



**Grażyna Trzpiot** 

University of Economics in Katowice, Faculty of Informatics and Communication

Department of Demography and Economic Statistics, [grazyna.trzpiot@ue.katowice.pl](mailto:grazyna.trzpiot@ue.katowice.pl)

## Quantile Non-parametric Additive Models

**Abstract:** Quantile regression allows us to assess different possible impacts of covariates on different quantiles of a response variable. Additive models for quantile functions provide an attractive framework for non-parametric regression applications focused on functions of the response instead of its central tendency. Total variation smoothing penalties can be used to control the smoothness of additive components. We write down a general approach to estimation and inference for additive models of this type. Quantile regression as a risk measure has been applied in sector portfolio analysis for a data set from the Warsaw Stock Exchange.

**Keywords:** Quantile regression, nonparametric regression, additive model

**JEL:** G11, C19

# 1. Introduction

Methods of supervised learning – generalised additive model (GAM) methods – are one of the most comprehensive procedures for nonparametric regression models. The idea is based on greater flexibility than traditional parametric modelling methods such as linear models or generalised linear models.

Models with additive nonparametric effects offer a valuable dimension reduction device throughout applied statistics. In this paper, we describe estimation methods for additive quantile regression models. The methods employ the total variation smoothing penalties introduced by Koenker (Koenker, Ng, Portnoy, 1994) for univariate components and next for bivariate components (Koenker, Mizera, 2004). These methods focus on selection of smoothing parameters including lasso-type selection of parametric components, and post-selection inference methods, particularly confidence bands for nonparametric components of the model.

The main goal of this paper is an application of the quantile regression additive model as a risk measure in sector portfolio analysis of data from the Warsaw Stock Exchange. The evaluation of empirical results is conducted to determine the existing gap between the subadditivity and robustness of risk measurement procedures.

## 2. Generalised Linear Models (GLM)

The linear regression function has the form:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki} + \varepsilon_i \quad (1)$$

where  $Y_i$  for  $i = 1, \dots, N$  are random explanatory variables,  $(x_{1i}, \dots, x_{ki})$  for  $i = 1, \dots, N$  are observed values of  $N$  – observations for  $k$  explanatory variables,  $\varepsilon_i$  are errors that for  $i = 1, \dots, N$  are by definition independent random variables with the same distribution with zero mean and constant variance.

The estimation using the least squares method determines the estimator of  $\beta$  coefficients, but it does not allow us to carry out significance tests of these coefficients. An additional assumption should be added about the distribution of random errors:  $\varepsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_N)$ , where  $\mathbf{I}_N$  is the  $N \times N$  identity matrix. The regression model can be equivalently saved in the matrix form. The linear regression model can be given by:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \text{ where } \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_N). \quad (2)$$

Taking into account the fact that the linear function of a random variable with a normal distribution is a random variable with a normal distribution, you can calculate the expected values in equation (2) and save model (2) as:

$$\mathbf{y} \sim N(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_N), \text{ where } \boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}. \quad (3)$$

This version of the model is a starting point for further work with the regression model, as it is easier to extend the model to distributions other than a normal distribution. It is worth noting that the expected values of  $Y$  are conditioned by the observed values of the explanatory variables, that is, the response variable is modelled with data contained in the known matrix  $\mathbf{X}$ . Based on equation (3), it can be shown that a linear relationship will measure the expected value  $Y$ ,  $E(Y) = \mu$ , and the values of the explanatory variables.

In practice, you can encounter the following problems when choosing a linear model:

- 1) the relationship between the response variable and explanatory variables is not linear,
- 2) random variables  $\varepsilon_i$ , and consequently the response variable, have no normal distribution,
- 3) random variables  $\varepsilon_i$  are not independent,
- 4) variance  $\varepsilon_i$  (i.e. the variance of the explanatory variable) is not constant for all observations.

Generalised linear models (GLM) and generalised additive models (GAM) are a partial solution to the first and second problem (Lindsey, 1997).

The equation (3) of the linear model can be extended to generalised linear models (GLM). In the generalised linear model in formula (3), the normal distribution of the variable  $Y_i$  is replaced by the exponential family of distributions. In addition, a monotonic link function  $g(\times)$  is introduced describing the relationship of the expected value of the response variable  $Y_i$  designated  $\mu_i$  with the linear predictor  $\eta_i$  being a linear combination of explanatory variables:  $g(\mu_i) = \eta_i = \mathbf{x}_i^T \boldsymbol{\beta}$ .

In the vector form, the above-presented record takes the form:

$$g(\boldsymbol{\mu}) = \boldsymbol{\eta} = \mathbf{X}^T \boldsymbol{\beta}. \quad (4)$$

### 3. Generalised Additive Models (GAM)

Traditional linear models and generalised linear models in many situations turn out to be an insufficient tool, as in situations describing reality many phenomena have a more complex character. An alternative for linear models and GLM mod-

els will be presented – generalised additive models (GAM). GAM models were developed in 1986 by Trevor Hastie and Rob Tibshirani. They proposed estimations for a multidimensional set of variables by means of additive approximation of the regression function, replacing the linear functions of explanatory variables with additive “non-parametric” functions which can be estimated by smoothed cubic spline functions.

In terms of regression, GAM have the following form:

$$E(Y_j|x_{i1}, x_{i2}, \dots, x_{ik}) = \alpha + f_1(x_{i1}) + f_2(x_{i2}) + \dots + f_k(x_{ik}), \quad (5)$$

where  $\alpha$  is a constant influence effect,  $f_j, j = 1, 2, \dots, k$  are unknown functions of the  $j$ -th explaining variable estimated, among others, using locally polynomial 1 or smoothed cubic spline functions<sup>1</sup>.

The estimation of the function  $f_j$  takes place jointly for  $j = 1, \dots, k$  using a certain iterative procedure – the *back fitting* algorithm.

In GAM models, the mean of the response variable  $Y$ , conditioned by the explanatory variables,  $\mu = E(Y|X)$ , is modelled using the additive functions  $f_j, j = 1, 2, \dots, k$  of explanatory variables. Similarly to GLM, we can specify  $g$  functions which will link  $\mu$  with additive functions of explanatory variables:

$$g(\mu) = \alpha + f_1(x_1) + f_2(x_2) + \dots + f_k(x_k). \quad (6)$$

Well-known examples of link functions include:

- 1)  $g(\mu) = \mu$  is the identity link function used in linear and additive models for a Gaussian result variable;
- 2)  $g(\mu) = \text{logit}(\mu)$  for models with the binary variable  $Y$ ;
- 3)  $g(\mu) = \text{probit}(\mu)$  for models with the binomial probability distribution of the variable  $Y$ ; *probit* is an inverse function to a Gaussian distribution:  $\text{probit}(\mu) = \Phi^{-1}(\mu)$ ;
- 4)  $g(\mu) = \log(\mu)$  for log-linear or log-additive models, for models with the Poisson distribution of the variable  $Y$ . The above-presented distributions belong to the exponential distribution family. In GAM models, similarly to GLM, the variable  $Y$  belongs to the exponential distribution family. In order to simplify the algorithm scheme, we can assume that the response variable under consideration has a Gaussian distribution and the  $g(x)$ . The function is an identical link function. Additionally, as the estimation method, we choose smoothed cubic spline functions.

<sup>1</sup> The matching function for smooth cubic spline functions in the R programme corresponds to the function `smooth.splines` (`splines`).

Taking into account the above-presented assumptions, the additive model has the following form:

$$Y_i = \alpha + \sum_{i=1}^k f(x_{ij}) + \varepsilon_i, \quad (7)$$

where  $\varepsilon_i$  is a random error with an average of 0.

## 4. Additive Models for Quantile Regression (AMQR)

Models with additive nonparametric effects offer a valuable dimension reduction device throughout applied statistics. Additive models have been introduced by Breiman and Friedman (1985) and Hastie and Tibshirani (1986; 1990). They provide a pragmatic approach to nonparametric regression modelling; by restricting nonparametric components so that they are composed of low-dimensional additive pieces, we can omit some of the worst aspects of the notorious curse of dimensionality. Additive models for quantile regression, and especially our implementation of methods in R, have been heavily influenced by Wood (2006; 2010).

In some fundamental respects, the approaches are quite distinct:

- 1) Gaussian likelihood is replaced by (Laplacean) quantile fidelity,
- 2)  $L^2$  norms used as measures of the roughness of fitted functions are replaced by corresponding  $L^1$  norms measuring total variation,
- 3) truncated basis expansions are supplanted by sparse algebra as a computational expedient.

In many other respects, however, the structure of the models is quite similar to the conditional mean model. We write down models for conditional quantiles indexed by  $\tau \in (0; 1)$  in the general form:

$$Q_{Y_i|x_i,z_i}(\tau|x_i,z_i) = x_i^T \theta_0 + \sum_{j=1}^J g_j(z_{ij}). \quad (8)$$

The nonparametric components  $g_j$  will be assumed to be continuous functions, either univariate,  $\mathbb{R} \rightarrow \mathbb{R}$ ; or bivariate,  $\mathbb{R}^2 \rightarrow \mathbb{R}$ . We will denote the vector of these functions as  $g = (g_1, \dots, g_J)$ . The task is to estimate these functions together with the Euclidean parameter  $\theta_0$ , by solving:

$$\min_{(\theta_0, g)} \sum \rho_\tau(y_i - x_i^T \theta_0 - \sum g_j(z_{ij})) + \lambda_0 \|\theta_0\|_1 + \sum_{j=1}^J \lambda_j \vee(\nabla g_j), \quad (9)$$

where  $q\tau(u) = u(\tau - I(u < 0))$  is the usual quantile objective function,  $\|\theta_0\|_1 = \sum_{k=1}^K |\theta_{0k}|$  and  $v(\nabla g_j)$  denotes the total variation of the derivative or gradient of the function  $g$ .

Solutions to this variational problem are piecewise linear functions with knots at the observed  $z_i$  in the univariate case and piecewise linear functions on a triangulation of the observed  $z_i$  in the bivariate case. It can be written as a linear programme with (typically) a very sparse constraint matrix consisting mostly of zeros. This sparsity greatly facilitates an efficient solution of the resulting problem, as described in Koenker and Ng (2005). Such problems are efficiently solved by modern interior point methods for linear programming. Back fitting is not required.

For use in practice, additive quantile regression methods must have several properties:

- 1) the range of model structures available for modelling quantiles must be comparable to that available for modelling the mean in conventional GAMs, otherwise the benefits of modelling quantiles may be offset by the disbenefits of insufficient model flexibility;
- 2) smoothing parameters must be estimated automatically, otherwise the modelling process becomes too labour intensive and subjective for widespread operational use;
- 3) uncertainty estimation has to be part of model estimation, since knowing forecast uncertainty is essential for operational use, and
- 4) methods must be sufficiently numerically efficient and robust for routine deployment.

Therefore, in the application part, we use the version of the algorithm<sup>2</sup> based on using smooth relationships between regressors and the quantile of interest using spline basis expansions, and we impose Gaussian smoothing priors to control model complexity. Random effects and parametric terms are not an additional complication (Wood, 2017). This algorithm performs the computations required for belief updating of priors using the loss and to estimate smoothing parameters using the general smooth modelling methods of Wood, Pya and Säfken (2016).

## 5. Application of AMQR in sector portfolio analysis

The WIG Food Index is a sector index listed on the Warsaw Stock Exchange, containing companies that participate in the WIG Index. Sector portfolio analysis was dedicated to the food sector. The base date for the WIG Food Index was set as 31<sup>st</sup> December, 1998. The sub index is characterised by the same methodology

2 <https://mfasiolo.github.io/qgam/articles/qgam.html> (accessed: 5.11.2018).

as the main WIG Index. This means that it is an income index, and when calculating it, you should take into account the prices of the shares it contains as well as the right to collect and the income from dividends. The WIG-Food Index consists of 23 companies, of which 15 were selected for analysis, which brings together 86.37% of the total shares in the portfolio and almost total shares in the market, as their sum amounts to 97.80%. The surveyed period from 18<sup>th</sup> February 2016 to 19<sup>th</sup> February 2018 consisted of 503 observations of closing prices for each of the companies.

The use of risk measures requires examining the types of rates of return distributions. The consistency of the distribution of rates of return with the hypothetical distribution was checked, which is necessary when using quantile risk measures. For this purpose, the Kolmogorov-Smirnov test was used, with the help of which the hypothesis on the compatibility of the distributions of rates of return with both normal and lognormal distributions was verified. In each case, the significance of the test is less than the assumed level of significance of 0.05. This means that the distributions of the rates of return are consistent with a normal or lognormal distribution. Next the parameters of rate of return distribution for each of the companies were calculated (Table 2), especially the third central moment, which is a measure of the asymmetry of empirical rates of return.

Table 1. Parameters of rate of return distribution

	<b>ASTARTA</b>	<b>COLIAN</b>	<b>GOBARTO</b>	<b>HELIO</b>	<b>IMCOMPA</b>
<i>R</i>	0.001	0.000	0.001	0.003	0.002
<i>V</i>	0.000	0.000	0.000	0.001	0.000
<i>S</i>	0.019	0.013	0.022	0.032	0.016
Skewness	0.512	0.571	-1.838	1.966	0.362
Kurtosis	1.892	4.344	46.205	10.51	3.385
	<b>INDYKPOL</b>	<b>KANIA</b>	<b>OTMU</b>	<b>MILKILAND</b>	<b>MBWS</b>
<i>R</i>	0.000	0.000	-0.001	0.001	-0.001
<i>V</i>	0.000	0.000	0.001	0.001	0.001
<i>S</i>	0.022	0.019	0.025	0.034	0.023
Skewness	0.931	0.327	0.819	1.200	-1.407
Kurtosis	6.676	1.406	6.553	6.158	10.87
	<b>MILKILAND</b>	<b>MBWS</b>	<b>OTMU</b>	<b>PEPEES</b>	<b>WAWEL</b>
<i>R</i>	0.001	-0.001	-0.001	0.003	0.000
<i>V</i>	0.001	0.001	0.001	0.001	0.000
<i>S</i>	0.034	0.023	0.025	0.028	0.017
Skewness	1.200	-1.407	0.819	2.243	0.159
Kurtosis	6.158	10.87	6.553	10.941	7.036

Source: own calculation

The optimal portfolio was built, it was assumed that the expected rate of return on the portfolio had to be greater than or equal to 0.001. In addition, it was assumed that the number of companies in the portfolio should be in the range from 5 to 7, and therefore simulations were carried out to choose the optimal solution. Portfolio received the optimal parameters according to the classical portfolio assumption. The following companies were included in the portfolio companies: IMCOMPANY (43.3%), ASTARTA (28%), PEPEES (14.8%), HELIO (0.77%), KSGAGRO (0.61%). Portfolio parameters  $E(R_p) = 0.001860$ ,  $V(R_p) = 0.000115$ ,  $S(R_p) = 0.010713$ .

As the second step, AMQR was applied for collecting additional information about potential investments. The results of the AMQR estimation are written below (for  $\tau = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$ ). The formula of the model for conditional quantiles indexed by  $\tau \in (0; 1)$  was:

$$\text{Portfolio}_{\text{return}} = g_1(\text{ASTARTA}) + g_2(\text{HELIO}) + g_3(\text{IMCOMPANY}) + g_4(\text{KSGAGRO}) + g_5(\text{PEPEES}).$$

Table 2. Results of estimation:  $q = 0.1$ , Link function: identity<sup>3</sup>

	<b>Coefficients</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt;  z )</b>
(Intercept)	-0.0130902	0.0008444	-15.501	< 2e-16***
ASTARTA	0.1922540	0.0441682	4.353	1.34e-05***
<b>HELIO</b>	0.0272125	0.0358617	0.759	0.448
<b>IMCOMPANY</b>	0.0563223	0.0409519	1.375	0.169
<b>KSGAGRO</b>	0.0141618	0.0162212	0.873	0.383
PEPEES	0.0635841	0.0285604	2.226	0.026*

R-sq. (adj) = 0.107; Deviance explained = 60.3%.

Source: own calculation

Table 3. Results of estimation:  $q = 0.2$ , Link function: identity

	<b>Coefficients</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt;  z )</b>
(Intercept)	-0.0080524	0.0006274	-12.835	< 2e-16***
ASTARTA	0.1719455	0.0321830	5.343	9.16e-08***
<b>HELIO</b>	0.0164149	0.0218579	0.751	0.45266
<b>IMCOMPANY</b>	0.0706967	0.0395272	1.789	0.07369
<b>KSGAGRO</b>	0.0018480	0.0160067	0.115	0.90809
PEPEES	0.0538913	0.0182244	2.957	x 0.00311**

R-sq. (adj) = 0.105; Deviance explained = 36.6%.

Source: own calculation

3 Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1, so significant coefficients are in bold.

Table 4. Results of estimation:  $q = 0.3$ , Link function: identity

	<b>Coefficients</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt;  z )</b>
(Intercept)	-0.0048054	0.0005224	-9.200	< 2e-16***
ASTARTA	0.1822405	0.0246876	7.382	1.56e-13***
<b>HELIO</b>	0.0225030	0.0149391	1.506	0.1320
<b>IMCOMPANY</b>	0.0608357	0.0395564	1.538	0.1241
KSGAGRO	0.0114467	0.0136948	0.836	0.4032
PEPEES	0.0390539	0.0161592	2.417	0.0157*

R-sq. (adj) = 0.115; Deviance explained = 21.2%.

Source: own calculation

Table 5. Results of estimation:  $q = 0.4$ , Link function: identity

	<b>Coefficients</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt;  z )</b>
(Intercept)	-0.002395	0.000484	-4.949	7.46e-07***
ASTARTA	0.184251	0.022627	8.143	3.86e-16***
<b>HELIO</b>	0.023606	0.016943	1.393	0.1635
<b>IMCOMPANY</b>	0.058168	0.031354	1.855	0.0636
<b>KSGAGRO</b>	0.015265	0.014225	1.073	0.2832
<b>PEPEES</b>	0.029889	0.016878	1.771	0.0766

R-sq. (adj) = 0.118; Deviance explained = 12.2%.

Source: own calculation

Table 6. Results of estimation:  $q = 0.5$ , Link function: identity

	<b>Estimate</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt;  z )</b>
(Intercept)	-0.0002535	0.0004502	-0.563	0.5734
ASTARTA	0.1843099	0.0229714	8.023	1.03e-15***
<b>HELIO</b>	0.0287113	0.0162737	1.764	0.0777
<b>IMCOMPANY</b>	0.0507559	0.0285765	1.776	0.0757
<b>KSGAGRO</b>	0.0182206	0.0123474	1.476	0.1400
<b>PEPEES</b>	0.0274408	0.0169367	1.620	0.1052

R-sq. (adj) = 0.12; Deviance explained = 9.21%.

Source: own calculation

Table 7. Results of estimation:  $q = 0.6$ , Link function: identity

	<b>Coefficients</b>	<b>Std. Error</b>	<b>z value</b>	<b>value Pr(&gt;  z )</b>
(Intercept)	0.0018161	0.0004333	4.192	2.77e-05***
ASTARTA	0.1785970	0.0226231	7.894	2.92e-15***
HELIO	0.0327112	0.0127013	2.575	0.0100*
<b>IMCOMPANY</b>	0.0442493	0.0242495	1.825	0.0680

	<b>Coefficients</b>	<b>Std. Error</b>	<b>z value</b>	<b>value Pr(&gt;  z )</b>
<b>KSGAGRO</b>	0.0207066	0.0121453	1.705	0.0882
<b>PEPEES</b>	0.0254548	0.0143774	1.770	0.0766

R-sq. (adj) = 0.12; Deviance explained = 11.8%.

Source: own calculation

Table 8. Results of estimation:  $q = 0.7$ , Link function: identity

	<b>Coefficients</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt;  z )</b>
(Intercept)	0.004172	0.000492	8.480	$< 2e-16^{***}$
ASTARTA	0.171091	0.026080	6.560	$5.37e-11^{***}$
HELIO	0.035320	0.013345	2.647	$0.00813^{**}$
<b>IMCOMPANY</b>	0.031043	0.024751	1.254	0.20977
<b>KSGAGRO</b>	0.028896	0.015243	1.896	0.05799
<b>PEPEES</b>	0.018844	0.014030	1.343	0.17924

R-sq. (adj) = 0.12; Deviance explained = 20.1%.

Source: own calculation

Table 9. Results of estimation:  $q = 0.8$ , Link function: identity

	<b>Coefficients</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt;  z )</b>
(Intercept)	0.0074191	0.0006572	11.289	$< 2e-16^{***}$
ASTARTA	0.1650370	0.0332550	4.963	$6.95e-07^{***}$
HELIO	0.0439175	0.0215212	2.041	$0.04128^*$
<b>IMCOMPANY</b>	0.0069497	0.0327173	0.212	0.83178
KSGAGRO	0.0456528	0.0176976	2.580	$0.00989^{**}$
<b>PEPEES</b>	0.0057250	0.0170957	0.335	0.73771

R-sq. (adj) = 0.111; Deviance explained = 34.8%.

Source: own calculation

Table 10. Results of estimation:  $q = 0.9$ , Link function: identity

	<b>Coefficients e</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt;  z )</b>
(Intercept)	0.0128307	0.0008865	14.474	$< 2e-16^{***}$
ASTARTA	0.2158058	0.0576068	3.746	$0.00018^{***}$
HELIO	0.0693276	0.0313564	2.211	$0.02704^*$
<b>IMCOMPANY</b>	0.0227034	0.0706839	0.321	0.74806
KSGAGRO	0.0584183	0.0277683	2.104	$0.03540^*$
<b>PEPEES</b>	-0.0238170	0.0187928	-1.267	0.20503

R-sq. (adj) = 0.0844; Deviance explained = 58.9%.

Source: own calculation

The estimated models have a diversified assessment of the significance of parameters. Additionally, for the chosen value of the quantile, we received different models selected by the estimation process, the equation has a different structure due to the lack of significance of the influence of subsequent explanatory variables. There is also a different assessment of the quality of the models determined using two measures R-sq. (adj) and explained deviance.

The usefulness of these models for the investor in the assessed portfolio is as follows: instead of measuring the portfolio return rate using the weighted average proposed by the classic Markowitz approach, we obtain an assessment of the variability of portfolio return ratios depending on the variability of the distribution of subsequent assets at the level of the determined level of the quantile. Presented portfolio results (Tables 2–10) inform the investor about the significantly greater variability of the extreme values of the rate of return (the tails of the portfolio distribution:  $q = 0.1$ ,  $q = 0.2$  and  $q = 0.8$  and  $q = 0.9$ ).

The level of risk aversion is always related to the equity of the investor as well as the capital invested in a given portfolio. The investor assesses the course of variability of the multidimensional distribution, obtaining information from the model for the set of different quantiles.

A very important element of the portfolio risk measurement is compliance with capital market requirements and the level of eventual capital collateral in the event of loss. Such assessments are related to the risk measure (VaR and CVaR) defined in the connection with the set of quantiles. Therefore, to be able to determine these measures, AMQR was applied for high (tail) quantile values. For the presented portfolio, results are shown in Tables 11–13.

Table 11. Results of estimation:  $q = 0.95$ , Link function: identity

	<b>Coefficients</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt;  z )</b>
(Intercept)	0.01714	0.00105	16.326	< 2e-16***
ASTARTA	0.19754	0.05660	3.490	0.000483***
HELIO	0.07493	0.03545	2.114	0.034521*
IMCOMPANY	0.08119	0.08191	0.991	0.321569
KSGAGRO	0.06638	0.02818	2.355	0.018510*
PEPEES	-0.04243	0.01692	-2.507	0.012172*

R-sq. (adj) = 0.0592; Deviance explained = 75.6%.

Source: own calculation

Table 12. Results of estimation:  $q = 0.98$ , Link function: identity

	<b>Coefficients</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt;  z )</b>
(Intercept)	0.022997	0.001598	14.394	< 2e-16***
ASTARTA	0.281302	0.064519	4.360	1.3e-05***

	<b>Coefficients</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt;  z )</b>
HELIO	0.104361	0.030227	3.453	0.000555***
IMCOMPANY	0.258274	0.092403	2.795	0.005189**
<b>KSGAGRO</b>	0.033125	0.041917	0.790	0.429382
PEPEES	-0.071385	0.023527	-3.034	0.002412**

R-sq. (adj) = -0.11; Deviance explained = 88.4%.

Source: own calculation

Table 13. Results of estimation:  $q = 0.99$ , Link function: identity

	<b>Coefficients</b>	<b>Std. Error</b>	<b>z value</b>	<b>Pr(&gt;  z )</b>
(Intercept)	0.027561	0.002494	11.052	< 2e-16***
ASTARTA	0.340449	0.088455	3.849	0.000119***
HELIO	0.113836	0.040484	2.812	0.004925**
<b>IMCOMPANY</b>	0.218061	0.155693	1.401	0.161339
<b>KSGAGRO</b>	-0.028257	0.045004	-0.628	0.530092
PEPEES	-0.093764	0.034192	-2.742	0.006101**

R-sq. (adj) = -0.169; Deviance explained = 93.6%.

Source: own calculation

There is an assessment of the quality of the models determined using two measures R-sq. (adj) and explained deviance. For the investor, this is information about high volatility derived from the tail of rate of return distribution of the portfolio, especially from assets significant in AMQR.

We based the presented application of AMQR in the analysed portfolios on a selected sector. The portfolios from the selected sector were analysed and the variability of the distribution of the rates of return in the audited period was not particularly significant. In the surveyed sector, all returns of the rate of return were characterised by a significant asymmetry, which means volatility in the tail of the returns. These properties can have a strong impact on the final results.

## 6. Conclusions

In this paper, we work with additive regression models. Additive models for quantile functions provide an attractive framework for non-parametric regression applications focused on functions of the response instead of its central tendency. Quantile regression provides a workable method for estimating effects of explanatory variables on different conditional quantiles of an outcome variable. In the application part, we focus on portfolio analysis based on the WIG Food Index, which is a sector index listed on the Warsaw Stock Exchange containing companies that participate in the WIG Index. Based on results of AMQR estimation, we claim usefulness of AMQR as a risk measure.

## References

- Breiman L., Friedman J. (1985), *Estimating optimal transformations for multiple regression and correlation*, “Journal of the American Statistical Association”, vol. 80, no. 391, pp. 580–598.
- Hastie T., Tibshirani R. (1986), *Generalized Additive Models*, “Statistical Science”, no. 1, pp. 297–310.
- Hastie T., Tibshirani R. (1990), *Generalized Additive Models*, Chapman-Hall, New York.  
<https://mfasiolo.github.io/qgam/articles/qgam.html> (accessed: 5.11.2018).
- Koenker R., Mizera I. (2004), *Penalized triograms: total variation regularization for bivariate smoothing*, “Journal of the Royal Statistical Society” (B), no. 66, pp. 145–163.
- Koenker R., Ng P. (2005), *A Frisch-Newton Algorithm for Sparse Quantile Regression*, “Mathematicae Applicatae Sinica”, no. 21, pp. 225–236.
- Koenker R., Ng P., Portnoy S. (1994), *Quantile smoothing splines*, “Biometrika”, no. 81, pp. 673–680.
- Lindsey J.K. (1997), *Applying Generalized Linear Model*, Springer, Berlin.
- Wood S. (2006), *Generalized Additive Models: An Introduction with R*, Chapman-Hall, New York.
- Wood S. (2010), *Mixed GAM Computation Vehicle with Automatic Smoothness Estimation*, <https://cran.r-project.org/web/packages/mgcv/mgcv.pdf> (accessed: 12.12.2019).
- Wood S.N. (2017). *Generalized additive models: an introduction with R*, CRC press, New York.
- Wood S.N., Pya N., Säfken B. (2016), *Smoothing parameter and model selection for general smooth models*, “Journal of the American Statistical Association”, vol. 111(516), pp. 1548–1575.

## Kwantylowe nieparametryczne modele addytywne

**Streszczenie:** Regresja kwantylowa jest narzędziem analitycznym, które pozwala na ocenę oddziaływania zmiennych wyjaśniających, współzależnych na różne kwantyle zmiennej wyjaśnianej. Addytywne modele funkcji kwantylowych stanowią atrakcyjne ramy dla nieparametrycznych aplikacji regresji skoncentrowanych na funkcjach kwantyli zamiast na ich centralnej tendencji. W celu kontrolowania gładkości składników dodatkowych można zastosować kary za całkowite wygładzanie zmian. W artykule przedstawiono ogólne podejście do estymacji i wnioskowania dla modeli addytywnych tego typu. Regresja kwantylowa wykorzystywana jako miara ryzyka została zastosowana w analizie portfela sektorowego dla zbioru danych z Giełdy Papierów Wartościowych w Warszawie.

**Słowa kluczowe:** regresja kwantylowa, regresja nieparametryczna, model addytywny

**JEL:** G11, C19

	<p>© by the author, licensee Łódź University – Łódź University Press, Łódź, Poland.          This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license CC-BY          (<a href="http://creativecommons.org/licenses/by/3.0/">http://creativecommons.org/licenses/by/3.0/</a>)</p> <p>Received: 2019-01-05; verified: 2019-03-27. Accepted: 2020-02-05</p>
 <p>Member since 2018 JM13714</p>	<p>This journal adheres to the COPE's Core Practices  <a href="https://publicationethics.org/core-practices">https://publicationethics.org/core-practices</a></p>