L^{''}(\theta_0)$/n converges to a position constant; the central limit theorem will apply to $L^{'}(\theta_0)$ so $\sqrt{nL^{'}}(\theta_0)/n$ converges to a multivariate normal distribution with mean 0, in particular, the randomness in $L^{'}$ is of order $\sqrt{n}$. So is the randomness in $-L^{''}$, but that can safely be ignored when computing the asymptotic distribution of $[-L^{''}(\theta_0)]^{-1} L^{'}(\theta_0)^T$, because $-L^{''}(\theta_0)$ is of order *n*.

Robust standard errors

We turn now to the case where the model is wrong. We continue to assume the Y_i are independent. The density of Y_i, however, is φ_i-which is not our parametric family. In other words, there is specification error in the model, so the likelihood function is in error too. The sandwich estimator (3.29) is held to provide standard error that are "robust to specification error.

Key Assumption.

(A possible extension will be mentioned, below). Equation (3.18) may look questionable in this new context. But

$$E_0[gi(Y_i \mid \theta)] = \int \left(\frac{\partial}{\partial \theta} f_i(y \mid \theta) \right) \frac{1}{f_i(y \mid \theta)} \varphi_i(y)\, dx$$

$$= 0 \ at\, \theta = \theta_0 \tag{3.35}$$

This is because θ_0 minimize the Kullback-Leibler,

$$E_0[gi(Y_i \mid \theta)] = \theta \to \int \log \left[\frac{\varphi_i(y)}{f_i(y \mid \theta)} \right] \varphi_i(y)\, dy. \tag{3.36}$$

By the key assumption, we get the same θ_0 for every *i*.

Under suitable conditions, the MLE will converge to θ_0. Furthermore, $\hat{\theta}$ - θ_0 will be asymptotically normal, with mean 0 and covariance $\hat{V}$ that is,

$$\hat{V}^{1/2}(\hat{\theta} - \theta_0) \to N(0_{p,} I_{p \times p}). \tag{3.37}$$

By definition, $\hat{\theta}$ is the θ that maximizes θ → $\prod_i f_i(Y_i \mid \theta)$ – although it is granted that Y_i does not have the density $f_i(. \mid \theta)$. In short, it is pseudo-likelihood that is being maximized, not a true likelihood.

3.5 Estimation

The quantile regression estimator can be obtained as the solution to a linear programming problem. EViesws uses a modified version of the Koenker and D'Orey (1987) version of the Barrodale and Roberts (1973) simplex algorithm.

Our experience with our optimized version of the BR algorithm is that its performance is certainly better than commonly portrayed. Using various subjects of the low-birth weight data described in Koenker and Hallock (2001), we find that while certainly not as fast as Cholesky-based linear regression (and possibly not as fast as interior point methods), the estimation times for the modified algorithm are quite reasonable.

Overall, our experience is that estimation times for the modified BR algorithm are roughly linear in the number of observations through a broad range of sample sizes. While our results are not definitive, we see no real impediment to using this algorithm for virtually all practical problems.

3.5.1 Asymptotic Distributions

Under mild regularity conditions, quantile regression coefficients may be shown to be asymptotically normally distributed (koenker, 2005) with different forms of the asymptotic covariance matrix depending on the model assumptions.

Computation of the coefficient covariance matrices occupies an important place in quantile regression analysis. In large part, this importance stems from the fact that the covariance matrix of the estimates depends on one or more nuisance quantities which must be estimated. Accordingly, a large literature has developed to consider the relative merits of various approaches to estimating the asymptotic variances.

We use direct methods for estimating the covariance matrix in i.id. settings and bootstrap resampling methods for i.id setting.

3.6 Sparsity Estimation

We have seen the importance of the sparsity function in the formula for the asymptotic covariance matrix of the quantile regression estimates for i.i.d. data. Unfortunately, the sparsity is a function of the unknown distribution F, and therefore is a nuisance quantity which must be estimated.

Eviews provides three methods for estimating the scalar sparsity $s(\tau)$: two Siddiqui (1960) difference quotient methods (koenker, 1994; Bassett and koenker (1982) and one density estimator (powell, 1986; Jones, 1992; Buchinsky 1995).

3.6.1 Siddiqui Difference Quotient

The first two methods are variants of a procedure originally proposed by Siddiqui (1960; see koenker, 1994), where we compute a simple difference quotient of the empirical quantile function.

$$s'(\tau)=[F^{-1}(\tau+h_n)-F^{-1}(\tau-h_n)]/(2h_n) \qquad (3.38)$$

For some bandwidth h_n tending to zero as the sample size $n\to\infty$. $s'(\tau)$ is in essence computed using a simply two-sided numeric derivative of the quantile function. To make this procedure operational we need to estimates of the empirical quantile function $F^{-1}{}_{(\tau)}$ at the two evaluation points, and what bandwidth to empty.

The first approach to evaluating the quantile functions, which EViews terms Siddiqui (mean fitted), is due to Bassett and koenker (1982). The approach involves estimating two additional quantile regression models for $\tau - h_n$ *and* $\tau + h_n$, and using the estimated coefficients to compute fitted quantile. Substituting the fitted quantiles into the numeric derivative expression yields:

$$s'(\tau) = X^{*'}(\beta(\tau + h_n) - \beta(\tau - h_n)] / (2h_n) \tag{3.39}$$

For an arbitrary X*, while the i.i.d. assumption implies that X* mat be set to any value, Bassett and Koenker using the mean value of X, noting that mean possesses two very desirable properties: the precision of the estimate is maximized at that point, and the empirical quantile function is monotone in τ when evaluated at X* $= \overline{X}$, so that $s'(\tau)$ will always yield a positive value for suitable h_n.

A second, less computationally intensive approach to evaluating the quantile functions computes the $\tau + h$ and $\tau - h$ empirical quantiles of the residuals from the original quantile regression equation, as in Koenker (1994). Following Koenker, we compute quantiles for the residuals excluding the p residuals that are set to zero in estimation, and interpolating values to get a piecewise linear version of the quantile. Eviews refers to this method as **Siddiqui (residual).**

$$h_n = n^{-1/5}\left(\frac{4.5(\Phi(\Phi^{-1}(\tau)))^4}{[2(\Phi^{-1}(\tau))^2 + 1]^2}\right)^{1/5} \tag{3.40}$$

(approximately) minimizes the mean square error (MSE) of the sparsity estimates.

Hall-Sheather proposed n alternative bandwidth that is specifically for testing. The hall-Sheather bandwidth is given by:

$$h_n = n^{-1/5} z_a^{2/5} \left(\frac{1.5(\Phi(\Phi^{-1}(\tau)))^2}{2(\Phi^{-1}(\tau))^2 + 1} \right)^{1/3} \quad (3.41)$$

$$z_\alpha = \Phi^{-1}(1-\alpha/2)$$

Where $z_\alpha = \Phi^{-1}(1-\alpha/2)$, for α the parameter controlling the size of the desired 1- α confidence intervals.

A similar testing motivation underlies the chamberlain bandwidth:

$$h_n = z_\alpha \sqrt{\frac{\tau(1-\tau)}{n}} \quad (3.42)$$

Which is derived using the exact and normal asymptotic confidence intervals for the order statistics (Buchinsky, 1995).

3.6.2 Kernel Density

Kernel density estimators of the sparsity offer an important alternative to the Siddiqui approach. Most of the attention has focused on kernel methods for estimating the derivative

$F^{-1'}(\tau)$ directly (Falk, 1988; Welsh, 1988), but one may also estimate $s(\tau)$ using the inverse of a kernel density function estimator (Powell, 1986; Jones, 1992; Buchinsky 1995).

Where $\upsilon(\tau)$ are the residuals from the quantile regression fit. Eviews supports the latter density function approach, which is termed the Kernel (residuals) methods, since it is closely related to the more commonly employed Powell (1984, 1989).

Kernel estimation of the density function requires specification of a bandwidth C_n, we follow koenker (2005, p. 81) in choosing:

$$C_n = k(\Phi^{-1}(\tau + h_n) - \Phi^{-1}(\tau - h_n) \qquad (3.43)$$

Where, k = min (s, IQR/1.34) is the Silverman (1986) robust estimate of scale (where s the sample standard deviation and IQR the interquartile range) and h_n is the Siddiqui bandwidth.

3.7 Bootstrapping

The direct methods of estimating the asymptotic covariance matrices of the estimates requires the estimation of the sparsity nuisance parameter, either at a single point, or conditionally for each observation. One method of avoiding this cumbersome estimation is to employ bootstrapping techniques for the estimation of the covariance matrix.

Some types of bootstrap methods are: the residual bootstrap (Residual), the design, or XY-pair, bootstrap (XY-pair), and two variants of the Markov chain Marginal bootstrap (MCMB and MBMB-A).

The following discussion provides a brief overview of the various bootstrap methods. For additional detail, see Buchinsky (1995, He and Hu (2002) and Kocherginsky, He, and Mu (2005)).

3.8 Residual Bootstrap

The residual bootstrap, is constructed by resampling (with replacement) separately from the residual $u'_i(\tau)$ and X_i.

Let u^* be an m-vector of resampled residuals, and let X* be an $m \times p$ matrix of independent resampled X. Note that m need not be equal to the original sample size n.) we form the dependent variable using the resampled residuals, resampled data, and estimated coefficients, Y* = $X^* \beta^{'}(\tau) + u^*$, and then construct a bootstrap estimate of $\beta(\tau)$ using Y* and X*.

This procedure is repeated for M bootstrap replications, and the estimator of the asymptotic covariance matrix is formed.

Where $\beta(\tau)$ is the mean of the bootstrap element. The bootstrap covariance matrix V(β) is simply a (scaled) estimate of the sample variance of bootstrap estimates of $\beta(\tau)$.

Note that the validity of using draws from $n_i(\tau)$ and X_i requires independence of the u and then X.

3.8.1 XY-pair (design) Bootstrap

The XY-pair bootstrap is the most natural form of bootstrap resampling, and is valid in settings where u and X are not independent. For the XY-pair bootstrap, we simply form B randomly drawn (with replacement) subsamples of size m from the original data, then compute estimates of $\beta(\tau)$ using the (y^*, X^*) for each subsample. The asymptotic covariance matrix is then estimated from sample variance of the bootstrap results.

3.8.2 Markov chain Marginal Bootstrap

The primary disadvantage to the residual and design bootstrapping methods is that they are computationally intensive, requiring estimation of a relatively difficult *p*-dimensional linear programming problem for each bootstrap replication.

He and Hu (2002) proposed a new method for constructing bootstrap replications that each *p*-dimensional bootstrap optimization to a sequence of *p* easily solved one-dimensional linear problems. The sequence of one –dimensional solutions forms a markov chain whose sample variance, computed using consistently approximates the true covariance for large n and M.

One problem with the MCMB is that high autocorrelations in the MCMB sequence for specific coefficients will result in poor estimates for practical length.

Kocherginsky, He and Mu (KHM, 2005) propose a modification to MCMB, which alleviates autocorrection problems by transforming the parameter space prior to performing the MCMB algorithm, and then transforming the result back to the original space. Note that the resulting MCMB-A algorithm requires the i.i.d. assumption, through the authors suggest that the method is robust against heteroskedasticity.

Practical recommendations for the MCMB-A are provided in KHM. Summarizing, they recommend that the methods be applied to problems where $n, \min(\tau, 1-\tau) > 5p$ M between 100 and 200 for relatively small problems $(n \leq 1000, p \leq 10)$. for moderately large problems with np between 10,000 and 20,000,000, they recommend M between 50 and 200 depending on one's level of patience.

3.9 Model Evaluation and Testing

Evaluation of the quality of a quantile regression model may be conducted using goodness-0f-fit criteria, as well formal testing using quasi-likelihood ratio and Wald's tests.

3.9.1 Goodness-Of-Fit

Koenker and Machado (1999) defines a goodness-of-fit statistic for quantile regression that is analogous to the R^2 from conventional regression analysis. We begin by recalling our linear quantile specification, $Q(\tau - / X_i, \beta(\tau)) = X_i^{'} \beta(\tau)$ and assume that we may partition the data and coefficient vector as $X_i = (1, X_{i1}^{'})^{'}$ and = $\beta_0(\tau), (\beta_1(\tau)^{'})^{'}$, so that

$$Q(\tau - / X_i, \beta(\tau)) = \beta_0(\tau) + X_{i1}^{'} \beta_1(\tau) \qquad (3.44)$$

The minimized unrestricted intercept-only functions. The koenker and Machado goodness-of-fit criterion is given by:

$$R^1(\tau) = 1 - V(\tau) / V(\tau) \tag{3.45}$$

Equation (3.30) is a natural analog of R^2 of the conventional R^2. $R^1(\tau)$ lies between 0,1, and measures the relative success of the model.

3.9.2 Quasi-Likelihood Ratio Tests

Koenker and Machado (1999) proposed two types of quasi-likelihood ratio tests for quantile regression where the error of the distribution is flexible but not limited to the asymmetric laplace distribution, he also describe quasi-likelihood ratio tests based on the change in the optimized value of the objective function after relaxation of the restriction imposed by the null hypothesis. They offer two test statistics which they term quantile-p test through as Koenker (2005) points out, they may also be thought of as quasi-likelihood ratio tests:

$$L_n(\tau) = \frac{2(D_1 - (D_2))}{\tau(1-\tau)\hat{s}}$$

$$L_n(\tau) = \frac{2D_2(\log(D_1(\tau)) - \log D_2(\tau)))}{\tau(1-\tau)\hat{s}} \tag{3.46}$$

Where $D_1(\tau) = \sum \rho_\tau (y_i - x\beta_1(\tau))$ is the sum of check losses for the reduced model

$D_2(\tau) = \sum \rho_\tau (y_i - x\beta_1(\tau))$ is the sum of check losses for extended model, and $\hat{s}$ is the estimated sparsity function. Which are both asymptotically X_q^2 where q is the number of restrictions by the null hypothesis.

You should note the presence of the sparsity term $s(\tau)$ in the denomination of both expressions. Any of the sparsity estimator outline in "sparsity estimation" may be employed for either the null or alternative specification; EViews uses the sparsity estimation under the alternative. The presence of $s(\tau)$ should be a tip off this test statistic require that the quartile density function does not depend on x, as in the pure location-shift model.

Note that EViews will always compute an estimate of the scale sparsity, even when you specify a Huber sandwich covariance method. This value of the sparsity will be used to compute QLR test statistics which may be less robust than the corresponding Wald counterparts.

3.9.3 Coefficient Tests

Given estimates of the asymptotic covariance matrix for the quantile regression estimates, you may construct Wald-type test of hypotheses and construct coefficient confidence ellipses in coefficient diagnostics.

3.10 Quantile Process Testing

The focus of our analysis thus far has been on the quantile regression model for a single quantile, τ. in a number of cases, we may instead be interested in forming joint hypotheses using coefficients for more one quantile. We may for example, be interested in evaluating whether the location-shift model is appropriate by testing for equality of slopes across quantile values. Consideration of more than one quantile regression at the same time comes under the general category of quantile process analysis.

While the EViews equation object in set up to consider only one quantile at a time, specializes tools allow you to perform the most commonly performed quantile process analyses.

Before proceeding to the hypothesis tests of interest, we must first outline the distributional theory. Define the process coefficient vector:

$$\beta = (\beta(\tau_1)', \beta(\tau_2)', ..., \beta(\tau_K)')' \qquad (3.47)$$

3.10 Slope Equality Testing

Koenker and Bassett (1982) propose testing for slope equality quantile as a robust test of heteroskedasticity. The null hypothesis is given by

$$H_0 : \beta_1(\tau_1) = \beta_1(\tau_K) = ... = \beta_1(\tau_K) \qquad (3.48)$$

Which imposes (p-1) (k-1) restriction on the confidents. We may form the corresponding Wald statistic, which is distributed as a $\chi^2_{(p-1)(K-1)}$

3.11 Symmetry Testing

Newey and Powell (1987) construct a test of the less restrictive hypothesis of symmetry, for example least squares estimators, but the approach may easily be applied to the quantile regression case.

The premise of the Newey and Powell test is that distribution of Y given X is symmetric, then:

$$\frac{\beta(\tau) + \beta(1-\tau)}{2} = \beta(1/2) \qquad (3.49)$$

We may evaluate this restriction using Wald tests on the quantile process. Suppose that there is an odd number, K, of sets of estimated coefficients ordered by τ_k. The middle value $\tau_{(K+1)/2}$ is assumed to be equal to 0.5, and the remaining τ are symmetric around 0.5, with $\tau_{j=1-j+1,}$ for j=1,….(k-1)/2. Then the Newey and Powell test null is the joint hypothesis that:

$$H_0 : \frac{\beta(\tau_j) + \beta(\tau_{K-j-1})}{2} = \beta(1/2) \tag{3.50}$$

For J= 1, …, (K-1)/2.

The Wald test formed for this null is zero under the null hypothesis of symmetry. The null has p(K-1)/2 restrictions, so the wald statistic is distributed as a $\chi 2_{\text{p(k-1)-2}}$. Newey and Powell point out that it is known prior that the errors are i.i.d., but possibly asymmetric, one can restrict the null to only examine the restriction for the intercept. This restricted null imposes only (K-1)/2 restrictions on the process coefficients.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Introduction

Criteria used for the goodness of fit of the model is Wald test. All test of significance were conducted at 5% level using a statistical software package EVIEWS.

4.2 Numerical Presentation and Discussion of the result

The researcher discusses estimation methods for quantile regression and quantify the relationship between a response variable Y and covariates **x**, the basic quantile regression model specifies the linear dependence of the conditional quantiles. The results from the QR are shown in Tables.

Table 4.1 Quantile Regression at 0.25th quantile using Huber sandwich method

Variable	Coefficient	Standard Error	t-Statistic	P-value
C	0.099041	0.15690	0.63121	0.5289
AGRICULTURE	0.906218	0.24503	3.69839	0.0003
INDUSTRY	1.013734	0.039685	25.543955	2.3354e-54
BUILDING_AND_CONSTRUCTION	2.039748	0.398841	5.114189	1.0252e-06
WHOLESALE_AND_RETAIL	1.601430	0.371834	4.306837	3.1098e-05
Pseudo R-squared	0.811878			
Adjusted R-squared	0.866465			

Table 4.2 Quantile Regression at 0.5th quantile Huber sandwich method

Variable	Coefficient	Standard Error	t-Statistic	P-value
C	0.130350	0.257815	0.505594	0.6139
AGRICULTURE	1.014445	0.443193	2.288946	0.0236
INDUSTRY	1.061494	0.057278	18.53232	0.0000
BUILDING_AND_CONSTRUCTION	1.632614	0.616481	2.648278	0.0090
WHOLESALE_AND_RETAIL	2.181310	0.358935	6.077167	0.0000
SERVICES	0.130350	0.257815	0.505594	0.6139
Pseudo R-squared	0.805537			
Adjusted R-squared	0.799941			

In table 4.1 & table 4.2 the output shows the results of fitting a quantile regression model at 0.25 quantile using Huber Sandwich S.E, methods and 0.50 quantile to describe the relationship between GDP and other four independent variables which gives the result of the fitted model as;

__GDP_GROWTH = 0.13035 + C*AGRICULTURE + C*INDUSTRY + C*BUILDING_AND_CONSTRUCTIO + C*WHOLESALE_AND_RETAIL

The top portion of the output displays the estimation settings. Here we see that our estimates use the Huber sandwich method for computing the covariance matrix, with individual estimates.

The results show the coefficients, along with standard errors, t-statistics and associated p-values. We see all coefficients are statistically significantly different from zero at conventional levels.

The bottom portion of the output reports the Koenker and Machado (1999) goodness-of-fit measure (pseudo R-squared), and adjusted version of the statistic, as well as the scalar estimate of the sparsity using the kernel method. Note that this scalar estimate is not used in the computation of the standard errors in this case since we are employing the Huber sandwich method.

Also reported are the minimized value of the objective function, the minimized constant-only version of the objective (“Restr. objective”), the constant-only coefficient estimate (“Quantile dependent var”), and the corresponding $L_n(\tau)$ form of the Quasi-LR statistic and associated probability for the difference between the two specifications (Koenker and Machado, 1999). Note that despite the fact that the coefficient covariances are computed using the robust Huber Sandwich, the QLR statistic assumes i.i.d. errors and uses the estimated value of the sparsity. In determining whether the model can be simplified, we can notice that the P-value on the independent variables exhibit statistical significance at 95.0%.

The closer to 1 R^2 is the better the fit, therefore could be interpreted as the fitted model is a good model because R^2 = 0.81187 is an evidence also, from the result of the analysis in table 4.2, the $pseudoR^2$ which represent the relevant goodness of fit measure in quantile regression, shows an improvement on the fit of the model.

Table 4.3 Quantile Regression at 0.25th quantile using XY-pair Bootstrap method

Variable	Coefficient	Standard Error	t-Statistic	P-value
C	0.099041	0.119336	0.829936	0.4080
AGRICULTURE	0.906218	0.094170	9.623190	0.0000
INDUSTRY	1.013735	0.029491	34.37452	0.0000
BUILDING_AND_CONSTRUCTION	2.039749	0.320235	6.369541	0.0000
WHOLESALE_AND_RETAIL	1.601430	0.122936	13.02657	0.0000
SERVICES	0.099041	0.119336	0.829936	0.4080
Pseudo R-squared	0.811879			
Adjusted R-squared	0.806465			

Table 4.4 Quantile Regression at 0.5th quantile using XY-pair Bootstrap method

Variable	Coefficient	Standard Error	t-Statistic	P-value
C	0.130350	0.103352	1.261224	0.2093
AGRICULTURE	1.014445	0.098637	10.28458	0.0000
INDUSTRY	1.061494	0.041281	25.71398	0.0000
BUILDING_AND_CONSTRUCTION	2.181310	0.280679	7.771548	0.0000
WHOLESALE_AND_RETAIL	1.632614	0.113712	14.35750	0.0000
SERVICES	0.130350	0.103352	1.261224	0.2093
Pseudo R-squared	0.805537			
Adjusted R-squared	0.799941			

Comparing table 4.2 and 4.4 the most part the results are quite similar. The header information shows the different method of computing coefficient covariances and sparsity estimates. The Huber Sandwich and bootstrap standard errors are not close. There are moderate differences between the two sparsity estimates, with the kernel using residual estimator of the sparsity roughly, but this difference has no substantive impact on the probability of the QLR statistic. From our results the quantile regression coefficients indicate that as the value of independent variable increases the conditional quantile also increases and all the economic indicators are more significant using the huber sandwich standard errors but less significant using the bootstrapping estimates at 95% confidence level looking at the p-values.

Table 4.5 Quantile Slope Equality Test for 0.5th quantile

Test Summary	Chi sq Statistic	Chi sq Stat df	P-value
Wald Test	17.94586	8	0.0216
Quantiles	**Variable**	**Standard Error**	**P- value**
0.25,0.5	Agriculture	0.362769	0.7654
	Industry	0.048066	0.3204
	Building & Construction	0.354427	0.6896
	Wholesale & Retail	0.503949	0.9507
0.5, 0.75	Agriculture	0.435010	0.9516
	Industry	0.050444	0.5006
	Building & Construction	0.300353	0.6113
	Wholesale & Retail	0.612818	0.9279

The results for the slope equality test for a median regression. We compare the slope coefficient for the median against those estimated at the upper and lower quartile. The top portion of the output shows the equation specification, and the Wald test summary. Not surprisingly, we see that the $\square^2$-statistic value of 17.95 > 2.17 is statistically significant at conventional test levels. We conclude that coefficients are not the same across quantile values and that the conditional quantiles are not identical.

$$H_0 : \beta_1(\tau_1) = \beta_1(\tau_K) = ... = \beta_1(\tau_K)$$

$$H_1 : \beta_1(\tau_1) \neq \beta_1(\tau_K) \neq .... \neq \beta_1(\tau_K)$$

The p-value 0.0216 < α = 0.05 we therefore fail to accept the null hypothesis.

Table 4.6 Symmetric Quantile Test

Test Summary	Chi sq Statistic	Chi sq Stat df	P-value
Wald Test	19.44197	6	0.0035
Quantiles	Variable	**Standard Error**	**P-value**
0.25, 0.75	C	0.034241	0.0721
	Agriculture	0.450721	0.5218
	Industry	0.168957	0.7099
	Building & Construction	0.033314	0.3284
	Wholesale & Retail	0.383567	0.6868

We see that the test compares estimates at the first and third quartile with the median specification. While earlier we saw strong evidence that the slope coefficients are constant across quantiles, we now see that there is evidence of departures from symmetry. The overall p-value for the test is 0.0035, and the individual coefficient restriction test values show even more evidence of asymmetry.

$$H_0 : \frac{\beta(\tau_j) + \beta(\tau_{K-j-1})}{2} = \beta(1/2)$$

$$H_1 : \frac{\beta(\tau_j) + \beta(\tau_{K-j-1})}{2} \neq \beta(1/2)$$

We therefore fail to accept the null hypothesis of symmetry because 0.0035 < α=0.05

The Quantile regression estimation equation is represented as follows

__GDP_GROWTH = C(1) + C(2)*AGRICULTURE + C(3)*INDUSTRY + C(4)*BUILDING_AND_CONSTRUCTIO + C(5)*WHOLESALE_AND_RETAIL

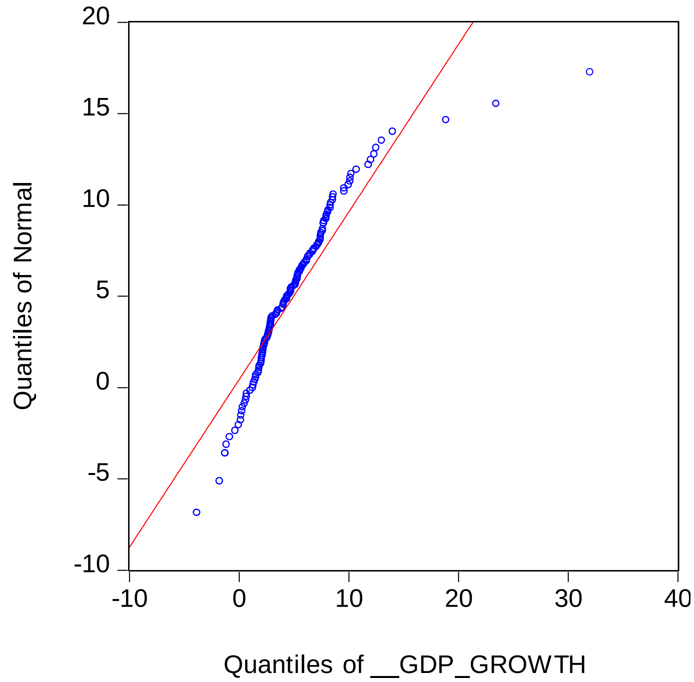

Figure 4.1 Quantile Plots

From our quantile-quantile plots we can assess the data in a single series and our data is normally distributed because the two distributions are the same,from the QQ-plot points forms a line that is almost straight which means quantiles are from a common distribution. This has help to check the assumption that our dependent variable is normally distributed.

CHAPTER FIVE

SUMMARY CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

In this chapter, we present summary, conclusion and recommendations based on the results obtained in the preceding chapter.

5.2 Summary

Quantile regression is emerging as a comprehensive approach to the statistical analysis of linear and non-linear response models, partly because classical linear theory is essentially a theoryjust for models of conditional expectations.

The primary aim in this work as stated in one of the previous chapters is to model quantile regression using time series data. In this thesis we generally discussed quantile regression model, its features and robustness in treating time series data. The research dissertation have been able to show that quantile regression offers an extension of univariate quantile estimation to estimation of conditional quantile functions thereby complementing the established mean regression methods by improving efficiency and flexibility in the estimations as well as it is more robust particularly in non-Gaussian distribution circumstances. We have also been able to see how quantile regression can help in mitigating the adverse consequence of outliers such as how it affects estimates. For example, with small data sets, one big outlier whereas quantile regression technique has proven to be a good option in cases involving such instances since LAD (quantile) regressions, are less sensitive to outliers.

The research has also been able to present a general overview of the quantile regression method, and a non-technical introduction to the basic model where we have seen that under normality. But, under non-normality, nonlinear estimators may be better than Least square estimators since the sample mean is the MLE under the Normal distribution; while the sample median is the MLE . Therefore, if we do not know which distribution is more likely, we say the median is more robust (better). We have also seen that quantile regression methods offer a useful deconstruction of conditional mean models by focusing on local parts of the conditional distribution, unlike conventional techniques.

5.3 Conclusion

Quantile regression is offering a comprehensive strategy for completing the regression picture as it goes beyond this primary goal of determining only the conditional mean, this is against the background that it enables researchers to efficiently manage the question of relationship between the response variable and covariate at any quantile of the conditional distribution function.

This case study illustrates that the quantile test is easy to apply and can be a useful tool to determine the statistical significance and its versatility when talking about residuals and constant variance assumption. Also, when dealing with non-normality and non-constant variance assumption, quantile regression is a more reliable model estimation of derivative. The common assumption of a normally distributed sample is also applicable to the quantile test.

Furthermore, Quantile regression helps one to overcome various problems that OLS is confronted with frequently; linearity of the coefficients, normality or Gaussian distribution for the response errors; error terms are not constant across a distribution and so forth by not imposing much stringent assumptions. Also, by focusing on the mean as a measure of location, information about the tails of a distribution is lost. This is as revealed by the dataset on the contribution of selected economic variables to the growth of the GDP.

In this research we have presented a general review of quantile regression models and we have illustrated that quantile regression has strong links to three very useful statistical concepts: regression,robustness and extreme value theory.

5.4 Recommendation

After analyzing and understanding the discussions, it is recommended that the assumptions of the model are fulfilled. The researcher recommends that Huber sandwitch quantile regression can be used to analyze Gross Domestic Product data to achieve better policies.

The research also recommends the involvement of the quantile regression model in the management of economic dataset intended for policy purposes so as to help boost the quality of output for better inference and help improve efficiency of decisions.

5.5 Contribution to Knowledge

This research has contributed to knowledge in various ways, the major contributions are

1) We illustrated applications of economic Gross Domestic Product data using quantile regression. In non-parametric case selecting bandwidth parameter was of utmost important.

2) The study has equally highlighted how QR provides researchers with an opportunity to examine causality beyond the mean of the distribution. By examining the lower and upper parts of a distribution which may provide significant insight for identifying process variables that influence their failure or extreme strength.

5.6 Areas of Further Studies

From the publicly available data set on Nigeria GDP we were able to model QR. We know that we can also model QR using the parametric approach, perhaps we can try forecasting using QR model too.

Extending the quantile regression approach to spatial and random-effects models is another area of interest.

References

Alexander et al., (2011). Quantile Regression for Time-series Cross-sectional data International. Journal of Statistics and Management. 6(1-2), 47-72.

Andreas, (2017). Cluster Robust Inference in Quantile Regresssion Models. Journal of American Statistical Association, Vol. 112 (517), 446-456.

Anirban, and Bijan (2013). Highlighting the Differences Between Conditional and Unconditional Quantile Regression Approaches Through an Application to Access Medication Adherence. National Centre for Biotech Information U.S. 22(9), 1052-1070.

Arellano and Bonhomme (2013). Random Effects Quantile Regression C (23): 1- 30

Asteriou and Stephen (2011). Applied Econometrics. Second Edition.

Bollerslev, (1986). Generalized Aotogressive Conditional Heteroskedasticity. Journal of Econometrics, 31, 307-327

Buchinsky, (1995). Quantile Regression, Box cox Transformation model and US Wage Structure. Journal of Econometrics, 65, 105-154.

Buchinsky, (1998). Recent Advances in Quantile Regression. The journal of Human Rsources, 33 (1), 88-126.

Buhai (2004). Quantile Regressions: Overview and Selected applications. Rotterdam Tinbergan Institute and Erasmus University.

Cade and Noon.,(2003). A gentle Itroduction to Quantile Regression for Ecologists. Frontiers in Ecology and Enviroment, 1(8), 412-420

Cade et al.,(2008). Modelling Special Distribution Using Regression Quantile. Journal of Applied Ecology, 45(1), 204-217

Cai et al (2013). Quantile Double Autoregressive Time Series. Jornal of Forecasting, 32(6), 551-560.

Cai, and John, (2015). Estimation of Non Crossing Quantile Regression Curves. Australian and Newzealand Journal of Statistics, 57 (1), 139-162.

Chernozhukov, et al (2015). Quantile Regression with Censoring and Endogeneity. Journal of Economics. 186(1), 201-221

Colin,(2004). An Introduction to Quantlle Regression and the QUANTREG Proceedure.SUG, 30, 213-230

Doguwa, (2015). Time Series Modelling of Nigeria's Internal Reserves. CBN Journal of Applied Statistics 6 (1a)

Fitzenbarger, et al (2002). Economics Application of Quantile Regression, New York, NY: Physica-Verlag Heidelberg, (1-5).

Fitzenberger, (1997). A guide to Censored Quantile Regressions. In: Handbook of Statistics, volume 15: Robust inference (Eds. G.S. Maddala and C.R. Rao.), (405- 457) Amsterdam: North Holland.

Gorr, et al (1985). An Adaptive Filtering Procedure for Estimating Regression Quantiles, Management Science, 31 (8), 1019-1029.

Gujurati, (2003). Basic Econometrics, International Edition McGraw Hill

Hao, (2007). Applied Quantile Regression (2nd ed.) John Wiley and Sons.

He, (1997). Quantile Curves without Crossing. The American statistician, 51, 186-192

Hou, et al (2012). A statistical Test for 85th and 25th Percentile Speed Using the Asymptotic Distribution of Sample Quantiles. Journal of Transportation Research Board 2279(6) 415-421

Iwok, (2016). Method of Estimating Missing Values in a Stationary AR Process. Asian Journal of Mathematics and Statistics 9 (1-3): 6-10

Iwok, (2017). Handling Seasonal Autoregressive Moving Average Model with Correlated Residuals. American Journal of Mathematics and Statistics, 7(1): 1-6

John, (2015). Robustness of Quantile Regression to Outliers. American Journal of Applied Mathematics and Statistics, 3 (2), 86-88.

Koenker, (2005). Quantile Regression . New York,NY: Cambridge University Press.

Koenker, (2008). Censored Quantile Regression Redux. Journal of Statistical Software, Vol. 27, Issue 6.

Koenker and Hillock., (2001). Quantile Regression, Journal of Economic Perspective, 15 (4), 143-156

Koenker and Orey, (1987). Algorithm AS229: Computing Regression Quantiles. Applied statistics, 36 (3): 383-393.

Koenker and Machado, (1999). Goodness of fit and Related Inference Processes for Quantile Regression. Journal of American Statistical Association, 94, 1296-1310.

Koenker and Bassett., (1978). Regression Quantiles. Econometrika, 46 (l): 33-50.

Kristinn, et al (2017). Interactive Q-learning for Quantiles, Journals of the American Statistical Association, 112, 638-649.

Manfred and Stanley., (2008). Notes for Data Analysis (11^{th}ed.). California State University Sacramento.

Mosteller and Turkey,(1977). Data Anaslysis and Regression. A Second Course in Statistics. Addisonj-Wesley: Reading MA.

Scott, (2012). Illusions in Regression Analysis. International Journal of Forecasting.3(3), 313-232

Sheather, and Marron, (1990). karnelQuantile Estimators, Journal of The American Statistical Association, 85, 410-416.

Speekman, (1998). Regression Analysis for Partially Linear Models. Journal of the Royal Statistical Society, Series B, 413-436.

Victor and Domiguez(1987). Computing Regression Quantiles. Applied Statistics, 36 (3), 383-393.

Victor and Domiguez, (2013). New Inference Methods for Quantile Regression Based on Resampling. The Econometrics journal, 16(2), 278-283.

Wagner (1959). Linear Programming Technique for Regression Analysis. Journal of the American Statistical Association, 5, 206-212.

Wunsch (2007). Confounding and Control. Demographic Research 16 (4): 97-120

Yebin, and Dawin (2014). A Quantile Regression Model for Time Series Data in the Presence of Additive Components. Journal of Communication in Statistics Theory and Methods 44(20), 20-24

Yu et al, (2003). Quantile Regression: Applications and Current Research Areas. The Statistician, 52 (3): 331-350.

APPENDIX A

Years	Quarter	% GDP Growth	Agriculture	Industry	Building and Construction	Wholesale and retail	Services
1980	**1**	2.93	1.34	-2.5	1.09	2.09	0.92
	2	5.31	1.64	0.25	0.94	1.58	0.91
	3	5.3	1.79	0.08	0.81	1.89	0.86
	4	7.85	1.5	3	0.87	1.74	0.74
1981	**1**	557.84	141.72	255.01	20.26	84.17	56.71
	2	552.19	171.27	244.97	17.04	62.22	56.68
	3	569.61	188.72	246.72	14.8	68.26	51.12
	4	525.12	153.82	244.47	15.24	66.94	44.64
1982	**1**	-3.4	0.63	-3.4	-0.98	0.71	-0.36
	2	-2.68	0.73	-2.79	-0.83	0.53	-0.32
	3	-1.8	0.78	-2.12	-0.7	0.57	-0.33
	4	-1.66	0.69	-1.86	-0.78	0.59	-0.3
1983	**1**	-6.31	-0.05	-4.39	-0.48	-0.44	-0.48
p	**2**	-7	-0.2	-5.62	-0.4	-0.32	-0.46
	3	-6.86	-0.36	-5.38	-0.33	-0.34	-0.44
	4	-8.51	-0.27	-7.11	-0.39	-0.36	-0.4
1984	**1**	-0.29	-1.49	4.24	-0.7	-1.43	-0.92
	2	-1.22	-1.76	3.1	-0.6	-1.07	-0.9
	3	-1.61	-1.81	3.09	-0.5	-1.14	-0.75
	4	-1.21	-1.62	2.29	0.02	-1.2	-0.72
1985	**1**	7.09	4.21	3.62	-0.96	0.68	0.5
	2	8.8	5.53	3.92	-0.82	0.51	0.48
	3	9.58	6.29	3.66	-0.69	0.54	0.47
	4	8.27	5.47	4.51	-1.36	0.58	0.43
1986	**1**	2.16	2.62	-1	-0.002	0.58	-0.04
	2	2.55	3.3	-1.11	-0.002	0.43	-0.07
	3	2.84	3.63	-1.1	-0.002	0.46	-0.15
	4	2.308	3.19	-1.24	-0.002	0.48	-0.12
1987	**1**	-0.46	-1.01	-0.96	0.18	1.05	0.28
	2	-0.68	-1.27	-0.6	0.15	0.77	0.27
	3	-0.68	-1.4	-0.59	0.13	0.82	0.23
	4	-0.17	-1.24	-0.29	0.14	0.87	0.21
1988	**1**	6.8	2.96	1.33	0.22	1.62	0.67
	2	7.48	3.67	1.75	0.18	1.21	0.67
	3	7.49	3.92	1.62	0.15	1.27	0.53
	4	7.75	3.47	2.25	0.17	1.35	0.51
1989	**1**	8	1.64	4.53	0.09	0.52	1.22
	2	7.75	1.97	3.94	0.08	0.54	1.22

	3	7.5	2.06	3.84	0.06	0.57	0.97
	4	7.2	1.85	3.73	0.07	0.6	0.95
1990	**1**	14.06	1.26	10.08	0.1	0.52	2.1
	2	13.06	1.53	8.93	0.09	0.39	2.12
	3	12.55	1.63	8.75	0.07	0.41	1.69
	4	12.37	1.45	8.75	0.08	0.44	1.65
1991	**1**	-1.74	0.95	-3.63	0.08	0.51	0.35
	2	-0.72	1.22	-2.74	0.07	0.38	0.35
	3	-0.62	1.37	-2.73	0.05	0.4	0.29
	4	-0.13	1.2	-2.11	0.06	0.43	0.29
1992	**1**	2.56	0.53	0.94	0.08	0.51	0.5
	2	2.25	0.71	0.59	0.07	0.38	0.5
	3	2.32	0.83	0.6	0.06	0.4	0.43
	4	1.89	0.71	0.28	0.06	0.43	0.41
1993	**1**	1.41	0.28	0.05	0.1	0.5	0.48
	2	1.27	0.46	-0.12	0.09	0.37	0.47
	3	1.36	0.61	-0.11	0.07	0.39	0.4
	4	1.05	0.5	-0.33	0.08	0.42	0.38
1994	**1**	0.003	0.65	-1.04	0.06	0.003	0.33
	2	0.292	0.84	-0.93	0.05	0.002	0.33
	3	0.383	0.95	-0.9	0.05	0.003	0.28
	4	0.243	0.83	-0.9	0.05	0.003	0.26
1995	**1**	2.37	1.07	0.83	0.06	0.01	0.4
	2	2.239	1.28	0.5	0.05	0.009	0.4
	3	2.229	1.35	0.51	0.04	0.009	0.32
	4	1.8	1.22	0.2	0.05	0.01	0.32
1996	**1**	4.57	1.23	2.73	0.03	0.14	0.44
	2	4.4	1.48	2.37	0.02	0.1	0.43
	3	4.39	1.58	2.32	0.02	0.11	0.36
	4	4.18	1.41	2.28	0.02	0.12	0.35
1997	**1**	2.77	1.24	0.6	0.14	0.24	0.55
	2	2.89	1.52	0.53	0.12	0.18	0.54
	3	2.9	1.64	0.52	0.1	0.19	0.45
	4	2.73	1.47	0.51	0.11	0.2	0.44
1998	**1**	3.11	1.22	0.64	0.13	0.49	0.63
	2	2.96	1.48	0.39	0.11	0.35	0.63
	3	2.99	1.58	0.43	0.09	0.37	0.52
	4	2.69	1.42	0.25	0.11	0.4	0.51
1999	**1**	2.67	1.57	-2.81	0.09	0.39	3.43
	2	0.55	1.93	-2.29	0.07	0.29	0.55
	3	0.65	2.06	-2.25	0.06	0.31	0.47

	4	0.73	1.89	-2.03	0.07	0.34	0.46
2000	**1**	7.67	0.92	5.87	0.09	0.26	0.53
	2	5.41	1.12	3.5	0.08	0.19	0.52
	3	5.34	1.21	3.42	0.07	0.2	0.44
	4	5.27	1.09	3.45	0.08	0.22	0.43
2001	**1**	5.82	1.19	0.01	0.28	0.39	3.95
	2	8.59	1.44	1.99	0.23	0.29	4.64
	3	8.65	1.54	1.93	0.2	0.3	4.68
	4	8.38	1.39	2.03	0.22	0.33	4.41
2002	**1**	17.54	15.38	-1.96	0.11	0.96	3.05
	2	22.06	19.47	-1.32	0.09	0.7	3.12
	3	24.11	21.68	-1.32	0.07	0.74	2.94
	4	21.83	19.14	-0.83	0.08	0.81	2.63
2003	**1**	10.74	2.61	6.99	0.19	0.77	0.18
	2	10	3.14	5.89	0.15	0.54	0.28
	3	10.17	3.4	5.65	0.12	0.57	0.43
	4	10.18	3.28	5.77	0.14	0.63	0.36
2004	**1**	-3.8	-3.13	-3.57	-0.41	2.77	0.54
	2	32.04	-14.59	10.52	-2.34	24.14	14.31
	3	18.91	7.24	2.93	0.07	4.54	4.13
	4	23.49	8.06	4.99	0.11	5.59	4.74
2005	**1**	4.75	2.39	-1.02	0.21	1.93	1.24
	2	4.09	2.88	-1.74	0.17	1.54	1.24
	3	8.13	3.22	1.92	0.15	1.67	1.17
	4	8.39	2.98	2.23	0.18	1.85	1.15
2006	**1**	7.1	2.56	0.18	0.24	2.66	1.46
	2	5.19	3.14	-1.48	0.2	1.86	1.47
	3	5.57	3.28	-1.08	0.17	1.89	1.31
	4	6.35	3.12	-0.41	0.19	2.07	1.38
2007	**1**	8.43	2.18	2.18	0.28	2.17	1.62
	2	5.43	2.71	-0.89	0.2	1.82	1.59
	3	6.49	3.04	-0.43	0.14	2.23	1.51
	4	7.75	3.82	-0.42	0.23	2.75	1.37
2008	**1**	5.78	2.16	-0.81	0.31	2.29	1.83
	2	5.21	2.51	-0.78	0.22	1.57	1.69
	3	6.01	2.58	-0.44	0.15	2.13	1.59
	4	8.19	3.53	-0.18	0.24	2.93	1.67
2009	**1**	5.01	1.87	-0.7	0.27	1.98	1.59
	2	7.45	3.59	-1.12	0.31	2.25	2.42
	3	7.3	3.13	-0.53	0.18	2.59	1.93
	4	7.67	3.31	-0.17	0.22	2.74	1.56

2010	**1**	7.32	2.79	-1.02	0.39	2.90	2.32
	2	7.79	3.75	-1.17	0.33	2.35	2.53
	3	7.96	3.42	-0.58	0.20	2.82	2.11
	4	8.6	3.71	-0.19	0.25	3.08	1.75
2011	**1**	6.89	0.47	4.4	0.17	1.15	0.7
	2	6.36	0.63	2.83	0.06	1.2	1.63
	3	3.6	0.91	0.92	0.28	1.14	0.34
	4	4.69	0.74	-1.43	1.19	1.25	2.94
2012	**1**	3.46	1.13	0.17	0.3	0.27	1.59
	2	4.11	1.49	0.63	1.14	0.29	1.81
	3	-2.17	-6.55	2.4	0.32	0.4	1.25
	4	3.64	2.37	-0.98	-0.48	0.5	2.14
2013	**1**	4.44	0.5	0	0.48	1.06	2.4
	2	5.41	0.57	-0.39	0.52	0.95	3.76
	3	5.16	0.94	-0.56	0.4	1.06	3.32
	4	6.77	0.76	0.81	0.48	1.28	3.44
2014	**1**	6.21	1.09	0.65	0.65	1.09	2.73
	2	6.55	0.79	1.92	0.45	0.88	2.51
	3	6.22	1.21	1.04	0.36	1.08	2.53
	4	5.95	0.89	1.3	0.44	0.87	2.45
2015	**1**	4	0.9	-1	0.5	1.1	2.5
	2	2.4	0.7	-1.1	0.3	0.9	1.6
	3	2.81	0.9	-0.1	0.01	0.7	1.3
	4	2.09	0.8	-0.7	-0.01	0.8	1.2
2016	**1**	0.2	0.6	-1.3	-0.2	1.2	-0.1
	2	-1.49	1	-1.5	-0.3	0.01	-0.7
	3	-2.6	1.2	-2.8	-0.2	-0.4	-0.4
	4	-1.6	1	-1.6	-0.2	-0.2	-0.6
2017	**1**	-0.89	0.7	-1.4	0.01	-0.6	0.4
	2	0.71	0.7	0.4	0.01	-0.3	-0.1
	3	2.11	0.9	1.6	0.01	-0.3	-0.1
	4	2.1	1.1	0.8	0.1	0.3	-0.2