Institute for Operations Research
and Computational Finance

University of St.Gallen

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

Vienna, 02/09/2015

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**.
  - Rossi & Harvey (2006), Gerlach et al. (2011): The time-varying feature comes from the assumption of autoregressive quantiles/time-varying volatility
  - 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:

$$y_t = f_t(\beta, x_t) + u_t \tag{1}$$

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

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

$$q_\alpha(y_t|\beta^\alpha, x_t) = f_t(\beta^\alpha, x_t) \tag{2}$$

- For any $\alpha \in (0, 1)$ the distance from y_t to a given quantile level q_α 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_α (see Hao & Naiman (2007)).

$$d_\alpha(y_t, q_\alpha) = \left\{ \begin{array}{ll} (\alpha - 1)|y_t - q_\alpha|, & y_t < q_\alpha \\ \alpha|y_t - q_\alpha|, & y_t \geq q_\alpha \end{array} \right\}. \tag{3}$$

The Skewed-Laplace connection

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

$$\min_\beta \sum_t \rho_\alpha(y_t - f_t(\beta, x_t)), \tag{4}$$

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

$$\rho_\alpha(u_t) = u_t \cdot (\alpha - \mathbf{1}(u < 0)) \tag{5}$$

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:

$$Ł_\alpha(\beta, \tau; y, x) \propto \tau^{-n} \exp\left\{-\tau^{-1} \sum_{t=1}^{n} (y_t - f_t(\beta, x_t)) \times [\alpha - \mathbf{1}_{(-\infty,0)}(y_t - f_t(\beta, x_t))]\right\}. \tag{6}$$

- 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 β in (6) is equivalent to the quantile estimator in (4).**

The time-varying coefficients feature

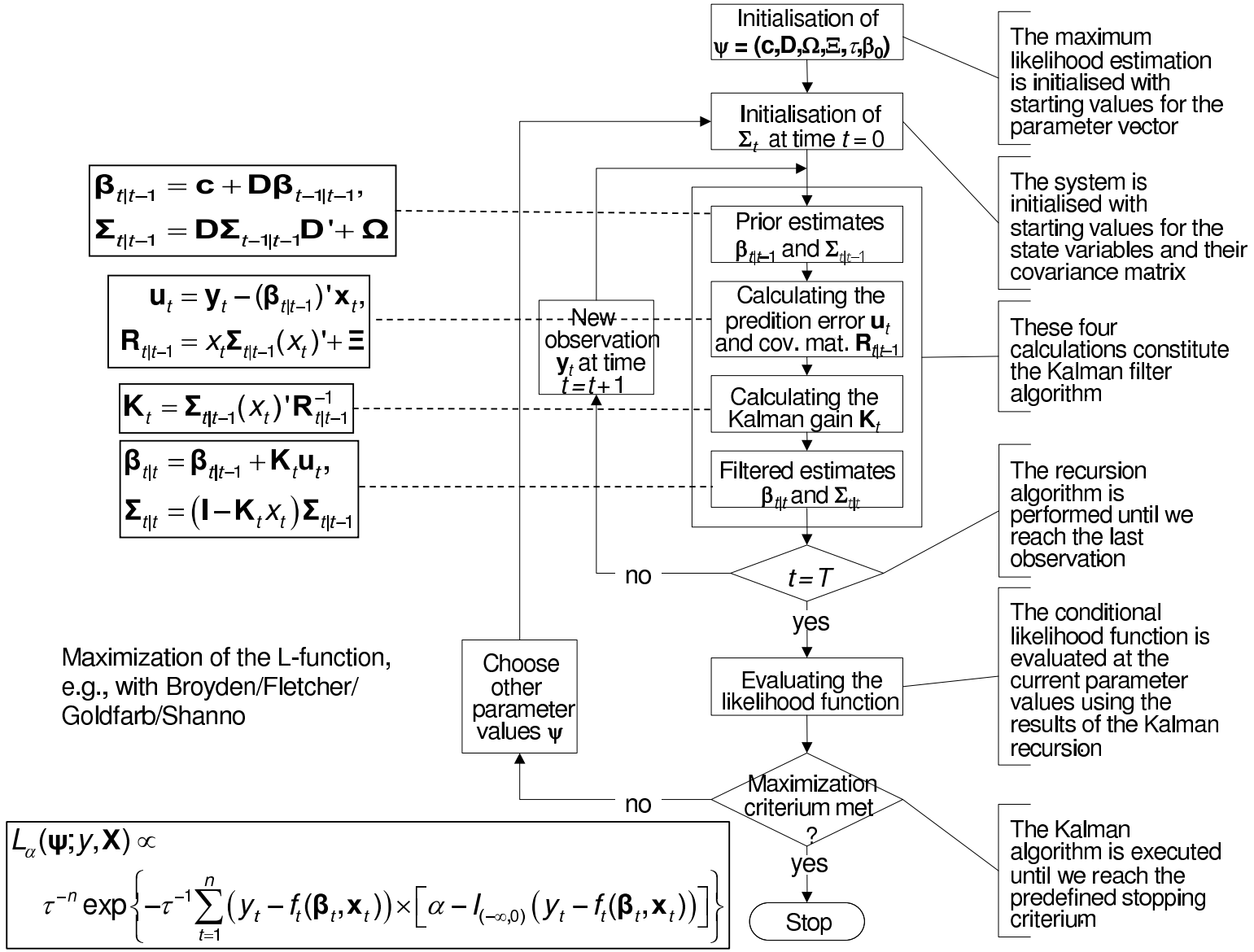
- We extend the model version in Equation (1) by introducing time-varying coefficients.
- We formulate a **state space model with time-dependent coefficients recursively filtered with Kalman Filter** (Kalman (1960)) and model parameters are estimated with maximum likelihood.
- The assumption that u_t 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 β_t.
- The state space formulation reads:

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

$$\beta_t^{\alpha} = c + D\beta_{t-1}^{\alpha} + w_t \tag{8}$$

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

Algorithm estimation procedure



Challenges for the optimisation

- In each iteration of the optimization procedure the likelihood function is updated with estimates of the time dependent state β_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:

$$\begin{aligned}
\ln Ł_\alpha(\Psi; y, x) \propto & - n \ln \tau - \frac{1}{\tau} \sum_{t=1}^{n} u_t \times \left[\alpha - \mathbf{1}_{(-\infty,0)}(u_t)\right] \\
& = - n \ln \tau - \frac{1}{\tau} \left[\alpha \sum_{t=1}^{n} \{u_t | u_t \geq 0\} + (\alpha - 1) \sum_{t=1}^{n} \{u_t | u_t < 0\}\right] \\
& = - n \ln \tau - \frac{1}{\tau} \left[\alpha \sum_{t=1}^{n} \{u_t | u_t \geq 0\} + (1 - \alpha) \sum_{t=1}^{n} \{-u_t | u_t < 0\}\right] \\
& = - n \ln \tau - \frac{\alpha}{\tau} \sum_{t=1}^{n} v_t^1 - \frac{(1-\alpha)}{\tau} \sum_{t=1}^{n} v_t^2 \qquad (9)
\end{aligned}$$

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

$$\min n \ln \tau + \frac{\alpha}{\tau} \sum_{t=1}^{n} v_t^1 + \frac{(1-\alpha)}{\tau} \sum_{t=1}^{n} v_t^2 \qquad (10)$$

$$\begin{aligned}
\text{s.t.} \quad & v_t^1 \geq y_t - (\beta_{t|t-1})' x_t, && t = 1, ..., n \\
& v_t^2 \geq (\beta_{t|t-1})' x_t - y_t, && t = 1, ..., n \\
& v_t^1, v_t^2 \geq 0, && t = 1, ..., n
\end{aligned}$$

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 Ψ 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 $t = 1, ..., n$, 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

$$\min n \ln \tau + \frac{\alpha}{\tau} \sum_{t=1}^{n} v_t^1 + \frac{(1-\alpha)}{\tau} \sum_{t=1}^{n} v_t^2 \tag{11}$$

$$\begin{aligned} \text{s.t.} \quad & -v_t^1 + v_t^2 + y_t - (\beta_{t|t-1})' x_t = 0, && t = 1, ..., n \\ & v_t^1, v_t^2 \geq 0, && t = 1, ..., n \end{aligned}$$

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

$$c_t(\widetilde{\Psi}) := -v_t^1 + v_2^1 + y_t - \left(\beta_{t|t}(\widetilde{\Psi})\right)' x_t$$

$$\min_{\widetilde{\Psi}} f(\widetilde{\Psi}) \quad \text{s.t.} \quad c_t(\widetilde{\Psi}) = 0; \quad v_t^1, v_t^2 \geq 0, \quad t = 1, ..., n. \tag{12}$$

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:

$$\Phi(\widetilde{\Psi}, \Lambda; \mu) := f(\widetilde{\Psi}) + \frac{\mu_k}{2} \sum_{t=1}^{n} \left(c_t(\widetilde{\Psi})\right)^2 - \sum_{t=1}^{n} \lambda_t c_t(\widetilde{\Psi}) \tag{13}$$

Case study: Fundamental model for quantiles of electricity prices

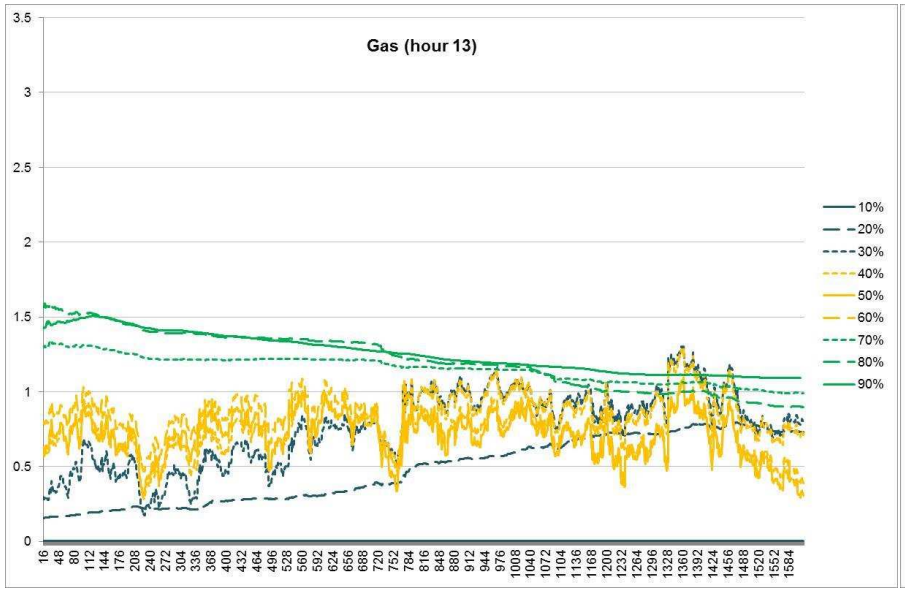
- Electricity day-ahead prices in Germany are explained by the following fundamental variables: demand, reserve margin, coal, gas, oil, CO2 certificate prices, wind, PV.
- 24 data sets have been constructed for each hour of a day between 01/01/2010–31/05/2014 to take into account the typical intra-day seasonality of electricity prices.
- We assume that renewable energies have a direct and indirect effect on electricity prices:

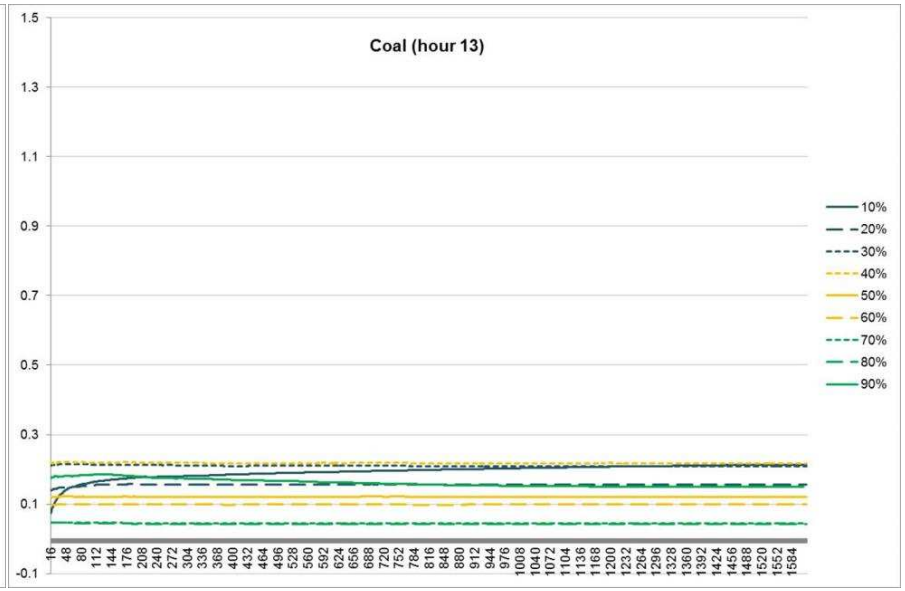
$$P_t^\alpha = \beta_{0t}^\alpha + \beta_{1t}^\alpha P_{t-1} + \beta_{2t}^\alpha Wind_t + \beta_{3t}^\alpha PV_t + \beta_{4t}^\alpha Demand_t + \\ + \beta_{5t}^\alpha Coal_t + \beta_{6t}^\alpha Gas_t + \beta_{7t}^\alpha Oil_t + \beta_{8t}^\alpha CO2_t + \beta_{9t}^\alpha Reserve_t \quad (14)$$

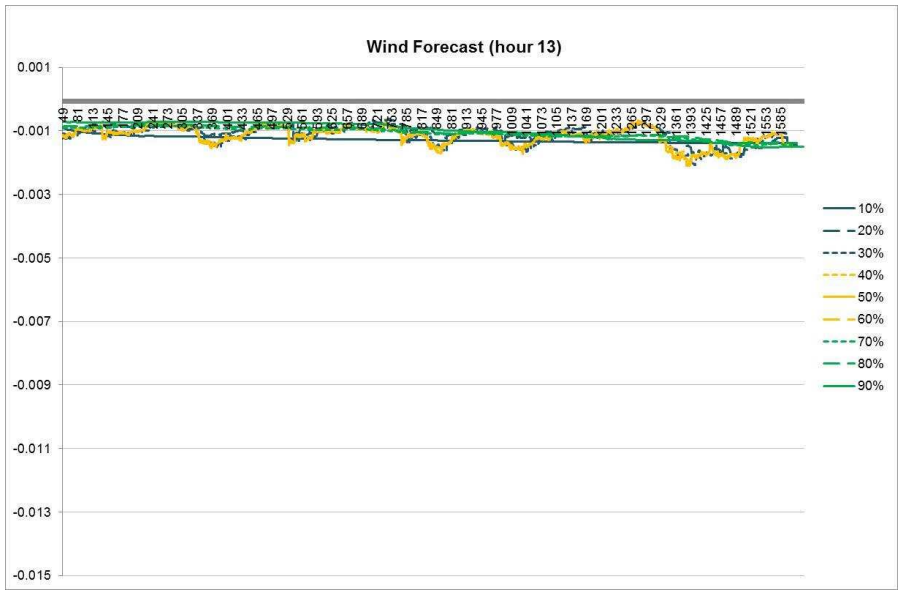
$$\beta_{jt}^\alpha = c + D\beta_{j(t-1)}^\alpha + \varepsilon_j^\alpha Wind_t + \gamma_j^\alpha PV_t + \phi_j^\alpha Reserve_t + w_t \quad \text{for} \quad j = \{1, ..., 9\}; \quad (15)$$

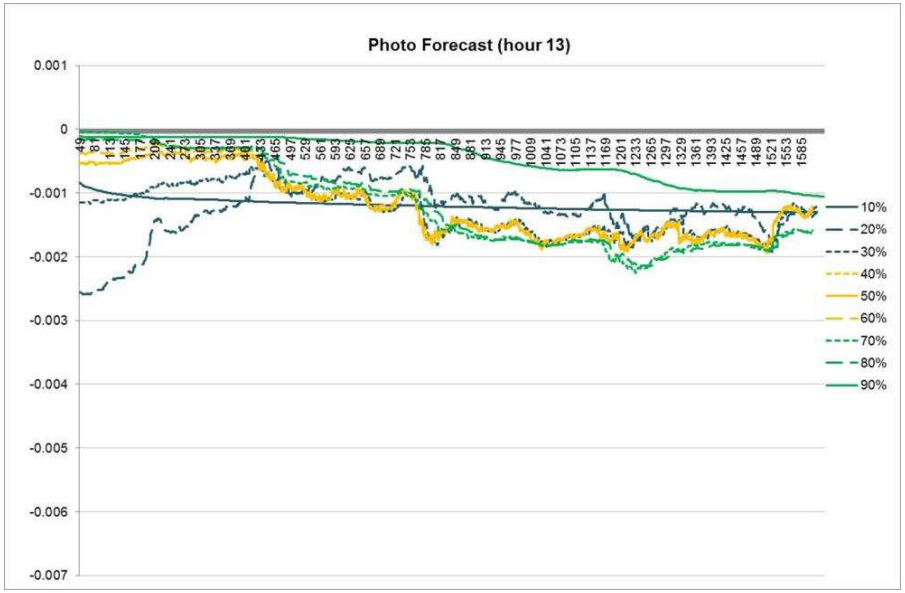
where for $j = \{1, ..., 4; 7, 8, 9\}$ ε_j, γ_j and ϕ_j take the value 0.

Results hour 13

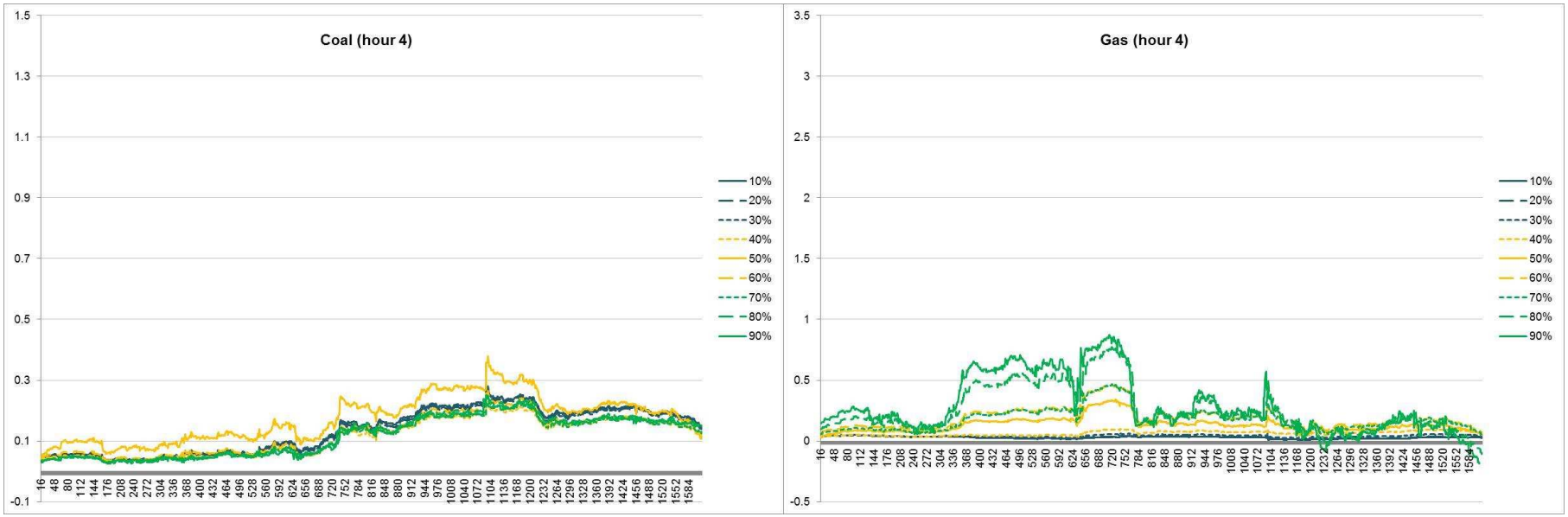


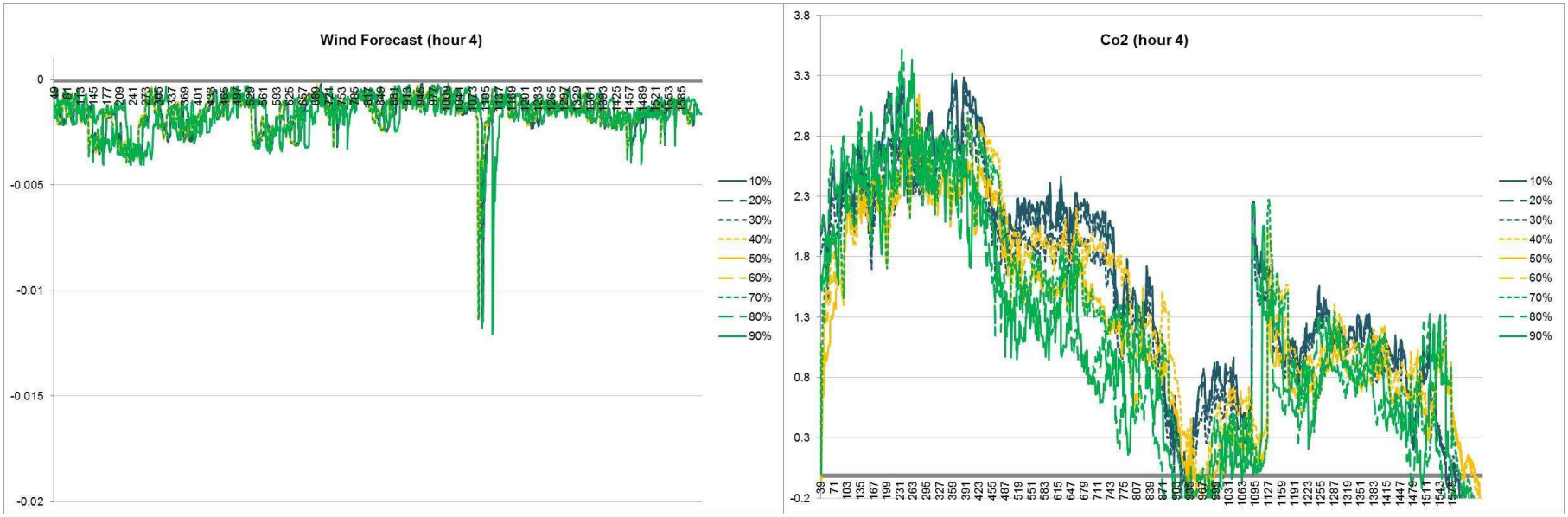




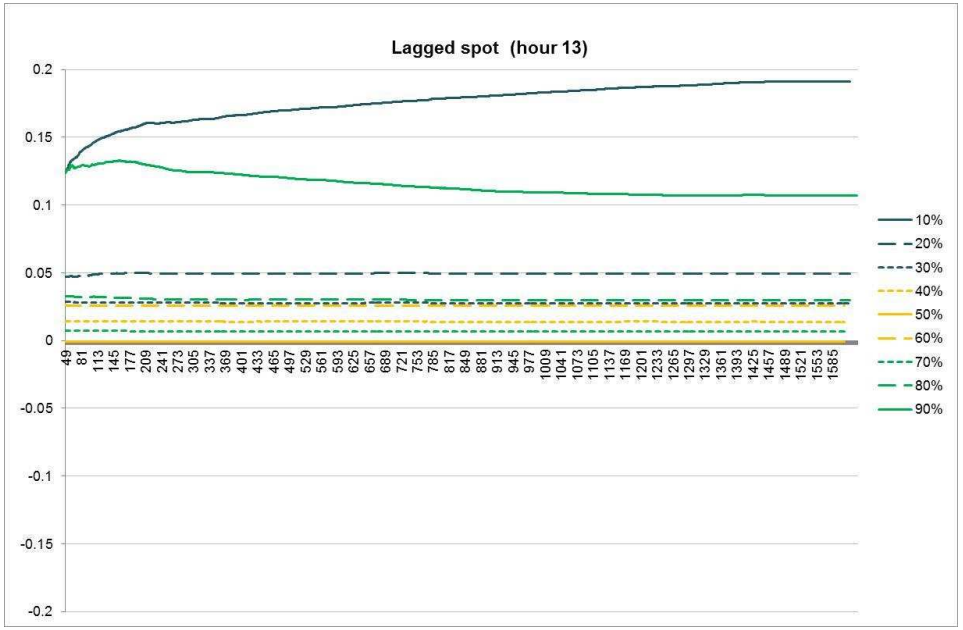


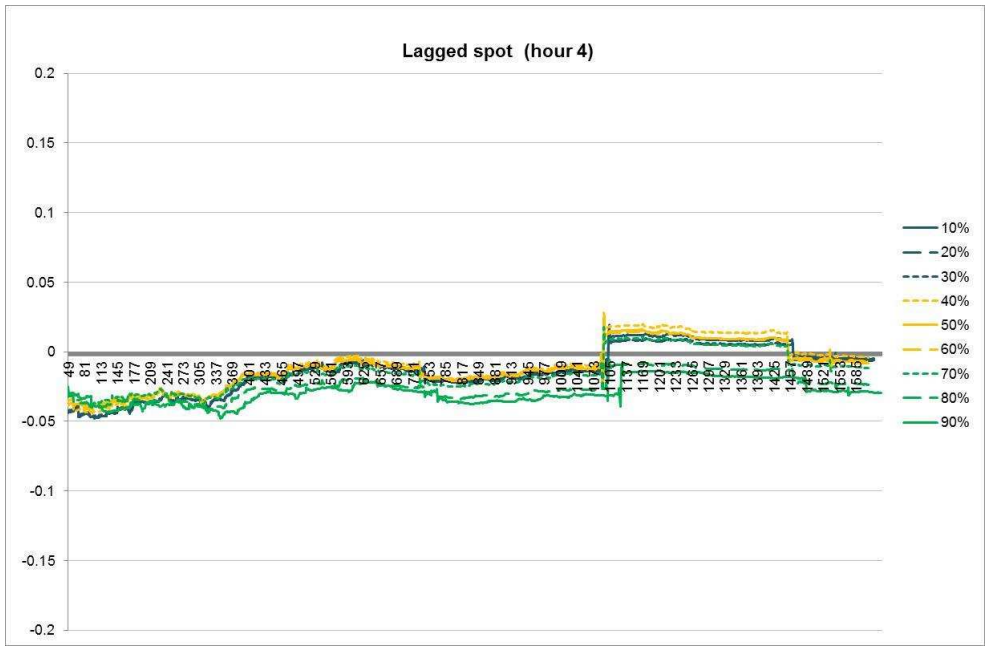
Results hour 4





Coefficients lagged spot prices hours 13 and 4





Indirect effect of renewables on electricity prices

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

Table 1: Parameter estimates for the transition equation including exogenous variables, for Hour 13, having as dependent variable the coefficient β for coal. Lower panel: test for omitted variables.

Results hour 13

- The marginal effects of gas on electricity prices generally increase with quantiles.
- 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.
- 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**.
- **Renewable energies reduce the marginal effects of traditional fuels**: coal and gas prices on the electricity prices (indirect effect).

Results hour 4

- There is a **continuous adaption of electricity prices to coal prices at all quantiles**, but highest marginal effects on central expectations: $Q50\%$: 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.

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 accross 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

- Gerlach, R., Chen, C. W. S. & Chan, N. C. Y. (2011), "Bayesian time-varying quantile forecasting for value-at-risk in financial markets", Journal of Business & Economic Statistics 29, 481–492.
- Honda, T. (2004), "Quantile regression in varying coefficient models", Journal of Statistical Planning and Inference 121, 113–125.
- Kim, M. O. (2007), "Quantile regression with varying coefficients", The Annals of Statistics 35(1), 92–108.
- Rossi, G. D. & Harvey, A. (2006), "Time-varying quantiles", Working paper University of Cambridge CWPE 0649.