

Supplementary Material for "Comparing flexible modelling approaches: the varying-thresholds model versus quantile regression"

This document is the Supplementary Material for the paper "Comparison flexible modelling approaches: the varying-thresholds model versus quantile regression". Apart from the graph of Section [S1](#), the material refers to Monte Carlo simulation results that are not included in the paper. Specifically, Section [S2](#) shows the results about 80% prediction intervals for all the considered sample sizes, namely 1000, 500 and 100, except for the two settings with full results already reported in the paper. Section [S3](#) shows the results of the simulation studies in terms of 95% prediction intervals instead of 80%, with sample size $n = 1000$. Then, Section [S4](#) compares the results of *Model 1* with the covariate having a uniform instead of a normal distribution. Finally, Section [S5](#) reports the results of an additional simulation for the case of non-monotone relationship.

S1 Graph to select the number of values for the threshold

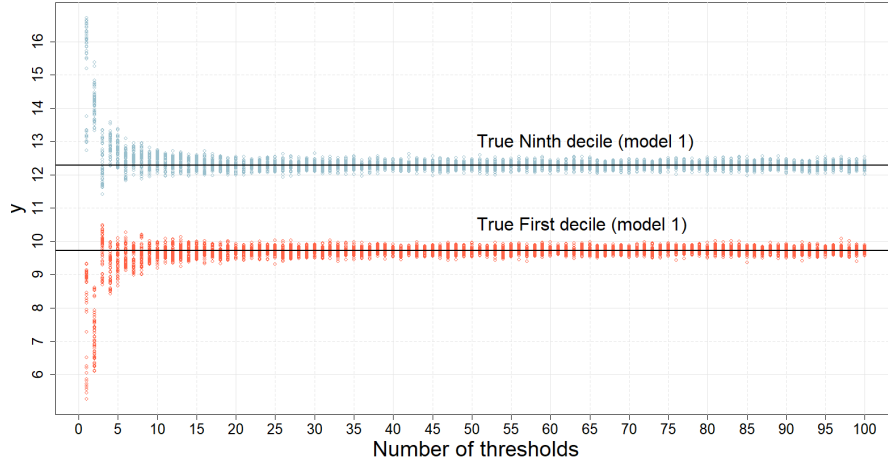


Fig. S1 Estimates of the first and ninth decile of the conditional distribution of $Y|X = 5$ as a function of the number of values for the threshold. For each threshold and for each decile (in orange and in light blue), 50 estimates are obtained by simulating from *Model 1* (linear relationship, Normal errors). The horizontal lines represent the true first and the ninth deciles. It appears that the number of values for the threshold adopted in the simulation study (52) is largely adequate.

S2 Performance of prediction intervals at 80% for sample sizes 1000, 500 and 100

Table S1 Empirical coverage, average width, and interval score of prediction intervals at level 80% for $Y|X = x$ with x at quantiles 0.1, 0.5 and 0.9, based on 1000 simulations from *Model 3* (linear relationship, heteroscedastic errors).

$X = x$ ($n=1000$)	<i>Quantile Regression</i>			<i>Varying-Thresholds Model</i>		
	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>
$q_x(0.1)$	0.967	1.214	1.245	0.999	3.116	3.117
$q_x(0.5)$	0.926	3.759	4.066	0.930	4.124	4.470
$q_x(0.9)$	0.626	6.306	13.318	0.665	7.005	12.899

$X = x$ ($n=500$)	<i>Quantile Regression</i>			<i>Varying-Thresholds Model</i>		
	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>
$q_x(0.1)$	0.969	1.209	1.241	0.998	3.111	3.112
$q_x(0.5)$	0.931	3.764	4.043	0.929	4.108	4.425
$q_x(0.9)$	0.613	6.315	14.142	0.658	6.943	13.805

$X = x$ ($n=100$)	<i>Quantile Regression</i>			<i>Varying-Thresholds Model</i>		
	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>
$q_x(0.1)$	0.923	1.198	1.316	0.965	3.496	3.559
$q_x(0.5)$	0.926	3.813	4.146	0.914	3.971	4.335
$q_x(0.9)$	0.620	6.477	14.359	0.733	10.010	15.358

Table S2 Empirical coverage, average width, and interval score of prediction intervals at level 80% for $Y|X = x$ with x at quantiles 0.1, 0.5 and 0.9, based on 1000 simulations from *Model 4* (linear relationship, heteroscedastic errors).

$X = x$ ($n=1000$)	Quantile Regression			Varying-Thresholds Model		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	0.598	6.301	14.031	0.644	6.987	13.600
$q_x(0.5)$	0.937	3.755	4.037	0.939	4.121	4.363
$q_x(0.9)$	0.972	1.216	1.243	0.996	3.124	3.129

$X = x$ ($n=500$)	Quantile Regression			Varying-Thresholds Model		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	0.594	6.337	13.893	0.658	6.979	13.545
$q_x(0.5)$	0.931	3.767	4.059	0.941	4.105	4.380
$q_x(0.9)$	0.959	1.213	1.260	0.995	3.116	3.127

$X = x$ ($n=100$)	Quantile Regression			Varying-Thresholds Model		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	0.598	6.492	14.640	0.711	9.812	15.434
$q_x(0.5)$	0.928	3.836	4.129	0.931	4.010	4.340
$q_x(0.9)$	0.928	1.206	1.312	0.965	3.451	3.542

Table S3 Empirical coverage, average width, and interval score of prediction intervals at level 80% for $Y|X = x$ with x at quantiles 0.1, 0.5 and 0.9, based on 1000 simulations from *Model 5* (quadratic relationship, Normal errors).

$X = x$ ($n=500$)	Quantile Regression			Varying-Thresholds Model		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	0.845	4.657	5.436	0.741	2.477	3.919
$q_x(0.5)$	0.704	4.651	6.582	0.865	3.306	3.956
$q_x(0.9)$	0.822	4.655	5.616	0.802	2.587	3.535

$X = x$ ($n=500$)	Quantile Regression			Varying-Thresholds Model		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	0.841	4.655	5.546	0.738	2.473	4.019
$q_x(0.5)$	0.692	4.654	6.588	0.852	3.290	4.008
$q_x(0.9)$	0.834	4.662	5.638	0.774	2.549	3.672

$X = x$ ($n=100$)	Quantile Regression			Varying-Thresholds Model		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	0.811	4.624	5.748	0.745	2.858	4.367
$q_x(0.5)$	0.697	4.621	6.645	0.837	3.091	3.880
$q_x(0.9)$	0.820	4.616	5.616	0.668	2.584	4.710

Table S4 Empirical coverage, average width, and interval score of prediction intervals at level 80% for $Y|X = x$ with x at quantiles 0.1, 0.5, and 0.9, based on 1000 simulations from *Model 6* (cubic relationship, Normal errors).

$X = x$ ($n=1000$)	Quantile Regression			Varying-Thresholds Model		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	0.991	6.651	6.677	0.890	6.712	7.283
$q_x(0.5)$	1.000	8.559	8.559	0.883	3.354	3.877
$q_x(0.9)$	1.000	10.469	10.469	0.783	2.790	3.841

$X = x$ ($n=500$)	Quantile Regression			Varying-Thresholds Model		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	0.987	6.621	6.661	0.892	6.643	7.253
$q_x(0.5)$	1.000	8.533	8.533	0.884	3.364	3.898
$q_x(0.9)$	1.000	10.460	10.460	0.719	2.755	4.288

$X = x$ ($n=100$)	Quantile Regression			Varying-Thresholds Model		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	0.986	6.476	6.536	0.712	5.901	8.538
$q_x(0.5)$	0.999	8.382	8.384	0.867	3.301	3.901
$q_x(0.9)$	0.994	10.325	10.376	0.414	2.206	7.577

Table S5 Empirical coverage, average width, and interval score of prediction intervals at level 80% for $Y|X = x$ with x at quantiles 0.1, 0.5, and 0.9, based on 1000 simulations from *Model 7* (linear relationship, Contaminated Normal errors).

$X = x$ ($n=1000$)	Quantile Regression			Varying-Thresholds Model		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	0.790	32.868	63.703	0.789	19.349	62.294
$q_x(0.5)$	0.801	33.221	60.177	0.844	37.694	59.912
$q_x(0.9)$	0.812	33.708	59.212	0.892	46.268	64.053

$X = x$ ($n=500$)	Quantile Regression			Varying-Thresholds Model		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	0.802	32.797	58.476	0.791	19.045	57.029
$q_x(0.5)$	0.801	33.226	59.219	0.840	37.541	60.258
$q_x(0.9)$	0.811	32.674	57.494	0.902	45.471	61.207

$X = x$ ($n=100$)	Quantile Regression			Varying-Thresholds Model		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	0.783	34.449	57.430	0.752	22.593	60.677
$q_x(0.5)$	0.803	33.559	57.452	0.838	36.027	57.568
$q_x(0.9)$	0.783	33.653	61.559	0.907	49.075	64.206

Table S6 Empirical coverage, average width, and interval score of prediction intervals at level 80% for $Y|X = x$ with x at quantiles 0.1, 0.5, and 0.9, based on 1000 simulations from *Model 8* (quadratic relationship, Student t errors).

$X = x$ ($n=1000$)	VTM <i>probit</i>			VTM <i>robit link $\nu = 3$</i>		
	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>
$q_x(0.1)$	0.798	3.340	5.837	0.747	2.956	5.896
$q_x(0.5)$	0.914	5.171	6.343	0.810	3.442	5.604
$q_x(0.9)$	0.870	4.730	6.107	0.793	3.375	5.627

$X = x$ ($n=500$)	VTM <i>probit</i>			VTM <i>robit link $\nu = 3$</i>		
	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>
$q_x(0.1)$	0.786	3.360	5.820	0.745	2.938	5.895
$q_x(0.5)$	0.874	4.936	6.961	0.773	3.367	6.462
$q_x(0.9)$	0.841	4.426	6.288	0.757	3.306	5.871

$X = x$ ($n=100$)	VTM <i>probit</i>			VTM <i>robit link $\nu = 3$</i>		
	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>
$q_x(0.1)$	0.818	4.583	6.823	0.801	4.272	6.801
$q_x(0.5)$	0.696	3.701	6.892	0.652	2.948	6.762
$q_x(0.9)$	0.568	3.659	9.224	0.528	3.155	9.488

Table S7 Empirical coverage, average width, and interval score of prediction intervals at level 80% for $Y|X = x$ with x at quantiles 0.1, 0.5, and 0.9, based on 1000 simulations from *Model 2* (linear relationship, Chi-squared errors). VTM with probit and cloglog link functions.

$X = x$ ($n=1000$)	VTM probit link			VTM cloglog link		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	0.693	4.675	8.483	0.730	5.110	8.359
$q_x(0.5)$	0.822	5.788	7.349	0.806	5.582	7.306
$q_x(0.9)$	0.903	6.292	8.290	0.833	5.809	8.010

$X = x$ ($n=500$)	VTM probit link			VTM cloglog link		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	0.709	4.648	7.938	0.743	5.069	7.814
$q_x(0.5)$	0.798	5.796	7.889	0.780	5.597	7.905
$q_x(0.9)$	0.893	6.291	8.798	0.823	5.797	8.558

$X = x$ ($n=100$)	VTM probit link			VTM cloglog link		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	0.799	5.300	9.026	0.839	5.779	8.827
$q_x(0.5)$	0.739	5.688	8.713	0.727	5.515	8.759
$q_x(0.9)$	0.882	7.451	9.318	0.814	6.781	9.158

S3 Performance of prediction intervals at 95% for sample size 1000

Table S8 Empirical coverage, average width, and interval score of prediction intervals at level 95% for $Y|X = x$ with x at quantiles 0.1, 0.5, and 0.9, based on 1000 simulations from *Model 1* (linear relationship, Normal errors).

$X = x$ ($n=1000$)	<i>Quantile Regression</i>			<i>Varying-Thresholds Model</i>		
	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>
$q_x(0.1)$	0.953	3.923	4.643	0.951	4.056	4.803
$q_x(0.5)$	0.940	3.912	4.954	0.941	3.923	4.914
$q_x(0.9)$	0.938	3.906	4.880	0.941	4.049	5.013

Table S9 Empirical coverage, average width, and interval score of prediction intervals at level 95% for $Y|X = x$ with x at quantiles 0.1, 0.5, and 0.9, based on 1000 simulations from *Model 2* (linear relationship, Chi-squared errors).

$X = x$ ($n=1000$)	<i>Quantile Regression</i>			<i>Varying-Thresholds Model</i>		
	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>
$q_x(0.1)$	0.947	9.125	12.466	0.954	8.698	12.731
$q_x(0.5)$	0.968	9.129	9.870	0.988	9.646	10.364
$q_x(0.9)$	0.943	9.162	11.501	0.990	12.207	13.139

Table S10 Empirical coverage, average width, and interval score of prediction intervals at level 95% for $Y|X = x$ with x at quantiles 0.1, 0.5, and 0.9, based on 1000 simulations from *Model 3* (linear relationship, heteroscedastic errors).

$X = x$ ($n=1000$)	<i>Quantile Regression</i>			<i>Varying-Thresholds Model</i>		
	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>
$q_x(0.1)$	1.000	3.569	3.569	1.000	6.670	6.670
$q_x(0.5)$	1.000	7.959	7.959	0.998	7.392	7.412
$q_x(0.9)$	0.923	12.303	16.311	0.899	18.487	24.403

Table S11 Empirical coverage, average width, and interval score of prediction intervals at level 95% for $Y|X = x$ with x at quantiles 0.1, 0.5, and 0.9, based on 1000 simulations from *Model 4* (linear relationship, heteroscedastic errors).

$X = x$ ($n=1000$)	<i>Quantile Regression</i>			<i>Varying-Thresholds Model</i>		
	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>
$q_x(0.1)$	0.908	12.271	17.757	0.902	18.480	24.914
$q_x(0.5)$	1.000	7.928	7.928	0.999	7.400	7.429
$q_x(0.9)$	1.000	3.563	3.563	1.000	6.822	6.822

Table S12 Empirical coverage, average width, and interval score of prediction intervals at level 95% for $Y|X = x$ with x at quantiles 0.1, 0.5, and 0.9, based on 1000 simulations from *Model 5* (quadratic relationship, Normal errors).

$X = x$ ($n=1000$)	<i>Quantile Regression</i>			<i>Varying-Thresholds Model</i>		
	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>
$q_x(0.1)$	0.984	6.247	6.518	0.886	3.367	5.385
$q_x(0.5)$	0.827	6.253	9.903	0.947	4.190	5.090
$q_x(0.9)$	0.979	6.263	6.536	0.928	3.807	4.777

Table S13 Empirical coverage, average width, and interval score of prediction intervals at level 95% for $Y|X = x$ with x at quantiles 0.1, 0.5, and 0.9, based on 1000 simulations from *Model 6* (cubic relationship, Normal errors).

$X = x$ ($n=1000$)	<i>Quantile Regression</i>			<i>Varying-Thresholds Model</i>		
	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>
$q_x(0.1)$	0.999	8.331	8.338	0.933	7.740	8.926
$q_x(0.5)$	1.000	10.777	10.777	0.949	4.211	5.042
$q_x(0.9)$	1.000	13.213	13.213	0.948	10.997	11.724

Table S14 Empirical coverage, average width, and interval score of prediction intervals at level 95% for $Y|X = x$ with x at quantiles 0.1, 0.5, and 0.9, based on 1000 simulations from *Model 7* (linear relationship, Contaminated Normal errors).

$X = x$ ($n=1000$)	<i>Quantile Regression</i>			<i>Varying-Thresholds Model</i>		
	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>	<i>Coverage</i>	<i>Avg. Width</i>	<i>Interval Score</i>
$q_x(0.1)$	0.914	61.248	69.888	0.920	60.457	93.511
$q_x(0.5)$	0.930	61.281	68.867	0.952	63.802	69.163
$q_x(0.9)$	0.928	61.310	68.386	0.986	73.030	74.497

Table S15 Empirical coverage, average width, and interval score of prediction intervals at level 95% for $Y|X = x$ with x at quantiles 0.1, 0.5, and 0.9, based on 1000 simulations from *Model 8* (quadratic relationship, Student t errors).

$X = x$ ($n=1000$)	<i>VTM probit</i>			<i>VTM</i> <i>robit link</i> $\nu = 3$		
	<i>Coverage</i>	<i>Avg.</i> <i>Width</i>	<i>Interval</i> <i>Score</i>	<i>Coverage</i>	<i>Avg.</i> <i>Width</i>	<i>Interval</i> <i>Score</i>
$q_x(0.1)$	0.939	7.250	10.766	0.917	5.999	10.273
$q_x(0.5)$	0.958	7.232	9.653	0.932	5.678	9.673
$q_x(0.9)$	0.939	7.064	9.776	0.924	5.586	9.361

Table S16 Empirical coverage, average width, and interval score of prediction intervals at level 95% for $Y|X = x$ with x at quantiles 0.1, 0.5, and 0.9, based on 1000 simulations from *Model 2* (linear relationship, Chi-squared errors).

$X = x$ ($n=1000$)	<i>VTM probit link</i>			<i>VTM cloglog link</i>		
	<i>Coverage</i>	<i>Avg.</i> <i>Width</i>	<i>Interval</i> <i>Score</i>	<i>Coverage</i>	<i>Avg.</i> <i>Width</i>	<i>Interval</i> <i>Score</i>
$q_x(0.1)$	0.954	8.698	12.731	0.949	9.192	12.585
$q_x(0.5)$	0.988	9.646	10.364	0.982	9.325	10.062
$q_x(0.9)$	0.990	12.207	13.139	0.989	11.395	12.408

S4 Performance of prediction intervals at 80% for sample sizes 1000, 500 and 100 from *Model 1* with the covariate having a uniform instead of a Normal distribution

Table S17 Empirical coverage, average width, and interval score of prediction intervals at level 80% for $Y|X = x$ with x at quantiles 0.1, 0.5, and 0.9, based on 1000 simulations from *Model 1* (linear relationship, Normal errors) with normal and uniform distributed covariate.

$X = x$ ($n=1000$)	VTM - uniform covariate			VTM - Normal covariate		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	0.842	2.860	3.553	0.798	2.552	3.467
$q_x(0.5)$	0.814	2.721	3.529	0.783	2.567	3.643
$q_x(0.9)$	0.775	2.433	3.532	0.775	2.545	3.634

$X = x$ ($n=500$)	VTM - uniform covariate			VTM - Normal covariate		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	0.844	2.811	3.557	0.776	2.524	3.593
$q_x(0.5)$	0.805	2.674	3.604	0.801	2.551	3.430
$q_x(0.9)$	0.750	2.340	3.642	0.785	2.523	3.569

$X = x$ ($n=100$)	VTM - uniform covariate			VTM - Normal covariate		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	0.817	2.990	3.875	0.812	2.983	4.098
$q_x(0.5)$	0.733	2.395	3.761	0.773	2.476	3.664
$q_x(0.9)$	0.789	2.746	3.874	0.788	2.964	3.978

S5 Non-monotone relationship

In this Section S5 we show the results for a DGM with a non-monotone relationship, specifically a quadratic function whose minimum lies in the range of X :

$$\begin{aligned} \text{Non-monotone Model: } Y &= 0.1 + 0.1X + 0.2X^2 + \varepsilon_N \\ \text{with } \varepsilon_N &\sim N(0, 1) \text{ and } X \sim U(-10, 10) \end{aligned} \quad (1)$$

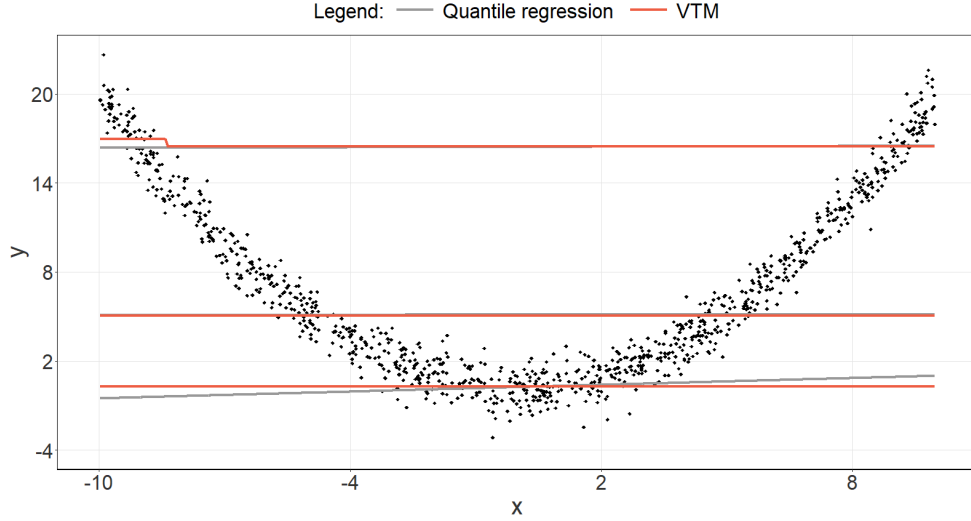


Fig. S2 Scatterplot of 1000 observations randomly drawn from *Non-monotone Model* (non-monotone relationship, Normal errors), with both regression methods using a linear functional form for estimating conditional quantiles, Linear Quantile Regression (grey lines) and Varying-Thresholds Model (red curves) at quantiles 0.1, 0.5 and 0.9.

Table S18 Empirical coverage, average width, and interval score of prediction intervals at level 80% for $Y|X = x$ with x at quantiles 0.1, 0.5, and 0.9, based on 1000 simulations from *Non-monotone Model* (non-monotone relationship, Normal errors), with both regression methods using a linear functional form for estimating conditional quantiles.

$X = x$ ($n=1000$)	Quantile Regression			Varying-Thresholds Model		
	Coverage	Avg. Width	Interval Score	Coverage	Avg. Width	Interval Score
$q_x(0.1)$	1.000	16.094	16.094	1.000	16.127	16.127
$q_x(0.5)$	0.483	16.114	20.500	0.477	16.109	20.491
$q_x(0.9)$	0.999	16.133	16.134	0.996	16.089	16.098

It is possible to change the functional form in both methods; using $\eta(\theta, x) = \beta_0(\theta) + \beta_1(\theta)x + \beta_2(\theta)x^2$ for the VTM to obtain a quadratic functional form for the conditional quantiles estimation, as shown in Figure S3. While, for a comparison, a scatterplot of 1000 observations randomly drawn from the *Non-monotone Model* (1) is shown in Figure S2, with both regression methods using a linear functional form for estimating the conditional quantiles.

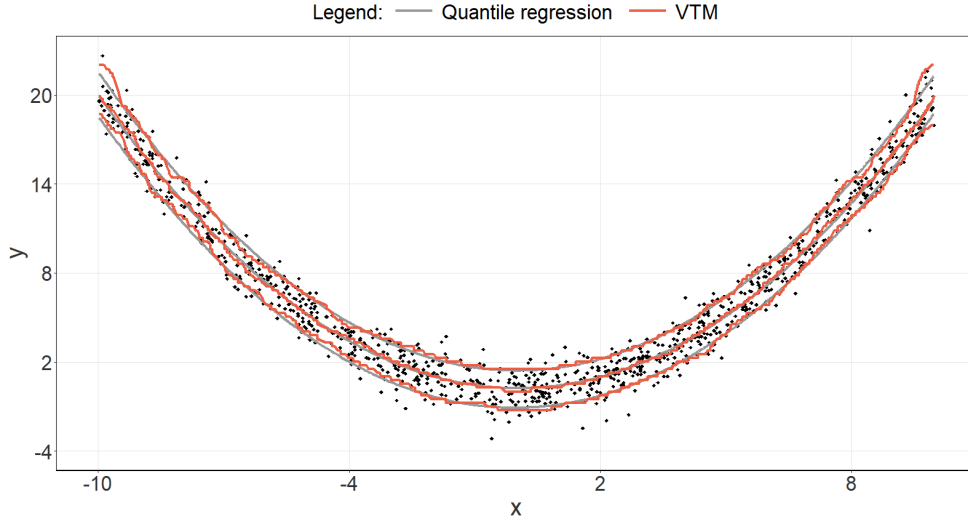


Fig. S3 Scatterplot of 1000 observations randomly drawn from *Non-monotone Model* (non-monotone relationship, Normal errors), with both regression methods using a quadratic functional form for estimating conditional quantiles, Linear Quantile Regression (grey lines) and Varying-Thresholds Model (red curves) at quantiles 0.1, 0.5 and 0.9.