Conditional Quantile Polynomial Distributed Lag Model with an Application to Rubber Price Returns

Kwadwo Agyei Nyantakyi
Ghana Institute of Management and Public Administration (GIMPA), Greenhill College, Business School, Box AH 50, Achimota
Accra, Ghana

Abstract: Impacts of returns on assets are not instantaneously felt, since there is a lag period. In this paper we consider the problem of developing a model for the conditional QPDL (quantile polynomial distributed lag) and investigate the influences of the conditioning variables on the location, scale and shape parameters of the QPDL model. As an economic application, we consider the production of rubber and its price returns of Sri Lanka. From the analysis we observed that the QPDL model applications were better estimators than the PDL (Polynomial Distributed Lag) models.

Key words: Asset returns, percentiles, parameter estimators, production.

1. Introduction

Impacts of returns on assets are not instantaneous, due to the lag effect. Some of agricultural prices have some evidences of lag behavior [1-3]. Various economic models such as geometric distributed lag model [4-6], adaptive expectations and partial adjustment models [7], rational expectation model [8], polynomial distributed lag model [9, 10], have being used for modelling these lags. Generally the conditional mean, used in theses estimations, provides only single summary measure for the conditional distribution of response (predictand) - the mean - given the predictor whereas quantile regression provides a more complete picture of the conditional distribution, thus, there is variation across the quantiles which exhibit different behavior [11, 12]. Also using conditional means for modelling price returns are not accurate since they present evidence of symmetric and cyclic responses [13, 14]. In the light of these, estimation of the QAR (Quantile Autoregressions) which focuses on the evolution of conditional quantiles in time instead of conditional means as in the usual AR (autoregression) becomes very important [15]. Quantile estimates in modelling are very important because it models the relation between a set of predictor variables and specific percentiles (or quantiles) of the response variable [16]. In order to address some of the issues raised above, Koenker and Ziao [17] developed a method for the Estimation of QAR models. We therefore in this research propose a conditional QPDL (quantile polynomial distributed lag) model to study the structural changes in the time series model by estimating the parameters of the QPDL model. As an economic application, we consider Sri Lanka rubber production and its price returns.

The objectives of this study are to develop a new method for the estimation of distributed lags using quantiles, to study the effects of rubber production on its price returns, and to compare the QPDL model and the PDL model

2. Materials and Methods

2.1 Data Source

In order to answer the issues raised above, secondary annual data was collected from FAOSTAT, food balance sheet, price statistics, available with the
Department of Census and Statistic Sri Lanka [18], and the World Bank (pink sheet) [19]. These data comprise of the production, imports, exports and prices of rubber. The rubber data ranges from 1961-2011.

2.2 Statistical Software

The R software, with the package “Quantreg” was used in fitting the conditional quantile polynomial distributed lag models.

2.3 The Model and Assumption

Given the autoregressive distributed lag

\[ y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + \ldots + \beta_n X_{t-k} + \epsilon_t \] (1)

With \( y_t \) dependent variable and \( X_t \) independent variable and \( \{\epsilon_t, 1 \leq t \leq k\} \) are independent identically distributed random errors.

The polynomial distributed lag by Almon [9] can be written as

\[ y_t = \varphi + \sum_{i=0}^{k} \beta_i X_{t-i} + \epsilon_t \] (2)

With \( k \) number of lags, and the \( \beta_i \)'s can be approximated by suitable polynomials. That is

\[ \beta_i = a_0 + a_1 i + a_2 i^2 + a_3 i^3 + \ldots + a_n i^n \]

Thus, for the \( n^{th} \) degree polynomial with \( k \) number of lags we have

\[ y_t = \varphi + \sum_{i=0}^{k} (a_0 + a_1 i + a_2 i^2 + a_3 i^3 + \ldots + a_n i^n) X_{t-i} + \epsilon_t \]

as defined by Chen, et al [20], Koenker and Bassett [21].

Given a uniformly standard random variable identically independently distributed sequence of random variables \( \{\psi_t\} \), we define the new model as conditional QPDL for the \( n^{th} \) degree polynomial as given below:

\[ y_t = \varphi(\psi_t) + \sum_{i=0}^{k} a_0 \psi_t X_{t-i} + \sum_{i=0}^{k} a_1 \psi_t(i) X_{t-i} + \sum_{i=0}^{k} a_2 \psi_t(i^2) X_{t-i} + \sum_{i=0}^{k} a_3 \psi_t(i^3) X_{t-i} + \ldots + \sum_{i=0}^{k} a_n \psi_t(i^n) X_{t-i} + \epsilon_t \] (4)

Assuming \( \{\epsilon_t, 1 \leq t \leq n\} \) are independent identically distributed random errors. We can write the lag model as:

\[ y_t = \varphi(\psi_t) + a_0 \psi_t \sum_{i=0}^{k} X_{t-i} + a_1 \psi_t \sum_{i=0}^{k} (i) X_{t-i} + a_2 \psi_t \sum_{i=0}^{k} (i^2) X_{t-i} + a_3 \psi_t \sum_{i=0}^{k} (i^3) X_{t-i} + \ldots + a_n \psi_t \sum_{i=0}^{k} (i^n) X_{t-i} + \epsilon_t \] (5)

letting

\[ Z_{ot} = \sum_{i=0}^{k} X_{t-i} \]
\[ Z_{1t} = \sum_{i=0}^{k} (i) X_{t-i} \]
\[ Z_{2t} = \sum_{i=0}^{k} (i^2) X_{t-i} \]
\[ Z_{3t} = \sum_{i=0}^{k} (i^3) X_{t-i} \]
\[ Z_{nt} = \sum_{i=0}^{k} (i^n) X_{t-i} \]

Then the QPDL can be simplified as:

\[ y_t = \varphi(\psi_t) + a_0(\psi_t) Z_{ot} + a_1(\psi_t) Z_{1t} + a_2(\psi_t) Z_{2t} + a_3(\psi_t) Z_{3t} + \ldots + a_n(\psi_t) Z_{nt} + \epsilon_t \] (6)

The conditional \( \alpha - quantile \) function (QPDL) can be written as

\[ Q_y(\alpha|Z_{ot} = \ldots, Z_{nt}) = \varphi(\alpha) + a_0(\alpha) Z_{ot} + a_1(\alpha) Z_{1t} + a_2(\alpha) Z_{2t} + a_3(\alpha) Z_{3t} + \ldots + a_n(\alpha) Z_{nt} + Q_y(\epsilon_t) \] (7)

where \( Q_y(\epsilon_t) = u_t \)

This can simply be written as
Conditional Quantile Polynomial Distributed Lag Model with an Application to Rubber Price Returns

\[ Q_{y_t}(\alpha | \xi) = Z_{it}^T \alpha(\alpha) \]

where

\[ Z_{it}^T = (1, Z_{0t}, - - - , Z_{nT})^T \]

Considering a second degree polynomial, we have

\[ Q_{y_t}(\alpha | Z_{it}) = \varphi(\alpha) + a_0(\alpha)Z_{0t} + a_1(\alpha)Z_{1t} + a_2(\alpha)Z_{2t} + \varepsilon_t \]

The conditional cumulative probabilities of \((Y_t)\).

This is given by

\[ P(Y_t \leq q(Z_{it}) | Z_{it} = z) = \alpha. \]

We solve the minimization problem

\[ S(\varphi, a_0, a_1 a_2) = \sum_{t|Y_t \geq aZ_{it}} \alpha(1 - \alpha) + \sum_{t|Y_t < aZ_{it}} \alpha \]

For a QPDL (2) we have

\[ Y_t = \varphi(\alpha) + a_0(\alpha)Z_{0t} + a_1(\alpha)Z_{1t} + a_2(\alpha)Z_{2t} + \varepsilon_t \]

Therefore, to solve the minimization problem for the estimates for \(\varphi, a_0, a_1, a_2\) we have

\[ S(\varphi, a_0, a_1 a_2) = \sum_{t|Y_t \geq aZ_{it}} \alpha(1 - \alpha) + \sum_{t|Y_t < aZ_{it}} \alpha \]

\[ (1 - \alpha)(1 - \alpha) = \sum_{t|Y_t \geq aZ_{it}} \alpha(1 - \alpha) + \sum_{t|Y_t < aZ_{it}} \alpha \]

where \(\omega = (\varphi(\alpha), a_0(\alpha), a_1(\alpha), a_2(\alpha))\) denotes the coefficients that can be explained by the independent variables with an AIC = 75.79 and BIC = 85.148.

We therefore proceed with the QPDL model estimation.

### 3. Results and Discussion

#### 3.1 Rubber Production

The summary statistics of rubber production (in tonne) and returns on rubber between 1964–2011 are given in Table 1.

From Table 1 we observe that the average price of rubber between 1964-2011 was USD 1.7029 per year with a minimum price of USD 0.75 and maximum of USD 4.43, respectively. Also it could be seen that the average production of rubber between 1964-2011 was 124,831 tonne with a minimum production of 86,230 and maximum of 159,158 tonne.

The polynomial distributed lag (3-lags) parameter estimates with Proz0 the first polynomial transformation (rubber production lag one), Proz1 the second polynomial transformation and Proz2 the third polynomial transformation, the standard errors and confidence intervals are given in Table 2.

From Table 2, we observed that only the parameter estimates of the second and the third lags are significant and the rest of the estimates are not significant. The R-squared value is 0.52, that is 52% of the dependant variable can be explained by the independent variables with an AIC = 75.79 and BIC = 85.148.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observations</th>
<th>Mean</th>
<th>Std. Dev</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rubber price</td>
<td>48</td>
<td>1.7029</td>
<td>0.7033</td>
<td>0.75</td>
<td>4.43</td>
</tr>
<tr>
<td>Rubber production</td>
<td>48</td>
<td>124,831</td>
<td>21,735.91</td>
<td>86,230</td>
<td>159,158</td>
</tr>
</tbody>
</table>

| Price | Coef. | Std. Err. | t      | P > |t|  | [95% Conf. Interval] |
|-------|-------|-----------|--------|------|-----|---------------------|
| Proz0 | -0.0000137 | 8.34e-06  | -1.64 | 0.108 | -0.0000305 | 0.0000031 |
| Proz1 | 0.000057   | 9.73e-06  | 5.86   | 0.000 | 0.0000374 | 0.0000766 |
| Proz2 | -0.0000198 | 3.24e-06  | -6.11  | 0.000 | -0.0000263 | -0.0000132 |
| Constant | 0.3184857 | 0.4604236 | 0.69   | 0.493 | -0.6094371 | 1.246408 |

R-squared = 0.52, RMSE = 0.5015, Obs = 48, Prob > F = 0.001, AIC = 75.79, BIC = 85.148.
The conditional QPDL with 3-lags, regression parameter estimates with Proz0 the first polynomial transformation for rubber production, Proz1 the second polynomial transformation and Proz2 the third polynomial transformation, the standard errors and confidence intervals are given in Table 3, 4, 5 and 6 below.

From Table 3, the results show that all parameter estimates are significant except Proz0 and the constant with an R-square of 0.67.

The resulting equation from Table 3 is:

\[ \hat{Y}_t = 0.167703 - 0.00000459Z_{0t} + 0.0000349Z_{1t} - 0.000013Z_{2t} \]  

(3.1)

Estimating the coefficients of the original variables, we have

\[ \hat{\beta}_0 = \hat{\alpha}_0 = -4.59 \times 10^{-6} \]
\[ \hat{\beta}_1 = \hat{\alpha}_0 + \hat{\alpha}_1 + \hat{\alpha}_2 = -4.59 \times 10^{-6} + 0.000035 - 0.000013 = 0.00001731 \]

\[ \hat{\beta}_2 = \hat{\alpha}_0 + 2\hat{\alpha}_1 + 4\hat{\alpha}_2 = -4.59 \times 10^{-6} + (2 \times 0.000035) \]
\[ - (4 \times 0.000013) = 0.0000132 \]
\[ \hat{\beta}_3 = \hat{\alpha}_0 + 3\hat{\alpha}_1 + 9\hat{\alpha}_2 = -4.59 \times 10^{-6} + (3 \times 0.000035) \]
\[ - (9 \times 0.000013) = -0.0000169 \]

Therefore, we have the QPDL estimated model as:

\[ Y_t = 0.1677 - 0.00000459X_{0t} + 0.00001733X_{t-1} + 0.0000132X_{t-2} - 0.0000169X_{t-3} \]  

(3.2)

The resulting equation from Table 4 is:

\[ Y_t = 1260894 - 0.0000147Z_{0t} + 0.0000626Z_{1t} - 0.0000218Z_{2t} \]

Although at the 50 percentile from Table 4, all the parameter estimates of the 1st, 2nd, and the 3rd lags are significant except the constant with R-square of 84% and AIC = 66.28824. Hence we proceeded to Table 5,

The resulting equation from Table 5 is:

Table 3  Conditional quantile polynomial distributed lag (QPDL) 0.25 quantile regression parameter estimates, standard errors and confidence intervals for production and price of rubber.

| Quantile | Price | Coef  | Std. Err. | t    | P > |t|  | [95% Conf. Interval] |
|----------|-------|-------|-----------|------|-----|---|----------------------|
| 0.25     | Proz0 | -4.59e-06 | 4.51e-06 | -1.02| 0.315|   | -0.0000137 0.0000045 |
|          | Proz1 | 0.0000349 | 5.96e-06 | 5.85 | 0.000|   | 0.0000229 0.0000469 |
|          | Proz2 | -0.000013 | 1.74e-06 | -7.46| 0.000|   | -0.0000165 -0.0000095 |
|          | Constant | 0.167703 | 0.2572442 | 0.65 | 0.518|   | -0.3506713 0.6862118 |

Pseudo R-squared = 0.67, Number of obs = 48, Prob > F = 0.000, AIC = 54.81562.

Table 4  Conditional quantile polynomial distributed lag (QPDL) 0.50 quantile regression parameter estimates, standard errors and confidence intervals for production and price of rubber.

| Quantile | Price | Coef  | Std. Err. | t    | P > |t|  | [95% Conf. Interval] |
|----------|-------|-------|-----------|------|-----|---|----------------------|
| 0.50     | Proz0 | -0.0000147 | 5.42e-06 | -2.71| 0.009|   | -0.0000256 -3.78e-06 |
|          | Proz1 | 0.0000626 | 6.32e-06 | 9.91 | 0.000|   | 0.0000499 0.0000754 |
|          | Proz2 | -0.0000218 | 2.11e-06 | -10.34| 0.000|   | -0.0000261 -0.0000176 |
|          | Constant | 0.1260894 | 0.2857902 | 0.44 | 0.661|   | -0.449883 0.7020617 |

Pseudo R-square = 0.84, Number of obs = 48, Prob > F = 0.000, AIC = 66.28824.

Table 5  Conditional quantile polynomial distributed lag (QPDL) 0.75 quantile regression parameter estimates, standard errors and confidence intervals for production and price of rubber.

| Quantile | Price | Coef  | Std. Err. | t    | P > |t|  | [95% Conf. Interval] |
|----------|-------|-------|-----------|------|-----|---|----------------------|
| 0.75     | Proz0 | -0.0000259 | 0.0000105 | -2.48| 0.017|   | -0.000047 -4.84e-06 |
|          | Proz1 | 0.0000731 | 0.0000157 | 4.66 | 0.000|   | 0.0000415 0.0001047 |
|          | Proz2 | -0.0000225 | 5.23e-06 | -4.31| 0.000|   | -0.0000331 -0.000012 |
|          | Constant | -0.5072473 | 0.5434615 | -0.93| 0.356|   | -1.602702 0.5878474 |

Pseudo R-square = 0.86, Number of obs = 48, Prob > F = 0.001, AIC = 88.64059.
$Y_t = -0.507 - 0.0000259Z_{xt} + 0.0000731Z_{1t} - 0.0000225Z_{2t}$

Although at the 75th percentile from Table 5, all the parameter estimates of the 1st, 2nd, and the 3rd lags are significant except the constant with R-square of 86% and AIC = 88.64. Hence, we proceeded to Table 6.

3.2 Test of the Parameter Estimates.

Khmaladze [22] tests whether the location-shift of the three QPDL lag coefficients are constant with respect to tau. The function Khmaladze Test computes both a joint test that all the covariate effects satisfy the null hypothesis (Tn), and a coefficient by coefficient version of the test.

$\text{Tn} = 4.314855$

We fail to reject the joint test statistic of the hypothesis that all the slope parameters of the model satisfy the hypothesis at critical values found in Ref. [17, 23] which are:

1\% critical value 5.350 and 5\% critical value 4.523 with $x_1 = \text{proz0}$ parameter estimate, $x_2 = \text{proz1}$ and $x_3 = \text{proz2}$. The test statistic estimates for $x_1$ to $x_3$ are given below:

$\text{Tn} = x_1$ $x_2$ $x_3$

$1.113720$ $1.346683$ $1.724839$

Hence the individual slope parameters satisfy the null hypothesis.

3.3 Analysis of Variance

Joint test of equality of slopes: tau in \{0.25 0.5 0.75\} is summarised in Table 7.

From the analysis of variance, we reject that there is no differences between the slopes of all the tau values at 5\% significance level.

The AIC’s for the QPDL rubber estimates were 54.81562, 66.28824, 88.64059 and 101.849 at tau =

### Table 6 Conditional quantile polynomial distributed lag (QPDL) 0.95 quantile regression parameter estimates, standard errors and confidence intervals for production and price of rubber.

<table>
<thead>
<tr>
<th>Quantile</th>
<th>Price</th>
<th>Coef</th>
<th>Std. Err.</th>
<th>t</th>
<th>P &gt;</th>
<th>95% Conf. Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.95</td>
<td>Proz0</td>
<td>-0.0000435</td>
<td>4.91e-06</td>
<td>-8.86</td>
<td>0.000</td>
<td>-0.0000534-0.0000336</td>
</tr>
<tr>
<td></td>
<td>Proz1</td>
<td>0.0000941</td>
<td>5.02e-06</td>
<td>18.75</td>
<td>0.000</td>
<td>0.000084-0.0001043</td>
</tr>
<tr>
<td></td>
<td>Proz2</td>
<td>-0.0000257</td>
<td>1.17e-06</td>
<td>-21.97</td>
<td>0.000</td>
<td>-0.0000281-0.0000234</td>
</tr>
<tr>
<td></td>
<td>Constant</td>
<td>-1.303292</td>
<td>0.4239428</td>
<td>-3.07</td>
<td>0.004</td>
<td>-2.157693-0.4488914</td>
</tr>
</tbody>
</table>

Pseudo R-square = 0.96, Number of obs = 48, Prob > F = 0.0031, AIC = 101.84897.

### Table 7 Quantile regression analysis of variance table.

<table>
<thead>
<tr>
<th>Df</th>
<th>ResidDf</th>
<th>F value</th>
<th>Pr(&gt; F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>138</td>
<td>2.3588</td>
<td>0.0336</td>
</tr>
</tbody>
</table>
0.25, 0.50, 0.75 and 0.95, respectively, and the PDL has AIC = 75.79 and BIC = 85.148. The AIC values for the QPDL at lower quantiles were smaller than the AIC for the PDL and higher than the PDL at higher quantiles.

3.4 Graphical Interpretation

Fig. 1 gives the QPDL parameter estimates of rubber production which enables us to compare and make inference, on the fitted model in the coefficient plots. The solid black line in these plots is the point estimate of the respective quantile regression fits, and the lighter grey region indicates a 95% confidence region. The solid (horizontal) black line in some of the plots indicates a null effect. The solid straight horizontal line in each of the plots indicates the estimated OLS effects with a red dashed lines around it representing its 95% confidence intervals.

From the intercept we could observe that quantile estimates differ significantly from the OLS and lies below the OLS line, and also decreases along higher quantiles.

From the production lag (proz0), the estimates decreases across the quantiles, and differ significantly from the OLS since it lies far above, and below the 95% confidence interval at both lower and higher quantiles.

From (proz1) there is an increase across the quantiles, and this differs significantly from the OLS since it lies below the OLS at lower quantile and above it at higher quantiles.
Although (proz2) decreases across the quantiles, it is not significantly different from the OLS estimates since it lies within the 95% confidence interval of the OLS estimates.

4. Conclusions

The hypothesis that there are no differences between the estimated slopes of all the tau values at 5% Significance level was rejected. Thus, there is a significant difference between all the slopes. The AIC’s for the QPDL rubber estimates were 54.81562, 66.28824, 88.64059 and 101.849 at tau = 0.25, 0.50, 0.75 and 0.95, respectively, and the PDL has AIC = 75.79 and BIC = 85.148.

Hence, we observe that there is an impact of both current production and three past productions on the pricing of rubber as suggested by lag effects in economic theory.

From the above analysis it could be observed that the QPDL models are better estimators than the PDL models, since the R-square measures of the QPDL models are higher than that of the PDL models and also the parameter estimates of the QPDL models are significant unlike that of the PDL models which were not significant.

We also observed that the “Ordinary” regression provides only a single summary measure for the conditional distribution of response (predictand) - the mean - given the predictor whereas quantile regression provides a more complete picture of the conditional distribution.

This idea naturally lends itself to all the generalizations that have followed ordinary regression. Many problems would actually benefit more from the prediction of extreme values rather than the mean.

From the graphical presentation, the true differences could be observed.

References

Conditional Quantile Polynomial Distributed Lag Model with an Application to Rubber Price Returns


