Universidade Federal do Rio Grande do Sul
Instituto de Matemática e Estatística
Programa de Pós-Graduação em Estatística

Regressão quantílica suavizada: Uma aplicação a séries temporais

Miguel Jandrey Natal

Porto Alegre, Setembro de 2021.

Dissertação submetida por Miguel Jandrey Natal[1] como requisito parcial para a obtenção do título de Mestre em Estatística pelo Programa de Pós-Graduação em Estatística da Universidade Federal do Rio Grande do Sul.

Orientador:

Prof. Dr. Eduardo de Oliveira Horta (UFRGS)

Comissão Examinadora:

Profa. Dra. Fernanda Maria Müller (UFRGS)

Prof. Dr. Flávio Augusto Ziegelmann (UFRGS)

Prof. Dr. Guilherme Pumi (UFRGS)

Data de Apresentação: 30 de Setembro de 2021

[1]Bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES

Agradecimentos

Agradeço:

- primeiramente à minha mãe, Circe, e ao meu irmão, Pedro, por serem o meu alicerce nos momentos mais difíceis e a inspiração para os momentos de conquista;
- ao meu pai, Rafael, ao meu avô, Honorival e ao meu padrasto, Cícero, pelo companheirismo e força paternal;
- ao meu orientador, Eduardo, pelo enorme auxílio, pela paciência e por ter sido verdadeiramente um mentor na minha trajetória acadêmica ao longo dos últimos dois anos;
- aos meus colegas de PPGEst, por toda ajuda, troca de ideias e companheirismo ao longo do Mestrado - em algum nível, espero ter conseguido ajudá-los como me senti ajudado;
- aos demais professores e funcionários do PPGEst, pelo empenho e dedicação para tornar o Programa um sucesso;
- aos meus líderes e colegas na GAVB, com quem tanto aprendo diariamente, das mais distintas formas;
- aos demais amigos e parentes fora da academia, me abstendo de citar nomes, para não correr o risco de deixar alguém de fora;
- ao Felipe e à Carmen, por terem me ensinado tanto e serem parte dessa conquista;
- à banca, pela disponibilidade;
- à CAPES, pela bolsa que financiou meus estudos e me possibilitou adquirir o conhecimento necessário para escrever esta dissertação;
- ao Artur, à Helen e à Débora pela cessão do modelo de dissertação.

CIP - Catalogação na Publicação

Natal, Miguel Jandrey
Regressão quantílica suavizada: Uma aplicação a séries temporais / Miguel Jandrey Natal. -- 2021.
67 f.
Orientador: Eduardo de Oliveira Horta.

Dissertação (Mestrado) -- Universidade Federal do Rio Grande do Sul, Instituto de Matemática e Estatística, Programa de Pós-Graduação em Estatística, Porto Alegre, BR-RS, 2021.

1. Regressão quantílica. 2. Quantil condicional. 3. Largura de banda baseada nos dados. 4. Simulação de Monte Carlo. 5. Séries temporais. I. Horta, Eduardo de Oliveira, orient. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

Resumo

A regressão quantílica modela quantis condicionais da variável resposta e traz o conceito de quantil para a estrutura de modelos lineares generalizados. Embora a regressão quantílica – tal como a conhecemos hoje – tenha sido introduzida há mais de quarenta anos, apenas recentemente tornou-se praticável para grandes volumetrias de dados, devido aos avanços computacionais. Como a função objetivo que o estimador canônico para os coeficientes em modelos de regressão de quantílica visa minimizar não é suave, a inferência estatística não é direta. Recentemente, uma nova proposta de estimação para tais coeficientes foi incorporada à literatura de regressão quantílica: o estimador suavizado para regressão quantílica do tipo de convolução via *kernel*. Com base nesta abordagem alternativa para a modelagem de regressão de quantílica, este trabalho visa implementar este estimador suavizado em um contexto de séries temporais. Uma vez que a teoria do estimador foi inicialmente formalizada considerando estruturas de dados *cross-section*, o objetivo aqui é tentar uma nova etapa, expandindo o seu estudo em uma estrutura de séries temporais. Além disso, investigamos caracterizações analíticas para o processo de geração de dados das dinâmicas *Quantile Autoregressive Distributed Lag* (QADL). Por meio de simulações de Monte Carlo, exploramos essas dinâmicas ao passo que avaliamos o desempenho do estimador suavizado em uma classe de modelos de regressão quantílica para séries temporais.
Palavras-chave: regressão quantílica, *quantil condicinal*, *QADL*, *séries temporais*, *largura de banda baseada nos dados*, *simulação de Monte Carlo*

Abstract

Quantile regression fits quantiles of the response variable and brings the concept of a quantile into the framework of general linear models. Although quantile regression was first introduced more than forty years ago, only recently it became practicable for large data due to computational advances. As the objective function that the standard quantile regression estimator aims to minimize is not smooth, statistical inference is not straightforward. Recently, a new estimation proposal for such coefficients was incorporated into the quantile regression literature: the convolution-type kernel smoothed quantile regression estimator. Based on this alternative approach to quantile regression modeling, this work aims to implement this smoothed estimator in a time series context. Since the estimation theory was formalized considering cross-sectional data, the goal here is to try a new step by expanding the study into a time series framework. Also, we investigate analytical characterizations for the data generating process of Quantile Autoregressive Distributed Lag (QADL) dynamics. Through Monte Carlo simulations, we explore these dynamics while evaluating the performance of the smoothed estimator in a class of time series quantile regression models.

Keywords: quantile regression, *conditional quantile*, *convolution-based smoothing*, *QADL*, time series, *data-driven bandwidth*, *Monte Carlo simulation*

Índice

Capítulo 1

Introdução

Muitos cenários nas ciências aplicadas revelam, ou pelo menos lançam alguma luz sobre as relações existentes entre variáveis distintas. Em tais cenários, pesquisadores podem necessitar de ferramentas específicas para completar os seus trabalhos – nomeadamente, a obtenção de uma boa descrição da estrutura de dependência entre estas variáveis de interesse.

Por um lado, o estimador de mínimos quadrados ordinários (MQO) para os coeficientes do modelo de regressão linear clássico (regressão na média) representa um dos fundamentos da teoria estatística. Para além do fato de a regressão na média ser um modelo parcimonioso e facilmente interpretável, o estimador de MQO goza de certa otimização em alguns casos, de modo que a combinação destas ferramentas - isto é, a estimação dos coeficientes de um modelo de regressão linear através do estimador de MQO - é muitas vezes um caminho desejável quando se procede à análise estatística.

Por outro lado, a modelagem de uma variável aleatória através da regressão linear clássica traz à tona algumas implicações restritivas. Por exemplo, imaginemos um modelo simples de regressão na média onde a variável dependente, digamos Y, tem a sua estrutura de dependência descrita por $\mathbb{E}[Y|X = x]$. A última expressão - que denota o valor esperado de Y dado $X = x$, ou seja, a sua média condicional - pode ser um pouco restritiva porque especifica apenas como a média de Y responde a variações em x. Nesse caso, não é uma surpresa que tais modelos sejam inadequados quando a média (condicional) não é um fim satisfatório por si mesma. Mosteller e Tukey (2005) vão além desta ideia, observando que, apesar de o modelo clássico de regressão linear fornecer um resumo razoável para as médias das distribuições correspondentes ao conjunto de x's, poderíamos ir além, incorporando quantis à análise, com o objetivo de obter uma imagem mais completa da variável resposta.

A teoria de regressão quantílica surge em um contexto no qual são necessárias ferramentas adequadas para realizar uma análise de dependência entre uma variável resposta e demais covariáveis. Embora os modelos clássicos de regressão linear sirvam bem em muitos problemas nas ciências aplicadas, não raramente ter um quadro completo da distribuição condicional da variável resposta pode incorporar informações valiosas à análise. Ao trazer o conceito de um quantil para a estrutura dos modelos lineares generalizados, o método de regressão quantílica expandiu a maneira de visualizar as relações transversais entre os dados. Enquanto a regressão linear clássica com covariáveis estritamente exógenas modela a média condicional de uma variável dependente, a regressão quantílica modela seus quantis condicionais, de modo que possamos recuperar sua distribuição condicional inteiramente, não apenas a média condicional. Nesse sentido, vale dizer que a regressão quantílica não deve ser interpretada como um modelo "melhor" quando comparada à regressão na média, mas sim como uma abordagem alternativa para modelar a relação entre uma variável dependente e demais covariáveis – particularmente quando informações importantes encontram-se na cauda da distribuição da variável resposta.

Desde o artigo seminal de Koenker e Bassett (1978), o qual introduziu a regressão de quantílica em sua forma moderna,[1] uma atenção considerável foi devotada ao estudo e desenvolvimento de modelos associados. Apesar de ter sido proposta há mais de quarenta anos, apenas recentemente a regressão quantílica tornou-se praticável para grandes volumetrias de dados, devido aos recentes avanços computacionais. Esse fato certamente ajudou a impulsionar sua popularização. A literatura agora tem uma ampla gama de aplicações nas quais a regressão quantílica tem sido útil. Pode-se citar várias áreas onde a regressão quantílica é contemplada, como economia Galvão et al. (2013), finanças (Engle e Manganelli (2004); Xu et al. (2016)), *machine learning* (Meinshausen (2006); Takeuchi et al. (2005)), medicina Hong et al. (2019), entre outras. Técnicas como *bootstrapping* Horowitz (1998) e métodos *Markov Chain Monte Carlo* (MCMC) Chernozhukov e Hong (2003) também são exemplos de tais aplicações dentro da própria Estatística. Na verdade, essa técnica de regressão chamou a atenção não apenas da comunidade de Estatística-Matemática, mas também daqueles cujas preocupações eram especificamente as aplicações do método.

Como a função objetivo (amostral) que o estimador canônico de regressão quantílica visa minimizar não é suave, a inferência estatística não é direta. Fernandes et al. (2021) propõe suavizar esta função objetivo, apresentando assim um estimador alternativo: o estimador suavizado de regressão quantílica do tipo convolução via *kernel*. Com base nesta abordagem alternativa para estimar o parâmetro funcional no modelo de regressão de quantílica, este trabalho visa implementar o estimador suavizado de regressão quantílica em um contexto de séries temporais. Uma vez que os autores formalizaram a teoria do estimador suavizado considerando dados em uma estrutura *cross-section*, o objetivo aqui é tentar uma nova etapa, expandindo o seu estudo em uma classe de modelos de séries temporais. Por meio de simulações de Monte Carlo, avaliamos o desempenho do estimador em cenários novos e existentes. Considerando que o primeiro conjunto de simulações foi replicado a partir dos cenários encontrados em Galvão et al. (2013), é importante dizer que eles foram adaptados para atender aos nossos propósitos, principalmente por incluir na análise a avaliação em mais níveis de quantílicos. Quanto aos novos cenários, eles são pequenas adaptações de algumas novas contribuições na literatura de regressão quantílica fornecidas por Horta (2021) – realizamos um estudo de Monte Carlo para esses cenários, que foram construídos com base em suas idéias. A partir desses novos cenários, pretendemos oferecer novas contribuições sobre dinâmicas *Quantile Autoregressive Distributed Lag* (QADL), notadamente investigando caracterizações analíticas para os processos de

[1] Como nos referimos atualmente, a regressão quantílica foi oficialmente introduzida por Koenker e Bassett (1978). No entanto, Koenker et al. (2018) salientam que sua verdadeira origem pode ser traçada a partir de Boscovich e Laplace através da "regressão na mediana", no século XVIII.

geração de dados dessas dinâmicas. Além disso, avaliamos o desempenho do estimador suavizado de Fernandes et al. (2021) em um contexto de séries temporais, particularmente na classe de modelos QADL de Galvão et al. (2013), e o comparamos com os resultados da estimação obtidos através do estimador canônico de Koenker e Bassett (1978).

Objetivo

Nesta dissertação, temos dois objetivos principais: primeiro, investigar caracterizações analíticas para o processo de geração de dados de dinâmicas QADL; segundo, conduzir um estudo de Monte Carlo massivo explorando essas dinâmicas – e, concomitantemente, avaliar o desempenho do estimador suavizado de Fernandes et al. (2021) em um contexto de série temporais, particularmente na classe de modelos QADL de Galvão et al. (2013).

Novidades do artigo

A relevância do estudo dá-se, primeiramente, pela implementação do estimador suavizado de regressão quantílica em um contexto de séries temporais – algo que não havia sido feito até então. Ainda, entendemos que as propostas de caracterizações analíticas para o processo de geração de dados de dinâmicas QADL representam uma novidade dentro da literatura de regressão quantílica. Ademais, nossos resultados indicam que o estimador suavizado de Fernandes et al. (2021) tem um bom desempenho em um contexto de séries temporais – e que pode ser considerado uma boa alternativa ao se estimar coeficientes em modelos de regressão quantílica de séries temporais.

Suporte computacional

Realizamos estudos de Monte Carlo para avaliar o estimador de Fernandes et al. (2021) – em um contexto de séries temporais – em termos de erro quadrático médio, viés e variância. Nestes estudos, foram considerados cenários novos e cenários já existentes (levemente adaptados). Toda a parte computacional desta dissertação foi realizada utilizando o software `R`, versão 4.0.3 (R Core Team, 2021). *Scripts* e tabelas das simulações realizadas podem ser encontrados em `https://github.com/migueljnatal/SQR-with-time-series-data`.

Compõem esta dissertação: uma pequena revisão bibliográfica, onde introduzimos conceitos básicos de regressão quantílica em séries temporais, bem como os modelos e estimadores estudados neste trabalho; a conclusão acerca do trabalho desenvolvido e alguns possíveis caminhos futuros; por último e mais importante, o artigo anexo.

Capítulo 2

Revisão Bibliográfica

Este capítulo é dedicado à especificação teórica de modelos de regressão quantílica em um contexto de séries temporais. Primeiramente, apresentamos um conjunto de aspectos teóricos, que incluem definições básicas, notação matemática e suposições. Em seguida, descrevemos o modelo de regressão quantílica e suas equações análogas considerando um contexto de séries temporais, além algumas considerações adicionais sobre a dinâmica de QADL, onde fornecemos algumas novas contribuições no que diz respeito aos seus aspectos teóricos. Por fim, apresentamos o procedimento de estimação canônico para regressão regressão quantílica proposto por Koenker e Bassett (1978), bem como a abordagem alternativa de Fernandes et al. (2021).

Na configuração mais básica, para uma variável escalar aleatória Y com função de distribuição cumulativa F_Y, definimos o τ**-ésimo quantil de** Y, com $0 < \tau < 1$, como sendo o número real $Q_Y(\tau) := \inf\{y \in \mathbb{R} \colon F_Y(y) \geqslant \tau\}$. O mapa $\tau \mapsto Q_Y(\tau)$ é chamado de função de quantilíca **de** Y ou o **inversa generalizada de** F_Y. A importância das funções quantílicas é a de que elas caracterizam inteiramente a distribuição de Y, já que F_Y é recuperável a partir de Q_Y. Na presença de um vetor de covariáveis exógenas X, de dimensão $\mathbb{R}^{\dim_X}$, definimos de maneira similar o τ**-ésimo quantil condicional de** Y **dado** $X = x$ como sendo o número real $Q_{Y|X}(\tau|x)$ definido por $x \in \operatorname{support}(X)$ e $\tau \in (0, 1)$, através de

$$Q_{Y|X}(\tau|x) = \inf\{y \in \mathbb{R} \colon F_{Y|X}(y|x) \geqslant \tau\}, \tag{2.1}$$

onde $F_{Y|X}(\cdot|x)$ é a função de distribuição acumulada condicional de Y dado $X = x$. O mapa $\tau \mapsto Q_{Y|X}(\tau|x)$ é chamado de **função quantílica condicional de** Y **dado** $X = x$. Em uma estrutura de séries temporais, por sua vez, alguns ajustes na notação e nas definições são necessários, conforme explicamos a seguir.

Regressão quantílica em séries temporais

No que segue, $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\in\mathbb{Z}}, \mathbf{P})$ denota um espaço de probabilidade filtrado e $(X_t, Y_t)_{t\in\mathbb{Z}}$ é um processo estocástico com espaço de estado $\mathbb{R} \times \mathbb{R}^{\dim_X}$, o qual assumimos estar adaptado a $\{\mathscr{F}_t\}_{t\in\mathbb{Z}}$. Para fins de identificação, também assumimos no restante do texto que, para todo $t \in \mathbb{Z}$, os componentes de X_t são variáveis aleatórias não degeneradas e linearmente independentes, bem como linearmente independentes de $\{Y_s\colon \leqslant t\}$. Para cada $t \in \mathbb{Z}$ e $0 < \tau < 1$, o **τ-ésimo quantil condicional de Y_t dado $\mathscr{F}_{t-1}$** é a variável aleatória $\mathscr{F}_{t-1}$-mensurável definida através de

$$Q_{Y_t}(\tau \,|\, \mathscr{F}_{t-1}) := \inf \big\{ y \in \mathbb{R}\colon\ \mathbf{P}(Y_t \leqslant y \,|\, \mathscr{F}_{t-1}) \geqslant \tau \big\}. \tag{2.2}$$

Assumimos, sem perda de generalidade, que uma versão regular de $y \mapsto \mathbf{P}(Y_t \leqslant y \,|\, \mathscr{F}_{t-1})$ está disponível. Neste caso, o processo estocástico $\{Q_{Y_t}(\tau \,|\, \mathscr{F}_{t-1})\colon\ 0 < \tau < 1\}$, que chamamos de **função quantílica condicional de Y_t dado $\mathscr{F}_{t-1}$**, é $\omega - wise$ a inversa generalizada da função de distribuição acumulada $\mathbf{P}(Y_t \leqslant \cdot \,|\, \mathscr{F}_{t-1})_\omega$.

Apresentamos agora o modelo *Quantile Autoregressive Distributed Lag* (QADL), adaptando ligeiramente a definição apresentada por Galvão et al. (2013).

Diz-se que um processo $(X_t, Y_t)_{t\in\mathbb{Z}}$ satisfaz um **Quantile Autoregressive Distributed Lag Model** de ordens lag_Y e lag_X, que escrevemos QADL $(\mathrm{lag}_Y, \mathrm{lag}_X)$ por simplicidade, se a relação

$$Q_{Y_t}(\tau|\mathscr{F}_{t-1}) = \alpha_0(\tau) + \sum_{j=1}^{\mathrm{lag}_Y} \alpha_j(\tau) Y_{t-j} + \sum_{\ell=1}^{\mathrm{lag}_X} X'_{t-\ell}\theta_\ell(\tau), \tag{2.3}$$

vale para todos $t \in \mathbb{Z}$ e todos $0 < \tau < 1$, para algumas funções $\alpha_j\colon (0,1) \to \mathbb{R}$ e $\theta_\ell\colon (0,1) \to \mathbb{R}^{\dim_X}$ (com $0 \leqslant j \leqslant \mathrm{lag}_Y$ and $1 \leqslant \ell \leqslant \mathrm{lag}_X$).

Dinâmica QADL: considerações adicionais

Nesta seção, exploramos o *design* do modelo QADL já estudado por Galvão et al. (2013), usando sua arquitetura bem definida como um modelo para tentar ir um pouco além no que concerne à dinâmica QADL. A especificação linear implica um forte vínculo na interação entre a forma funcional dos coeficientes de regressão, por um lado, e o suporte das covariáveis, por outro; não se pode ter funções de coeficientes flexíveis *e* um suporte de covariáveis não especificado/genérico. Mas não é surpresa que todo ganho tem um custo, e, neste caso, isso significa que é necessário impor restrições ao suporte conjunto de (X_t, Y_t) para desfrutar de toda a flexibilidade da regressão de quantílica – a flexibilidade que decorre precisamente de funções cujos coeficientes não são constantes em τ.

Agora podemos apresentar um dos cenários principais que serão contemplados em nossos estudos de Monte Carlo da próxima seção. O resultado a seguir é uma ligeira adaptação da Proposição 4.1 em Horta (2021), que estabelece condições suficientes que permitem especificar a função de quantil condicional de Y_t dado $\mathscr{F}_{t-1}$ por meio de um *família de hiperplanos afins determinados por "funções vertex"*. Para simplificar, nos restringimos à configuração com $\mathrm{lag}_Y = \mathrm{lag}_X = \dim_X = 1$. Uma generalização multidimensional pode ser obtida como discutido na observação seguinte à Proposição 4.1 em Horta (2021).

Proposição 1. Sejam $v_{ij}\colon (0,1) \to [0,1]$, $i,j \in \{0,1\}$ funções não-decrescentes e contínuas à esquerda com $v_{11} + v_{00} = v_{10} + v_{01}$. Adicionalmente, sejam Y_0 and X_0 variáveis aleatórias com suporte no intervalo unitário, e sejam $U_1, U_2, \ldots$ and $V_1, V_2, \ldots$

iid sequências de variáveis aleatórias no intervalo unitiário, com ambas as sequências mutuamente independentes e também independentes de Y_0 and X_0. Além disso, defina, para $t \geqslant 1$,

$$\begin{aligned} Y_t &\coloneqq \alpha_0(U_t) + \alpha_1(U_t)Y_{t-1} + \theta_1(U_t)X_{t-1} \\ X_t &\coloneqq Q_t(V_t, X_{t-1}, \ldots, X_0, Y_t, Y_{t-1}, \ldots, Y_0), \end{aligned} \tag{2.4}$$

onde, para todo $t \geqslant 1$, a função mensurável $Q_t \colon (0,1) \times [0,1]^{2t+1} \to [0,1]$ é uma função quantílica no seu primeiro argumento, e onde $\alpha_0 \coloneqq v_{00}$, $\alpha_1 \coloneqq v_{10} - v_{00}$ e $\theta_1 \coloneqq v_{01} - v_{00}$. Então o processo $(X_t, Y_t)_{t \in \mathbb{Z}}$ tem espaço de estados contido no quadrado unitário $[0,1]^2$ e satisfaz

$$Q_{Y_t}(\tau|\mathscr{F}_{t-1}) = \alpha_0(\tau) + \alpha_1(\tau)Y_{t-1} + \theta_1(\tau)X_{t-1} \tag{2.5}$$

para todo $0 < \tau < 1$ e todo $t \geqslant 1$, onde $\mathscr{F}_t \coloneqq \sigma\{(X_s, Y_s) \colon\ s \leqslant t\}$.

Condições de compatibilidade de vértices

Esta seção é importante porque fornece um modelo para nosso estudo de simulação, com base nas idéias apresentadas na Proposição 1. Suponha que já tenhamos escolhido as funções v_{10} e v_{01} – por exemplo, ambas podem ser funções quantílicas da distribuição Beta, possivelmente com parâmetros distintos. Nosso objetivo é selecionar uma função $v_{00} \colon (0,1) \to \mathbb{R}$ de modo que o seguinte seja válido:

1. v_{00} é não-decrescente e contínua à esquerda, com $\mathrm{range}(v_{00}) \subseteq [0,1]$;
2. $v_{11} \coloneqq v_{10} + v_{01} - v_{00}$ é não-decrescente e contínua à esquerda, e mantém $\mathrm{range}(v_{11}) \subseteq [0,1]$.

É importante ressaltar que se nos restringirmos ao cenário onde v_{10}, v_{01} e v_{00} são diferenciáveis, então a monotonicidade de v_{11} é equivalente ao requisito de $v'_{10} + v'_{01} \geqslant v'_{00}$. O próximo resultado, que é uma contribuição original, oferece condições suficientes para atingir esse objetivo. Por conveniência de notação, escrevemos $v(\tau-) \coloneqq \lim_{u \uparrow \tau} v(u)$ e, similarmente, $v(\tau+) \coloneqq \lim_{u \downarrow \tau} v(u)$.

Proposição 2. Sejam $v_{10}, v_{01} \colon (0,1) \to \mathbb{R}$ quaisquer duas funções não-negativas e continuamente diferenciáveis com $v_{10}(0+) = v_{01}(0+) = 0$ e $v_{10}(1-) = v_{01}(1-) = 1$. Além disso, seja $\lambda \colon (0,1) \to \mathbb{R}$ qualquer função contínua e não-negativa satisfazendo a restrição $0 < \int_0^1 (v'_{10}(u) + v'_{01}(u))\lambda(u)\,\mathrm{d}u < \infty$. Então, temos que:

1. A função $v_{00} \colon (0,1) \to \mathbb{R}$ definida através de

$$v_{00}(\tau) \coloneqq \frac{\int_0^\tau (v'_{10}(u) + v'_{01}(u))\lambda(u)\,\mathrm{d}u}{\int_0^1 (v'_{10}(u) + v'_{01}(u))\lambda(u)\,\mathrm{d}u}, \quad 0 < \tau < 1, \tag{2.6}$$

   é continuamente diferenciável e não-negativa, com $v_{00}(0+) = 0$ e $v_{00}(1-) = 1$.

2. A função $v_{11} := v_{10}+v_{01}-v_{00}$ é continuamente diferenciável e satisfaz $v_{11}(0+) = 0$, $v_{11}(1-) = 1$ e, para $0 < \tau < 1$,

$$v'_{11}(\tau) = (v'_{10}(\tau) + v'_{01}(\tau)) \left(1 - \frac{\lambda(\tau)}{\int_0^1 (v'_{10}(u) + v'_{01}(u))\lambda(u)\,\mathrm{d}u}\right). \tag{2.7}$$

Em particular, para que v_{11} seja não-decrescente, é necessário e sufucente que a condição

$$\lambda(\tau) \leqslant \int_0^1 (v'_{10}(u) + v'_{01}(u))\lambda(u)\mathrm{d}u$$

se mantenha para cada τ tal que $v'_{10}(\tau) + v'_{01}(\tau) \neq 0$.

Técnicas de estimação

Suponha que $(X_t, Y_t)_{t\in\mathbb{Z}}$ satisfaça a dinâmica QADL (2.3). Uma propriedade importante da regressão quantílica é que, para cada $0 < \tau < 1$, mantém que

$$\beta(\tau) = \underset{b\in\mathbb{R}^{\dim Z}}{\arg\min}\, \mathbb{E}\,\rho_\tau(Y_0 - Z'_0 b). \tag{2.8}$$

Onde $2\rho_\tau(v) := (2\tau - 1)v + |v|$ é a *check function* usual.

O fato de que, para cada $0 < \tau < 1$, o parâmetro $\beta(\tau)$ é um $\arg\min$ sugere fortemente que ele pode ser estimado considerando as contrapartes da amostra para o problema de otimização na equação (2.8). Nesta seção, apresentamos o *estimador canônico* de Koenker e Bassett (1978), bem como o *estimador suavizado de regressão quantílica do tipo convolução via kernel* de Fernandes et al. (2021).

O estimador canônico de regressão quantílica

Em seu artigo seminal, Koenker e Bassett (1978) propôs o que nos referimos neste artigo como o **estimador padrão (ou canônico) de regressão quantílica**, definido como

$$\widehat{\beta}_n(\tau) := \underset{b\in\mathbb{R}^{\dim Z}}{\arg\min}\, \frac{1}{n}\sum_{t=1}^{n} \rho_\tau(Y_t - Z'_t b), \quad 0 < \tau < 1.$$

Também é sabido pelo literatura estabelecida que $\sqrt{n}\left(\widehat{\beta}_n(\tau) - \beta(\tau)\right)$ é assintoticamente normal-centrado com matriz de covariância $\Sigma(\tau)$.

Essas noções nos permitem inferir algumas desvantagens. Primeiro, a representação Bahadur-Kiefer para o estimador padrão tem uma taxa baixa. Em segundo lugar, $\Sigma(\tau)$ é difícil de estimar. Na verdade, ambas resultam da falta de suavidade da função objetivo. Esta falta de suavidade é um conceito muito importante para compreender a seguinte técnica de estimação.

Estimador suavizado de regressão quantílica do tipo convolução via kernel

Como a função objetivo (amostral) que o estimador canônico de regressão quantílica visa minimizar não é suave, a inferência estatística não é direta. Fernandes et al. (2021) propõe suavizar sua função objetivo, apresentando assim um estimador alternativo: o **estimador suavizado de regressão quantílica do tipo de convolução via kernel**. Aplicando uma convolução na *check function* em relação a uma função *kernel* padronizada, os autores obtiveram um estimador de regressão quantílica suavizado, que é dado por

$$\widehat{\beta}_n^{\zeta}(\tau) \coloneqq \underset{b \in \mathbb{R}^{\dim Z}}{\arg\min} \frac{1}{n} \sum_{t=1}^{n} k_{\zeta(\tau)} * \rho_\tau (Y_t - Z_t' b), \quad 0 < \tau < 1, \tag{2.9}$$

onde $k_h(v) \coloneqq k(v/h)/h$ é uma função *kernel*, $\zeta \equiv \zeta_n = (\zeta_n(\tau)\colon 0 < \tau < 1)$ é um processo estocástico com espaço de estados $[h_n, h^n]$ que captura o grau de suavização e $f * g(v) \coloneqq \int_{-\infty}^{\infty} f(u) g(v-u) \, du$ denota a convolução de f com g. Podemos considerar $\widehat{\beta}_n^{\zeta}(\tau)$ como um estimador do parâmetro auxiliar suavizado

$$\beta^{\zeta}(\tau) \coloneqq \underset{b \in \mathbb{R}^{\dim Z}}{\arg\min} \, \mathbb{E} \, k_{\zeta(\tau)} * \rho_\tau (Y_0 - Z_0' b), \quad 0 < \tau < 1,$$

que é uniformemente próximo de $\beta(\tau)$ sob suposições moderadas, como mostrado em Fernandes et al. (2021) – resolvendo assim a questão do viés introduzido pela suavização.

Capítulo 3

Conclusões e trabalhos futuros

Nesta dissertação, fornecemos duas contribuições principais para a literatura de regressão quantílica. Em primeiro lugar, investigamos caracterizações analíticas para o processo de geração de dados de dinâmicas QADL. Em segundo lugar, conduzimos um estudo de Monte Carlo massivo explorando essas dinâmicas, ao passo que avaliamos o desempenho do estimador de regressão quantílica suavizado de Fernandes et al. (2021) em um contexto de série temporais, particularmente na classe de modelos QADL de Galvão et al. (2013). Além de seguir os procedimentos necessários para estudar o desempenho do estimador em um contexto de série temporais, estabelecemos proposições a respeito dessas novas caracterizações para dinâmicas QADL. Em relação ao estudo de Monte Carlo, obtivemos resultados numéricos que corroboram a maioria dos pressupostos teóricos referentes ao estimador suavizado de Fernandes et al. (2021). Neste estudo, definimos o processo ζ 's como uma série de larguras de banda proporcionais à largura de banda baseada nos dados ζ^* de Silverman (1986), o que funcionou bem. Confirmando nossas suspeitas, em termos de MSE e variância, dada uma escolha adequada de largura de banda, $\beta_n^{\zeta^*}(\tau)$ sistematicamente superou $\beta_n(\tau)$ na maioria dos cenários explorados – exceto para as versões *location-scale* no primeiro conjunto de simulações, cada cenário apresentou pelo menos uma largura de banda ótima, na qual o estimador suavizado superou o canônico. O último resultado se estende para os *cenários existentes* e *novos cenários*. O preço a pagar por $\beta_n^{\zeta^*}(\tau)$ ter um MSE menor e uma variância menor, na maioria dos cenários estudados, foi um viés maior. Este fato era esperado, no entanto.

Em síntese, nossos resultados indicam que o estimador suavizado tem um bom desempenho em um contexto de séries temporais – e que pode ser considerado uma boa alternativa ao se estimar coeficientes em modelos de regressão quantílica de séries temporais. Além disso, as novas propostas para o processo de geração de dados da dinâmica QADL – compiladas pelas Proposições 1 e 2 – forneceram resultados satisfatórios no que concerne às simulações. Nesse sentido, nossa principal conclusão é a de que, dada uma largura de banda ideal – $2\zeta^*$, em nossos casos – $\beta_n^{\zeta^*}(\tau)$ superou globalmente $\beta_n(\tau)$ para todos os cenários simulados.

Um primeiro passo adicional para este estudo seria replicar a "*Application to House Price Returns*" implementada por Galvão et al. (2013) – desta vez incluindo o estimador suavizado de Fernandes et al. (2021) como um procedimento de estimativa alternativo para os coeficientes em modelos QADL.

Outro objetivo ambicioso que nos ocorre é compreender totalmente as ideias propostas por Yang e Yu (2007), para que possamos explorá-las trazendo o conceito de convergência uniforme quantílica amostrais em uma estrutura de regressão quantílica que inclua covariáveis exógenas – isto é, não apenas no caso univariado. Finalmente, qualquer trabalho futuro considerando alguma das caracterizações analíticas de dinâmicas QADL que foram discutidas neste trabalho é de nosso interesse.

Referências Bibliográficas

V. Chernozhukov e H. Hong. An MCMC approach to classical estimation. *Journal of Econometrics*, 115:293–346, 2003. URL https://doi.org/10.1016/S0304-4076(03)00100-3. 3

R. F. Engle e S. Manganelli. Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, 22(4):367–381, 2004. URL https://doi.org/10.2307/1392044. 3

M. Fernandes, E. Guerre, e E. Horta. Smoothing quantile regressions. *Journal of Business & Economic Statistics*, 39(1):338–357, 2021. URL https://doi.org/10.1080/07350015.2019.1660177. 3, 4, 5, 8, 9, 10

A. F. Galvão, G. Montes-Rojas, e S. Y. Park. Quantile autoregressive distributed lag model with an application to house price returns. *Oxford Bulletin of Economics and Statistics*, 13:307–321, 2013. URL https://doi.org/10.1111/j.1468-0084.2011.00683.x. 3, 4, 6, 10

H. G. Hong, D. C. Christiani, e Y. Li. Quantile regression for survival data in modern cancer research: expanding statistical tools for precision medicines. *Precision Clinical Medicines*, 2(2):90–99, 2019. URL https://doi.org/10.1093/pcmedi/pbz007. 3

J. L. Horowitz. Bootstrap Methods for Median Regression Models. *Econometrica*, 66(6):1327–1352, November 1998. URL https://ideas.repec.org/a/ecm/emetrp/v66y1998i6p1327-1352.html. 3

E. Horta. *Quantile Regression: Facts and digressions*. Department of Statistics, UFRGS, Porto Alegre, Brazil, 1 edition, 2021. URL https://github.com/eduardohorta/QuantRegFacts. 3, 6

R. Koenker e G. Bassett. Regression quantiles. *Econometrica*, 46(1):33–50, 1978. URL http://www.jstor.org/stable/1913643. 3, 4, 5, 8

R. Koenker, V. Chernozhukov, X. He, e L. Peng. *Handbook of Quantile Regression*. CRC Press, 2018. 3

N. Meinshausen. Quantile regression forests. *Journal of Machine Learning Research*, 7(35):983–999, 2006. URL http://jmlr.org/papers/v7/meinshausen06a.html. 3

F. Mosteller e Tukey. *Data Analysis and Regression: A Second Course in Statistics*. Taylor & Francis, 2005. URL https://ui.adsabs.harvard.edu/abs/1977dars.book.....M. 2

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021. URL https://www.R-project.org/. 4

B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, London, 1986. URL https://doi.org/10.1002/bimj.4710300745. 10

I. Takeuchi, Q. V. Le, T. Sears, e A. J. Smola. Nonparametric quantile regression. *Journal of Machine Learning Research*, 7:1001–1022, 2005. URL https://dl.acm.org/doi/10.5555/1248547.1248592#sec-ref. 3

Q. Xu, X. Lu, C. Jiang, e K. Yu. Quantile autoregression neural network model with applications to evaluating value at risk. *Applied Soft Computing*, 13:3120–3160, 2016. URL http://dx.doi.org/10.1016/j.asoc.2016.08.003. 3

S. Yang e K. Yu. *A note on the Bahadur representation of sample quantiles for mixing random variables*. Department of Mathematical Sciences, Brunel University, UK, 1 edition, 2007. URL https://bura.brunel.ac.uk/bitstream/2438/12676/3/FullText.pdf. 11

Anexo A

Artigo Natal e Horta

Smoothing quantile regressions with time series data

Miguel Jandrey Natal Eduardo Horta

Abstract

Quantile regression fits quantiles of the response variable and brings the concept of a quantile into the framework of general linear models. Although quantile regression was first introduced more than forty years ago, only recently it became practicable for large data due to computational advances. As the objective function that the standard quantile regression estimator aims to minimize is not smooth, statistical inference is not straightforward. Recently, a new estimation proposal for such coefficients was incorporated into the quantile regression literature: the convolution-type kernel smoothed quantile regression estimator. Based on this alternative approach to quantile regression modeling, this work aims to implement this smoothed estimator in a time series context. Since the estimation theory was formalized considering cross-sectional data, the goal here is to try a new step by expanding the study into a time series framework. Also, we investigate analytical characterizations for the data generating process of Quantile Autoregressive Distributed Lag (QADL) dynamics. Through Monte Carlo simulations, we explore these dynamics while evaluating the performance of the smoothed estimator in a class of time series quantile regression models.
Keywords: quantile regression, conditional quantile, convolution-based smoothing, QADL, time series, data-driven bandwidth, Monte Carlo simulation

1 Introduction

Many scenarios in the applied sciences reveal, or at least shed some light on existing relations between distinct variables. In such scenarios, researchers might need proper tools for completing their jobs, namely obtaining a good description of the dependence structure between these variables of interest. Concerning Statistics, Koenker (2005) argues that:

> *"Much of applied Statistics may be viewed as an elaboration of the linear regression model and associated estimation methods of least squares.* [...] *Quantile regression is intended to offer a comprehensive strategy for completing the regression picture".*

On the one hand, the least-squares estimation of the (mean) linear regression model represents one of the foundations of statistical theory. In addition to the fact that mean regression is parsimonious and easily interpretable model, the ordinary least squares (OLS) estimator enjoys certain optimality in some cases, so that combining these tools—i.e., estimating the coefficients of a linear regression model through the OLS estimator—is often a desirable path when proceeding with statistical analysis.

On the other hand, modeling a random variable through linear (mean) regression brings up some restrictive implications. For instance, let us imagine a simple mean regression model where the dependent variable, say Y, has its structure of dependence described by $\mathbf{E}[Y|X = x]$. The latter expression—which denotes the expected value of Y given $X = x$, that is, its conditional mean—might be a bit restrictive because it only specifies how the mean of Y responds to variations on x. In that case, it is not a surprise that such models are inadequate when the (conditional) mean is not a satisfactory end in itself. Mosteller and Tukey (2005) go beyond this idea by remarking that, in spite of the fact that the regression model gives a reasonable summary for the averages of the distributions corresponding to the set of x's, we could go further by incorporating quantiles to the analysis, aiming to obtain a complete picture of the response.

The theory of quantile regression arises in a context where one requires proper tools to perform an analysis of dependence between a response and covariates. Although classical linear regression models serve well in many problems in the applied sciences, not rarely having a complete picture of the conditional distribution of the response variable can incorporate valuable information. By bringing the concept of a quantile into the framework of general linear models (GLM), quantile regression expanded the way of visualizing the cross-sectional relations between data. While classical linear regression with strictly exogenous covariates models the conditional mean of a dependent variable, quantile regression fits quantiles of the response, so that we can recover its entire conditional distribution, not just the conditional mean. In this sense, it is worth saying that quantile regression should not be interpreted as a "better" model when compared to mean regression, but instead as an alternative approach for modeling the relation between a dependent variable and covariates—particularly when reliable information lies on the tails of the regressand's distribution.

Since the seminal paper by Koenker and Bassett (1978), which introduced quantile regression in its modern form,[1] considerable attention has been devoted to the study and development of associated models. Despite being proposed more than forty years ago, only recently quantile regression became more practicable for large data, due to computational advances. This fact certainly helped to bolster its popularization. The literature now has a wide range of applications in which quantile regression has been useful. One can mention several areas where quantile regression is contemplated, such as economics (Galvão et al., 2013), finance ((Engle and Manganelli, 2004); (Xu et al., 2016)), machine learning ((Meinshausen, 2006); (Takeuchi et al., 2005)), medicine (Hong et al., 2019), among others. Techniques such as bootstrapping (Horowitz, 1998) and Markov Chain Monte Carlo (MCMC) methods (Chernozhukov and Hong, 2003) are also examples of such applications within Statistics itself. Indeed, it has drawn the attention not only of the mathematical-statistics community, but also of those whose concerns were mainly applications of the method.

As the sample objective function that the standard quantile regression estimator aims to minimize is not smooth, statistical inference is not straightforward. Fernandes et al. (2021) propose to smooth this objective function, thus presenting an alternative estimator: the convolution-type kernel

[1] As we currently refer, quantile regression was officially introduced by Koenker and Bassett (1978). However, according to Koenker et al. (2018) its true origin can be traced from Boscovich and Laplace through the "median regression" in the eighteenth century.

smoothed quantile regression estimator. Based on this alternative approach for estimating the functional parameter in the quantile regression model, this work aims to implement the convolution-type kernel quantile regression estimator in a time series context. Since the authors have formalized the theory of the smoothed estimator considering cross-sectional data, the goal here is to try a new step by expanding their study into a class of time series models. Through Monte Carlo simulations, we evaluate the estimator's performance in new and existing scenarios. Whereas the first ensemble of simulations was replicated from the scenarios found in Galvão et al. (2013), it is important to say that they were adapted to fit our purposes, mainly by evaluating more quantile levels. As for the new scenarios, they are slight adaptations from some novel contributions on the quantile regression literature given by Horta (2021)—we perform a Monte Carlo study on these scenarios, which were constructed based on his ideas. With these new scenarios, we aim to offer novel contributions concerning QADL dynamics, notably by investigating analytical characterizations for the data generating processes of these dynamics.

Section 1 ends with this brief introduction for the quantile regression historical and our main goals throughout this work. In Section 2, we properly describe quantile regression and its associated methods, as well as some further considerations on the QADL dynamics, which aim to be novel contributions. Then, in Section 3, we conduct a Monte Carlo study to numerically evaluate both the canonical and the smoothed quantile regression estimators in a class of time series quantile regression models. Finally, in Section 4, we conclude with a general analysis of what is shown in this article and suggestions for further steps and possible applications of the theoretical contributions this work aims to provide.

2 Methodology

This section is dedicated to the theoretical specification of quantile regression models in a time series context. First, we present a set up of the theoretical aspects, which include basic definitions, mathematical notation, and assumptions. Then, we describe the quantile regression model and its analogous equations considering a time series context. We introduce some further considerations on the QADL dynamics, where we provide some novel contributions with regards to its theoretical aspects. Lastly, we present the canonical estimation procedure for quantile regression proposed by Koenker and Bassett (1978) and the alternative approach of Fernandes et al. (2021).

Definitions and introductory concepts

In the most basic setting, for a scalar random variable Y with cumulative distribution function F_Y, we define the **τth quantile of Y**, with $0 < \tau < 1$, to be the real number $Q_Y(\tau) \coloneqq \inf\{y \in \mathbb{R} \colon F_Y(y) \geqslant \tau\}$. The mapping $\tau \mapsto Q_Y(\tau)$ is called the **quantile function of Y** or the **generalized inverse of F_Y**. The importance of quantile functions is that they entirely characterize the distribution of Y, as F_Y is recoverable from Q_Y. In the presence of a vector of covariates X of dimension $\mathbb{R}^{\dim_X}$, we similarly define the **conditional τth quantile of Y given $X = x$** to be the real number $Q_{Y|X}(\tau|x)$ defined, for $x \in \operatorname{support}(X)$ and $\tau \in (0, 1)$, through

$$Q_{Y|X}(\tau|x) = \inf\{y \in \mathbb{R} \colon F_{Y|X}(y|x) \geqslant \tau\}, \tag{1}$$

where $F_{Y|X}(\cdot|x)$ is the conditional cumulative distribution function of Y given $X = x$. The mapping $\tau \mapsto Q_{Y|X}(\tau|x)$ is called the **conditional quantile function of Y given $X = x$**. In a time series framework, in turn, some adjustments in notation and definitions are required, as we explain below.

Quantile regression with time series data

In what follows, $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\in\mathbb{Z}}, \mathbf{P})$ denotes a filtered probability space, and $(X_t, Y_t)_{t\in\mathbb{Z}}$ is a stochastic process with state space $\mathbb{R} \times \mathbb{R}^{\dim_X}$, which we assume is adapted to $\{\mathscr{F}_t\}_{t\in\mathbb{Z}}$. For identification purposes, we also assume throughout the remainder of the text that, for all $t \in \mathbb{Z}$, the components of X_t are non-degenerate and linearly independent random variables, as well as linearly independent from $\{Y_s \colon s \leqslant t\}$. For each $t \in \mathbb{Z}$ and $0 < \tau < 1$, the **conditional τth quantile of Y_t given $\mathscr{F}_{t-1}$** is the $\mathscr{F}_{t-1}$-measurable random variable defined through

$$Q_{Y_t}(\tau \,|\, \mathscr{F}_{t-1}) \coloneqq \inf\Big\{y \in \mathbb{R}\colon \; \mathbf{P}(Y_t \leqslant y \,|\, \mathscr{F}_{t-1}) \geqslant \tau\Big\} \tag{2}$$

We assume, without loss of generality, that a regular version of $y \mapsto \mathbf{P}(Y_t \leqslant y \,|\, \mathscr{F}_{t-1})$ is at hand. In this case, the stochastic process $\{Q_{Y_t}(\tau \,|\, \mathscr{F}_{t-1})\colon\; 0 < \tau < 1\}$, which we call the **conditional quantile function of Y_t given $\mathscr{F}_{t-1}$**, is ω-wise the generalized inverse of the cumulative distribution function $\mathbf{P}(Y_t \leqslant \cdot \,|\, \mathscr{F}_{t-1})_\omega$.

Example 1 (Classical autoregressive process of order 1)**.** Let $(U_t\colon t \in \mathbb{Z})$ be a doubly infinite sequence of *iid* Gaussian random variables, and let α be a real number with $|\alpha| < 1$. Define $Y_t \coloneqq \lim_{p\to\infty} \sum_{\ell=0}^{p} \alpha^\ell U_{t-\ell}$, $t \in \mathbb{Z}$. Clearly the sequence $(Y_t\colon\; t \in \mathbb{Z})$ satisfies the stochastic difference equations

$$Y_t = \alpha Y_{t-1} + U_t, \qquad t \in \mathbb{Z}. \tag{3}$$

We call $(Y_t\colon\; t \in \mathbb{Z})$ an **autoregressive process of order 1** (without an intercept, or "trend"). It is easy to see that in this context the conditional quantile function of Y_t given the σ-field $\mathscr{F}_{t-1}$ generated by $\{Y_s\colon\; s < t\}$ is of the form $\tau \mapsto \alpha Y_{t-1} + Q_U(\tau)$ for all t, where Q_U denotes the unconditional quantile function of U_0.

We now introduce the Quantile Autoregressive Distributed Lag Model (QADL) model, slightly adapting the definition put forth by Galvão et al. (2013).

Definition 1. A process $(X_t, Y_t)_{t\in\mathbb{Z}}$ is said to satisfy a **Quantile Autoregressive Distributed Lag Model** of orders lag_Y and lag_X, which we write $\mathrm{QADL}(\mathrm{lag}_Y, \mathrm{lag}_X)$ for short, iff the relation

$$Q_{Y_t}(\tau|\mathscr{F}_{t-1}) = \alpha_0(\tau) + \sum_{j=1}^{\mathrm{lag}_Y} \alpha_j(\tau) Y_{t-j} + \sum_{\ell=1}^{\mathrm{lag}_X} X'_{t-\ell}\theta_\ell(\tau), \tag{4}$$

holds for all $t \in \mathbb{Z}$ and all $0 < \tau < 1$, for some functions $\alpha_j\colon (0,1) \to \mathbb{R}$ and

$\theta_\ell\colon (0,1) \to \mathbb{R}^{\dim_X}$ (with $0 \leqslant j \leqslant \text{lag}_Y$ and $1 \leqslant \ell \leqslant \text{lag}_X$).

Remark 1. Galvão et al. (2013) consider a setting in which a "contemporaneous" X_t appears in the right-hand-side of equation (4)—that is, the indices in the second summand run from 0 through lag_X—, and with $\mathscr{F}_{t-1}$ redefined to be the σ-field generated by $\{(X_s, Y_{s-1})\colon\ s \leqslant t\}$. Mathematically the difference is merely cosmetic, as one can easily substitute each X_s by the random variable $W_s \coloneqq X_{s-1}$, $s \in \mathbb{Z}$. In a purely predictive model, the specification given in equation (4) is convenient. However, if a structural model (economic, for instance) establishes a contemporaneous relationship between Y_t and certain exogenous covariates, then the suggested reparametrization may be more suitable.

Remark 2. The system of functional equations (4) entirely determines the conditional distribution of Y_t given $\mathscr{F}_{t-1}$, for all $t \in \mathbb{Z}$, but it is important to stress that it does not determine the multivariate dynamics of $(X_t, Y_t)_{t\in\mathbb{Z}}$. Indeed, nothing is being said about the law of $(X_t)_{t\in\mathbb{Z}}$. Put another way, the QADL model is in fact a *partially specified model*, in the sense that there might be more than one probability law for $(X_t, Y_t)_{t\in\mathbb{Z}}$ according to which (4) holds. A sufficient condition to recover the law of $(X_t, Y_t)_{t\in\mathbb{Z}}$ is that, additionally to (4), one provides the marginal distribution of, say, (X_0, Y_0), together with the conditional distribution of X_t given $(\mathscr{F}_{t-1}, Y_t)$ for each $t \in \mathbb{Z}$.

Remark 3. It is convenient, especially when it comes to computational implementations, to introduce the notation

$$Z_t' = \begin{pmatrix} 1 & Y_{t-1} & \cdots & Y_{t-\text{lag}_Y} & X_{t-1}' & \cdots & X_{t-\text{lag}_X}' \end{pmatrix}, \quad t \in \mathbb{Z}$$

and, for $0 < \tau < 1$,

$$\beta(\tau)' = \begin{pmatrix} \alpha_0(\tau) & \alpha_1(\tau) & \cdots & \alpha_{\text{lag}_Y}(\tau) & \theta_1(\tau)' & \cdots & \theta_{\text{lag}_X}(\tau) \end{pmatrix}',$$

(notice that $\beta(\tau) \in \mathbb{R}^{1+\text{lag}_Y+\dim_X\cdot\text{lag}_X}$) so that we can rewrite (4) more compactly as

$$Q_{Y_t}(\tau \mid \mathscr{F}_{t-1}) = Z_t'\beta(\tau). \tag{5}$$

The above notation also makes clear that the cross sectional setting can be accommodated—albeit with a somewhat cumbersome notation—as a special case of QADL dynamics: this corresponds to the requirement that (Z_t, Y_t) is independent of $\sigma\{(Z_s, Y_s)\colon\ s < t\}$ for each $t \in \mathbb{Z}$. In this case, since $Q_{Y_t}(\cdot|\mathscr{F}_{t-1}) = Q_{Y_t}(\cdot|\sigma(Z_t))$, we forcibly have that $\alpha_j(\cdot) = |\theta_\ell(\cdot)| \equiv 0$, for all $1 \leqslant j \leqslant \mathrm{lag}_Y$ and $2 \leqslant \ell \leqslant \mathrm{lag}_X$.

Example 2 (Quantile autoregressive models)**.** Assume $(X_t, Y_t)_{t\in\mathbb{Z}}$ satisfies the relation in (4) for all $t \in \mathbb{Z}$ and $0 < \tau < 1$, with $\theta_j(\cdot) \equiv 0 \in \mathbb{R}^{\dim_X}$ for all $j \in \{1, \ldots, \mathrm{lag}_X\}$. This is the **quantile autoregressive model of order** lag_Y, written QAR(lag_Y) for short, introduced by Koenker and Xiao (2006). Classical autoregressive models are a special case of the QAR model where the coefficients of lagged Y's are constant in τ. Let us compute the first two correlation coefficients ρ_1 and ρ_2 of a stationary QAR(1) model: if $Q_{Y_t}(\tau|\mathscr{F}_{t-1}) = \alpha_0(\tau) + \alpha_1(\tau)Y_{t-1}$, then, using the conditional version of the Fundamental Theorem of Simulation, we obtain $\mathbf{E}(Y_t|\mathscr{F}_{t-1}) = \int_0^1 \alpha_0(u)\,\mathrm{d}u + Y_{t-1}\int_0^1 \alpha_1(u)\,\mathrm{d}u$. Thus,

$$\mathbf{E}(Y_t) = \int_0^1 \alpha_0(u)\,\mathrm{d}u + \mathbf{E}(Y_{t-1})\int_0^1 \alpha_1(u)\,\mathrm{d}u = \int_0^1 \alpha_0(u)\,\mathrm{d}u + \mathbf{E}(Y_t)\int_0^1 \alpha_1(u)\,\mathrm{d}u$$

and $\mathbf{E}(Y_t) = \int_0^1 \alpha_0(u)\,\mathrm{d}u/(1-\int_0^1 \alpha_1(u)\,\mathrm{d}u)$. An easy computation using $\mathbf{E}(Y_{t-1}Y_t) = \mathbf{E}(Y_{t-1}\,\mathbf{E}(Y_t|\mathscr{F}_{t-1}))$ yields $\mathrm{Cov}(Y_t, Y_t - 1) = \mathbf{Var}(Y_0)\int_0^1 \alpha_1(u)\,\mathrm{d}u$. Similarly, $\mathrm{Cov}(Y_t, Y_{t-2}) = \mathbf{Var}(Y_0)(\int_0^1 \alpha_1(u)\,\mathrm{d}u)^2$. This tells us that $\rho_1 = \int_0^1 \alpha_1(u)\,\mathrm{d}u$ and $\rho_2 = (\int_0^1 \alpha_1(u)\,\mathrm{d}u)^2$.

Coefficient and covariate restrictions

It is well known in the quantile regression literature that, under the linear specification (4), monotonicity (w.r.t. τ) of the conditional quantile functions imposes some discipline on the functional forms taken by the coefficients, or (unless the coordinates of the dependent variables are themselves functionally dependent) on the support of the covariate vectors Z_t (Koenker, 2005; Galvão et al., 2013). One possible workaround is to take linearity as an approximation of the "true" model, for example, only requiring that the function $\tau \mapsto Z_t'\beta(\tau)$ be monotone in a relevant region of the $\mathscr{F}_{t-1}$-space. This solution is valid if the goal is estimation—for instance, as noted in Koenker and Xiao (2006), the estimated conditional quantile function $\widehat{Q}_{Y_t}(\tau|\mathscr{F}_t) = Z_t'\widehat{\beta}(\tau)$ is ensured to be monotone in τ at $Z_t = \bar{Z}$—, but not so much if one is interested in simulating QADL models with non-constant coefficients. In this section, we briefly discuss some of the restrictions implied by the requirement of linearity, after which we put forth a new approach entailing sufficient conditions under which the functional parameter $\beta(\cdot)$ is ensured to yield non-decreasing maps $\tau \mapsto z'\beta(\tau)$ for all $z \in \text{support}(Z_t)$. Alternative approaches are found, for example, in Koenker and Xiao (2006), Gourieroux et al. (2008) and Galvão (2009).

Example 3 (Quantile crossing)**.** Suppose $(Y_t)_{t\in\mathbb{Z}}$ follows a quantile autoregressive dynamics of order 1, that is, it holds that

$$Q_{Y_t}(\tau|\mathscr{F}_{t-1}) = \alpha_0(\tau) + \alpha_1(\tau)Y_{t-1}, \qquad 0 < \tau < 1,\, t \in \mathbb{Z}. \tag{6}$$

Assume further that the coefficient α_1 is not constant in τ, and without loss of generality suppose there are two distinct quantile levels $0 < \hat{\tau} < \tilde{\tau} < 1$ such that $\alpha_1(\hat{\tau}) > \alpha_1(\tilde{\tau})$. Then $\text{support}(Y_0)$ is necessarily bounded from above. Indeed, if this were not the case, one would have $Q_{Y_1}(\hat{\tau}|\mathscr{F}_0)_\omega > Q_{Y_1}(\tilde{\tau}|\mathscr{F}_0)_\omega$ for a set of ω's having positive probability. But this is forbidden since $\tau \mapsto Q_{Y_1}(\tau|\mathscr{F}_0)_\omega$ is monotonically increasing except for ω lying in a $\mathbf{P}$-null set. This example (which can be easily extended to the general QADL framework) illustrates the fact that the support of Z_t strongly restricts the functional form of the coefficients $\tau \mapsto \beta(\tau)$, or equivalently that the requirement that $\tau \mapsto z'\beta(\tau)$ be non-decreasing strongly restricts the set of "valid" z's; informally (and not quite rigorously), one can say that *if the regressors in a quantile regression model are unbounded, then some coefficients are necessarily constant.* Given this, it is customary to restrict attention to covariate vectors with bounded support, although such a restriction is not strictly necessary.

Example 4 (Sign restrictions)**.** Assume equation (5) holds "without an intercept", that is, with $\alpha_0(\cdot) \equiv 0$. In this setting, the sign of the covariates cannot vary freely (except if all other coefficients in β are constant functions of τ), in the following sense: assume there is a point $z \in \mathbb{R}^{\dim z - 1}$ such that both vectors $\hat{z} := (1 \quad z)'$ and $\tilde{z} := (1 \quad -z)'$ are "possible outcomes" of, say, Z_0. Then, as the functions $\tau \mapsto \hat{z}'\beta(\tau)$ and $\tau \mapsto \tilde{z}'\beta(\tau) = -\hat{z}'\beta(\tau)$ are both non-decreasing (since they are quantile functions), the only possibility left is that $\tau \mapsto \beta(\tau)$ is a constant function. Therefore, whenever α_0 is identically zero, linear quantile regression is incompatible with, say, multivariate Gaussian regressors. As in the above example, we see once more that the support of Z_t and the functional form of the coefficients $\tau \mapsto \beta(\tau)$ are mutually restrictive. Again, one can informally say that *if the regressors in a quantile regression model change sign freely, then some coefficients are necessarily constant.*

In view of the preceding examples, we may conclude that—in the (linear) quantile regression framework at least—it is convenient to restrict attention to covariate vectors whose support is a bounded subset of $\mathbb{R}_+^{\dim z} := [0, +\infty)^{\dim z}$, especially if the interest lies in studying a richer scenario where the structure of dependence in (4) allows for non-constant coefficients besides α_0. Importantly, as mentioned by (Horta, 2021, section 4.1), *imposing such constraints on the covariates may demand a restriction on the response as well.* This is the case, for example, when the response in (4) is equal in distribution to some of the regressors in the covariate vector. One ought to notice that this is precisely what happens when (X_t, Y_t) is (strongly) stationary with a non-zero autoregressive component (say, $\alpha_1(\cdot)$ is not identically zero).

In the remainder of this paper, we intend to study the class of stationary stochastic processes $(X_t, Y_t)_{t \in \mathbb{Z}}$ having a probability law for which the time homogeneous quantile regression equation (4) holds for $0 < \tau < 1$ and $t \in \mathbb{Z}$. In view of the preceding examples and discussion, we restrict our attention to dynamics that have as state space a bounded subset of

$[0, +\infty)^{\dim_X+1}$—as a matter of fact, we lose no generality in assuming that $\text{support}(X_t, Y_t) \subseteq [0, 1]^{\dim_X+1}$. Considering all these restrictions, how to proceed in such a challenging scenario?

QADL dynamics: further considerations

In this section, we explore the QADL model design already studied by Galvão et al. (2013), using its well defined architecture as a template for trying to go a bit further on QADL dynamics. As we have discussed above, the linear specification implies a strong bind in the interplay between the functional form of the regression coefficients on the one hand, and the support of the covariates on the other; one cannot have flexible coefficient functions *and* a unspecified/generic covariate support. But it is no surprise that every yield comes at a cost, and presently this means that it is necessary to impose restrictions on the joint support of (X_t, Y_t) in order to enjoy the full flexibility of quantile regression—a flexibility that stems precisely from coefficient functions that are not constant in τ.

We are now able to introduce one of the core scenarios that will be contemplated in our Monte Carlo studies of the next section. The following result is a slight adaptation of Proposition 4.1 in Horta (2021), which establishes sufficient conditions enabling one to specify the conditional quantile function of Y_t given $\mathscr{F}_{t-1}$ via an underlying *family of affine hyperplanes determined by "vertex functions"*. For simplicity, we restrict ourselves to the setting with $\text{lag}_Y = \text{lag}_X = \dim_X = 1$. A multidimensional generalization can be obtained as discussed in the remark following Proposition 4.1 in Horta (2021).

Proposition 1. Let $v_{ij} \colon (0, 1) \to [0, 1]$, $i, j \in \{0, 1\}$ be non-decreasing, left-continuous functions with $v_{11} + v_{00} = v_{10} + v_{01}$. Let additionally Y_0 and X_0 be random variables supported on the unit interval, and let $U_1, U_2, \ldots$ and $V_1, V_2, \ldots$ be *iid* sequences of random variables in the unit interval, with both sequences mutually independent and also independent from Y_0 and X_0.

Define moreover, for $t \geqslant 1$,

$$\begin{aligned} Y_t &:= \alpha_0(U_t) + \alpha_1(U_t)Y_{t-1} + \theta_1(U_t)X_{t-1} \\ X_t &:= Q_t(V_t, X_{t-1}, \ldots, X_0, Y_t, Y_{t-1}, \ldots, Y_0), \end{aligned} \tag{7}$$

where, for all $t \geqslant 1$, the measurable mapping $Q_t\colon (0,1) \times [0,1]^{2t+1} \to [0,1]$ is a quantile function of its first argument, and where $\alpha_0 := v_{00}$, $\alpha_1 := v_{10} - v_{00}$ and $\theta_1 := v_{01} - v_{00}$. Then the process $(X_t, Y_t)_{t\in\mathbb{Z}}$ has state space contained in the unit square $[0,1]^2$ and satisfies

$$Q_{Y_t}(\tau|\mathscr{F}_{t-1}) = \alpha_0(\tau) + \alpha_1(\tau)Y_{t-1} + \theta_1(\tau)X_{t-1} \tag{8}$$

for all $0 < \tau < 1$ and all $t \geqslant 1$, where $\mathscr{F}_t := \sigma\{(X_s, Y_s)\colon\ s \leqslant t\}$.

Sketch of proof. The idea of the proof is that the "vertex functions" v_{ij} specify the conditional quantile function of Y_t given $\mathscr{F}_{t-1}$ (which by construction depends only on Y_{t-1} and X_{t-1}) at the four vertices of the unit square, namely at $Y_{t-1} = i, X_{t-1} = j$ for $i, j \in \{0,1\}$. The requirements that each v_{ij} is non-decreasing, left-continuous, and with $\mathrm{range}(v_{ij}) \subseteq [0,1]$ then ensure that, for any pair (y, x) lying in the unit square, the mapping

$$\tau \mapsto v_{00}(\tau) + (v_{10}(\tau) - v_{00}(\tau)) \cdot y + (v_{01}(\tau) - v_{00}(\tau)) \cdot x$$

is non-decreasing and left-continuous, with range contained in $[0,1]$. A simple application of the conditional version of the Fundamental Theorem of Simulation then yields the stated representation (8). For the details of the proof, see Horta (2021). ■

Remark 4. As we have highlighted in the proof of Proposition 1, for each fixed $0 < \tau < 1$, the points $v_{00}(\tau)$, $v_{01}(\tau)$, $v_{10}(\tau)$ and $v_{11}(\tau)$ can be identified with the "unit square vertices" of a uniquely determined affine hyperplane in $\mathbb{R}^2$. In fact, one of the vertices is redundant—it is implied by the other three—and this is encoded in the requirement that $v_{11} + v_{00} = v_{10} + v_{01}$. In practice, from a modeling perspective, the researcher has the freedom to choose any three vertex functions that are non-decreasing, left-continuous and with range contained in the unit interval, as long as the remaining one satisfies these requirements as well. Notice, however, that the condition is sufficient but not necessary, as shown in Example 5 below.

Remark 5. The above proposition allows one to obtain a stochastic process with index set $\mathbb{N} \cup \{0\}$ and not $\mathbb{Z}$ as we had previously defined. In practice, especially with regard to simulation and inference, this distinction is immaterial.

Remark 6. The process $(X_t, Y_t)_{t\geqslant 0}$ constructed in Proposition 1 may or may not be stationary, Markovian, ergodic, etc. Actually, all these stochastic properties depend on the initial distribution of the pair (Y_0, X_0) and on the "transition quantile functions" Q_t; in fact, as mentioned by Horta (2021), despite the possibility of X_t displaying a quite arbitrary dynamics, the equation for Y_t strongly suggests the Markov property, which is attained whenever one can find a function $Q^*\colon (0,1) \times [0,1]^3$ such that the identity

$$Q_t(\tau, x_{t-1}, \ldots, x_0, y_t, y_{t-1}, \ldots, y_0) = Q^*(\tau, x_{t-1}, y_t, y_{t-1})$$

holds for every $0 < \tau < 1$, every $t \geqslant 1$ and every $y_0, \ldots, y_t, x_0, \ldots, x_{t-1} \in [0,1]$.

Example 5. For $0 < \tau < 1$, let $v_{00}(\tau) \coloneqq \tau$ and $v_{10}(\tau) = v_{01}(\tau) \equiv 0$. Let (X_0, Y_0) be a random vector supported on the set

$$\Delta \coloneqq \Big\{(x, y) \in [0,1]^2 \colon x + y \leqslant 1\Big\},$$

and define $(X_t, Y_t)_{t\geqslant 1}$ recursively via (7) with $Q_t(\tau, x_{t-1}, \ldots, x_0, y_t, \ldots, y_0) + \leqslant 1 - y_t$. Then the process $(X_t, Y_t)_{t\geqslant 0}$ has state space contained in Δ, and satisfies (8) for all $0 < \tau < 1$ and $t \geqslant 1$, even though $v_{11} \equiv -v_{00}$ is strictly decreasing. Notice that, conditionally on $\mathscr{F}_t$, Y_{t+1} is uniformly distributed on $(0, 1 - (X_t + Y_t)]$.

Vertex compatibility conditions

This section is important as it provides a template for our simulation study, building on the ideas put forth in Proposition 1. Suppose we have already chosen the functions v_{10} and v_{01}—for example, both can be quantile functions of the Beta distribution, possibly with distinct parameters. Our goal is to select a function $v_{00}\colon (0,1) \to \mathbb{R}$ such that the following holds:

1. v_{00} is non-decreasing and left-continuous, with $\operatorname{range}(v_{00}) \subseteq [0,1]$;
2. $v_{11} \coloneqq v_{10} + v_{01} - v_{00}$ is non-decreasing and left-continuous, and it holds that $\operatorname{range}(v_{11}) \subseteq [0,1]$.

Importantly, if we restrict ourselves to the scenario where v_{10}, v_{01} and v_{00} are differentiable, then the monotonicity of v_{11} is equivalent to the requirement that $v'_{10} + v'_{01} \geqslant v'_{00}$. The next result, which is an original contribution, provides sufficient conditions for achieving said goal. For notational convenience, we write $v(\tau-) \coloneqq \lim_{u\uparrow\tau} v(u)$ and similarly $v(\tau+) \coloneqq \lim_{u\downarrow\tau} v(u)$.

Proposition 2. Let $v_{10}, v_{01}\colon (0,1) \to \mathbb{R}$ be any two non-decreasing, continuously differentiable functions, with $v_{10}(0+) = v_{01}(0+) = 0$ and $v_{10}(1-) = v_{01}(1-) = 1$. Also, let $\lambda\colon (0,1) \to \mathbb{R}$ be any non-negative, continuous function satisfying the requirement that $0 < \int_0^1 (v'_{10}(u) + v'_{01}(u))\lambda(u)\,\mathrm{d}u < \infty$. Then the following holds:

1. The function $v_{00}\colon (0,1) \to \mathbb{R}$ defined through

$$v_{00}(\tau) \coloneqq \frac{\int_0^\tau (v'_{10}(u) + v'_{01}(u))\lambda(u)\,\mathrm{d}u}{\int_0^1 (v'_{10}(u) + v'_{01}(u))\lambda(u)\,\mathrm{d}u}, \quad 0 < \tau < 1, \tag{9}$$

   is continuously differentiable and non-decreasing, with $v_{00}(0+) = 0$ and $v_{00}(1-) = 1$.
2. The function $v_{11} \coloneqq v_{10} + v_{01} - v_{00}$ is continuously differentiable, satisfies $v_{11}(0+) = 0$, $v_{11}(1-) = 1$ and, for $0 < \tau < 1$,

$$v'_{11}(\tau) = (v'_{10}(\tau) + v'_{01}(\tau))\left(1 - \frac{\lambda(\tau)}{\int_0^1 (v'_{10}(u) + v'_{01}(u))\lambda(u)\,\mathrm{d}u}\right). \tag{10}$$

In particular, in order that v_{11} be non-decreasing, it is necessary and sufficient that the bound

$$\lambda(\tau) \leqslant \int_0^1 (v'_{10}(u) + v'_{01}(u))\lambda(u)\mathrm{d}u$$

holds for each τ such that $v'_{10}(\tau) + v'_{01}(\tau) \neq 0$.

Proof. It is immediate that both v_{00} and v_{11} are continuously differentiable. The remainder of the proof is also straightforward: clearly $v_{00}(0+) = 0$ and $v_{00}(1-) = 1$. Also, since the assumptions imply $v'_{10} \geqslant 0$ and $v'_{01} \geqslant 0$, and as v'_{00} is proportional to $(v'_{10} + v'_{01})\lambda \geqslant 0$, it follows that v_{00} is non-decreasing. This establishes the first stated item. For the second item, a direct computation yields $v_{11}(0+) = 0$ and $v_{11}(1-) = 1$. It remains to prove the asserted equivalence. That the derivative of v_{11} is given by (10) is just a matter of verification, so v_{11} is non-decreasing whenever the required bound holds. By the same token, if this inequality fails at some τ for which $v'_{10}(\tau)+v'_{01}(\tau) > 0$, then the continuity of λ, v'_{10} and v'_{01} implies $v'_{11}(\tau) < 0$ in a neighborhood of τ—and, *a fortiori*, v_{11} decreases in this neighborhood. ■

Remark 7. The above proposition has been stated in a way that all the vertex functions satisfy $\mathrm{cl}(\mathrm{range}(v_{ij})) = [0,1]$ and $\mathrm{int}(\mathrm{range}(v_{ij})) = (0,1)$. This can be relaxed at the expense of clarity of exposition. Also, it is important to notice that this method for obtaining the vertex functions precludes scenarios where Y_t can be discrete, although allowing for mixed distributions.

Remark 8. It is important to have at hand, especially for computational implementations, the formula $\int_0^\tau \lambda(u)v'_{ij}(u)\mathrm{d}u = \int_0^{v_{ij}(\tau)} \lambda(F_{ij}(x))\mathrm{d}x$ where F_{ij} is the right-continuous generalized inverse of v_{ij} (that is, the distribution function whose quantile function is v_{ij}). This fact can be easily established by applying the Inverse Function Theorem and integration by substitution.

Remark 9. Proposition 2 tells us that, once v_{10} and v_{01} have been chosen, the problem of finding functions v_{00} and v_{11} satisfying the required constraints can be equivalently posed as the problem of finding a suitable "wheigh function" λ. Here is a partial converse to this result: suppose already given four non-decreasing, continuously differentiable "vertex functions" v_{ij} with $v_{ij}(0+) = 0$ and $v_{ij}(1-) = 1$ for $i, j \in \{0, 1\}$, satisfying the compatibility condition $v_{11}+v_{00} = v_{10}+v_{01}$. Also assume, for simplicity, that $v'_{10}(\tau)+v'_{01}(\tau) > 0$ for all $0 < \tau < 1$. Then, by tautology, we have

$$v'_{00} = (v'_{10} + v'_{01}) \times \frac{v'_{00}}{v'_{10} + v'_{01}}.$$

Thus, setting $\lambda \coloneqq v'_{00}/(v'_{10} + v'_{01})$, we conclude that $\lambda(\tau) \leqslant 1 = \int_0^1 (v'_{10}(u) + v'_{01}(u))\lambda(u)\, \mathrm{d}u$ for all τ, as $v'_{11} = (v'_{10} + v'_{01}) \cdot (1 - \lambda) \geqslant 0$.

The results and proposals presented above are implemented and numerically evaluated in the next section. In the last few pages, we intended to give a reasonable summary about what the entitled "further considerations" are. Before moving on to the more pragmatic part of this work, where we present our ensembles of simulations and study them within the latter "deeper specifications", we do need to formalize the estimation techniques for quantile regression that are implemented in this article.

Estimation procedures

Suppose $(X_t, Y_t)_{t\in\mathbb{Z}}$ satisfies the QADL dynamics (4), recalling the notation introduced in equation (5). An important property of quantile regression is that, for each $0 < \tau < 1$, it holds that

$$\beta(\tau) = \underset{b\in\mathbb{R}^{\dim Z}}{\arg\min}\, \mathbf{E}\, \rho_\tau(Y_0 - Z'_0 b). \tag{11}$$

where $2\rho_\tau(v) \coloneqq (2\tau - 1)v + |v|$ is the usual check function. This follows from the fact that $Q_Y(\tau) = \arg\min_{y\in\mathbb{R}} \rho_\tau(Y - y)$, which is well known in the literature, albeit customarily one only finds its proof for continuous Y. The general case is tackled in Hunter and Lange (2000).

The fact that, for each $0 < \tau < 1$, the parameter $\beta(\tau)$ is an arg min strongly suggests that it can be estimated by considering sample counterparts to the optimization problem in equation (11). In this section, we present the *standard quantile regression estimator* of Koenker and Bassett (1978), as well as the *convolution-type kernel quantile regression estimator* of Fernandes et al. (2021).

The standard QR estimator

In their seminal paper, Koenker and Bassett (1978) proposed what we refer in this article as the **standard (or canonical) quantile regression estimator**, defined as

$$\widehat{\beta}_n(\tau) \coloneqq \underset{b \in \mathbb{R}^{\dim Z}}{\arg\min} \; \frac{1}{n} \sum_{t=1}^{n} \rho_\tau(Y_t - Z_t' b), \quad 0 < \tau < 1.$$

It is also known through the established literature that $\sqrt{n}\big(\widehat{\beta}_n(\tau) - \beta(\tau)\big)$ is asymptotically centered normal with covariance matrix $\Sigma(\tau)$. These notions allow us to infer some drawbacks. First, the Bahadur-Kiefer representation for the standard QR estimator has a poor rate. Second, $\Sigma(\tau)$ is hard to estimate. In fact, both stem from lack of smoothness of objective function. This lack of smoothness is a very important concept for comprehending the following estimation procedure.

Convolution-type kernel smoothed QR regression estimator

As the sample objective function that the standard quantile regression estimator aims to minimize is not smooth, statistical inference is not straightforward. Fernandes et al. (2021) propose to smooth its objective function, thus presenting an alternative estimator: the **convolution-type kernel smoothed quantile regression estimator**. By applying a convolution on the check function against a standardized kernel function, the authors obtained a smoothed quantile regression estimator, which is given by

$$\widehat{\beta}_n^{\zeta}(\tau) \coloneqq \underset{b \in \mathbb{R}^{\dim Z}}{\arg\min} \; \frac{1}{n} \sum_{t=1}^{n} k_{\zeta(\tau)} * \rho_\tau(Y_t - Z_t' b), \quad 0 < \tau < 1, \tag{12}$$

where $k_h(v) \coloneqq k(v/h)/h$ is a kernel function, $\zeta \equiv \zeta_n = (\zeta_n(\tau) \colon 0 < \tau < 1)$ is a stochastic process with state space $[h_n, h^n]$ which captures the degree of smoothing, and $f * g(v) \coloneqq \int_{-\infty}^{\infty} f(u) g(v-u) \, \mathrm{d}u$ denotes the convolution of f

against g. We may consider $\widehat{\beta}_n^{\zeta}(\tau)$ as an estimator of the auxiliary smoothed parameter

$$\beta^{\zeta}(\tau) \coloneqq \underset{b\in\mathbb{R}^{\dim Z}}{\arg\min}\, \mathbf{E}\, k_{\zeta(\tau)} * \rho_\tau(Y_0 - Z_0' b), \quad 0<\tau<1,$$

which is uniformly close to $\beta(\tau)$ under mild assumptions, as shown in Fernandes et al. (2021)—thus settling the issue of bias introduced by smoothing.

Considering a cross-sectional context—that is, when $(Y_t, Z_t)_{t\in\mathbb{Z}}$ is *iid*—and under a set of mild regularity conditions, the authors obtained theoretical properties for this estimator. The first one is that its Bahadur-Kiefer representation has a better linearization rate than the standard quantile regression estimator: it holds that the statistic $\sqrt{n}\big(\widehat{\beta}_n^{\zeta}(\tau) - \beta^{\zeta}(\tau)\big)$ can be expressed as a sum of centered, *iid* random variables, uniformly bounded (in τ), plus a remainder term whose stochastic order dominates the $\mathcal{O}_{\mathbf{P}}$ rate in the Bahadur-Kiefer representation of standard quantile regression, uniformly for $\tau \in [\underline{\tau}, \overline{\tau}]$, as long as η has its sample paths lying in $[h_n, h^n]$ with a high probability, and under the assumption that $h_n, h^n \to 0$ at appropriate rates. Other important result concerns the asymptotic variance of the estimator. Again under regularity assumptions, it holds that the variance of the leading term in the Bahadur-Kiefer representation of $\sqrt{n}\big(\widehat{\beta}_n^{\zeta}(\tau) - \beta^{\zeta}(\tau)\big)$ is equal to $\Sigma(\tau) - c\eta(\tau)\,\mathbf{E}(XX'f(X'\beta(\tau)|X))^{-1}$ plus a remainder term that is small-oh in h^n, and this holds uniformly for $\tau \in [\underline{\tau}, \overline{\tau}]$. In the latter equation, $c>0$ is a constant depending only on the kernel k. Thus, as $\Sigma(\tau)$ is the covariance matrix of standard quantile regression estimator, it follows that the asymptotic variance of any linear combination of the components of $\widehat{\beta}_n^{\zeta}$ is smaller than the corresponding variance for $\widehat{\beta}_n$, and since the asymptotic bias is negligible, one can conclude that the Asymptotic Mean Squared Error (AMSE) of $\lambda'\widehat{\beta}_n^{\zeta}(\tau)$, for any $\lambda \in \mathbb{R}^{\dim Z}$, is smaller than $\mathrm{AMSE}\big(\lambda'\widehat{\beta}_n(\tau)\big)$. Last but not least, the authors derive an optimal bandwidth $(\eta(\tau)\colon\ 0<\tau<1)$ for which $\mathrm{AMSE}\big(\lambda'\widehat{\beta}_n^{\eta}(\tau)\big)$ is asymptotically minimal. What needs to be highlighted here is the fact that the AMSE of the smoothed estimator is lower than the AMSE of the canonical estimator by Koenker and Bassett (1978). This holds for cross sectional data, and one of our aims is to investigate whether these advantages remain valid in a time series context. In fact, as the estimation procedure for QADL models in Galvão et al. (2013) is based on the standard quantile regression estimator, one of the main goals of this work is to analyze the convolution-type kernel smoothed quantile regression estimator in a class of time series models. The next section is dedicated to this analysis. Through Monte Carlo simulations, we perform a numerical evaluation of both the standard and the smoothed estimators. Although our focus is explicit on QADL models, we try to go further on the simulated scenarios by incrementing more quantile levels and by studying alternative kinds of dynamics than those presented in the original QADL paper.

3 Monte Carlo study

As previously mentioned, this work focuses on the analysis and simulation studies of the QADL model. We start this section by exploring the data generating processes already studied in Galvão et al. (2013). In order to assess the performance of the smoothed QR estimator, we replicate some simulated scenarios, as well as we introduce additional quantile levels on the numerical evaluation. Then, aiming to assess the performance of the competing estimators when considering alternative QADL dynamics, we go deeper into the "further considerations" presented in section 2. We computationally implement the ideas presented in Proposition 2 and numerically evaluate both the canonical and the smoothed QR estimators within six different scenarios. These scenarios are distinguishable through the choice of the weight function λ.

Experiment design

We conducted a comprehensive Monte Carlo study with two main goals: to investigate analytical characterizations for the data generating process of QADL dynamics, and to explore these dynamics while evaluating the performance of the smoothed estimator of Fernandes et al. (2021) in a time series framework. To evaluate the smoothed estimator's performance within quantile regression time series models, we first replicate the simulation found in Galvão et al. (2013) introducing additional quantile levels.[2] This first ensemble of Monte Carlo simulations—which are slight adaptations from those presented by Galvão et al. (2013)—compose what we refer to this article as *existing scenarios*.

As for the second ensemble of simulations, the reader might have guessed by now that it concerns the construction scheme introduced in Proposition 2. As a matter of fact, to numerically evaluate the estimation procedures previously discussed, we do the necessary computational implementations combining elements of both Propositions 1 and 2, with the latter representing the core for the construction of our simulated scenarios. These scenarios comprehend different choices of $\lambda(\tau)$—which we refer to as *new scenarios*. We describe the first and second ensemble of simulations in the following subsections.

Existing scenarios

Inspired by the data generating processes employed in Galvão et al. (2013), we generate data from the model

$$Y_t = a_0 + a_1 Y_{t-1} + \vartheta_1 X_{t-1} + \vartheta_2 X_{t-2} + (\delta + \gamma X_t)\epsilon_t. \tag{13}$$

We consider the following scenarios:

Location-shift model 1. We set the parameter vector to

$$\begin{bmatrix} a_0 & a_1 & \vartheta_1 & \vartheta_2 & \delta & \gamma \end{bmatrix}' = \begin{bmatrix} 0.0 & 0.5 & 0.5 & 0.5 & 1.0 & 0.0 \end{bmatrix}', \tag{14}$$

with the the *iid* disturbances $(\epsilon_t)_{t\in\mathbb{Z}}$ following a standard Gaussian distribution, and with ϵ_t generated independently from $\mathscr{F}_{t-1}$ for all t. Additionally, we generate the exogenous covariates $(X_t)_{t\in\mathbb{Z}}$ as an *iid* sequence having the same marginal distribution as the innovation terms, with $X_t \perp (\mathscr{F}_{t-1}, \varepsilon_t)$ for all t.

[2] In their paper, Galvão et al. (2013) consider $\tau \in \{0.1, 0.25, 0.5, 0.75, 0.9\}$. Our expanded grid considers $\tau \in \{0.01, 0.03, 0.05, \ldots, 0.95, 0.97, 0.99\}$, with a total of 50 quantile levels.

Location-shift model 2. Again we set the parameter vector as in eq. (14), but now with *iid* disturbances $(\epsilon_t)_{t\in\mathbb{Z}}$ following a t-Student distribution with 3 degrees of freedom, and with ϵ_t generated independently from $\mathscr{F}_{t-1}$ for all t. Moreover, we generate the exogenous covariates $(X_t)_{t\in\mathbb{Z}}$ as an *iid* sequence having the same marginal distribution as the innovation terms, with $X_t \perp (\mathscr{F}_{t-1}, \varepsilon_t)$ for all t.

Location-scale model 1. We set the parameter vector to

$$\begin{bmatrix}a_0 & a_1 & \vartheta_1 & \vartheta_2 & \delta & \gamma\end{bmatrix}' = \begin{bmatrix}0.0 & 0.5 & 0.5 & 0.5 & 0.0 & 0.2\end{bmatrix}', \tag{15}$$

with the the *iid* disturbances $(\epsilon_t)_{t\in\mathbb{Z}}$ following a standard Gaussian distribution, and with ϵ_t generated independently from $\mathscr{F}_{t-1}$ for all t. Additionally, we generate the exogenous covariates $(X_t)_{t\in\mathbb{Z}}$ as an *iid* sequence of χ^2_3 random variables (the non-negative distribution avoids the problem of quantile crossing), with $X_t \perp (\mathscr{F}_{t-1}, \varepsilon_t)$ for all t.

Location-scale model 2. Once more we set the parameter vector as in eq. (15), now with the the *iid* disturbances $(\epsilon_t)_{t\in\mathbb{Z}}$ following a t-Student distribution with 3 degrees of freedom, and with ϵ_t generated independently from $\mathscr{F}_{t-1}$ for all t. The sequence $(X_t)_{t\in\mathbb{Z}}$ is generated in the same manner as in the location-scale model 1.

Following Galvão et al. (2013), in every scenario we set $Y_{-100} = 0$ and generate Y_t for $t \in \{-99, \ldots, 1, \ldots, n\}$ according to equation (13), using observations from 1 to n for estimation. In the simulations, we considered sample sizes of $n = 100$ and $n = 1{,}000$, and set the number of replications to 5,000. It is important to notice that the coefficients in (13) do not coincide with the functional coefficients in the QADL representation (4). Indeed, for both the location-shift models, we have

$$Q_{Y_t}(\tau|\mathscr{F}_{t-1}) = [a_0 + Q_\epsilon(\tau)] + a_1 Y_{t-1} + \vartheta_1 X_{t-1} + \vartheta_2 X_{t-1}, \quad 0 < \tau < 1,$$

thus showing that the "functional intercept" in a QADL representation of these models is $\alpha_0(\cdot) = a_0 + Q_\epsilon(\cdot)$ and not a_0. Similarly, for the location-scale models we readily see that the identity

$$Q_{Y_t}(\tau|\mathscr{F}_{t-1}) = a_0 + a_1 Y_{t-1} + [\vartheta_1 + \gamma Q_\epsilon(\tau)]X_{t-1} + \vartheta_2 X_{t-1}$$

holds for each $0 < \tau < 1$, so the functional coefficient of X_{t-1} in the corresponding QADL representation is $\theta_1(\cdot) = \vartheta_1 + \gamma Q_\epsilon(\cdot)$ and not ϑ_1. In view this, when evaluating the estimators and computing the comparison metrics, in the location-shift scenarios we consider

- $\widehat{\alpha}_0(\cdot)$ as an estimator for $a_0 + Q_\epsilon(\cdot) = Q_\epsilon(\cdot)$;
- $\widehat{\alpha}_1(\cdot)$ as an estimator for $a_1 = 0.5$;
- $\widehat{\theta}_j(\cdot)$ as an estimator for $\vartheta_j = 0.5$, $j = 1, 2$.

On the other hand, for the location-scale scenarios we consider

- $\widehat{\alpha}_0(\cdot)$ as an estimator for $a_0 = 0.0$;
- $\widehat{\alpha}_1(\cdot)$ as an estimator for $a_1 = 0.5$;
- $\widehat{\theta}_1(\cdot)$ as an estimator for $\vartheta_1 + \gamma Q_\epsilon(\cdot) = 0.5 + 0.2Q_\epsilon(\cdot)$;
- $\widehat{\theta}_2(\cdot)$ as an estimator for $\vartheta_2 = 0.5$.

We compare the estimators, with the quantile level τ varying in the grid $\{0.01, 0.03, 0.05, \ldots, 0.95, 0.97, 0.99\}$, in terms of bias, variance, mean squared error (MSE), and—in order to provide a global picture—the mean integrated squared error (MISE), given by

$$\text{MISE} = \mathbf{E}\int_0^1 \left\|\widehat{\beta}(\tau) - \beta(\tau)\right\|^2 \mathrm{d}\tau,$$

where $\|\cdot\|$ is the Euclidian norm in $\mathbb{R}^{\dim Z}$, and where $\widehat{\beta}$ varies among the competing estimators.

New scenarios: novel characterizations of QADL dynamics

Establishing a link between the ideas proposed by Horta (2021) and the novel contributions presented in this paper, our second ensemble of simulations aimed to explore the aforementioned "further considerations" about QADL dynamics. We conducted a series of Monte Carlo experiments studying novel proposals of QADL dynamics, specifically those related to Proposition 2. In addition to that, we evaluated the performance of the smoothed estimator of Fernandes et al. (2021) in a time series framework where these dynamics were indeed contemplated. Lastly, we compared the smoothed QR estimator's results to those obtained via the standard QR estimator.

Remembering the ideas behind Propositions 1 and 2, we choose the functions v_{10} and v_{01}, for convenience, to be quantile functions of the Beta distribution, with parameters (a_{10}, b_{10}) and (a_{01}, b_{01}), respectively. As discussed in Remark 9, Proposition 2 tells us that, once we have v_{10} and v_{01}, the problem of finding functions v_{00} and v_{11} satisfying the required constraints can be equivalently posed as the problem of finding a suitable "weight function" λ, where $\lambda\colon (0,1) \to \mathbb{R}$ is any non-negative, continuous function satisfying $0 < \int_0^1 (v'_{10}(u) + v'_{01}(u))\lambda(u)\,\mathrm{d}u < \infty$. The core of our simulated scenarios is built upon the choice of λ. Actually, each scenario that we explored was obtained through a different choice for λ.[3] We then computationally implemented the necessary steps in Proposition 2 so that we could get the function $v_{00}\colon (0,1) \to \mathbb{R}$ defined through equation (9) and the function $v_{11} := v_{10} + v_{01} - v_{00}$. Still important, we also numerically guaranteed that, for every choice of λ we made, the condition $\lambda(\tau) \leqslant \int_0^1 (v'_{10}(u) + v'_{01}(u))\lambda(u)\mathrm{d}u$ holds for each τ such that $v'_{10}(\tau) + v'_{01}(\tau) \neq 0$. As mentioned above, we chose the functions v_{10} and v_{01} to be quantile functions of the Beta distribution. As these quantile functions generally do not admit a simple, closed form representation, in our implementations, we bypass the issue of computing v'_{10} and v'_{01} by numerically calculating the integrals $\int_0^\tau (v'_{10}(u) + v'_{01}(u))\lambda(u)$ and $\int_0^1 (v'_{10}(u) + v'_{01}(u))\lambda(u)$ through integration by parts, via the identity

$$\int_0^\tau \lambda(u)Q'(u)\mathrm{d}u = \lambda(\tau)Q(\tau) - \int_0^\tau \lambda'(u)Q(u)\mathrm{d}u, \quad 0 < \tau < 1$$

with $Q = (v_{10} + v_{01})$. We evaluated six different choices for the weight function $\lambda\colon (0,1) \to \mathbb{R}$, which are:

[3] One may choose any λ satisfying the constraints in Proposition 2 and it will imply a distinct dynamics. In this work, as the number of possible choices can be easily extrapolated, we focused on six different choices of λ.

Root functions. $\lambda(\tau) = \sqrt[r]{\tau}$, with $r = 2$, $r = 3$ and $r = 6$.
Sigmoid function. $\lambda(\tau) = (1 + \mathrm{e}^{-\tau})^{-1}$.
Cosine function. $\lambda(\tau) = \cos(\tau)$.
Hacovercosin function. $\lambda(\tau) = 2^{-1}(\sin(\tau) + 1)$.

We simulate the model in (8) for each of the six specifications described above, with sample sizes $n \in \{100, 500, 1{,}000\}$, resulting in a total of 18 scenarios. For the Beta quantile functions v_{10} and v_{01}, we set the parameters as $(a_{10}, b_{10}) = (1, 3)$ and $(a_{01}, b_{01}) = (5, 1)$ respectively. In all cases we initialize X_{-100} and Y_{-100} as independent, standard uniform random variables, and generate Y_t and X_t for $t = -99, \ldots, 1, \ldots, n$ according to (7),[4] using the observations from 1 to n for estimation. In order to evaluate the competing estimators, we generate 5,000 replications in each scenario. In each replication, we compute the smoothed QR estimator of Fernandes et al. (2021) and the canonical QR estimator, employing the same grid of τ's and the same metrics for comparison, as in the preceding section.

Bandwidth selection

As mentioned in Section 2, to compute $\widehat{\beta}_n^{\zeta}(\tau)$ in (12), one needs to select the "bandwidth process" ζ. Following Fernandes et al. (2021), we use Silverman's rule-of-thumb bandwidth (Silverman, 1986), which can be expressed, in the present context, through

$$\zeta^*(\tau) = \frac{1.06}{\sqrt[5]{n}}\, \widehat{s}(\tau), \quad \tau \in (0, 1) \tag{16}$$

where n is the sample size and, for each quantile level $\tau \in (0, 1)$, a different $\widehat{s}(\tau)$ is computed using the following algorithm:

1. Compute the canonical estimator $\widehat{\beta}_n(\tau)$ to obtain the corresponding "residuals" $\widehat{\epsilon}_i(\tau) := Y_i - Z_i' \widehat{\beta}_n(\tau)$, for $i \in \{1, \ldots, n\}$;

[4] A choice for Q_t must be made, and here we employed $Q_t(\cdot, x_{t-1}, \ldots, x_0, y_t, \ldots, y_0) = \alpha_0(\cdot) + \alpha_1(\cdot) x_{t-1} + \theta_1(\cdot) y_{t-1}$. Thus, for all t the random variable X_t is equal in distribution to Y_t; moreover, X_t and Y_t are independent conditionally on $\mathscr{F}_{t-1}$.

2. Calculate the sample interquartile range $\mathrm{iq}(\tau)$ corresponding to the residuals $\widehat{\epsilon}_1(\tau), \ldots, \widehat{\epsilon}_n(\tau)$;
3. Compute the sample standard deviation $\widehat{\sigma}(\tau)$, from $\widehat{\epsilon}_1(\tau), \ldots, \widehat{\epsilon}_n(\tau)$;
4. Obtain $\widehat{s}(\tau)$, given by $\widehat{s}(\tau) := \min\left\{0.7199528 \times \mathrm{iq}(\tau),\, \widehat{\sigma}(\tau)\right\}$.

Evidently, the bandwidth ζ^* in (16) is τ-dependent, i.e., we have a different smoothing parameter for each $\tau \in (0, 1)$. Furthermore, ζ^* is data-driven, that is, it varies with each replication, as it depends on the sample generated. We point out that, by construction, the rule-of-thumb bandwidth ζ^* is indicated for cases where there is normality in the data—see Silverman (1986)—which does not entirely extend for the cases presented in our simulations. Nevertheless, Fernandes et al. (2021) signal that this smoothing parameter can perform well even when the Gaussianity assumption is violated. In order to explore the estimator's performance for other smoothing parameters, we chose to evaluate a series of bandwidths proportional to ζ^*. For the *existing scenarios*, in the location-shift cases, we consider the $\widehat{\beta}_n^{\zeta}(\tau)$ estimator computed with $\zeta \in \{\zeta^*/4, \zeta^*/2, \zeta^*, 2\zeta^*, 4\zeta^*\}$, whereas in the location-scale versions we consider $\widehat{\beta}_n^{\zeta}(\tau)$ computed with $\zeta \in \{\zeta^*/16, \zeta^*/8, \zeta^*/4, \zeta^*/2, \zeta^*, 2\zeta^*, 4\zeta^*, \zeta^*, 8\zeta^*, 16\zeta^*\}$.[5] For the *new scenarios*, in all simulations we consider $\widehat{\beta}_n^{\zeta}(\tau)$ computed with $\zeta \in \{\zeta^*/4, \zeta^*/2, \zeta^*, 2\zeta^*, 4\zeta^*\}$.

Monte Carlo study results: existing scenarios

Location-shift models

Considering the first part of this ensemble of simulations, these scenarios provided the most straightforward conclusions. It can be seen in Figures 1 and 2 that $\widehat{\beta}_n^{\zeta^*}(\tau)$ systematically outperforms $\widehat{\beta}_n(\tau)$ in terms of MSE and variance for almost all levels of ζ^*—the exception being the α_0 coefficient. As it was expected due to the mathematical construction of the smoothed estimator, its variance seems to be uniformly lower than the variance of the canonical estimator, in the *location-shift* cases. When analyzing the global

[5] This distinction is done due to the results of preliminary simulations, which indicated that additional ζ's produced estimates with smaller MSE.

metric (MISE) in Table 1, it can be seen that, when jointly evaluating the estimates for all the coefficients in the simulated model, the smoothed QR estimator performs better than the canonical QR estimator for most levels of ζ^*. Also, as the sample size n increases, MSE gets lower and estimates get closer to their real values, what—numerically, at least—indicates a proper asymptotic behavior for $\widehat{\beta}_n^{\zeta^*}(\tau)$ in a time series data context. The optimal bandwidth was obtained with $\zeta^*/2$. There is a trade-off, however. As expected, biases for the smoothed estimator tend to be higher when compared to the canonical estimator.

Location-shift-scale models

These scenarios did not provide conclusions as straightforward as the preceding ones. However, some valuable insights were derived from them. First, although in most cases the MSE and the variance of the smoothed estimator are not uniformly lower when compared to the canonical estimator along the quantiles, it can be seen in Figures 3 and 4 that there is an optimal bandwidth ($\zeta^*/2$) in which such metrics are visually indistinguishable—which in turn indicates that, given a proper choice of (optimal) bandwidth, $\widehat{\beta}_n^{\zeta^*}(\tau)$ can perform similarly to $\widehat{\beta}_n(\tau)$. The global metrics (MISE) in Table 1 show that, for the *location-scale model 1*, $\widehat{\beta}_n^{\zeta^*}(\tau)$ in general does not outperform $\widehat{\beta}_n(\tau)$. For the *location-scale model 2*, analogous results hold. However, considering the *location-scale model 2*, in smaller sample sizes the smoothed estimator seems to behave as well as—or even better—than the canonical estimator, given a proper choice of bandwidth. As the sample size n increases, MSE gets lower and estimates get closer to their real values, what can indicate a proper asymptotic behavior for $\widehat{\beta}_n^{\zeta^*}(\tau)$ in a time series data context, even when the error term is related to a covariate such as is the case in (13) whenever $\gamma > 0$.

Table 1: Global metrics (MISE): $\widehat{\beta}_n(\tau)$ and $\widehat{\beta}_n^{\zeta^*}(\tau)$.

		Location-shift models		Location-scale models	
		Model 1	Model 2	Model 1	Model 2
$\widehat{\beta}_n$	$n = 100$	0.1024	0.7421	0.0407	0.3497
	$n = 1000$	0.0099	0.1180	0.0019	0.0125
$\widehat{\beta}_n^{\zeta^*}$	$n = 100$	0.0730	0.6464	0.0937	0.4523
	$n = 1000$	0.0085	0.1150	0.0207	0.0621
$\widehat{\beta}_n^{2\zeta^*}$	$n = 100$	0.1468	0.8687	0.3320	1.2805
	$n = 1000$	0.3615	0.1320	0.1065	0.2912
$\widehat{\beta}_n^{4\zeta^*}$	$n = 100$	0.9844	3.3552	1.4184	5.2720
	$n = 1000$	0.2063	0.3615	0.5408	1.4206
$\widehat{\beta}_n^{\frac{\zeta^*}{2}}$	$n = 100$	0.0810	0.6758	0.0463	0.3292
	$n = 1000$	0.0085	0.1154	0.0051	0.0202
$\widehat{\beta}_n^{\frac{\zeta^*}{4}}$	$n = 100$	0.0904	0.7057	0.0529	0.3348
	$n = 1000$	0.0091	0.1163	0.0064	0.0159
$\widehat{\beta}_n^{8\zeta^*}$	$n = 100$			6.0240	21.8913
	$n = 1000$			2.5314	6.4475
$\widehat{\beta}_n^{16\zeta^*}$	$n = 100$			24.7382	89.0743
	$n = 1000$			10.9189	27.3419
$\widehat{\beta}_n^{\frac{\zeta^*}{8}}$	$n = 100$			0.1074	0.3801
	$n = 1000$			0.0104	0.0191
$\widehat{\beta}_n^{\frac{\zeta^*}{16}}$	$n = 100$			0.1162	0.4035
	$n = 1000$			0.0104	0.0200

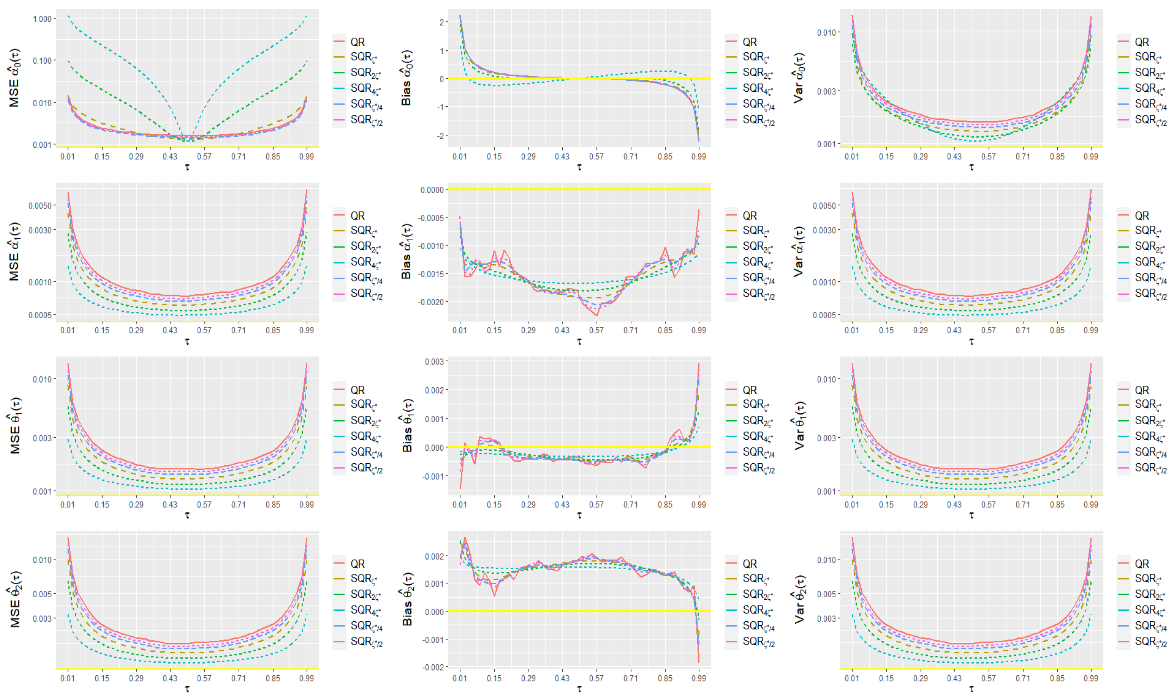

Figure 1: *Location-shift model 1*: Performance of both the canonical quantile regression estimator (QR) by Koenker and Bassett (1978) and the smoothed quantile regression estimator (SQR) by Fernandes et al. (2021) in terms of their MSE, bias and variance. Each row in the above table shows the mentioned metrics for the respective coefficients in the model. Sample size of $n = 1{,}000$. MSE and variance are on a logarithmic scale.

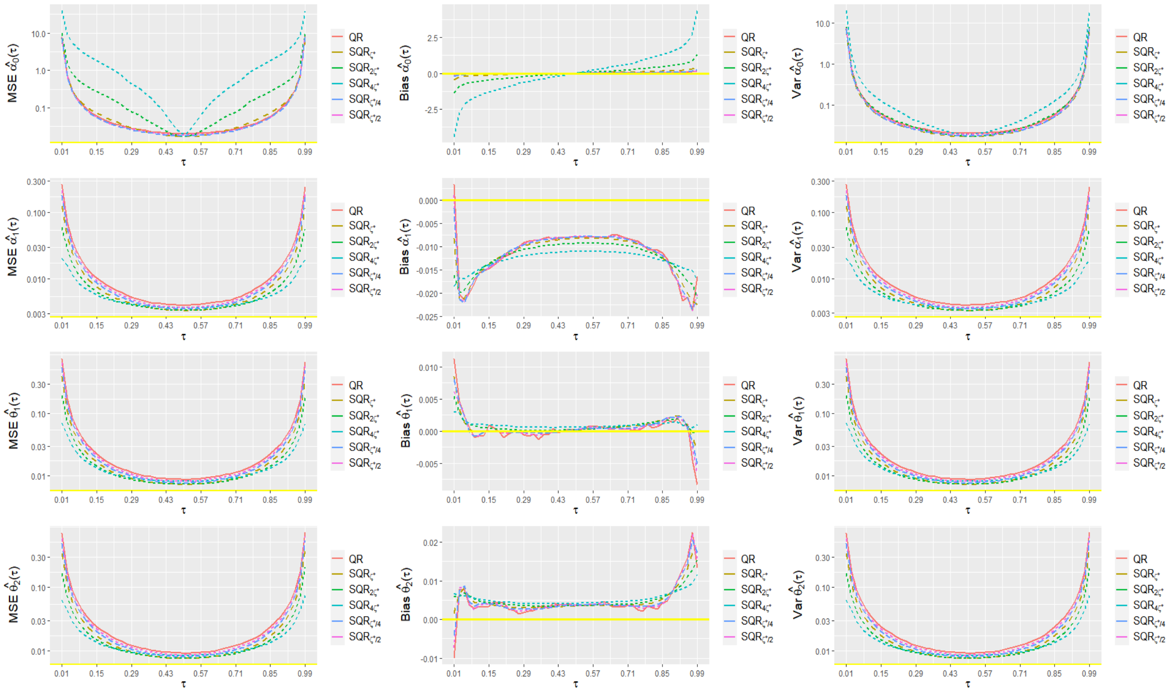

Figure 2: *Location-shift model 2*: Performance of both the canonical quantile regression estimator (QR) by Koenker and Bassett (1978) and the smoothed quantile regression estimator (SQR) by Fernandes et al. (2021) in terms of their MSE, bias and variance. Each row in the above table shows the mentioned metrics for the respective coefficients in the model. Sample size of $n = 1{,}000$. MSE and variance are on a logarithmic scale.

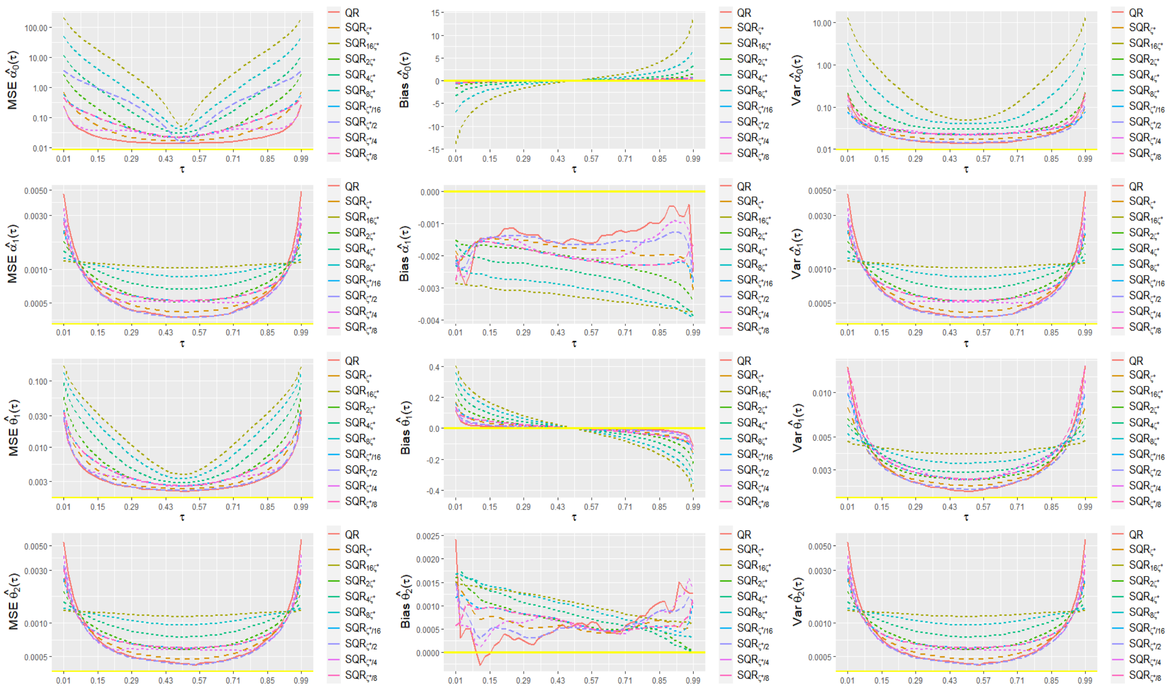

Figure 3: *Location-scale model 1*: Performance of both the canonical quantile regression estimator (QR) by Koenker and Bassett (1978) and the smoothed quantile regression estimator (SQR) by Fernandes et al. (2021) in terms of their MSE, bias and variance. Each row in the above table shows the mentioned metrics for the respective coefficients in the model. Sample size of $n = 100$. MSE and variance are on a logarithmic scale.

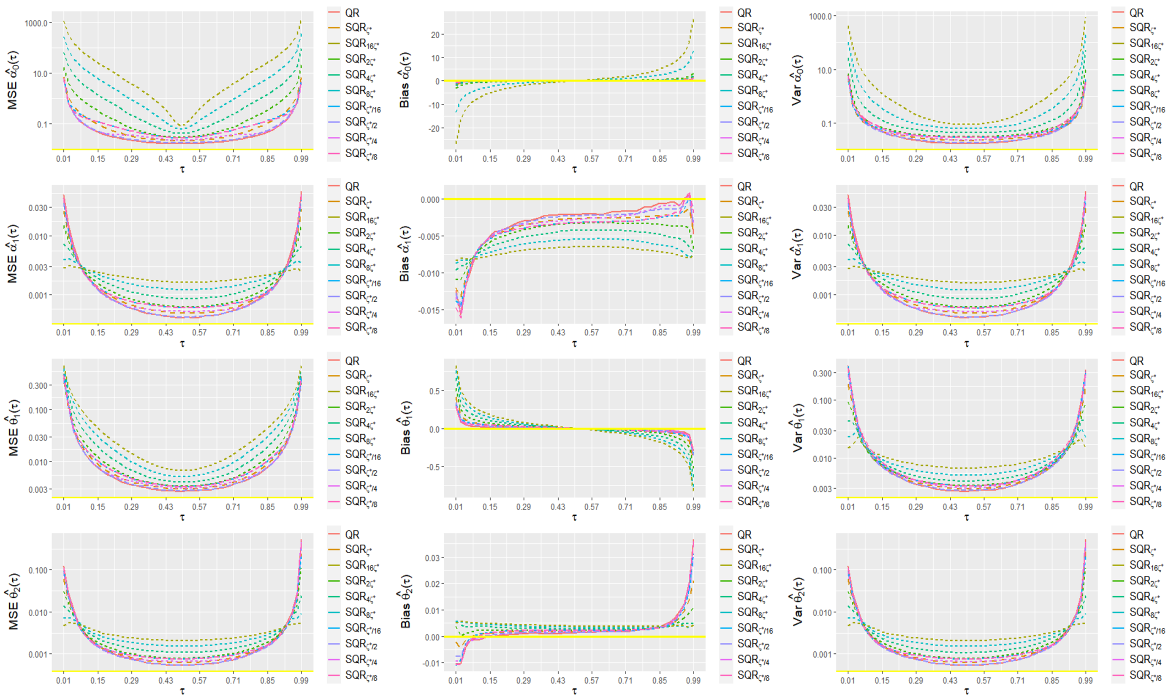

Figure 4: *Location-scale model 2*: Performance of both the canonical quantile regression estimator (QR) by Koenker and Bassett (1978) and the smoothed quantile regression estimator (SQR) by Fernandes et al. (2021) in terms of their MSE, bias and variance. Each row in the above table shows the mentioned metrics for the respective coefficients in the model. Sample size of $n = 100$. MSE and variance are on a logarithmic scale.

Monte Carlo study results: new scenarios

Root functions

In the second ensemble of simulations, six different choices λ were employed. The class of **root functions** is composed of three functions.[6] It can be seen in Figure 5 that $\widehat{\beta}_n^{\zeta^*}(\tau)$ uniformly outperforms $\widehat{\beta}_n(\tau)$ in terms of MSE and variance along the grid of τ's for almost all levels of ζ^*—the exception being, once again, the α_0 coefficient. When analyzing the global metric (MISE) in Table 1, it can be seen that there is only one bandwidth in which the smoothed QR estimator outperforms the canonical QR estimator in all scenarios. However, this can be interpreted as a consistent evidence that there might always be an optimal bandwidth able to offer such a performance. As a matter of fact, this conclusion about the optimal bandwidth was also suggested in the following simulation results. Still, as the sample size n increases, MSE gets lower and estimates get closer to their real values, it is possible to infer that the asymptotic behavior for $\widehat{\beta}_n^{\zeta^*}(\tau)$ in a time series data context is similar to what was expected: lower variance and lower MSE. The optimal bandwidth was obtained with $2\zeta^*$. The trade-off between the latter metrics being lower and the bias being higher persists.

Sigmoid function

When $n = 1{,}000$, the scenario defined by λ being the Sigmoid function provided very good results. As the sample gets larger, in general the better performance of an estimator becomes less visible in a plot. In this scenario, however, besides all the coefficients having a uniformly lower variance for all levels of ζ^*, it can be seen that two out of the three coefficients had MSE uniformly lower as well. The optimal bandwidth was obtained with $2\zeta^*$ and the results presented in the last scenario remained valid. As expected, biases for the smoothed estimator tend to be higher when compared to the canonical estimator.

[6] The root functions are $\sqrt{\tau}$, $\sqrt[3]{\tau}$, $\sqrt[6]{\tau}$. In the set of graphs, we brought only the plot of $\sqrt[6]{\tau}$, as results are pretty much the same for the three root functions we tested.

Cosine function

Results for the Cosine function were analogous to those mentioned in the preceding scenarios. The optimal bandwidth was obtained with $2\zeta^*$. Amongst the six functions that were tested, the cosine was the less computationally expensive one throughout the Monte Carlo experiments.

Hacovercosine function

As for the Hacovercosine function, it can be seen that results are analogous once more. The optimal bandwidth was $2\zeta^*$. Even though the plots for all the tested functions are similar, one can see how a different choice of λ might affect the evaluated metrics along the quantiles. As the functions must satisfy the same constraints, they share analytical properties, which is the reason why the graphs may be so similar.

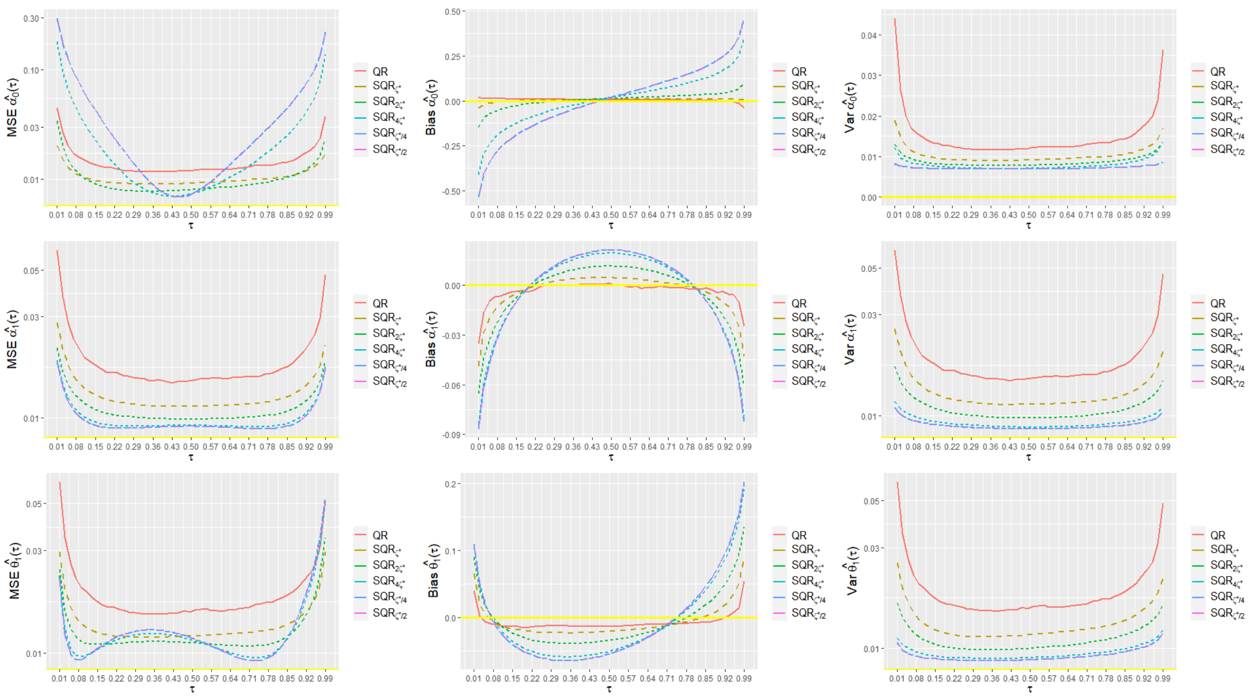

Figure 5: Performance of both the canonical quantile regression estimator (QR) of Koenker and Bassett (1978) and the smoothed quantile regression estimator (SQR) of Fernandes et al. (2021) in terms of their MSE, bias and variance. Each row in the above table shows the mentioned metrics for the respective coefficients in the model (8) defined through Proposition 2 when $\lambda(\tau) = \sqrt[6]{\tau}$. Sample size of $n = 100$. MSE and variance are on a logarithmic scale.

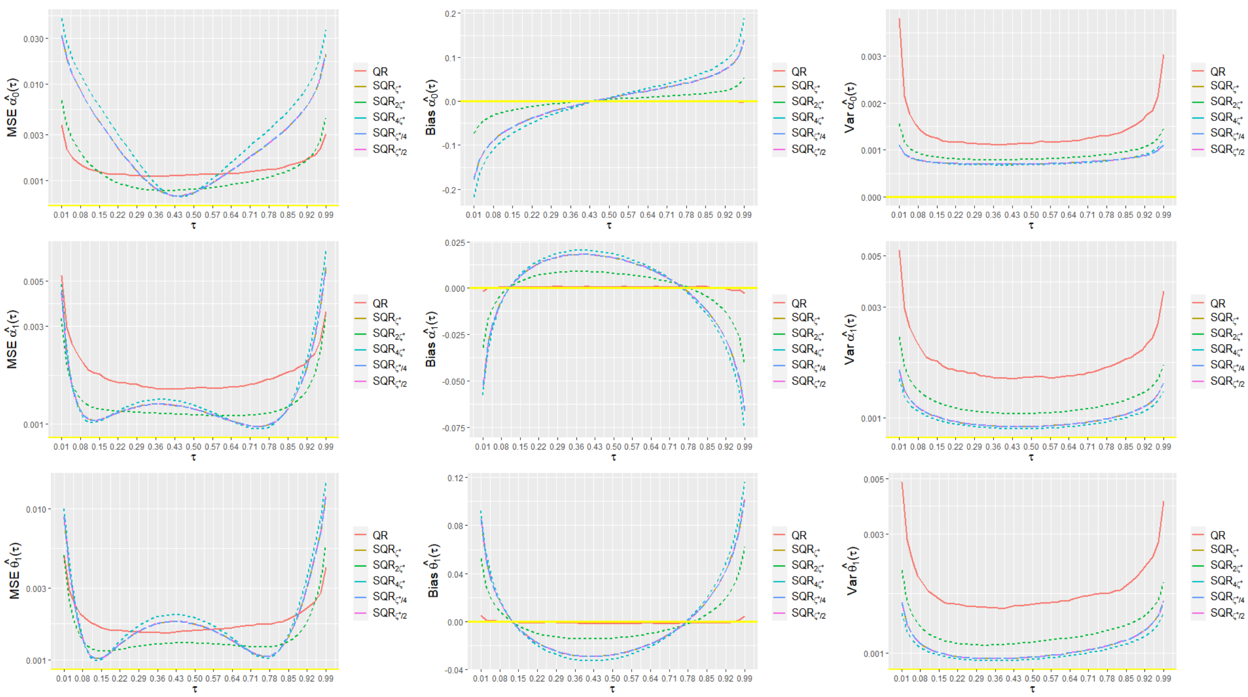

Figure 6: Performance of both the canonical quantile regression estimator (QR) of Koenker and Bassett (1978) and the smoothed quantile regression estimator (SQR) of Fernandes et al. (2021) in terms of their MSE, bias and variance. Each row in the above table shows the mentioned metrics for the respective coefficients in the model (8) defined through Proposition 2 when $\lambda(\tau) = (1 + e^{-\tau})^{-1}$. Sample size of $n = 1{,}000$. MSE and variance are on a logarithmic scale.

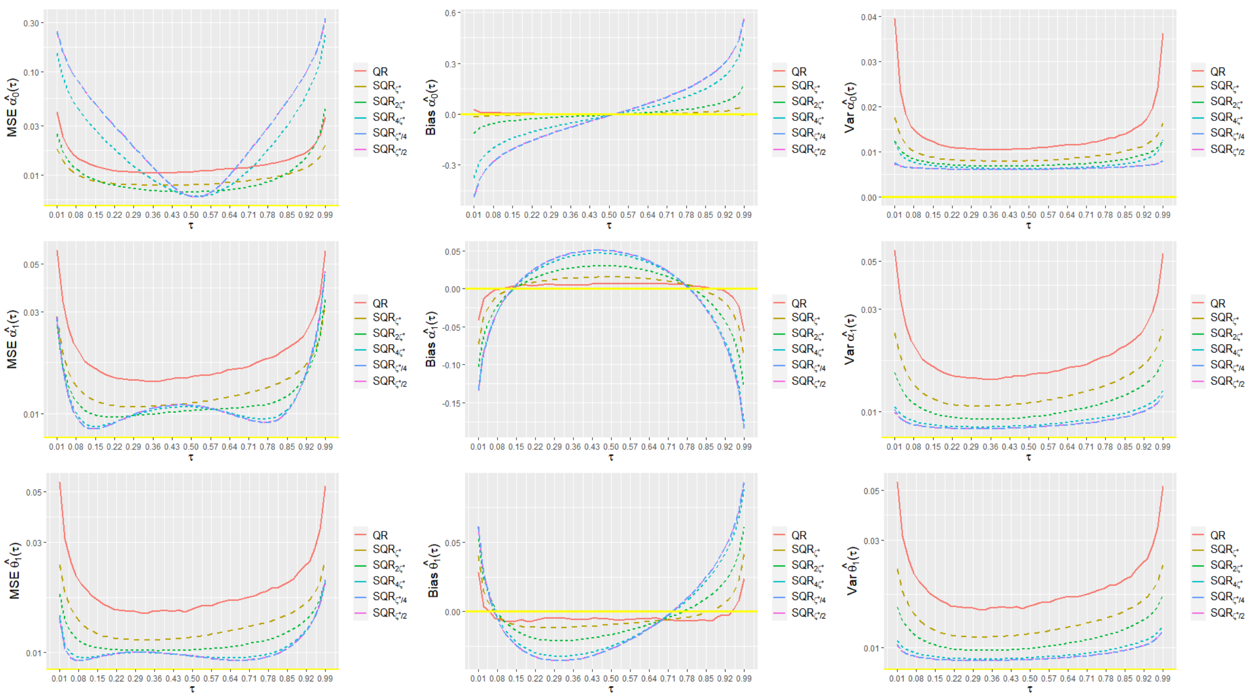

Figure 7: Performance of both the canonical quantile regression estimator (QR) of Koenker and Bassett (1978) and the smoothed quantile regression estimator (SQR) of Fernandes et al. (2021) in terms of their MSE, bias and variance. Each row in the above table shows the mentioned metrics for the respective coefficients in the model (8) defined through Proposition 2 when $\lambda(\tau) = \cos(\tau)$. Sample size of $n = 100$. MSE and variance are on a logarithmic scale.

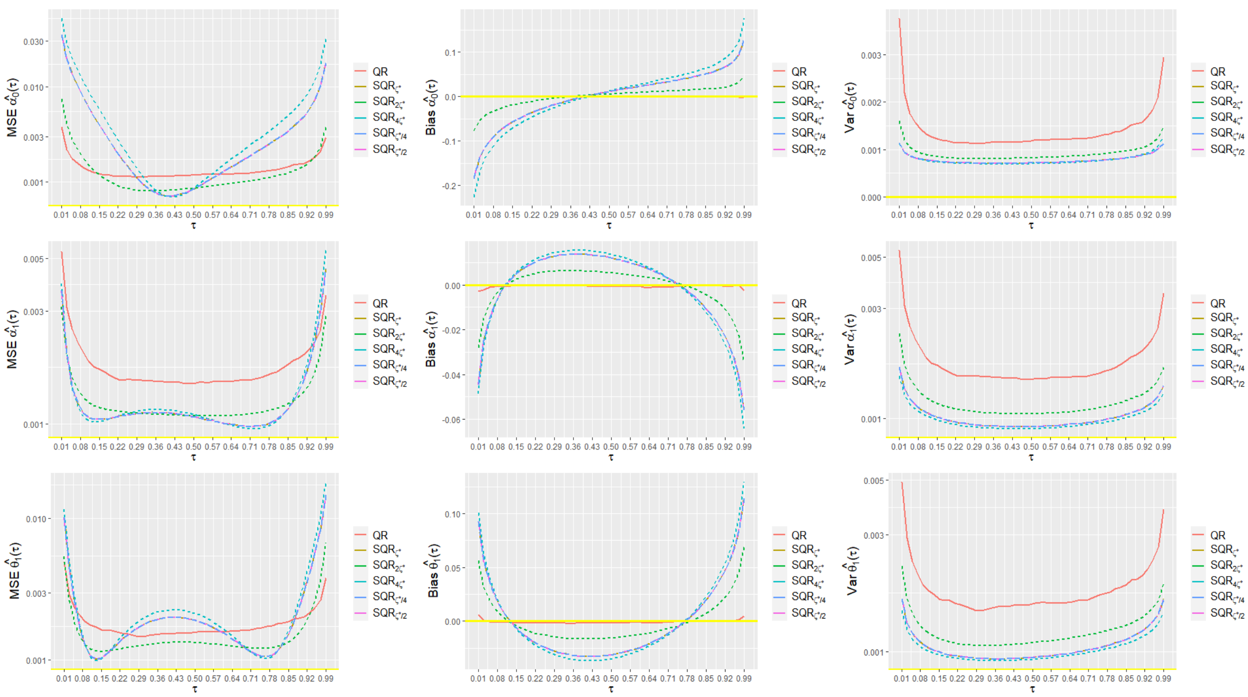

Figure 8: Performance of both the canonical quantile regression estimator (QR) of Koenker and Bassett (1978) and the smoothed quantile regression estimator (SQR) of Fernandes et al. (2021) in terms of their MSE, bias and variance. Each row in the above table shows the mentioned metrics for the respective coefficients in the model (8) defined through Proposition 2 when $\lambda(\tau) = 2^{-1}(\sin(\tau) + 1)$. Sample size of $n = 1{,}000$. MSE and variance are on a logarithmic scale.

Table 2: Global metrics (MISE): $\widehat{\beta}_n(\tau)$ and $\widehat{\beta}_n^{\zeta^*}(\tau)$.

		$\sqrt{\tau}$	$\frac{1}{1+\mathrm{e}^{-\tau}}$	$\sqrt[3]{\tau}$	$\cos(\tau)$	$\sqrt[6]{\tau}$	$\frac{\sin(\tau)+1}{2}$
$\widehat{\beta}_n$	$n = 100$	0.0486	0.0544	0.0507	0.0530	0.0538	0.0534
	$n = 500$	0.0088	0.0103	0.0095	0.0100	0.0100	0.0101
	$n = 1000$	0.0043	0.0051	0.0046	0.0050	0.0049	0.0050
$\widehat{\beta}_n^{\zeta^*}$	$n = 100$	0.0345	0.0374	0.0355	0.0366	0.0374	0.0370
	$n = 500$	0.0133	0.0118	0.0127	0.0112	0.0120	0.0118
	$n = 1000$	0.0083	0.0078	0.0076	0.0079	0.0073	0.0078
$\widehat{\beta}_n^{2\zeta^*}$	$n = 100$	0.0326	0.0345	0.0330	0.0339	0.0341	0.0342
	$n = 500$	0.0075	0.0079	0.0076	0.0077	0.0076	0.0078
	$n = 1000$	0.0041	0.0042	0.0041	0.0042	0.0041	0.0042
$\widehat{\beta}_n^{4\zeta^*}$	$n = 100$	0.0491	0.0535	0.0485	0.0536	0.0499	0.0530
	$n = 500$	0.0158	0.0165	0.0151	0.0166	0.0148	0.0164
	$n = 1000$	0.0102	0.0101	0.0094	0.0101	0.0091	0.0101
$\widehat{\beta}_n^{\frac{\zeta^*}{4}}$	$n = 100$	0.0676	0.0727	0.0667	0.0725	0.0680	0.0720
	$n = 500$	0.0156	0.0155	0.0155	0.0148	0.0143	0.0154
	$n = 1000$	0.0083	0.0078	0.0076	0.0078	0.0073	0.0078
$\widehat{\beta}_n^{\frac{\zeta^*}{2}}$	$n = 100$	0.0676	0.0727	0.0667	0.0725	0.0680	0.0720
	$n = 500$	0.0156	0.0155	0.0155	0.0148	0.0143	0.0154
	$n = 1000$	0.0083	0.0078	0.0076	0.0079	0.0073	0.0078

4 Final remarks

In this paper, we aimed to provide two main contributions to the quantile regression literature. First, we investigated analytical characterizations for the data generating process of QADL dynamics. Second, we conducted a massive Monte Carlo study exploring these dynamics while evaluating the performance of the smoothed quantile regression estimator of Fernandes et al. (2021) in a time series framework, particularly in the class of QADL models of Galvão et al. (2013). Besides following the standard procedures needed to study the estimator's performance in a time series context, we established propositions concerning these novel characterizations for QADL dynamics. As for the Monte Carlo study, we obtained numerical results that corroborate most of the theoretical assumptions concerning the smoothed QR estimator of Fernandes et al. (2021). In this study, we defined the process ζ's as a series of bandwidths proportional to the data-driven bandwidth ζ^* of Silverman (1986), which worked well. Confirming our suspicions, in terms of MSE and variance, given a proper choice of bandwidth, $\beta_n^{\zeta^*}(\tau)$ systematically outperformed $\beta_n(\tau)$ in most of the explored scenarios—except for the *location-scale* versions in the first ensemble of simulations, each scenario presented at least one optimal bandwidth, in which the smoothed estimator outperformed the canonical one. The latter result extends for both the *existing* and the *new* scenarios. The price to pay for $\beta_n^{\zeta^*}(\tau)$ having a smaller MSE and a lower variance was a higher bias in most of the studied scenarios. This fact was expected, though.

To sum up, our results indicate that the smoothed estimator indeed performs well in a time series context—and that it may be considered a good alternative when estimating coefficients in time series quantile regression models. Moreover, the new proposals for the data generating process of QADL dynamics—compiled by both Propositions 1 and 2—provided satisfactory results in simulations. In this sense, our main conclusion is that, given an optimal bandwidth—$2\zeta^*$ in our cases—$\beta_n^{\zeta^*}(\tau)$ globally outperformed $\beta_n(\tau)$ for all the simulated scenarios.

A first further step to this study would be to replicate the "Application to House Price Returns" implemented by Galvão et al. (2013)—this time including the smoothed QR estimator by Fernandes et al. (2021) as an alternative estimation procedure for the coefficients in QADL models.

Another ambitious goal that occurred to us is to fully comprehend the ideas proposed by Yang and Yu (2007), so that we may be able to explore them by bringing the concept of uniform convergence of sample quantiles in a quantile regression framework that includes exogenous covariates—that is, not just in the univariate case.

Finally, any future work considering some of the analytical characterizations of QADL dynamics discussed in this work is very much in our interest.

References

V. Chernozhukov and H. Hong. An MCMC approach to classical estimation. *Journal of Econometrics*, 115:293–346, 2003. URL https://doi.org/10.1016/S0304-4076(03)00100-3. 3

R. F. Engle and S. Manganelli. Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, 22(4):367–381, 2004. URL https://doi.org/10.2307/1392044. 3

M. Fernandes, E. Guerre, and E. Horta. Smoothing quantile regressions. *Journal of Business & Economic Statistics*, 39(1):338–357, 2021. URL https://doi.org/10.1080/07350015.2019.1660177. 3, 5, 17, 18, 19, 23, 24, 25, 28, 29, 30, 31, 34, 35, 36, 37, 39, 40

A. F. Galvão. Unit root quantile autoregression testing using covariates. *Journal of Econometrics 152(2)*, 152(2):165–178, 2009. URL https://doi.org/10.1016/j.jeconom.2009.01.007. 9

A. F. Galvão, G. Montes-Rojas, and S. Y. Park. Quantile autoregressive distributed lag model with an application to house price returns. *Oxford Bulletin of Economics and Statistics*, 13:307–321, 2013. URL https://doi.org/10.1111/j.1468-0084.2011.00683.x. 3, 4, 6, 7, 9, 11, 19, 20, 21, 39, 40

C. S. Gourieroux, J. Jasiak, and S. Hara. Dynamic quantile models. *Journal of Econometrics*, 147(1):198–205, 2008. URL https://doi.org/10.1214/19-EJS1607. 9

H. G. Hong, D. C. Christiani, and Y. Li. Quantile regression for survival data in modern cancer research: expanding statistical tools for precision medicines. *Precision Clinical Medicines*, 2(2):90–99, 2019. URL https://doi.org/10.1093/pcmedi/pbz007. 3

J. L. Horowitz. Bootstrap Methods for Median Regression Models. *Econometrica*, 66(6):1327–1352, November 1998. URL https://ideas.repec.org/a/ecm/emetrp/v66y1998i6p1327-1352.html. 3

E. Horta. *Quantile Regression: Facts and digressions*. Department of Statistics, UFRGS, Porto Alegre, Brazil, 1 edition, 2021. URL https://github.com/eduardohorta/QuantRegFacts. 4, 10, 11, 12, 13, 22

D. R. Hunter and K. Lange. Quantile regression via an MM algorithm. *Journal of Computational and Graphical Statistics*, 9(1):60–77, 2000. URL https://doi.org/10.2307/1390613. 16

R. Koenker. *Quantile Regression*. Econometric Society Monographs. Cambridge University Press, 2005. URL https://doi.org/10.1017/CBO9780511754098. 2, 9

R. Koenker and G. Bassett. Regression quantiles. *Econometrica*, 46(1):33–50, 1978. URL http://www.jstor.org/stable/1913643. 3, 5, 16, 17, 18, 28, 29, 30, 31, 34, 35, 36, 37

R. Koenker and Z. Xiao. Quantile autoregression. *Journal of the American Statistical Association*, 101(475):980–990, 2006. URL https://doi.org/10.1198/016214506000000672. 8, 9

R. Koenker, V. Chernozhukov, X. He, and L. Peng. *Handbook of Quantile Regression*. CRC Press, 2018. 3

N. Meinshausen. Quantile regression forests. *Journal of Machine Learning Research*, 7(35):983–999, 2006. URL http://jmlr.org/papers/v7/meinshausen06a.html. 3

F. Mosteller and Tukey. *Data Analysis and Regression: A Second Course in Statistics*. Taylor & Francis, 2005. URL https://ui.adsabs.harvard.edu/abs/1977dars.book.....M. 2

B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, London, 1986. URL https://doi.org/10.1002/bimj.4710300745. 24, 25, 39

I. Takeuchi, Q. V. Le, T. Sears, and A. J. Smola. Nonparametric quantile regression. *Journal of Machine Learning Research*, 7:1001–1022, 2005. URL https://dl.acm.org/doi/10.5555/1248547.1248592#sec-ref. 3

Q. Xu, X. Lu, C. Jiang, and K. Yu. Quantile autoregression neural network model with applications to evaluating value at risk. *Applied Soft Computing*, 13:3120–3160, 2016. URL http://dx.doi.org/10.1016/j.asoc.2016.08.003. 3

S. Yang and K. Yu. *A note on the Bahadur representation of sample quantiles for mixing random variables*. Department of Mathematical Sciences, Brunel University, UK, 1 edition, 2007. URL https://bura.brunel.ac.uk/bitstream/2438/12676/3/FullText.pdf. 40