A fully parametric approach for solving quantile regressions with time-varying coefficients

Florentina Paraschiv, University of St. Gallen, ior/cf
with
Derek Bunn, London Business School
Sjur Westgaard, Norwegian University of Science and Technology

Paris, 17/12/2015
Outlook

- **Modelling and forecasting the risk quantiles of price distributions** is a more crucial activity than formulating central expectations: quantile regressions are a useful tool.
- State of the art literature on quantile regressions:
  - The quantile regression methodology is well developed for linear and nonparametric models (Honda (2004), Kim (2007)), but literature on non-linear models/fully parametric approaches is scarce.
  - There is no model in the literature for the situation where price quantiles are time-varying due to price adaption to fundamental variables.
- Our contribution is two-fold:
  - We propose a novel approach for solving quantile regressions by a **fully parametric approach**, as alternative to the standard semi-parametric approach of Koenker & Bassett (1978).
  - We propose a recursive algorithm for **filtering time-varying coefficients** in the context of quantile regressions.
Theoretical background

- The general specification of a dynamic time series quantile model reads:

\[ y_t = f_t(\beta, x_t) + u_t \] (1)

with \( t = 1, \ldots, n \) the time dimension; \( y_t \in \mathbb{R} \) is the dependent variable; \( x_t = (x_{1t}, \ldots, x_{pt}) \in \mathbb{R}^p \) are the explanatory variables; \( \beta' \) is a \( p \)-vector of unknown parameters and \( u_t \) is an error term.

- The conditional \( \alpha \in (0, 1) \) level quantile is then

\[ q_\alpha(y_t|\beta^\alpha, x_t) = f_t(\beta^\alpha, x_t) \] (2)

- For any \( \alpha \in (0, 1) \) the distance from \( y_t \) to a given quantile level \( q_\alpha \) is measured by the absolute distance, but different weights are applied depending on whether \( y_t \) is to the left or to the right of \( q_\alpha \) (see Hao & Naiman (2007)).

\[ d_\alpha(y_t, q_\alpha) = \begin{cases} (\alpha - 1)|y_t - q_\alpha|, & y_t < q_\alpha \\ \alpha|y_t - q_\alpha|, & y_t \geq q_\alpha \end{cases} \] (3)
The Skewed-Laplace connection

• We look for the value $q_\alpha$ that minimises the mean distance from $y_t$: $\mathbb{E}[d_\alpha(y_t, q_\alpha)]$. The minimum occurs when $q_\alpha$ is the $\alpha$’s quantile of $y_t$. $q_\alpha$ depends on $\beta^\alpha$ which is the solution to:

$$\min_{\beta} \sum_t \rho_\alpha(y_t - f_t(\beta, x_t)),$$

where $\rho(\cdot)$ is a loss function specified as:

$$\rho_\alpha(u_t) = u_t \cdot (\alpha - 1(u < 0))$$

with $u_t = y_t - f_t(\beta, x_t)$.

• We assume that residuals $u_t$ are SL-distributed. The $SL(\mu, \tau, \alpha)$ has a density function:

$$\mathcal{L}_\alpha(\beta, \tau; y, x) \propto \tau^{-n} \exp \left\{ -\tau^{-1} \sum_{t=1}^n (y_t - f_t(\beta, x_t)) \times [\alpha - 1_{(-\infty, 0)}(y_t - f_t(\beta, x_t))] \right\}.$$

• We observe that the sum which must be minimised in Equation (4) is contained in the exponent of the likelihood. Thus, the maximum likelihood estimation for $\beta$ in (6) is equivalent to the quantile estimator in (4).
The time-varying coefficients feature

- We extend the model version in Equation (1) by introducing time-varying coefficients.

- We formulate a state space model with time-dependent coefficients recursively filtered with Kalman Filter (Kalman (1960)) and model parameters are estimated with maximum likelihood.

- The assumption that $u_t$ follow a $SL(\mu, \tau, \alpha)$ distribution is used here to estimate parametrically the quantiles of the dependent variables $y_t$.

- The likelihood function is similar to the specification in (6) but we account in addition for time-varying $\beta_t$.

- The state space formulation reads:

  $$y_t^\alpha = (\beta_t^\alpha)'x_t + u_t$$  \hspace{1cm} (7)

  $$\beta_t^\alpha = c + D\beta_{t-1}^\alpha + w_t$$  \hspace{1cm} (8)

where $u_t \sim SL(0, \tau, \alpha)$, $\Xi$ is the variance of residuals $u_t$ and $w_t \sim N(0, \Omega)$. 

A fully parametric approach for solving quantile regressions – p.5
Algorithm estimation procedure

\[
\begin{align*}
\beta_{t|t-1} &= c + D\beta_{t-1}, \\
\Sigma_{t|t-1} &= D\Sigma_{t-1} D' + \Omega, \\
\mathbf{u}_t &= y_t - (\beta_{t|t-1})' \mathbf{x}_t, \\
\mathbf{R}_{t|t-1} &= \mathbf{x}_t \Sigma_{t|t-1} (\mathbf{x}_t)' + \Xi, \\
\mathbf{K}_t &= \Sigma_{t|t-1} (\mathbf{x}_t)' \mathbf{R}_{t|t-1}^{-1}, \\
\beta_{t|t} &= \beta_{t|t-1} + \mathbf{K}_t \mathbf{u}_t, \\
\Sigma_{t|t} &= (I - \mathbf{K}_t \mathbf{x}_t) \Sigma_{t|t-1}, \\
\end{align*}
\]

Maximization of the L-function, e.g., with Broyden/Fletcher/Goldfarb/Shanno

\[
L_t(\psi; y, \mathbf{x}) \propto \exp\left\{-\sum_{t=1}^{n} (y_t - f_t(\beta_t, \mathbf{x}_t)) x_t \left[\alpha - l_{\psi,0}(y_t - f_t(\beta_t, \mathbf{x}_t))\right]\right\}
\]

**Algorithm steps:**

1. **Initialization:**
   - Initial values for \(\psi\), \(\beta_{t|t-1}\), and \(\Sigma_{t|t-1}\)

2. **Update:**
   - New observation \(y_t\) at time \(t+1\)
   - Prediction error \(\mathbf{u}_t\)
   - Covariance matrix \(\mathbf{R}_{t|t-1}\)
   - Kalman gain \(\mathbf{K}_t\)
   - Updated \(\beta_{t|t}\) and \(\Sigma_{t|t}\)
   - Check if \(t = T\) or maximization criterion met?

3. **Evaluation:**
   - Evaluate new likelihood function
   - Maximize function

4. **Recursion Algorithm:**
   - Perform recursion until last observation
   - Conditional likelihood function evaluated

5. **Final Steps:**
   - Stop when Kalman algorithm is executed or predefined stopping criterion met
Challenges for the optimisation

- In each iteration of the optimization procedure the likelihood function is updated with estimates of the time dependent state $\beta_t$ which requires a new run through the Kalman Filter.

- The shape of the likelihood function is piecewise (V-shape) and nonlinear due to the indicator function.

- How do we handle this:
  - We firstly proceed with the elimination of the non-differentiability of likelihood function caused by the indicator function.
  - We reformulate the problem as a non-linear optimization with constraints
  - We further relax the problem by moving the constraints into the objective: solved numerically by the Augmented Lagrangian Method.

- Advantage of a fully parametric approach: by filtering coefficients with Kalman Filter, the estimates are robust with respect to the variance and distribution of residuals.
Estimation procedure

- Rewrite the problem as a non-linear minimisation with linear constraints:

\[
\ln \mathbf{L}_\alpha(\Psi; y, x) \propto - n \ln \tau - \frac{1}{\tau} \sum_{t=1}^{n} u_t \times [\alpha - 1_{(-\infty, 0)}(u_t)]
\]

\[
= - n \ln \tau - \frac{1}{\tau} \left[ \alpha \sum_{t=1}^{n} \{u_t | u_t \geq 0\} + (\alpha - 1) \sum_{t=1}^{n} \{u_t | u_t < 0\} \right]
\]

\[
= - n \ln \tau - \frac{1}{\tau} \left[ \alpha \sum_{t=1}^{n} \{u_t | u_t \geq 0\} + (1 - \alpha) \sum_{t=1}^{n} \{-u_t | u_t < 0\} \right]
\]

\[
= - n \ln \tau - \frac{\alpha}{\tau} \sum_{t=1}^{n} v_t^1 - \frac{(1 - \alpha)}{\tau} \sum_{t=1}^{n} v_t^2
\]

(9)

where \(v_t^1 := \max\{u_t, 0\}, v_t^2 := \max\{-u_t, 0\}\) are introduced as linear constraints.

\[
\min n \ln \tau + \frac{\alpha}{\tau} \sum_{t=1}^{n} v_t^1 + \frac{(1 - \alpha)}{\tau} \sum_{t=1}^{n} v_t^2
\]

(10)

s.t. \(v_t^1 \geq y_t - (\beta_{t|t-1})' x_t, \quad t = 1, \ldots, n\)

\(v_t^2 \geq (\beta_{t|t-1})' x_t - y_t, \quad t = 1, \ldots, n\)

\(v_t^1, v_t^2 \geq 0, \quad t = 1, \ldots, n\)
Problem with embedded betas

• We cannot pass the problem defined in (10) to a solver yet, since the coefficients $\beta_{t|t-1}$ are not constant.

• Instead, the time-varying coefficients are functions of the parameter vector $\Psi$ that should be estimated, i.e., $\beta_{t|t-1} := \beta_{t|t-1}(\Psi)$.

• As a consequence, the estimates $\beta_{t|t-1}$ for $t = 1, \ldots, n$, which are now part of the constraints, must always be updated with the Kalman Filter when the parameter vector changes in one iteration of the optimisation algorithm before the log-likelihood function (i.e., the objective) is evaluated.

• For that reason, we aim at integrating the constraints into the objective.
Problem with embedded betas

\[
\min n \ln \tau + \frac{\alpha}{\tau} \sum_{t=1}^{n} v_t^1 + \frac{(1 - \alpha)}{\tau} \sum_{t=1}^{n} v_t^2
\]

(11)

\[
\text{s.t.} \quad -v_t^1 + v_t^2 + y_t - (\beta_{t|t-1})' x_t = 0, \quad t = 1, \ldots, n
\]

\[
v_t^1, v_t^2 \geq 0, \quad t = 1, \ldots, n
\]

\[
f(\tilde{\Psi}) := n \ln \tau + \frac{\alpha}{\tau} \sum_{t=1}^{n} v_t^1 + \frac{(1 - \alpha)}{\tau} \sum_{t=1}^{n} v_t^2
\]

\[
c_t(\tilde{\Psi}) := -v_t^1 + v_t^2 + y_t - \left(\beta_{t|t}(\tilde{\Psi})\right)' x_t
\]

(12)

By moving the equality constraints into the objective as a penalty term, we can solve the problem with the Augmented Lagrangian Method (e.g., see Nocedal/Wright, 2006, ch 17). The objective of the relaxed problem reads as:

\[
\min_{\tilde{\Psi}} f(\tilde{\Psi}) \quad \text{s.t.} \quad c_t(\tilde{\Psi}) = 0; \quad v_t^1, v_t^2 \geq 0, \quad t = 1, \ldots, n.
\]

(13)
Case study: Fundamental model for quantiles of electricity prices

- Electricity day-ahead prices in Germany are explained by the following fundamental variables: demand, reserve margin, coal, gas, oil, CO2 certificate prices, wind, PV.

- 24 data sets have been constructed for each hour of a day between 01/01/2010–31/05/2014 to take into account the typical intra-day seasonality of electricity prices.

- We assume that renewable energies have a direct and indirect effect on electricity prices:

\[
P_{t}^{\alpha} = \beta_{0t}^{\alpha} + \beta_{1t}^{\alpha} P_{t-1} + \beta_{2t}^{\alpha} Wind_{t} + \beta_{3t}^{\alpha} PV_{t} + \beta_{4t}^{\alpha} Demand_{t} + \\
+ \beta_{5t}^{\alpha} Coal_{t} + \beta_{6t}^{\alpha} Gas_{t} + \beta_{7t}^{\alpha} Oil_{t} + \beta_{8t}^{\alpha} CO_{2t} + \beta_{9t}^{\alpha} Reserve_{t} \\
\beta_{jt} = c + D\beta_{j(t-1)}^{\alpha} + \varepsilon_{j}^{\alpha} Wind_{t} + \gamma_{j}^{\alpha} PV_{t} + \phi_{j}^{\alpha} Reserve_{t} + w_{t} \text{ for } j = \{1, ..., 9\}; \tag{15}
\]

where for \( j = \{1, ..., 4; 7, 8, 9\} \) \( \varepsilon_{j}, \gamma_{j} \) and \( \phi_{j} \) take the value 0.
Results hour 13. Time varying betas for gas
Results hour 13. Time varying betas for coal

Coal (hour 13)
Results hour 13. Time varying betas for wind

Wind Forecast (hour 13)
Results hour 13. Time varying betas for PV

Photo Forecast (hour 13)
Correlations betas wind and gas

Correlation Wind/Gas = -0.256
Results hour 13

- The marginal effects of gas on electricity prices generally increase with quantiles.
- **Renewable energies reduce the marginal effects of traditional fuels:** coal and gas prices on the electricity prices (indirect effect).
- There is a **more obvious time-variability pattern in the coefficients for gas and PV**, especially at the 30% and 50% quantiles, **than in the case of coal and wind:** in hours with high demand gas is burned in addition and as the most expensive technology in use, it sets the electricity price. This can still be substituted with Wind and PV, and thus price adaption.
- We observe less adaption of electricity prices to gas and PV at extremely high quantiles (90%). Extremely high electricity prices usually occur when there is excess residual demand in the market for a certain hour. PV will not substitute the gas in production, but it will have a balancing effect for the excess demand.
- The coefficients of the lagged spot prices have a positive sign at all quantiles, but no time-variability: in the case of our peak hour 13, **producers exercise market power.**
Results hour 4. Time-varying betas for gas

Gas (hour 4)
Results hour 4. Time-varying betas for wind

A fully parametric approach for solving quantile regressions – p.20
Results hour 4. Time-varying betas for carbon prices
Results hour 4

- There is a **continuous adaption of electricity prices to coal prices at all quantiles**, but highest marginal effects on central expectations: $Q_{50\%}$: production is mainly coal based in the night.

- There is **price adaption to wind at all quantiles**.

- The price adaption to coal reflects moreover changes in the fuel prices than substitution with wind (high administrative costs).

- **Marginal effects of gas prices on electricity prices increase with quantiles**: gas still sets the price. The adaption can be due to a substitution with wind in production (flexible plants).

- **Price adaption to carbon prices at all quantiles**: coal is a carbon-intense technology.
Fitted price quantiles
## Indirect effect of renewables on electricity prices

### Table 1: Parameter estimates for the transition equation including exogenous variables, for Hour 13, having as dependent variable the coefficient $\beta$ for gas. Lower panel: test for omitted variables.

<table>
<thead>
<tr>
<th>Dependent variable: $\beta_{gas}$</th>
<th>Method: Least Squares</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>$Q_{10%}$</td>
<td>$Q_{30%}$</td>
<td>$Q_{50%}$</td>
<td>$Q_{70%}$</td>
<td>$Q_{90%}$</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0.0058</td>
<td>0.1685</td>
<td>0.1013</td>
<td>0.0335</td>
<td>0.0442</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.7713)</td>
<td>(164.2877)</td>
<td>(247.3429)</td>
<td>(83.4187)</td>
<td>(24.4604)</td>
<td></td>
</tr>
<tr>
<td>coal(-1)</td>
<td>0.9689</td>
<td>0.2047</td>
<td>0.1603</td>
<td>0.2606</td>
<td>0.7477</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(113.8393)</td>
<td>(42.2338)</td>
<td>(47.1631)</td>
<td>(29.3623)</td>
<td>(72.2733)</td>
<td></td>
</tr>
<tr>
<td>$\varepsilon_{Wind}^{\alpha} \times 10^8$</td>
<td>1.7400</td>
<td>-6.4700</td>
<td>-0.3790</td>
<td>-1.9000</td>
<td>-16.9000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.6874)</td>
<td>(-9.1974)</td>
<td>(-1.2569)</td>
<td>(-6.8967)</td>
<td>(-7.4140)</td>
<td></td>
</tr>
<tr>
<td>$\gamma_{PV}^{\alpha} \times 10^8$</td>
<td>4.0000</td>
<td>-17.1000</td>
<td>-2.9500</td>
<td>-4.1000</td>
<td>-36.1000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.8491)</td>
<td>(-28.3929)</td>
<td>(-11.5297)</td>
<td>(-17.3053)</td>
<td>(-15.7340)</td>
<td></td>
</tr>
<tr>
<td>$\phi_{Reserve}^{\alpha} \times 10^8$</td>
<td>0.3860</td>
<td>-1.4100</td>
<td>-0.2090</td>
<td>0.2210</td>
<td>-3.3300</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.0177)</td>
<td>(-3.1759)</td>
<td>(-1.0939)</td>
<td>(1.2652)</td>
<td>(-2.3820)</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.9995</td>
<td>0.6497</td>
<td>0.5961</td>
<td>0.4708</td>
<td>0.8674</td>
<td></td>
</tr>
<tr>
<td>Test for omitted variables (wind, pv, reserves), p-values</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-statistic</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>Likelihood ratio</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td></td>
</tr>
</tbody>
</table>
Conclusion

- We propose a novel approach for quantile regressions with time-varying coefficients based on a fully parametric approach.
  - Time-varying coefficients are recursively filtered with Kalman Filter and model parameters are estimated by maximum likelihood.
  - To deal with the non-differentiability of the likelihood function, we formulated the problem as a non-linear optimisation with constraints.
  - We further moved the constraints into the objective and solved a relaxed problem with the Augmented Lagrangian Method.

- We showed the importance of modelling the time-varying feature in the context of a fundamental model for quantiles of electricity prices.
  - We found that marginal effects of market fundamentals on electricity prices differ across price quantiles and between different hours within one day.
  - A good understanding of the risk drivers of electricity prices is important for an optimal production planning, for the fine tuning of bids in electricity trading and generally for a lower exposure to risk.
References


Merry Energy Christmas!