Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2018 doi:10.20944/preprints201806.0198.v1

Peer-reviewed version available at *Risks* **2018**, *6*, 97; doi:10.3390/risks6030097

General Quantile Time Series Regressions for Applications in Population Demographics

Gareth W. Peters[1]

[1] Department of Actuarial Mathematics and Statistics, Heriot-Watt University, Scotland, UK

Keywords: quantile regression; quantile time series; demographics; mortality; longevity; modelling mortality projection

Abstract

This paper has three objectives, the first is to present a detailed overview in the form of a tutorial for the developments of several key quantile time series modelling approaches. The second objective is to develop a general framework to represent such quantile models in a unifying manner in order to easily develop extensions and connections between existing models that can then be developed to further extend these models in practice. In this regard, the core theme of the paper is to provide perspectives to a general audience of core components that go into construction of a quantile time series model and then to explore each of these core components in detail. The paper is not addressing the concerns of estimation of these models, as there is existing literature on these aspects in many settings, we provide references to relevant works on these aspects in several classes of model. Instead, the focus is rather to provide a unified framework to construct such models for practitioners, therefore the focus is instead on the properties of the models and links between such models from a constructive perspective. The third objective is to compare and discuss the application of the different quantile time series models on several sets of interesting demographic and mortality based time series data sets of relevance to life insurance analysis. The exploration included detailed mortality, fertility, births and morbidity data in several countries with more detailed analysis of regional data in England, Wales and Scotland.

1 Introduction

The theme of this tutorial is not to develop an analysis of all possible variants of quantile time series model: their model properties and relevant estimation approaches. Instead, the tutorial takes a considered perspective on popular classes of quantile time series model structure that have been developed in the literature. It first introduces the key papers in the literature that have led to developments in some important classes of model. The second aspect of this paper is to them provide a tutorial coverage of the different model components one can consider when developing a quantile times series model. This is the novelty introduced in this paper, where existing models are viewed from a different perspective and core components of relevance to general quantile time series modelling are decomposed and re-presented in components that allow development of new models and extensions of existing models. In this regard the theme of the tutorial is to decompose quantile time series modelling into a few core model components and then to go into detail on choices developed in the literature relating to these particular components. In viewing quantile time series modelling from this perspective we are able to easily introduce new aspects to the modelling such as modifications to the quantile error function as well as non-linear transformations of relevance to broader classes of quantile time series model.

In particular, in this paper we consider classes of univariate quantile regression models developed in the context of time series modelling. We begin by providing a brief overview of different classes of parametric quantile time series regression models. Quantile valued time series models can be of many different types with regard to their regression structure: function on scalar or vector regressions; function on function regressions; and scalar on function regressions.

We then explore novel applications of quantile time series modelling in applications of relevance to life insurance contexts. In particular we explore a range of mortality and demographic data sets via quantile time series regressions. The purpose of this is to illustrate for actuarial practitioners how one may utilise such time series modelling techniques to explore relevant demographic and mortality related time series data sets. The outputs of this analysis can be directly useful in insurance applications for instance in life insurance

 doi:10.20944/preprints201806.0198.v1

Peer-reviewed version available at Risks 2018, 6, 97; doi:10.3390/risks6030097

into detail on these particular areas of application, instead we focus on modelling and comparison of different quantile time series models on these real mortality and demographic data sets obtained for England, Wales, Scotland and Northern Ireland.

In general, when studying such data sets, we note that one can separate such data sets typically into four categories [1]:

- Demographic data: which includes factors such as age, sex, migration patterns, ethnicity and marital status in populations. Typically this comes from census type data sets.
- Health event data: this involves recordings of health events affecting individuals or populations which can include births, deaths, health conditions, primary care interactions, secondary care interactions and health hazards;
- Circumstantial data: focuses on aspects of individuals' and populations' circumstances that may affect the wider determinants of health, including socio-economic, lifestyle, and environmental data. Such data in this type of category can include education data; employment data; housing data and environmental data.;
- National reference data: which includes data not collected for the sole purpose of health analysis, however it can be used in connection with health data.

In this manuscript we will focus on data reflecting demographic data and health event data. In Section 9 we provide a detailed overview of which data sets are considered followed by the anlaysis of numerous quantile time series models, both linear and non-linear, applied to these data sets with detailed description of the fitted model properties which can act as a guide for practitioners.

Before we go into specific literature and background on quantile times series modelling, it would be remiss of us not to briefly mention some core developments of quantile regression contexts, prior to going into more focus on time series specifics applications. The growth in general quantile regression modelling goes back to influential works such as those developed by Koenker and Bassett Jr [1978], Koenker [2005], Gilchrist [2000], Buchinsky [1998], Koenker [2004], Yu and Jones [1998] and in Bayesian modelling settings in Yu and Moyeed [2001], Thompson et al. [2010] and Yu et al. [2003].

In addition, more recently, in Bernardi et al. [2016a] they extended the Bayesian family of Asymmetric Laplace distribution (ALD) quantile regression models to the family of Skewed Exponential Power (SEP) models, with the intention of accounting for application settings in which the assumption of ALD models, that produce non-heavy tails relative to the Gaussian hypothesis is generalized to allow for greater skewness and kurtosis range. In similar context, with the perspective of generalizing the distributional properties of the resulting quantile regression and its distributional tail properties, in Lancaster and Jae Jun [2010] they study applications of Bayesian exponentially tilted empirical likelihood to inference about quantile regressions.

From the perspective of covariate selection and shrinkage in Bayesian quantile regressions, there are also a sequence of papers on regularization methods for quantile regressions, such as Li et al. [2010]. In such works, the authors demonstrate that regularization methods such as lasso are effective in Bayesian quantile regression contexts. For a good overview of the past 40 years of quantile regression modelling over a wide spectrum of quantile models and modelling domains see the discussions in Koenker [2017].

Quantile regression has also begun to be explore in more general regression settings such as in panel data applications, see a recent study of such important applications where novel bootstrap procedures are developed for panel quantile regression settings in Galvao and Montes-Rojas [2015] and references therein.

The application of quantile regression models in financial risk and insurance has also recently begun to develop in works such as Dong et al. [2015], Peters et al. [2016] and discussions in Operational risk contexts in Cruz et al. [2015] and Peters and Shevchenko [2015]. Furthermore, there are Bayesian application works such as Hu et al. [2013] which develops a Bayesian partially collapsed Gibbs sampler approach to fitting single-index models in conditional quantile regression. In Bernardi et al. [2016b] they consider the challenge of model combining or model averaging in dynamic quantile regression settings which they termed the general dynamic model averaging DMA framework.

We now shift the focus from general quantile regression background to a particular focus on time series contexts. We first comment on the general specification aspects of a quantile time series regression structure, before providing an overview of different classes of quantile time series models that have been proposed in the

[1] see discussions in https://www.healthknowledge.org.uk/public-health-textbook/health-information

the quantile function time series and regressions has three main components:

1. The conditional distribution of the regression time series model which defines a conditional quantile function of the dependent variable, given the explanatory variables, primarily comprised in our case of time series past observations of the process but may also include additional exogenous covariates perhaps with distributed lags.;

2. The structural component of the regression model based on the link functions and imposed model structures linking the regression structures with the covariates to the location and scale of the conditional distribution which defines a conditional quantile functions of the response.;

3. The actual choice of independent variables (regressors), that is, the covariates in the regression model as well as some basis function regression structures in some of the models proposed.

We will go into specific details in each of these aspects of quantile time series modelling in the remainder of the tutorial paper.

1.1 Outline and Contributions

In Section 2 we overview important developments in the context of quantile regression modelling with a particular focus on quantile times series models.

In Section 3 we propose a general modelling framework for developing a wide class of quantile time series models that is detailed in a constructive manner involving four key ingredients. The first is the structural form of the quantile regression or time series model (linear or non-linear maps), the second is the choice of quantile error function, the third is the choice of lag structure for the endogenous variables that characterize the quantile time series structure and the fourth is the choice of endogenous covariates and their lag structures. Then we introduce examples in the case of non-parametric and parametric linear and non-linear quantile time series models.

In Section 4 we develop an overview of non-parametric quantile time series models in linear and non-linear settings. Then in Section 5 we overview parametric quantile time series models also according to linear and non-linear classes of time series model.

In Section 6, we characterize the class of quantile error functional families in categories of location-scale, shape-scale and some special heavy tailed families of quantile functions. We conclude this section with discussion on truncated quantile error models for time series constructions with restricted supports.

In Section 7 and Section 8 we detail how to develop transformation of these basic families of parametric quantile error models to other significantly more flexible families of quantile error models based on two classes of mappings: the Tukey Elongation Maps and secondly the Rank Transmutation Map framework. We note that these general transformations may also be used to construct other aspects of the quantile time series model, not just the quantile error function.

In Section 9 we will explore applications of the general quantile time series model constructions on a range of demographic and mortality data. The focus will be to explore with applications the properties of some of the models discussed in the tutorial overview and to explore their application relevance for future developments in new insurance domains such as life-insurance modelling. We then conclude.

1.2 Notation

In this section we briefly introduce some core notation used throughout the manuscript. We denote random variables by upper script and their realization by lower script. We denote vectors by bold and scalars by non-bold. Furthermore, we use the notation Q_{Y_t} to refer to the quantile function of a random variable Y_t obtained as the t-th time instance of a time series $\{Y_1, \ldots, Y_t, \ldots\}$. We will typically refer to the quantile level by variable $u \in [0, 1]$. The generic notation used for static model parameters in any of the introduced time series models will be denoted by vector $\boldsymbol{\theta}$ unless otherwise discussed. Furthermore, we will utilize the notation $\mathcal{F}_t$ to denote the natural sigma-algebra of the observed time series given by $\mathcal{F}_t = \sigma(Y_1, \ldots, Y_t)$. We will therefore denote by $Q_{Y_t}(u|\mathcal{F}_{t-1}; \boldsymbol{\theta})$ the conditional quantile function of random variable Y_t condition on the information set $\mathcal{F}_{t-1}$, i.e. the observations of the time series until time $t-1$ given by $\{y_0, y_1, \ldots, y_{t-1}\}$ and the static model parameters generically denoted by $\boldsymbol{\theta} \in \mathbb{R}^d$.

Furthermore, we will in general denote functional coefficients of the quantile time series structure by functions $\alpha_i(u)$ which will multiple lagged values of the time series. We will impose additional structure on represents white noise sequence ϵ_t with distributional static parameters generically denoted by $\gamma \in \mathbb{R}^{\cdot}$. Additional notation used in the paper is introduced and defined where required.

2 Developments of Quantile Time Series Models

In this section we begin by introducing a few key developments of quantile time series based modelling. We then develop a framework representation that is general and encapsulates structural representations that include each of these discussed time series developments. We believe that in doing this we will provide a clarity to practitioners around how to develop each of the core components of a quantile time series model in a general manner.

In its earliest form, quantile regression as introduced by Koenker and Bassett Jr [1978], generalizes the notion of sample quantiles to linear and nonlinear regression models including the least absolute deviation estimation as its special case. In developing such a regression framework, one is able to develop estimation methods for conditional quantile functions at any (or all) probability levels. Such conditional quantile regression structures are now increasingly studied as they have been shown to present new information, compared to classical generalized linear model (GLM) regressions or linear mean regressions, see discussions in Koenker [2000].

In addition to these works on quantile regression modelling, of direct relevance to this manuscript there are also a range of works that explore the development of time series based quantile regression models. There have been quantile model developments in both parametric, non-parametric, linear, non-linear, stationary and non-stationary time series contexts.

In the summary of Koenker [2017] the background of time series developments in quantile models is briefly over viewed where they discuss the development of the quantile Autoregressive (QAR) class of models discussed in Koenker and Xiao [2006] and given by parametric form, for a univariate time series $\{Y_t\}$ by:

$$Q_{Y_t}\left(u|\mathcal{F}_{t-1};\boldsymbol{\theta}\right) = \sum_{i=1}^{p} \alpha_i(u) Y_{t-i} + Q_{\epsilon}\left(u;\boldsymbol{\gamma}\right) \tag{1}$$

with $u \in [0,1]$ and where $Q_{Y_t}\left(u|\mathcal{F}_t;\boldsymbol{\theta}\right)$ denotes the conditional quantile function of random variable Y_t, $\mathcal{F}_t$ will denote the natural filtration generated by the sigma algebra of time series $\{Y_t\}$ and $\alpha_i(u)$ denotes the i-th lagged functional coefficient of the QAR time series model with $\boldsymbol{\theta}$ denoting generically the vector of all model parameters. Furthermore, we denote $Q_{\epsilon}\left(u;\boldsymbol{\gamma}\right)$ as the time series white noise i.i.d. error ϵ_t quantile function with parameters $\boldsymbol{\gamma}$, this is denoted by $\alpha_0(u)$ in the aforementioned reference.

To understand how to go from a time series model to a quantile time series model, consider the following relationship detailed in Example 2.1. This example illustrates the relationship between a functional, random coefficient AR time series model and its equivalent form expressed as a quantile time series model. Of course, any time series model, even without a structure of random functional coefficients can also admit a quantile time series model. However, in general the direct link between such a model and the coefficients of each model, the underlying time series model and the quantile time series model may not be easy to obtain in an explicit closed-form. However, as we will illustrate, we can construct several examples in the following sections that relate time series models to quantile time series models in a more explicit fashion. We begin with following AR example.

Example 2.1 (Relating Functional time series with Random Coefficients to Quantile time series)**.** *Consider a functional time series model with random coefficients in an AR structure given by*

$$\begin{aligned} y_t &= \sum_{i=1}^{p} \alpha_i(U_t) y_{t-i} + \alpha_0(U_t) \\ &= \sum_{i=1}^{p} \alpha_i(U_t) y_{t-i} + \epsilon_t \end{aligned} \tag{2}$$

with $U_t \sim U[0,1]$ for all t and where $\alpha_i : [0,1] \mapsto \mathbb{R}$. We need only then assume that

$$\sum_{i=1}^{p} \alpha_i(U_t) y_{t-i} + \alpha_0(U_t)$$

Note: one way to construct such a solution is to consider positive valued time series such that $Y_t \in \mathbb{R}^+$ then if each $\alpha_i(u) : [0,1] \mapsto \mathbb{R}^+$ and is a quantile function which is scale invariant, one has that the resulting linear combination $\sum_{i=1}^{p} \alpha_i(U_t) y_{t-i}$ is a quantile function. Furthermore, one can write the equivalent conditional quantile function time series model as follows:

$$Q_{Y_t}\left(u|\mathcal{F}_{t-1};\boldsymbol{\theta}\right) = \sum_{i=1}^{p} \alpha_i(u) y_{t-i} + \epsilon_t \tag{3}$$

which is obtained by use of the following general rule for any monotone increasing function g and standard uniform random variable U:

$$Q_{g(U)}\left(u\right) = g\left(Q_U(u)\right) = g(u) \tag{4}$$

where we use the fact that for a uniform random variable the distribution and quantile function satisfy the linear relationship, such that $F_U(u) = Q_U(u) = u$.

Such an example was illustrated in Koenker and Xiao [2006] where they point out that one can also consider, from a regression perspective, an alternative formulation of such functional regressions with scalar or vector on function regression. In particular, one can define a scalar (vector) on function regression version of the QAR model, with co-monotonic random functional coefficients, denoted as the random coefficients by

$$Y_t = \sum_{i=1}^{p} \alpha_i\left(U_t\right) Y_{t-i} + Q_{\epsilon}\left(U_t;\boldsymbol{\gamma}\right) \tag{5}$$

for i.i.d. $U_t \sim U(0,1)$. Such QAR models, in which the autoregressive coefficients are expressed as monotone functions of a single, scalar random variable are interesting as they allow one to

"*capture systematic influences of conditioning variables on the location, scale and shape of the conditional distribution of the response, and therefore constitute a significant extension of classical constant coefficient linear time series models in which the effect of conditioning is confined to a location shift.*"
(Koenker and Xiao [2006])

The motivation for such quantile time series models, in the literature has been discussed from numerous perspectives, since they not only provide new information not obtained from classical time series models such as Seasonal Autoregressive Moving Average (SARIMA) type models, these quantile time series models can also be more robust to outliers and heavy tailed noise, see discussions in Fitzenberger et al. [2013].

Furthermore, recent studies have begun to develop properties specifically for QAR models, such as the notion of the quantile correlation (QACF) and quantile partial correlation (QPACF), defined in Li et al. [2015] according to the natural quantile extension from standard time series settings for the quantile autocovariance and autocorrelations:

$$\mathrm{QACVF}_u\left\{Y_t, Y_{t-\tau}\right\} = \mathbb{E}\left[\left(u - \mathbb{I}\left[Y_t - Q_{Y_t}(u)\right]\right)\left(Y_{t-\tau} - \mathbb{E}\left[Y_{t-\tau}\right]\right)\right], \tag{6}$$

$$\mathrm{QACF}_u\left\{Y_t, Y_{t-\tau}\right\} = \frac{\mathrm{QACVF}_u\left\{Y_t, Y_{t-\tau}\right\}}{\sqrt{\mathbb{V}\mathrm{ar}\left(u - \mathbb{I}\left[Y_t - Q_{Y_t}(u)\right]\right)\mathbb{V}\mathrm{ar}\left(Y_{t-\tau}\right)}}, \tag{7}$$

where $Q_{Y_t}(u)$ denotes the unconditional quantile of Y_t with $u \in [0,1]$. Then the analogous QPACF can be obtained from the quantile based correlation and is given according to the expression in [Li et al., 2015, Equation 2.2]. Since this work, others such as Han et al. [2016] have further extended these notions to study multiple QAR time series models via ideas such as the cross-quantilogram which can measure the quantile dependence between two quantile time series.

Other variants of non-stationary QAR models have also been explored in Aue et al. [2017] where they consider locally stationary QAR models through piece-wise local in time constructions given by,

$$Q_{Y_t}\left(u|\mathcal{F}_{t-1};\boldsymbol{\theta}\right) = \sum_{i=1}^{p_k} \alpha_{k,i}(u) Y_{t-i} + Q_{\epsilon_k}\left(u;\boldsymbol{\gamma}\right), \tag{8}$$

and they extend the standard QAR model to have a definition specific to each of the k segments. They also discuss issues to do with model selection and segmentation, even developing a framework that performs local segmentation of the space and model selection per quantile level.

this is a topic we will discuss in significantly more detail in later sections of this manuscript. However, we note that there are many ways to construct and decide upon such a reference quantile function, we note recent developments in this space given by the flexible class of non-parametric estimators proposed in Stephanou et al. [2017]. They propose as simple non-parametric L-estimator class of kernel representations of the error quantile, based on the class of Hermite Askey-orthogonal polynomials. In other papers on non-parametric quantile time series literature there are also works such as those in Cai [2002]. In this work, the authors study nonparametric estimation of regression quantiles for time series data by inverting a weighted Nadaraya-Watson estimator of the conditional distribution function. Other non-parametric works on quantile regression time series include Cai and Xu [2008].

In addition to classes of QAR model, quantile time series regressions have been studied in both linear and nonlinear autoregressive settings in Bloomfield and Steiger [1983], Cai [2010a], Cai et al. [2013], Weiss [1991] and Davis and Dunsmuir [1997]. The development of autoregressive conditional heteroscedasticity ARCH and GARCH models in quantile time series settings has also been undertaken by Koenker and Zhao [1996] and Lee and Noh [2013]. Although, the focus in this paper does not concern statistical estimation of quantile regression models, rather the characterization of such models from a perspective of sufficient conditions for time series momentum and reversal, we believe it is still useful to mention that in frequentist estimation techniques for AR-ARCH type quantile models, identification in parameter estimation can be a challenge, see discussion in Noh and Lee [2015]. These authors demonstrate how a simple AR(1)-ARCH(1) model parametrized as follows,

$$\begin{aligned} Y_t &= \alpha_0 + \alpha_1 Y_{t-1} + \epsilon_t, \\ \epsilon_t &= \sigma_t \eta_t, \quad \sigma_t^2 = \omega + \beta \epsilon_{t-1}^2 \quad \text{for } t \in \mathbb{Z}, \end{aligned} \tag{9}$$

where $\{\eta_t\}$ are i.i.d. random variables with $\mathbb{E}[\eta_t] = 0$ and $\mathbb{E}\left[\eta_t^2\right] = 1$, and when re-expressed as a quantile regression form has parameter identification issues. To see this, consider the reformulated AR(1)-ARCH(1) model represented as a quantile regression time series form as follows

$$Q_{Y_t}(u|\mathcal{F}_{t-1};\boldsymbol{\theta}) = \alpha_0 + \alpha_1 Y_{t-1} + Q_\epsilon(u;\boldsymbol{\gamma})\left(\omega + \beta\left(Y_{t-1} - \alpha_0 - \alpha_1 Y_{t-2}\right)^2\right)^{1/2}, \tag{10}$$

where $Q_\epsilon(u;\boldsymbol{\gamma}) = \inf\{x : P(\epsilon_1 \leq x;\boldsymbol{\gamma}) \geq u\}$ and $\mathcal{F}_t = \sigma(Y_s : s \leq t)$ denotes the σ-field generated by $\{Y_s : s \leq t\}$. They then observe that since the u-th quantile of η_t is unknown, then the parameters in (10) are not identifiable. Fortunately, this issue can be overcome with appropriate re-parametrization of the model, see discussion in Lee and Noh [2013].

The re-parametrized form that is identifiable is obtained by setting $\epsilon_t = h_t u_t$, $h_t^2 = 1 + \gamma\epsilon_{t-1}^2$ with $h_t^2 = \sigma_t^2/\omega$, $u_t = \sqrt{\omega}\eta_t$ and $\gamma = \beta/\omega$. In such a re-parametrization one has re-expressed the ARCH model as a conditional scale model with no scale constraints on the i.i.d. innovations. The resulting conditional quantile time series model is then given by

$$Q_{Y_t}(u|\mathcal{F}_{t-1};\boldsymbol{\theta}) = \alpha_0 + \alpha_1 Y_{t-1} + Q_\epsilon(u;\boldsymbol{\gamma})\left(1 + \gamma\left(Y_{t-1} - \alpha_0 - \alpha_1 Y_{t-2}\right)^2\right)^{1/2}. \tag{11}$$

General results for identification parametrizations that extend beyond the AR-ARCH models to ARMA-AGARCH models are also developed in Noh and Lee [2015].

These developments were then presented in a general framework that is widely utilized, known as the Conditional Autoregressive VaR model (CAViaR) of Engle and Manganelli [2004]. This model class has been extended to study explicitly models which have both conditional location and scale components, see Noh and Lee [2015].

In Cai [2016] and then later in Noh and Lee [2015] the authors developed general classes of conditional location-scale quantile time series models based on time series models of the generic form:

$$Y_t = \mu_t(\alpha) + \sigma_t(\alpha)\epsilon_t \quad \text{for } t \in \mathbb{Z}, \tag{12}$$

where $\mu_t(\alpha)$ and $\sigma_t(\alpha)$ were functions they used to denote $\mu(Y_{t-1}, Y_{t-2}, \ldots; \alpha)$ and $\sigma(Y_{t-1}, Y_{t-2}, \ldots; \alpha)$ for some measurable functions $\mu, \sigma : \mathbb{R}^\infty \times \Theta_1 \to \mathbb{R}$; α denotes the true model parameter; Θ_1 is a model parameter space; $\{\epsilon_t\}$ are i.i.d. random variables with an unknown common distribution function F_ϵ.

In fact, such ideas for quantile regression were previously well developed and explained in the location and scale class of quantile regression models in Gilchrist [2000]. The equivalent quantile time series models may be presented according to the general location-scale generalized form given by

$$Q_{Y_t}(u|\mathcal{F}_{t-1},\boldsymbol{\theta}) = \mu\left(Y_{t-1:t-p},\boldsymbol{\alpha}\right) + \sigma\left(Y_{t-1:t-q},\boldsymbol{\beta}\right)Q_\epsilon(u;\boldsymbol{\gamma}) \tag{13}$$

error function denoted by $Q_\epsilon(u|\gamma)$.

Special cases of such models had previously also been considered in works such as Cai et al. [2013] who developed the quantile time series version of the double AR(p) models of Ling [2007] given by

$$Y_t = \alpha_0 + \alpha_1 Y_t + \cdots + \alpha_p Y_{t-p} + \epsilon_t \sqrt{\beta_0 + \beta_1 Y_{t-1}^2 + \cdots + \beta_q Y_{t-q}^2}, \tag{14}$$

where $\beta_i > 0$ for $i \in \{0, \ldots, q\}$, $\epsilon_t \overset{i.i.d.}{\sim} N(0,1)$ and Y_t is independent of ϵ_t for all t.

As noted in Cai et al. [2013] this is a special case of the ARMA-ARCH models proposed by Weiss [1984], however it is structurally distinct from the ARCH models proposed by Engle [1982] when one considers the settings with $\alpha_i = 0$. The Double-QAR(p,q) model proposed in Cai et al. [2013] is then given by

$$Q_{Y_t}(u|\mathcal{F}_{t-1}, \boldsymbol{\theta}) = \alpha_0 + \alpha_1 Y_t + \cdots + \alpha_p Y_{t-p} + \sqrt{\beta_0 + \beta_1 Y_{t-1}^2 + \cdots + \beta_q Y_{t-q}^2}\, Q_\epsilon(u; \boldsymbol{\gamma}) \tag{15}$$

where model parameters $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ with $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_p)$ and $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_q)$. The quantile error distribution can be formed from a range of different choices of quantile distribution, this will be discussed in further sections. We note that the choice adopted in Cai et al. [2013] was a flexible quantile sub-family of the generalized lambda family of distributions developed in Freimer et al. [1988].

Following the development of QAR time series models there were extensions to such models such as those proposed in Cai and Stander [2008] where they developed a class of quantile self-exciting threshold autoregressive time series models. Such models are the quantile time series extension of self-exciting threshold autoregressive time series (SETAR) often referred to as the class of QSETAR models which are generally characterized for instance by the QAR model extension given by:

$$Q_{Y_t}(u|\mathcal{F}_t) = \sum_{i=1}^{m+1} \left[\left(\sum_{j=1}^{p} \alpha_{i,j}(u) Y_{t-j} \right) + Q_{\epsilon_{i,t}}(u) \right] \mathbb{I}\{y_{t-d} \in (y_{i-1}, y_i]\} \tag{16}$$

where we consider the following state-space threshold partitions $-\infty = y_0 < y_1 < \cdots < y_{m+1} = \infty$ and denote by $\mathbb{I}\{\cdot\}$ the indicator function taking value one, when the arguments condition is satisfied and zero otherwise.

In addition to these models, there have been other classes of non-linear quantile time series models developed such as those studied in Peters et al. [2016], Dong et al. [2015] and Chen et al. [2017] where they proposed both non-linear quantile models of the form

$$Q_{Y_t}(u|\mathcal{F}_{t-1}, \boldsymbol{X}; \boldsymbol{\theta}) = T\left(\mathcal{F}_t, \mathcal{G}_t, Q_\epsilon(u; \boldsymbol{\gamma})\right) \tag{17}$$

which form an equivalent of distributed lag ARDL models with exogenous covariates. Furthermore, these models were also extended in Chen et al. [2017] to quantile state-space models (QSSM) that they termed the AQUA class. In Peters et al. [2016] several class of transform function $T(\cdot)$ were explored based on the Tukey class of elongation transforms, including popular sub-classes of G-and-H, G-and-K, G-and-J, G-G, H-H transforms. An efficient R-package for estimation and description of its functionality is provided by Prangle [2017].

In terms of forecasting of quantile time series models there are multiple approaches one can adopt, see discussions in Cai [2010b]. Furthermore, there is a branch of quantile time series models relating to extreme value theory, see a detailed discussion in McNeil and Frey [2000]. To complete this section we make the following example to illustrate a general mapping from a time series model to a quantile time series model

Example 2.2 (Relating General Non-Linear Stochastic Volatility time series Models to Quantile time series Models)**.** *Consider a functional time series model characterized by the general structural form*

$$\begin{aligned} Y_t &= \mu\left(Y_{t-p}, \ldots, Y_{t-1}; \boldsymbol{\theta}\right) + \sum_{j=1}^{q} \theta_j \sigma\left(Y_{t-s-q}, \ldots, Y_{t-1-q}, t; \boldsymbol{\gamma}\right) \alpha_0(U_{t-q}) + \sigma\left(Y_{t-s}, \ldots, Y_{t-1}, t\right) \alpha_0(U_t) \\ &= \mu\left(Y_{t-p}, \ldots, Y_{t-1}\right) + \sum_{j=1}^{q} \theta_j \sigma\left(Y_{t-s-q}, \ldots, Y_{t-1-q}, t\right) \epsilon_{t-q} + \sigma\left(Y_{t-s}, \ldots, Y_{t-1}, t\right) \epsilon_t \end{aligned} \tag{18}$$

for time t and where U_t is an i.i.d. sequence of uniform random variables $U_t \sim U[0,1]$ and where $\alpha_0 : [0,1] \mapsto \mathbb{R}$ which we assume satisfies that $\alpha_0(u)$ is a monotone increasing function in U_t.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2018 doi:10.20944/preprints201806.0198.v1

Peer-reviewed version available at Risks 2018, 6, 97; doi:10.3390/risks6030097

values of the process Y_t and time t for the trend such that $\mu : \mathbb{R}^p \times \mathbb{R}^+ \mapsto \mathbb{R}$ parametrized by model parameters generically denoted by $\boldsymbol{\theta}$ and a non-linear stochastic volatility structure given by functional form $\sigma(Y_{t-s}, \dots, Y_{t-1}; \boldsymbol{\gamma})$ such that $\sigma : \mathbb{R}^s \mapsto \mathbb{R}^+$ with generic model parameters $\boldsymbol{\gamma}$. Then in this case, one has the general equivalent conditional quantile function time series model can be obtained, under the restrictions that $\theta_i \geq 0$ for $i \in \{1, \dots, q\}$, given by

$$\begin{aligned} Q_{Y_t}(u|\mathcal{F}_{t-1}; \boldsymbol{\theta}) &= \mu(Y_{t-p}, \dots, Y_{t-1}) + \sum_{j=1}^{q} \theta_j \sigma(Y_{t-s-q}, \dots, Y_{t-1-q}, t)\, Q_{\epsilon_{t-q}}(u) + \sigma(Y_{t-s}, \dots, Y_{t-1}, t)\, Q_{\epsilon_t}(u) \\ &= \mu(Y_{t-p}, \dots, Y_{t-1}) + Q_{\tilde{\epsilon_t}}(u|\mathcal{F}_{t-1}; \boldsymbol{\theta}_{1:q},). \end{aligned} \tag{19}$$

To obtain this representation, we use the fact that a positive scaling of a quantile function is a quantile function, such that for all t, s, q on has that $\theta_j \sigma(Y_{t-s-q}, \dots, Y_{t-1-q}, t)\, Q_{\epsilon_{t-q}}(u)$ is a quantile function. Furthermore, any linear combination of quantile functions is also a quantile function and we apply the rule that a linear translation of a quantile is a quantile function. Finally, we are back to the condition that we may then apply again the following general rule for any monotone increasing function g and standard uniform random variable U:

$$Q_{g(U)}(u) = g(Q_U(u)) = g(u) \tag{20}$$

where we use the fact that for a uniform random variable the distribution and quantile function satisfy the linear relationship such that $F_U(u) = Q_U(u) = u$.

Remark 2.1. *We note, that in general the above example is supposed to be illustrative in the general approach one can adopt to moving from time series models to the equivalent relationship in a quantile time series model. In general, one typically starts with either a time series model or a quantile time series model, never particularly worrying about the relationships between these quantities. However, it is beneficial for aspects of results derived in future sections of this manuscript to explain this relationship explicitly. We also note that one must in general still be cautious to consider such general model structures from the perspective of estimation and appropriate identification considerations.*

3 General Construction of Quantile Time Series Regression Models

The goal of this section is to present a generic quantile time series framework for constructing flexible quantile time series models: parametric and non-parametric.

In the literature it is common to develop quantile time series models dependent only on endogenous variables for the regression structure. In other words, the vast majority of time series quantile regression models discussed are constructed from conditional distributions which are based on information filtration, denoted $\mathcal{F}_t = \sigma(Y_0, \dots, Y_t)$ constructed from the sigma-algebra naturally generated by the time series signal under consideration.

However, in many practical settings it will be of interest to extend the class of quantile time series models such as the QAR models to include lagged endogenous covariates of different forms. A first natural extension would be to begin by developing the conditional quantile function forming the time series to also be dependent explicitly on a second filtration of observed exogenous covariates, which will be denoted by $\mathcal{G}_t = \sigma(\boldsymbol{X}_0, \dots, \boldsymbol{X}_t)$, where $\boldsymbol{X}_t \in \mathbb{R}^d$.

In such cases one can construct a range of extensions of the QAR model for instance, such as the two illustrative examples below:

$$\begin{aligned} Q_{Y_t}(u|\mathcal{F}_{t-1}, \mathcal{G}_t; \boldsymbol{\theta}) &= \sum_{i=1}^{p} \alpha_i(u) Y_{t-i} + \sum_{j=1}^{d} \sum_{i=1}^{k} \beta_{j,i} X_{j,t-i} + Q_\epsilon(u; \boldsymbol{\gamma}) \\ Q_{Y_t}(u|\mathcal{F}_{t-1}, \mathcal{G}_t; \boldsymbol{\theta}) &= \sum_{i=1}^{p} \alpha_i(u) Y_{t-i} + \sum_{j=1}^{d} \sum_{i=1}^{k} \beta_{j,i} Q_{X_{j,t-i}|X_{j,1:t-i}}(u) + Q_\epsilon(u; \boldsymbol{\gamma}), \end{aligned} \tag{21}$$

where in the second case one may wish to impose that $\beta_{j,i} > 0$ for all i, j to ensure the resulting conditional quantile function is well defined. These illustrative models are based on linear relationships in the quantile

Peer-reviewed version available at Risks 2018, 6, 97; doi:10.3390/risks6030097

presented by the following generic parametric conditional quantile relationship

$$Q_{Y_t}\left(u|\mathcal{F}_{t-1},\mathcal{G}_t;\boldsymbol{\theta}\right) = T\left(\mathcal{F}_t,\mathcal{G}_t,Q_\epsilon(u)\right) \tag{22}$$

which involves some form of quantile preserving map defined in detail in the following sections where different classes of transformation will be developed, including classical approaches based on location-scale, shape-scale maps as well as more advanced approaches such as the Rank Transmutation Maps RTM and the Elongation transforms of Tukey.

Remark 3.1 (Characterizing Generalized Quantile time series Models)**.** *To characterize this general class of quantile time series models we will consider defining six attributes:*

1. *choice of mapping function $T(\cdot)$ which can be linear or non-linear on the quantile error function or on the quantile function time series "trend" structure.;*
2. *the choice of quantile error function $Q_\epsilon(u)$ - if in the parametric model context;*
3. *the inclusion or not of lagged observations of the time series of interest $(Y_t)_{t\geq 0}$, obtained by the natural filtration generated by the realization of the process (denoted $\mathcal{F}_t$), which enter in the model in either the location or the scale or both.;*
4. *the inclusion or not of lagged exogenous covariates generically denoted by set of vectors $\boldsymbol{x}_t = (x_{1,t},\dots,x_{d,t})$, obtained by the natural filtration generated by the realization of the process (denoted $\mathcal{G}_t$), which enter in the model in either the location or the scale or both.;*
5. *the choice of parametric vs non-parametric model, through either an explicit specification of a quantile error function $Q_\epsilon(u)$ for the model, or a non-parametric approach when no quantile function is explicitly considered.;*
6. *function on function regressions, when one models the entire quantile function by considering all quantile levels $u \in [0,1]$ or some sub-set of this range, versus individual quantile regressions for a specific target quantile level u.*

We briefly explore each of these quantile regression components in both parametric and non-parametric time series contexts in the following sections. In the cases that we focus on parametric families of quantile time series models, we will consider those that can be generated from the type of location scale regression quantile constructions considered in Gilchrist [2000].

In the following sub-sections we discuss each of these components in turn, starting with the distributional aspects of the quantile regression models we consider. In this regard, the models of particular focus in this tutorial are the following classes of quantile model:

1. the Asymmetric Laplace (AL) distribution;
2. the regularly varying and heavy tailed classes of power-law distributions which may for instance be characterized by models whose hazard rate $r(y)$ given by
$$r(y) = \frac{f_Y(y)}{\overline{F}_Y(y)} \tag{23}$$
satisfies the condition for instance in the right tail that
$$\lim_{y\to\infty} r(y) = 0.; \tag{24}$$
3. one, two, three and four parameter parametric distributional models often occurring as sub-members of the Pearson family and the Exponential family and is dispersion extensions.
4. the Rank Transmutation composite quantile function maps
5. the Tukey G-and-H elongation transform family;

that we extend to classes of quantile time series models.

We begin with an overview of some core examples of non-parametric quantile time series models and how these related to the Asymmetric Laplace parametric model. We then proceed with more detailed illustrations of parametric modelling of quantile time series with specific families of models outlined and explained as noted above.

Peer-reviewed version available at Risks 2018, 6, 97; doi:10.3390/risks6030097

In this case we will consider the sub-class of regression quantile time series models given by

$$Q_{Y_t}(u) = T\left(\mathcal{F}_t, \mathcal{G}_t\right) \tag{25}$$

where we drop from the transformation $T(\cdot)$ the component corresponding to the quantile error distribution specification $Q_\epsilon(u)$, hence making the model non-parametric in nature.

4.1 Examples of Linear Nonparametric Quantile Time Series Models

In a non-parametric quantile regression time series approach, one seeks to estimate regression coefficients without the need to make any assumptions on the distribution of the response, or equivalently the residuals. To understand this, we will first introduce a simple quantile AR process. We will focus on the family of models to begin with that have a linear transformation i.e. where the mapping $T(\cdot)$ is considered to be a simple linear function of the coefficients. Furthermore, we will consider a specific target quantile level u in the initial set-up below, not a functional regression structure.

Definition 4.1 (Non-Parametric Quantile Autoregressive QAR(p) Time Series)**.** *Consider the time series $(Y_t)_{t\geq 0}$, then the quantile time series model is defined according to the conditional quantile functions as follows:*

$$Q_{Y_t}(u|\mathcal{F}_t) = \alpha_{0,u} + \sum_{k=1}^{p} \alpha_{k,u}\, y_{t-k} \tag{26}$$

where $\mathcal{F}_t$ denotes information set or filtration that defines the time series dynamic. For instance, $\mathcal{F}_t$ may be the natural filtration generated by the observed time series $(Y_t)_{t\geq 0}$ up to the current time point, or it may contain additional structure such as lagged covariates. The notation, $u \in (0,1)$, corresponds to the quantile level, and location of the u-th quantile level is dictated by coefficient lagged previous values of the time series, where the coefficients can be quantile level specific, with $\boldsymbol{\alpha}_u = (\alpha_{0,u}, \ldots, \alpha_{p,u})$ the linear model coefficients for quantile level u. The coefficients are characterized as the solution to the system of equations, for all $s \in \{1, \ldots, t, \ldots\}$ given by:

$$\mathbb{P}\mathrm{r}\left[y_s \leq \alpha_{0,u} + \sum_{k=1}^{p} \alpha_{k,u}\, y_{s-k} | \mathcal{F}_s \right] = u, \tag{27}$$

when such a solution exists.

Remark 4.1. *Note, in this non-parametric specification, no distributional assumption is being made regarding the conditional distribution for $Y_t|\mathcal{F}_t$, so at this stage, one may wonder how can the coefficients $\boldsymbol{\alpha}_u$ be estimated. The answer involves reformulating the coefficients as the solution of a loss function minimization, which doesnâĂŹt require any distributional assumptions to be made on the time series marginal or conditional distributions.*

This reformulation is given by the quantile loss function. Hence, we may analogously obtain estimates of the quantile model parameters, non-parametrically by solving the following loss function minimization:

$$\min_{\alpha_{0,u},\ldots,\alpha_{p,u}} \sum_t \rho_u(\epsilon_t) = \sum_t \epsilon_t[u - I(\epsilon_t < 0)] \tag{28}$$

and $\epsilon_t = y_t - \alpha_{0,u} - \sum\limits_{k=1}^{p} \alpha_{k,u}\, y_{t-k}$.

Furthermore, it has been shown in Koenker and Hallock [2001], Koenker and Machado [1999] and Yu and Moyeed [2001] that under this loss function ρ_u for quantile regression, the parameter estimates of $\boldsymbol{\alpha}_u$, which may be obtained by minimizing the loss function in (28) will be equivalent to the maximum likelihood estimates of $\boldsymbol{\alpha}_u$ when the conditional distribution of $Y_t|\mathcal{F}_t$ follows the Asymmetric Laplace proxy distribution given in Definition 4.2.

Definition 4.2 (Asymmetric Laplace Distribution)**.** *A random variable $X \sim AL(\mu, \sigma, p)$ has an Asymmetric Laplace (AL) law if it has the following distribution and density*

$$\begin{aligned} F(x;\mu,\sigma,p) &= \begin{cases} \frac{\sigma}{p+1/p} \exp\left[\frac{\sigma}{p}(x-\mu)\right] & \text{if } x < \mu, \\ \frac{\sigma}{p+1/p} \exp\left[-\sigma p (x-\mu)\right] & \text{if } x \geq \mu, \end{cases} \\ f(x;\mu,\sigma,p) &= \left(\frac{\sigma}{p+1/p}\right) \exp\left[-(x-\mu)\sigma s p^s\right], \end{aligned} \tag{29}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2018 doi:10.20944/preprints201806.0198.v1

Peer-reviewed version available at Risks 2018, 6, 97; doi:10.3390/risks6030097

The mean, variance, skewness and kurtosis of this model are given by:

$$\mathbb{E}[X] = \mu + \frac{\sigma(1-2p)}{p(1-p)}, \qquad \mathbb{V}ar[X] = \frac{\sigma^2(1-2p+2p^2)}{(1-p)^2p^2}, \tag{30}$$

$$S[X] = \frac{2[(1-p)^3-p^3]}{((1-p)^2+p^2)^{3/2}}, \qquad K[X] = \frac{9p^4+6p^2(1-p)^2+9(1-p)^4}{(1-2p+2p^2)^2}. \tag{31}$$

Note, when $p = 1$ we have that the AL distribution simplifies to the well known Laplace distribution. Hence, we may now observe that this family of distributions contains, embedded in the exponential argument, exactly the component required for minimization in the quantile regression loss function. This allows us to write the problem of solving for the coefficients of the model in Definition 4.1 given by:

$$f(y_t|\mathcal{F}_t;\mu_t,\sigma^2,p) = \frac{p(1-p)}{\sigma}\exp\left(-\frac{(y_t-\mu_t)}{\sigma}[p - I(y_t \le \mu_t)]\right) \tag{32}$$

for the location parameter or mode μ_t, the scale parameter $\sigma > 0$ and the skewness parameter $p \in (0,1)$ equals to the quantile level u. Since the pdf (32) contains the loss function (28), it is clear that parameter estimates which maximize (32) will minimize (28).

In this formulation the AL distribution represents the conditional distribution of the observed dependent variables (responses) given the covariates. More precisely, the location parameter μ_t of the AL distribution links the coefficient vector $\boldsymbol{\alpha}_u$ and associated covariates in the linear time series regression model to the location of the AL distribution.

4.2 Examples of Non-Linear Nonparametric Quantile Time Series Models

A natural extension of the QAR(p) class of quantile time series models is to consider the non-linear class of non-parametric models. Here we are treating the mapping $T(\cdot)$ as comprised of a combination of non-linear and potentially also linear components.

Under the representation presented for the QAR(p) model and its embedding within the AL family for estimation convenience, then in this context it is straightforward to extend the quantile regression model to allow for heteroscedasticity in the response which may vary as a function of the quantile level u under study. To achieve this one can simply add a regression structure linked to the scale parameter σ_t in the same manner as was done for the location parameter.

This would correspond to what we will call the Dynamic Volatility QAR(p) time series model given in the following definition.

Definition 4.3 (Examples of Non-Parametric Dynamic Volatility Quantile Autoregressive DV-QAR(p,q) Time Series)**.** *Consider the time series $(Y_t)_{t\ge 0}$, then the DV-QAR(p) quantile time series model is defined according to the conditional quantile functions as follows:*

$$Q_{Y_t}(u|\mathcal{F}_t) = \alpha_{0,u} + \sum_{k=1}^{p} \alpha_{k,u}\, y_{t-k} \tag{33}$$

with

$$\mathbb{V}ar[Y_t|\mathcal{F}_t] = \frac{1+u^4}{\sigma_t^2 u^2} = \frac{1+u^4}{u^2}\left(\beta_{0,u} + \sum_{k=1}^{q}\beta_{k,u}\, y_{t-k}\right)^{-2} \tag{34}$$

where $\mathcal{F}_t$ denotes information set or filtration that defines the time series dynamic. For instance, $\mathcal{F}_t$ may be the natural filtration generated by the observed time series $(Y_t)_{t\ge 0}$ up to the current time point, or it may contain additional structure such as lagged covariates. The notation, $u \in (0,1)$, corresponds to the quantile level, and location of the u-th quantile level is dictated by coefficient lagged previous values of the time series, where the coefficients can be quantile level specific, with $\boldsymbol{\alpha}_u = (\alpha_{0,u},\ldots,\alpha_{p,u})$ and $\boldsymbol{\beta}_u = (\beta_{0,u},\ldots,\beta_{q,u})$ the linear model coefficients for quantile level u. The coefficients are characterized as the solution to the system of equations, for all $s \in \{1,\ldots,t,\ldots\}$ given by:

$$\mathbb{P}r\left[y_s \le \alpha_{0,u} + \sum_{k=1}^{p}\alpha_{k,u}\, y_{s-k}|\mathcal{F}_s\right] = u, \tag{35}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2018 doi:10.20944/preprints201806.0198.v1

Peer-reviewed version available at Risks 2018, 6, 97; doi:10.3390/risks6030097

$$\begin{aligned}\frac{1}{t-1}\sum_{s=1}^{t}\left\{y_s - \left[\alpha_{0,u} + \sum_{k=1}^{p}\alpha_{k,u}\, y_{s-k} + \frac{1-u^2}{u}\left(\beta_{0,u} + \sum_{k=1}^{q}\beta_{k,u}\, y_{t-k}\right)\right]\right\}^2 \\ = \frac{1+u^4}{u^2}\left(\beta_{0,u} + \sum_{k=1}^{q}\beta_{k,u}\, y_{t-k}\right)^{-2}\end{aligned} \tag{36}$$

when such a solution exists.

Remark 4.2. *This second constraint in the above definition of the DV-QAR(p) process imposes the restriction that the time series process will admit a representation in which the original time series* $(Y_t)_{t\geq 0}$ *will be heteroskedastic with a volatility given by the functional specification:*

$$\mathbb{V}ar\left[Y_t|\mathcal{F}_t\right] = \frac{1+u^4}{\sigma_t^2 u^2} = \frac{1+u^4}{u^2}\left(\beta_{0,u} + \sum_{k=1}^{q}\beta_{k,u}\, y_{t-k}\right)^{-2} \tag{37}$$

In terms of fitting such a DV-QAR(p) model, it will be convenient to observe the following relationship between this non-parametric model and its embedding withing a parametric AL distributional model.

Remark 4.3 (Embedding of the non-parametric DV-QAR(p) within Scale-Location Varying Asymmetric Laplace Model)**.** *Equivalently, we assume* $Y_t|\mathcal{F}_t$ *conditionally follows an AL distribution denoted by* $Y_t|\mathcal{F}_t \sim AL(\mu_t, \sigma_t^2, u)$*. Then*

$$Y_t = \mu_t + \epsilon_t \sigma_t \tag{38}$$

where $\epsilon_t \sim AL(0,1,u)$*, the location and scale dynamic functions are given by*

$$\begin{aligned}\mu_t &= \alpha_{0,u} + \sum_{k=1}^{p}\alpha_{k,u}\, y_{t-k}, \\ \sigma_t^2 &= \exp(\beta_{0,u} + \sum_{k=1}^{q}\beta_{k,u}\, y_{t-k}).\end{aligned} \tag{39}$$

Discussion on the parametric regression model, in particular, the choice of link function and structure of regression terms will be undertaken in later sections.
Note: *this representation has the following advantages:*

- *the parameters can be estimated by maximum-likelihood under the AL distribution family; and*
- *importantly, it links the quantile process to a linear (when* $\sigma_t = \sigma$*) AR process with a driving noise sequence given by an AL error with appropriately chosen asymmetry parameter for* $p = u$ *corresponding to the target quantile level.*

Other examples of non-linear time series models have been proposed in the literature such as the double-AR time series structures of Cai et al. [2013], which we modify below to the non-parametric specification, embedded within an AL distribution estimation framework as noted in the $DV - QAR(p,q)$ models above.

Definition 4.4 (Non-Parametric Dynamic Volatility Quantile Double Autoregressive DV-QDAR(p,q) Time Series)**.** *Consider the time series* $(Y_t)_{t\geq 0}$*, then the DV-QDAR(p) quantile time series model is defined according to the conditional quantile functions as follows:*

$$Q_{Y_t}(u|\mathcal{F}_t) = \alpha_{0,u} + \sum_{k=1}^{p}\alpha_{k,u}\, y_{t-k} \tag{40}$$

with

$$\mathbb{V}ar\left[Y_t|\mathcal{F}_t\right] = \frac{1+u^4}{\sigma_t^2 u^2} = \frac{1+u^4}{u^2}\left(\beta_{0,u} + \sum_{k=1}^{q}\beta_{k,u}\, y_{t-k}^2\right)^{-1} \tag{41}$$

where $\mathcal{F}_t$ *denotes information set or filtration that defines the time series dynamic. For instance,* $\mathcal{F}_t$ *may be the natural filtration generated by the observed time series* $(Y_t)_{t\geq 0}$ *up to the current time point, or it may contain additional structure such as lagged covariates. The notation,* $u \in (0,1)$*, corresponds to the quantile level, and location of the u-th quantile level is dictated by coefficient lagged previous values of the time series, where the coefficients can be quantile level specific, with* $\boldsymbol{\alpha}_u = (\alpha_{0,u}, \ldots, \alpha_{p,u})$ *and* $\boldsymbol{\beta}_u = (\beta_{0,u}, \ldots, \beta_{q,u})$*, such that* $\beta_{k,u} > 0$ *for all* $k \in \{0,1,\ldots,q\}$*, are the linear model coefficients for quantile level* u*.*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2018 doi:10.20944/preprints201806.0198.v1

Peer-reviewed version available at Risks 2018, 6, 97; doi:10.3390/risks6030097

solutions to the following system of equations, for all $s \in \{1, \ldots, t, \ldots\}$ given by:

$$\mathbb{P}\mathrm{r}\left[y_s \leq \alpha_{0,u} + \sum_{k=1}^{p} \alpha_{k,u}\, y_{s-k} | \mathcal{F}_s\right] = u, \tag{42}$$

subject to the constraint for the filtration $\mathcal{F}_t$ given by

$$\begin{aligned} &\frac{1}{t-1}\sum_{s=1}^{t}\left\{y_s - \left[\alpha_{0,u} + \sum_{k=1}^{p}\alpha_{k,u}\, y_{s-k} + \frac{1-u^2}{u}\left(\beta_{0,u} + \sum_{k=1}^{q}\beta_{k,u}\, y_{t-k}\right)\right]\right\}^2 \\ &= \frac{1+u^4}{u^2}\left(\beta_{0,u} + \sum_{k=1}^{q}\beta_{k,u}\, y_{t-k}^2\right)^{-1} \end{aligned} \tag{43}$$

when such a solution exists.

As demonstrated previously, the scale or volatility function has been specifically written in a form that naturally admits its embedding within an AL distribution family, this greatly facilitates the estimation, as one can directly avoid the estimation under the complicated non-linear coupled and constrained system of equations above, replacing this with standard maximum likelihood of the AL distribution family for two of its parameters μ and σ.

5 Parametric Quantile Time Series Models

In this case we will consider the sub-class of regression quantile time series models given by

$$Q_{Y_t}\left(u|\mathcal{F}_{t-1}, \mathcal{G}_t; \boldsymbol{\theta}\right) = T\left(\mathcal{F}_t, \mathcal{G}_t, Q_\epsilon(u)\right) \tag{44}$$

where we now include explicitly in the transformation $T(\cdot)$ the component corresponding to the quantile error distribution specification $Q_\epsilon(u)$, hence making the model parametric in nature.

5.1 Examples of Linear Parametric Quantile Time Series Models

In this section we discuss some core examples of different choices of parametric quantile time series models. Then in future sections we will jointly describe very general choices for functions $T(\cdot)$ and quantile error function classes $Q_\epsilon(u)$ that will jointly transform the parametric quantile error family into a conditional quantile function for Y_t. In terms of $T(\cdot)$ maps, the most common choice of transform classes that will be applicable will be the class of linear additive, non-linear multiplicative and non-linear Q-transform rules, see discussions in Gilchrist [2000].

To illustrate an example of the class of parametric quantile time series models that are naturally included in this general quantile family structure, we consider the generalized linear quantile seasonal autoregressive integrated model framework as specified in Definition 5.1. This class of models already includes many models proposed previously in the literature, and corresponds to simple linear form for the transform function $T(\cdot)$ in lags of the observed time series.

Definition 5.1 (Generalized Linear Quantile SARI (GL-QSARI) time series)**.** *We define the class of generalized linear quantile seasonal autoregressive integrated SARI models (GL-QSARI(p,d,P,D)) by the following transformation function $T(\cdot)$*

$$\begin{aligned} Q_{Y_t}\left(u|\mathcal{F}_{t-1}; \boldsymbol{\theta}\right) &= T\left(\mathcal{F}_t, Q_\epsilon(u)\right) \\ &= \phi(B,u)\,\Phi(B,u)\,\nabla^d\,\nabla_s^D Y_{t-1} + Q_\epsilon\left(u; \boldsymbol{\gamma}\right) \end{aligned} \tag{45}$$

with $u \in [0,1]$ and where $Q_{Y_t}\left(u|\mathcal{F}_t; \boldsymbol{\theta}\right)$ denotes the conditional quantile function of random variable Y_t and the generalized operators for the quantile function setting given by:

$$\begin{aligned} \phi(x,u) &= 1 + \phi_1(u)x + \phi_2(u)x^2 + \ldots + \phi_p(u)x^p \\ \Phi(x,u) &= 1 + \Phi_1(u)x^s + \Phi_2(u)x^{2s} + \ldots + \Phi_P(u)x^{Ps}. \end{aligned}$$

where $\phi_i(u)$ and $\Phi_i(u)$ each denote the i-th lagged functional coefficient of the AR and SAR time series model components with $\boldsymbol{\theta}$ denoting generically the vector of all model parameters. Furthermore, we denote $Q_\epsilon\left(u; \boldsymbol{\gamma}\right)$ as the time series white noise i.i.d. error ϵ_t quantile function with parameters $\boldsymbol{\gamma}$.

long as $\phi(B,u)\,\Phi(B,u)\,\nabla^d\,\nabla_s^D Y_{t-1}$ is a monotone increasing function of u.

In the following, we outline some examples of sub-models and model restrictions of the GL-QSARI framework that will produce a valid conditional quantile function.

Example 5.1. *There are multiple ways that one can achieve a valid conditional quantile function of the GL-QSARI model ($Q_{Y_t}(u|\mathcal{F}_{t-1};\boldsymbol{\theta})$) such as outlined below in different sub-model constructions:*

1. *Model example one can be constructed by assuming that each $\phi_i(u)$ and $\Phi_i(u)$ are monotone increasing functions of $u \in [0,1]$ and each coefficient function is a positive function such that $\phi_i(u) : [0,1] \mapsto \mathbb{R}^+$ and $\Phi_i(u) : [0,1] \mapsto \mathbb{R}^+$.;*
2. *Model example two can be constructed by assuming that each $\phi_i(u)$ are monotone increasing functions of $u \in [0,1]$ and each coefficient function $\Phi_i(u)$ is a constant function.;*
3. *Model example three can be constructed by assuming that each $\Phi_i(u)$ are monotone increasing functions of $u \in [0,1]$ and each coefficient function $\phi_i(u)$ is a constant function.*

Remark 5.2. *It is interesting to note that the above model can be specified in a quantile time series contexts, without ever having to specify explicitly the time series model underlying the random variable sequence Y_t. Furthermore, the estimation of such a model can be done independently of the estimation of the model parameters in the corresponding equivalent time series model for Y_t.*

5.2 Examples of Linear Parametric Quantile Time Series with Distributed Lags

Next one can readily extend this model to the class of distributed lag GL-QSARIDL models, where it is possible to incorporate a set of exogenous lagged observable covariates into the model structure. Again we illustrate this example with the linear class of function $T(\cdot)$, noting that in future sections of the manuscript we will describe more general classes of mapping function $T(\cdot)$ which can be applied to both the quantile error function and the structural components of the regression.

Definition 5.2 (Generalized Linear Quantile SARI Distributed Lag (GL-QSARIDL) time series)**.** *We define the class of generalized linear quantile seasonal autoregressive integrated SARI models (GL-QSARI(p,d,r,P,D)) by the following transformation function $T(\cdot)$*

$$\begin{aligned} Q_{Y_t}(u|\mathcal{F}_{t-1};\boldsymbol{\theta}) &= T(\mathcal{F}_t,\mathcal{G}_t,Q_\epsilon(u)) \\ &= \phi(B,u)\,\Phi(B,u)\,\nabla^d\,\nabla_s^D Y_{t-1} + \widetilde{\phi}(B,u)X_t + Q_\epsilon(u;\boldsymbol{\gamma}) \end{aligned} \tag{46}$$

with $u \in [0,1]$ and where $Q_{Y_t}(u|\mathcal{F}_t,\mathcal{G}_t;\boldsymbol{\theta})$ denotes the conditional quantile function of random variable Y_t, and the generalized operators for the quantile function setting given by:

$$\begin{aligned} \phi(x,u) &= 1+\phi_1(u)x+\phi_2(u)x^2+\ldots+\phi_p(u)x^p \\ \widetilde{\phi}(x,u) &= 1+\phi_1(u)x+\phi_2(u)x^2+\ldots+\phi_r(u)x^r \\ \Phi(x,u) &= 1+\Phi_1(u)x^s+\Phi_2(u)x^{2s}+\ldots+\Phi_P(u)x^{Ps}. \end{aligned}$$

Furthermore, we denote $Q_\epsilon(u;\boldsymbol{\gamma})$ as the time series white noise i.i.d. error ϵ_t quantile function with parameters $\boldsymbol{\gamma}$.

Remark 5.3. *Note, this model will of course be a well defined quantile time series model so long as $\phi(B,u)\,\Phi(B,u)\,\nabla^d\,\nabla_s^D Y_{t-1}$ and $\widetilde{\phi}(B,u)X_t$ are each monotone increasing functions of u.*

In future sections we will discuss other classes of model for which the mapping function $T(\cdot)$ is no longer selected as a linear map, but the resulting conditional function $Q_{Y_t}(u|\mathcal{F}_t,\mathcal{G}_t;\boldsymbol{\theta})$ will still represent a well defined conditional quantile function of random variable Y_t. In general, we will also spend time explaining different families of quantile error function $Q_\epsilon(u;\boldsymbol{\gamma})$ that can be considered.

We note that in this context of quantile transformations, we will generally consider specifically special classes of model that satisfy the condition that Gilchrist [2000] refers to as the class of non-linear transforms known as the P-Class. Special cases of such P-Class we will consider include the family of Tukey Elongation transforms above (see summary in Peters et al. [2016], Peters and Sisson [2006], Cruz et al. [2015] and Peters

[2009] in which they consider the special sub-set of models defined by

$$v = G\left[F^{1}(u)\right], \tag{47}$$

where F and G are cumulative distribution functions (CDFs). In addition, there are numerous authors who have studied the generalized properties of quantile-based functionals of asymmetry and kurtosis see Balanda and MacGillivray [1990], Rayner and MacGillivray [2002].

We begin by first exploring below the choice of quantile error functions $Q_\epsilon(u;\boldsymbol{\gamma})$ that are closed-form and flexible enough to be used in a range of parametric quantile time series modelling contexts. Following from these specifications, we then discuss different examples of transformations to obtain conditional quantile functions.

6 Parametric Quantile Time Series Models: Error Quantile Functions

In this section, we will introduce a key component of the the generic parametric quantile time series model framework we propose based on representation:

$$Q_{Y_t}(u|\mathcal{F}_{t-1},\mathcal{G}_t;\boldsymbol{\theta}) = T\left(\mathcal{F}_t,\mathcal{G}_t,Q_\epsilon(u)\right), \tag{48}$$

which focuses on the modelling choices of the quantile error function $Q_\epsilon(u)$.

In explaining different families of models for $Q_\epsilon(u)$ we will also introduce two highly flexible choices of mapping function $T(\cdot)$ that can either be applied to known parametric quantile error functions to obtain more flexible families of error quantile function or they can be applied to the quantile time series relationship $T\left(\mathcal{F}_t,\mathcal{G}_t,Q_\epsilon(u)\right)$ to produce non-linear quantile time series models.

It is also important to talk about families of quantile functions that admit parametric representations, as can be expected in many cases a random variable Y may have a well defined and closed form expression for its distribution function $F_Y(y;\boldsymbol{\theta})$, however its quantile function given by $Q_Y(u;\boldsymbol{\theta}) = F_Y^{-1}(y;\boldsymbol{\theta})$ may not be easily obtainable as a function in closed form. However, there are several important and practical cases for different classes of parametric models for which one can obtain both functions in closed form, these are discussed below and presented in the context of quantile error models. This is the analog in time series settings of thinking about the quantile function of ϵ_t the generic notation for the driving noise in the time series.

In this section we specifically consider a few flexible parametric models that can be used to act as reference error quantile functions that will form the core input to the mapping $T(\cdot)$ in the parametric class of quantile time series models we will consider.

We will separate the quantile error models into three categories:

- Location and Scale families of quantile function;
- Shape and Scale families of quantile function; and
- Heavy tailed families of quantile function.

In addition to these examples of parametric quantile error functions - we note that one may adopt a class of transformations of such quantile functions to construct different variants of these quantile models with additional properties and extra degrees of freedom. Such transformations are discussed below in more detail with regard to the Tukey Elongation class.

6.1 Location and Scale Quantile Error Families

In this section we discuss a few examples of quantile error model that practitioners can consider in the class of location scale models. Such models are useful for practitioners in time series contexts as they allow for direct interpretation of the parameters and the role such quantile error parameters have on adjustments to the conditional quantile function of the time series, as a result of perturbation of the error parameters. We will present four core models for practitioners to use which represent a range of light and heavy tailed structure as well as asymmetric structures around the mode of the error distribution. Some simple and practically useful examples of location scale quantile functions include the light tailed symmetric case of Gaussian and the heavy tailed case of Cauchy.

Peer-reviewed version available at Risks 2018, 6, 97; doi:10.3390/risks6030097

$N(\mu, \sigma)$ is given by

$$Q_\epsilon(u) = \mu + \sigma\sqrt{2}erf^{-1}(2u-1) \tag{49}$$

where the error function is given by

$$\begin{aligned} erf(u) &= \frac{1}{\sqrt{\pi}}\int_{-x}^{x} e^{-t^2}\,\mathrm{d}t \\ &= \frac{2}{\sqrt{\pi}}\int_{0}^{x} e^{-t^2}\,\mathrm{d}t. \end{aligned} \tag{50}$$

Such a model is of relevance if a practitioners believes there are no extreme observations in the time series that is being considered for fitting. In addition, this model has a symmetric consideration in the tails.

Definition 6.2 (Cauchy Quantile Function)**.** *The quantile function for a Cauchy random variable $\epsilon \sim Cauchy(\mu, \sigma)$ is given by*

$$Q_\epsilon(u) = \mu + \sigma\tan\left[\pi\left(u - \frac{1}{2}\right)\right]. \tag{51}$$

Unlike the Gaussian case, here we consider a heavy tailed error quantile function. Such a model is of relevance if a practitioners believes there is likely to be extreme observations in the time series that is being considered for fitting. In addition, this model has a symmetric consideration in the tails.

Definition 6.3 (Asymmetric Laplace Quantile Function)**.** *The quantile function for a Asymmetric Laplace random variable $\epsilon \sim AL(\mu, \sigma, p)$ is given by*

$$Q_\epsilon(u) = \begin{cases} \mu + \frac{\sigma}{1-p}\log(\frac{u}{p}), & \text{if } 0 \le u \le p, \\ \mu - \frac{\sigma}{p}\log(\frac{1-u}{1-p}), & \text{if } p < u \le 1. \end{cases} \tag{52}$$

Note that the shape parameter p of the AL distribution gives the magnitude and direction of skewness. AL distribution is skewed to left when $p > 0.5$ and skewed to right when $p < 0.5$ and hence it can model the left skewness of most log transformed loss data directly through this shape parameter p.

This model is a compromise between the two previous models. The ALD is popular in practice due to it convenient parametric structure for quantile loss functions, however from the perspective of distributional properties it allows for a light to intermediate tail behaviour. In addition, it allows for asymmetric distributional properties in the left and right tails of the error distribution. It does not however allow for heavy tailed features and often it may be beneficial to consider heavy tails as well as asymmetry. One way to achieve this was studied extensively in Zhu and Zinde-Walsh [2009] where they discuss the relationship between popular families of models the Exponential Power distributions (EPD), the Skewed Exponential Power distributions (SEPD) and the Asymmetric Exponential Power distributions (AEPD).

In Bernardi et al. [2016a] where they introduced the notion of the skewed exponential power (SEP) distribution to quantile regression settings. The SEP model has found wide application uptake in volatility modelling contexts, see examples in Marín and Sucarrat [2012] and DiCiccio and Monti [2004] and references therein.

Definition 6.4 (Skewed Exponential Power Quantile Function)**.** *If the error random variable is considered distributed according to a Skewed Exponential Power distribution, $\epsilon \sim SEP(\mu, \sigma, \alpha, \tau)$, then its density is given by*

$$f_{SEP}(\epsilon; \mu, \sigma, \alpha, \tau) = \begin{cases} \frac{1}{\sigma}\kappa_{EP}(\alpha)\exp\left[-\frac{1}{\alpha}\left(\frac{\mu-\epsilon}{2\tau\sigma}\right)^\alpha\right], & \text{if } \epsilon \le \mu, \\ \frac{1}{\sigma}\kappa_{EP}(\alpha)\exp\left[-\frac{1}{\alpha}\left(\frac{\epsilon-\mu}{2(1-\tau)\sigma}\right)^\alpha\right], & \text{if } \epsilon > \mu, \end{cases} \tag{53}$$

with $\mu \in \mathbb{R}$ the location parameter, $\sigma \in \mathbb{R}^+$ and $\alpha \in (0, \infty)$ the scale and shape parameters respectively. In addition, the parameter $\tau \in (0,1)$ controls the skewness of the distribution. Furthermore, we denote by κ_{EP} the function

$$\kappa_{EP} = \left[2\alpha^{\frac{1}{\alpha}}\Gamma\left(1 + \frac{1}{\alpha}\right)\right]^{-1}$$

where $\Gamma(\cdot)$ is the complete gamma function. When represented in this density form, one can observe that the location parameter μ will directly correspond to the τ level quantile. One can also express the quantile *given by*

$$Q_\epsilon\left(u;\alpha,p_1,p_2\right)=\begin{cases}-2\alpha^*\left[p_1G^{-1}\left(1-\frac{\nu}{\alpha};\frac{1}{p_1}\right)\right]^{1/p_1}, & \text{if } \nu\leq\alpha,\\ 2(1-\alpha^*)\left[p_2G^{-1}\left(1-\frac{1-\nu}{1-\alpha};\frac{1}{p_2}\right)\right]^{1/p_2}, & \text{if } \nu>\alpha,\end{cases} \tag{54}$$

where $\nu\in[0,1]$ *and*

$$G(\epsilon;\eta)=\frac{1}{\Gamma(\eta)}\int_0^\epsilon z^{\eta-1}\exp\left(-z\right)dz.$$

To obtain the SEP distribution from the AEP one selects $p_1=p_2=p$ *and* $\alpha^*=\alpha$.

We also note that practitioners can gain an intuition for this model through the fact that this model is related directly to sub-families of skewed Laplace distribution and skewed normal distributions.

6.2 Shape and Scale Quantile Error Families

The other class of models that are of relevance to practitioners include the shape and scale family of models. We consider some examples of shape scale families of models that can act as base quantile models for the error distribution in a quantile regression. This section will introduce light tailed through to ultra-heavy tailed models. The advantage of such heavy tailed models is that they may provide the ability to capture extreme observations in the observed time series more accurately. Another perspective on such models is that they may allow for a robustification to outliers of quantile time series modelling.

We begin with a simple model in the class of shape-scale quantile families that can be utilized as an error distribution, given by the Weibull distributions quantile function.

Definition 6.5 (Weibull Quantile Function)**.** *The quantile function for a Weibull random variable* $\epsilon\sim Weibull(\alpha,\beta)$ *with shape* $\alpha>0$ *and scale* $\beta>0$ *is given by*

$$Q_\epsilon(u;\alpha,\beta)=\begin{cases}\beta\left[-\ln\left(1-u\right)\right]^{\frac{1}{\alpha}}, & u\geq F(0),\\ 0, & u<F(0)\end{cases} \tag{55}$$

As a second case of shape-scale family of models we consider a well known transform of the location-scale Gaussian case, given by the Log-Normal model quantile error function.

Definition 6.6 (Log-Normal Quantile Function)**.** *The quantile function for a Log-Normal random variable* $\epsilon\sim LogNormal(\mu,\sigma)$ *with* $\mu>\mathbb{R}$ *and* $\sigma>0$ *is given by*

$$Q_\epsilon(u;\mu,\sigma)=\exp\left(\mu+\sigma\sqrt{2}erf^{-1}(2u-1)\right) \tag{56}$$

The next example of shape scale family of models involves variations of a power law error quantile function, given by modifications of the Pareto quantile function.

Definition 6.7 (Pareto Quantile Function)**.** *The quantile function for a Pareto random variable* $\epsilon\sim Pareto(x_m,\alpha)$ *with distribution given by*

$$F_\epsilon(\epsilon;x_m,\alpha)=\begin{cases}1-\left(\frac{x_m}{\epsilon}\right)^\alpha, & \epsilon\geq x_m,\\ 0, & \epsilon<x_m,\end{cases} \tag{57}$$

is given by

$$Q_\epsilon(u;x_m,\alpha)=\begin{cases}(1-u)^{-\frac{1}{\alpha}}x_m, & u\geq F_\epsilon(x_m),\\ 0, & u<F_\epsilon(x_m)\end{cases} \tag{58}$$

This general class of quantile error function has been previously extended to multi-parameter versions, for instance those studied in the works of Cai [2010a] and Dong et al. [2015]. Below, we present a simple example of such quantile models that one may adopt for a quantile error function given by the polynomial power-Pareto (PP) quantile error function model.

 doi:10.20944/preprints201806.0198.v1

Peer-reviewed version available at *Risks* **2018**, *6*, 97; doi:10.3390/risks6030097

function for a random variable $\epsilon \sim PP(\gamma_1, \gamma_2)$ with density given by

$$f_\epsilon(\epsilon_i|\mu,\sigma,\gamma_1,\gamma_2) \quad = \quad \frac{u_i^{1-\gamma_1}(1-u_i)^{\gamma_2+1}}{\sigma[\gamma_2 u_i + \gamma_1(1-u_i)]} \tag{59}$$

where u_i is an implicit function of the following structure which can be obtained by solving the system of equations defined for each observation

$$\epsilon_i = \mu + u_i^{\gamma_1}\,(1-u_i)^{-\gamma_2}\,\sigma \tag{60}$$

The resulting quantile distribution of this model is the combination of a power distribution with a Pareto distribution, which enables us to model both the main body and the tails of a distribution. In considering the PP model the quantile function of is comprised of two components:

- *component 1: a power distribution $F_1(\epsilon) = \epsilon^{\frac{1}{\gamma_1}}$ where $\epsilon \in [0,1]$ and $\gamma_1 > 0$ with a corresponding quantile function then given by $Q_1\,(u;\gamma_1) = u^{\gamma_1}$ for $u \in [0,1]$; and*

- *component 2: a Pareto distribution function $F_2(\epsilon) = 1 - \epsilon^{-\frac{1}{\gamma_2}}$ where $\epsilon \geq 1$ and $\gamma_2 > 0$ with a corresponding quantile function then given by $Q_2\,(u;\gamma_2) = (1-u)^{-\gamma_2}$.*

One may use the fact that the product of the two quantile functions will remain a strictly valid quantile function producing the new quantile function family known as the Polynomial-Power Pareto model. The resulting structural form given by the inverse cdf of the Pareto distribution with an additional polynomial power term:

$$Q_\epsilon(u) := F_\epsilon^{-1}(u|\gamma_1,\gamma_2) = u^{\gamma_1}(1-u)^{-\gamma_2}. \tag{61}$$

To complete the examples we show a very flexible class of shape scale model corresponding to the Generalized Beta Family (GB2) class of models. In the case of the GB2 class of models, we see that this family contains sub-families of models given by the Generalized Gamma (GG) family and the standard shape scale family of the Gamma distribution.

The type two generalized beta distribution (GB2) has attractive features for modelling, as it has a positive support $\mathbb{R}^+$ and nests a number of important distributions as its special cases. The GB2 distribution has four parameters, which allows it to be expressed in various flexible densities. See discussions in Dong and Chan [2013] for a more detailed description of GB2 distribution including its pdf and distribution family.

Definition 6.9 (Type 2 Generalized Beta Quantile Error Function)**.** *If $\epsilon \in \mathbb{R}^+$ follows a GB2 distribution, then it can be characterized by the density given by*

$$f_\epsilon(\epsilon|a,b,p,q) = \frac{\frac{a}{b}(\frac{\epsilon}{b})^{ap-1}}{B(p,q)[1+(\frac{\epsilon}{b})^a]^{p+q}}, \quad \text{for } \epsilon \geq 0 \tag{62}$$

where a, p and q are shape parameters and b is the scale parameter. We may rewrite the GB2 model as a generalized Beta distribution with pdf

$$f_B(\widetilde{\epsilon}|p,q) = \frac{1}{B(p,q)}\widetilde{\epsilon}^{p-1}(1-\widetilde{\epsilon})^{p+q} \tag{63}$$

via the transformation $\widetilde{\epsilon} = \dfrac{(\frac{\epsilon}{b})^a}{1+(\frac{\epsilon}{b})^a}$. The GB2 is directly relevant for quantile regression models since one may also find its quantile function in closed form according to the following expression:

$$Q_\epsilon(u) = \frac{\exp\left(\mu\right)B(p,q)}{B(p+1/a,q-1/a)}\left(\frac{F_B^{-1}(u|p,q)}{1-F_B^{-1}(u|p,q)}\right)^{\frac{1}{a}}. \tag{64}$$

where $\mu = \mathbb{E}[\epsilon]$.

We note that in general, when we know the mean function of the model as well as the quantile function, we may either perform a mean regression to estimate the parameters or a quantile regression - *these two different approaches will in general produce different results for the resulting parameter estimates of the model, except*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2018 doi:10.20944/preprints201806.0198.v1

Peer-reviewed version available at Risks 2018, 6, 97; doi:10.3390/risks6030097

mean and one performs quantile regression for the median. In this case one would obtain a more robust regression (less sensitive to outliers) than obtained form the mean regression. In all other cases, these results will differ, however having fit either the quantile model, or the mean regression model, we may reverse back to get the quantile model from a mean fit or the mean implied by the quantile regression fit and compare their differences, as indicated for the GB2 model below.

Remark 6.1 (Link Between GB2 Quantile Error Function and Mean Regressions)**.** *In mean regression, b can be linked to the mean μ of the distribution as follows:*

$$b = \frac{\mu B(p,q)}{B(p+1/a, q-1/a)} \tag{65}$$

where μ is for instance a log-link to a linear function of covariates μ in (66) according to the relationship:

$$\mathbb{E}\left[Y_i|\boldsymbol{x}_i\right] = \mu_i = \exp\left(\alpha_0 + \sum_{k=1}^{m} \alpha_k\, x_{i,k}\right). \tag{66}$$

Then the variance is given by:

$$\mathbb{V}ar\left[Y_i|\boldsymbol{x}_i\right] = \mu_i^2 \left\{\frac{B(p,q)B(p+2/a, q-2/a)}{[B(p+1/a, q-1/a)]^2} - 1\right\}. \tag{67}$$

There are many widely known and utilized sub-families of the GB2 family, see a detailed overview of these in Dong and Chan [2013], McDonald [1984] and Stacy [1962].

We would argue that from a parsimony perspective practitioners would be best suited to first try the simple 2 parameter families to assess the quality of their fitted quantile time series models, if the resulting model fit is adequate then these would suffice. If however, the fit is not adequate, then one may generalize to the 3 and 4 parameter families that also admit heavy tailed features and general skewness structure.

6.3 Truncated Error Quantile Functions

It will often in practice be beneficial to work with models for which the random variable of interest Y_t will be restricted to one of the possible domains $Y \in [L, \infty)$, $Y \in (-\infty, U]$ or $Y \in [L, U]$. In this case, the model constructed will require the quantile error function $Q_{\epsilon_t}(u)$ to also be restricted to this domain. This is easily achieved in many cases and we will illustrate this below with examples of truncated quantile error families.

To proceed, we consider a standard distribution $F(y)$ (such as LogNormal, Gamma, etc.) with a corresponding density function $f(y)$. However, one may be interested in modeling a time series restricted above some threshold $L \in \mathbb{R}$ only. Then, one can consider a distribution truncated below L formally defined as

$$F^{tr}(y) = \frac{F(y) - F(L)}{1 - F(L)} \mathbb{I}_{y \geq L} \tag{68}$$

with a corresponding truncated density function

$$f^{tr}(y) = \frac{f(y)}{1 - F(L)} \mathbb{I}_{y \geq L}. \tag{69}$$

Note that this truncated density is a proper density function, that is, $\int\limits_0^\infty f^{tr}(y)dy = 1$.

In principle, assuming the mapping $T(\cdot)$ is restricted in its range to the interval $Y_t \in [L, \infty)$, this would not necessarily require explicitly that $Q_{\epsilon_t}(u)$ be restricted to the same interval, however in practice it would be natural to consider such cases. Therefore, we briefly outline how this is easily achieved in parametric quantile error models generically as follows, for $u \in [0, 1]$ as follows

$$\begin{aligned} F^{tr}(\epsilon) &= \frac{F(\epsilon) - F(L)}{1 - F(L)} \mathbb{I}_{\epsilon \geq L} \\ \Rightarrow F(\epsilon)\mathbb{I}_{\epsilon \geq L} &= u\left(1 - F(L)\right) + F(L) \\ \Rightarrow F^{-1}(\epsilon)\mathbb{I}_{\epsilon \geq L} &= Q_\epsilon(\widetilde{u})\mathbb{I}_{\widetilde{u} \geq F(L)} \end{aligned} \tag{70}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2018 doi:10.20944/preprints201806.0198.v1

Peer-reviewed version available at Risks 2018, 6, 97; doi:10.3390/risks6030097

appropriate restriction from the indicator and adjustment to quantile level from the normalization from the truncation.

Similarly, one can model below L using distribution truncated above L:

$$F^{tr}(y) = \frac{F(y)}{F(L)}\mathbb{I}_{y \leq L}, \quad f^{tr}(y) = \frac{f(y)}{F(L)}\mathbb{I}_{y \leq L}. \tag{71}$$

As above, we can easily tackle this case also in parametric quantile error models generically as follows, for $u \in [0,1]$ as follows

$$\begin{aligned} F^{tr}(\epsilon) &= \frac{F(\epsilon)}{F(L)}\mathbb{I}_{\epsilon \leq L} \\ &\Rightarrow F(\epsilon)\mathbb{I}_{\epsilon \leq L} = uF(L) \\ &\Rightarrow F^{-1}(\epsilon)\mathbb{I}_{\epsilon \leq L} = Q_\epsilon(\widetilde{u})\mathbb{I}_{\widetilde{u} \leq F(L)} \end{aligned} \tag{72}$$

where $\widetilde{u} = uF(L)$.

If there is a need to model in a specific range $[L, U]$, one can use distribution $F(y)$ truncated below L and above U:

$$F^{tr}(y) = \frac{F(y) - F(L)}{F(U) - F(L)}\mathbb{I}_{L \leq y \leq U}, \quad f^{tr}(y) = \frac{f(y)}{F(U) - F(L)}\mathbb{I}_{L \leq y \leq U}. \tag{73}$$

The truncated quantile error model can be obtained as follows

$$\begin{aligned} F^{tr}(\epsilon) &= \frac{F(\epsilon) - F(L)}{F(U) - F(L)}\mathbb{I}_{L \leq \epsilon \leq U} \\ &\Rightarrow F(\epsilon)\mathbb{I}_{L \leq \epsilon \leq U} = u\left(F(U) - F(L)\right) + F(L) \\ &\Rightarrow F^{-1}(\epsilon)\mathbb{I}_{L \leq \epsilon \leq U} = Q_\epsilon(\widetilde{u})\mathbb{I}_{F(L) \leq \widetilde{u} \leq F(U)} \end{aligned} \tag{74}$$

where $\widetilde{u} = u\left(F(U) - F(L)\right) + F(L)$ and $Q_\epsilon(\cdot)$ is the same functional form as the

7 Generalized Elongation Deformation Quantile Error Families

In this section we discuss the family of quantile deformation models generally known in statistics as the family of Tukey elongation transforms, see detailed overview of such models in Peters et al. [2016] and Peters and Sisson [2006]. This family of models can be considered to be a generalization of the family of Rank Transmutation Maps (RTMs) discussed in Shaw and Buckley [2009]. Others who have addressed similar issues to do with distortion transforms to map quantile functions of a base distribution to another class of distributions include the early work of De Helguero [1908], there has since also been the related stream of works who studied distortions of density functions (as opposed to directly the quantile) developed by VICARI [2005], Azzalini [2005] and Genton [2005].

Here we discuss several distributional families relevant to modelling which can only be specified via the transformation of another standard random variable, for example a Gaussian. Examples of such models which are typically defined through their quantile functions include the Johnson family, with base distribution given by Gaussian or logisitic, and the Tukey family with base distribution typically given by a Gaussian or logistic. The concept of constructing skewed and heavy-tailed distributions through the use of a transformation of a Gaussian random variable was originally proposed in the work of Tukey [1977] and is therefore aptly named the family of Tukey distributions. This family of distributions was then extended by Hoaglin [1985a], Jorge and Boris [1984], Azzalini [1985] and Fischer et al. [2007]. The multivariate versions of these models have been discussed by Field and Genton [2012].

Within this family of distributions, two particular subfamilies have received the most attention in the literature; these correspond to the g-and-h and the g-and-k distributions. The first of these families the g-and-h has been studied in several contexts, see for instance the developments in the areas of risk and insurance modelling in Dutta and Perry [2006], Peters and Sisson [2006], Degen et al. [2007], Jiménez and Arunachalam [2011] and the detailed discussion in [Cruz et al., 2015, Chapter 9]. The second family of g-and-k models has been looked at in works such as Haynes et al. [1997] and Hossain and Hossain [2009].

The advantage of models such as the g-and-h family for modelling is the fact that they provide a very flexible range of skew, kurtosis, and heavy-tailed features while also being specified as a rather simple transformation of standard Gaussian random variates, making simulation under such models efficient and simple.

 doi:10.20944/preprints201806.0198.v1

Peer-reviewed version available at *Risks* **2018**, *6*, 97; doi:10.3390/risks6030097

Tukey suggested several nonlinear transformations of a reference random variable, typically considered to be symmetric and often selected to be a standard normal random variable in practical model applications, which will be denoted below by $W \sim \text{Normal}(0,1)$. There are then several sub-families of elongation transform that each produce different transformations of the reference quantile function of random variable W that induce specific skew and kurtosis features, relative to the base model.

One of the most well known of these classes of transformation is the g-and-h transformations which involve a skewness transformation of type g and a kurtosis transformation of type h. If one replaces the kurtosis transformation of the type h with the type k, one obtains the g-and-k family of distributions discussed by Rayner and MacGillivray [2002]. If the type h transformation is replaced by the type j transformation, one obtains the g-and-j transformations of Fischer and Klein [2004].

The generic specification of the Tukey transformation is provided in Definition 7.1. These types of transformations were labeled elongation transformations, where the notion of elongation was noted to be closely related to tail properties such as heavy-tailedness. See discussions by Hoaglin [1985b]. In considering such a class of elongation transformations to obtain a distribution, one is comparing the tail strength of the new distribution with that of the base distribution (such as a Gaussian or logistic). In this regard, one can think of tail strength or heavy-tailedness as an absolute concept, whereas the notion of elongation strength is a relative concept. In the following, we will first consider relative elongation compared to a base distribution for a generic random variable W. It should be clear that such a measure of relative tail behavior is independent of location and scale.

Remark 7.1 (Desirable Properties of Quantile Elongation Transformation)**.** *An elongation transformation $T(\cdot)$ should also satisfy the following properties:*

1. *Preservation of Symmetry: it is desirable that should one wish equi-probable tails on left and right, then the mapping should be able to preserve symmetry, say around the mode, such that $T(w) = T(-w)$ will hold under certain parameter settings;*

2. *Deformation Around the Mode Controlled: the base distribution for the random variable's quantile function being transformed should not be significantly transformed/deformed in the center, such that $T(w) = w + O(w^2)$ for w around the mode;*

3. *Additional Relative Skewness and Relative Kurtosis: to increase the heaviness of the tails of the resulting distribution relative to the base distribution, it is important to assume that $T(\cdot)$ is a strictly monotonically increasing transform that is convex, that is, one has the transform satisfying for $w > 0$ that $T'(w) > 0$ and $T''(w) > 0$.*

One such transformation family satisfying these properties is the Tukey transformations.

Definition 7.1 (Tukey transformations)**.** *Consider a Gaussian random variable $W \sim \text{Normal}(0,1)$ and transformation $X = r(W)$ then the resultant transformed error variable ϵ will be from a Tukey law if the corresponding transformation $r(W)$ is given by*

$$r(W) = WT(W)^{\theta}, \tag{75}$$

for a parameter $\theta \in \mathbb{R}$. Under this transformation we also have directly in closed form the quantile function of the error random variable ϵ in terms of the quantile function of the base random variable W as follows

$$Q_{\epsilon}(u) = a + bQ_W(u)T(Q_W(u))^{\theta} \tag{76}$$

with translation and scaling constants a, b for quantile levels $u \in [0,1]$.

Typically, in several applications, it will be desirable when working with such severity models to enforce a constraint that the tails of the resulting distribution after transformation are heavier than the Gaussian distribution. In this case, one should consider a transformation $T(w)$, which is positive, symmetric, and strictly monotonically increasing for positive values of $w \geq 0$. In addition, it will be desirable to obtain this property of heavy tails relative to the Gaussian to also consider setting the parameter $\theta \geq 0$. As discussed, a series of kurtosis transformations are proposed in the literature. The Tukey transformations of types h, k, and j are provided in Definition 7.2.

by $T_h(w)$, is given by

$$T_h(w) = \exp\left(w^2\right). \tag{77}$$

The k-type transformation, denoted by $T_k(w)$, is given by

$$T_k(w) = 1 + w^2. \tag{78}$$

The j-type transformation, denoted by $T_j(w)$, is given by

$$T_j(w) = \frac{1}{2}\left[\exp(w) + \exp(-w)\right]. \tag{79}$$

In addition to the kurtosis transformations, there are skewness transformations that have been developed in the Tukey family, such as the g-type transformation.

Definition 7.3 (Tukey's skewness transformation)**.** *The g-type transformation, denoted by $T_g(w)$, is given by*

$$T_g(w) = \frac{\exp(w) - 1}{w} \tag{80}$$

The generalized g-type transformation, denoted by $T_g^*(w)$, is given by

$$T_g^*(w) = \left[1 + c\frac{1 - \exp\left(-gW\right)}{1 + \exp\left(-gW\right)}\right] \tag{81}$$

To nest all these transformations within one class of transformations, the work of Fischer [2010] proposed a power series representation denoted by the subscript a given in Equation (82). This suggestion, though it nested the other families of distributions, is not practical for use as it involves the requirement of estimating a very large (infinite) number of parameters a_i to obtain the data-generating mechanism:

$$T_a(w) = \sum_{i=0}^{\infty} a_i w^{2i}. \tag{82}$$

It was further observed in Fischer [2010] that this nesting structure may be replaced with a different form, given by the general transformation taking the form given in Equation (83):

$$T_{hjk}(w;\alpha,\beta,\gamma) = \left(1 + \frac{\left(w^2+\gamma\right)^{\alpha} - \gamma^{\alpha}}{\beta}\right)^{\beta}, \quad \alpha > 0, \beta \geq 1, \gamma > 0. \tag{83}$$

Then it is clear that the original h-, k-, and j-type transformations are recovered with $T_h(w) = T_{hjk}(w; 1, \infty, \gamma)$, $T_k(w) = T_{hjk}(w; 1, 1, \gamma)$, and $T_j(w) \approx T_{hjk}(w; 0.5, \infty, 0.5)$. Further details of these transformations is provided in future sections where these classes of transformation as also applied to develop conditional quantile time series models.

7.1.1 Properties of the g-and-h Quantile Error Family

A detailed overview of the properties of these models is provided in Peters et al. [2016], we highlight a couple of them of relevance to understanding attributes of these models briefly below.

One can obtain the moments of Tukey family of distributions, with generically denoted Tukey quantile transform given by $r(W) = WT(W)^{\theta}$, as the solution to the following integrals, where the n-th moment is given with respect to the transformed moments of the base density as follows:

$$\mathbb{E}\left[\epsilon^n\right] = \mathbb{E}\left[r(W)^n\right] = \int_{-\infty}^{\infty} r(w)^n f_W(w)dw, \tag{84}$$

From such a result, one may now express the moments of the g-and-h distributed random variable according to the result in Proposition 7.1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2018 doi:10.20944/preprints201806.0198.v1

Peer-reviewed version available at *Risks* **2018**, *6*, 97; doi:10.3390/risks6030097

GH(a = 0, b = 1, g, h) with constant parameters g and h > 0. The n-th integer moment is given with respect to the standard Normal distribution and the n-th power of the transformed quantile function given by

$$r(W) = a + b\frac{\exp(gW) - 1}{g} \exp\left(\frac{hW^2}{2}\right). \tag{85}$$

to produce moments according to the relationship

$$\mathbb{E}\left[\epsilon^n\right] = \mathbb{E}\left[r(W)^n\right] \tag{86}$$

which will exist if $h \in \left[0, \frac{1}{n}\right)$. *One can also observe more generally that under the g-and-h transform the following identity holds with regard to powers of the standard Gaussian,* $W \sim \text{Normal}(0,1)$, *such that*

$$\begin{aligned} \epsilon^n &= r(W)^n = T_{g,h}(W; a, b, g, h)^n \\ &= \left(a + bT_{g,h}(W; a = 0, b = 1, g, h)\right)^n \\ &= \sum_{i=0}^{n} \frac{n!}{(n-i)!i!} a^{n-i} b^i T_{g,h}(W; a = 0, b = 1, g, h)^i, \end{aligned} \tag{87}$$

which will produce moments given by

$$\begin{aligned} \mathbb{E}\left[\epsilon^n\right] &= \mathbb{E}\left[\left(a + bT_{g,h}(W; a = 0, b = 1, g, h)\right)^n\right] \\ &= \sum_{i=0}^{n} \frac{n!}{(n-i)!i!} a^{n-i} b^i \mathbb{E}\left[T_{g,h}(W; a = 0, b = 1, g, h)^i\right]. \end{aligned} \tag{88}$$

Furthermore, it was shown by Dutta et al. [2002] that when it exists one can obtain the general expression

$$\mathbb{E}\left[T_{g,h}(W; a = 0, b = 1, g, h)^i\right] = \frac{\sum_{r=0}^{i} (-1)^r \frac{i!}{(i-r)!r!} \exp\left(\frac{(i-r)^2 g^2}{2(1-ih)}\right)}{\sqrt{(1-ih)} g^i}, \tag{89}$$

Proof. This result follows from direct application of the binomial series expansion result for polynomial integer powers, followed by the moment of the i-th integer order integration result derived in Dutta et al. [2002]. □

Remark 7.2. *We note the following properties of moments for the g (h=0) and the h (g=0) distributions respectively. In the case of g distribution, since the g-distribution is a horizontally shifted LogNormal distribution, then the moments of the g-distribution take the same form as those of a LogNormal model with appropriate adjustment for the translation. The h-distributional family is symmetric (except the double h-h family); consequently, all odd-order moments for the h-subfamily are zero, see further discussion in Dutta et al. [2002].*

Furthermore, using these moment identities one can easily then find the skew, kurtosis, and coefficient of variations for model families such as the g-and-h, the g-distributions and h-distributions. In addition to these simple population summaries of the g-and-h model, one could also consider other generalized properties of quantile-based functionals of asymmetry and kurtosis (see Balanda and MacGillivray [1990], Rayner and MacGillivray [2002], and Balanda and MacGillivray [1988]).

7.1.2 Tail Behaviour of the g-and-h Quantile Error Family

In terms of the tail behavior of the g-and-h family of distributions, the properties of such severity models have been studied by numerous authors such as Morgenthaler and Tukey [2000] and Degen et al. [2007]. In particular, the tail property (index of regular variation) for the g-and-h family of distributions was first studied for the h-distribution by Morgenthaler and Tukey [2000] and later for the g-and-h distribution by Degen et al. [2007] (see Proposition 7.2). In addition, the second-order regular variation properties of the g-and-h family of distributions was studied by Degen et al. [2007].

In order to study the properties of regular variation of the g-and-h family of loss distribution models it is first important to recall some basic definitions. First, we note that a positive measurable function $f(\cdot)$ is regularly varying if it satisfies the conditions in Definition 7.4, see discussion in Karatzas and Shreve [1991].

Peer-reviewed version available at *Risks* **2018**, *6*, 97; doi:10.3390/risks6030097

infinity) with an index $\alpha \in \mathbb{R}$ if it satisfies:

- *It is defined on some neighborhood $[x_0, \infty)$ of infinity; and*
- *It satisfies the following limiting relationship*

$$\lim_{x\to\infty} \frac{f(\lambda x)}{f(x)} = \lambda^{\alpha}, \quad \forall \lambda > 0. \tag{90}$$

We note that when $\alpha = 0$, then the function $f(\cdot)$ is said to be slowly varying (at infinity). From this definition one can show that a random variable has a regularly varying distribution if it satisfies the condition in Definition 7.5.

Definition 7.5 (Regularly varying random variable)**.** *A loss random variable X with distribution $F_X(x)$ taking positive support is said to be regularly varying with index $\alpha \geq 0$ if the right tail distribution $\overline{F}_X(x) = 1 - F_X(x)$ is regularly varying with index $-\alpha$.*

The following important features can be noted about regularly varying distributions as shown in Theorem 7.1, see detailed discussion in Bingham et al. [1989].

Theorem 7.1 (Properties of regularly varying distributions)**.** *Given a loss distribution $F_X(x)$ satisfying $F_X(x) < 1$ for all $x \geq 0$, the following conditions on $F_X(x)$ can be used to verify that it is regularly varying such that $F_X(x) \in RV_{\alpha}$:*

- *If $F_X(x)$ is absolutely continuous with density $f_X(x)$ such that for some $\alpha > 0$ one has the limit*

$$\lim_{x\to\infty} \frac{x f_X(x)}{\overline{F}_X(x)} = \alpha. \tag{91}$$

  Then $f_X(x)$ is regularly varying with index $-(1+\alpha)$ and consequently $\overline{F}_X(x)$ is regularly varying with index $-\alpha$;

- *If the density $f_X(x)$ for loss distribution $F_X(x)$ is assumed to be regularly varying with index $-(1+\alpha)$ for some $\alpha > 0$. Then the following limit,*

$$\lim_{x\to\infty} \frac{x f_X(x)}{\overline{F}_X(x)} = \alpha, \tag{92}$$

  will also be satisfied if $\overline{F}_X(x)$ is regularly varying with index $-\alpha$ for some $\alpha > 0$ and the density $f_X(x)$ will be ultimately monotone.

Proof. These results are derived in Bingham et al. [1989]. □

Many additional properties are described for such heavy tailed distribution and density functions. Here we will utilize the above stated conditions to assess the regular variation properties of the right tail of the g-and-h family of loss models. In particular we will see if a single distributional parameter characterizes the heavy tailed feature as captured by the notion of regular variation index, or if the relationship is more complex.

Proposition 7.2 (Index of regular variation of g-and-h distribution)**.** *Consider the random variable $W \sim \text{Normal}(0,1)$ and a random variable ϵ, which has severity distribution given by the g-and-h distribution with parameters $a, b, g, h \in \mathbb{R}$, denoted $\epsilon \sim \text{GH}(a,b,g,h)$, with $h > 0$ and density (distribution) $f(x)$ (and $F(x)$) . Then the index of regular variation is obtained by considering the following limit*

$$\lim_{x\to\infty} \frac{x f(x)}{\overline{F}(x)} = \lim_{x\to\infty} \frac{\phi(u)\left(\exp(gu) - 1\right)}{\left(1 - \Phi(u)\right)\left(g\exp(gu) + hu(\exp(gu) - 1)\right)} = \frac{1}{h} \tag{93}$$

for $u = k^{-1}(x)$ where the function $k(x)$ is given by

$$k(x) = \frac{\exp\left(gx\right) - 1}{g} \exp\left(\frac{hx^2}{2}\right). \tag{94}$$

Hence, one can state that $\overline{F}_{\epsilon} \in RV_{-\frac{1}{h}}$.

Peer-reviewed version available at Risks 2018, 6, 97; doi:10.3390/risks6030097

The asymptotic tail behavior of the h-family of Tukey distributions was studied by Morgenthaler and Tukey [2000] and is given in Proposition 7.4.

Proposition 7.4 (h-type tail behaviour). *Consider the h-type transformation, where* $W \sim \text{Normal}(0,1)$ *is a standard Gaussian random variable and the random variable* ϵ *has severity distribution given by the h-distribution with parameters* $a, b, h \in \mathbb{R}$*, denoted* $\epsilon \sim H(a,b,h)$ *according to*

$$\epsilon = T_h(W; a, b, h) := a + bW \exp\left(\frac{hW^2}{2}\right). \tag{95}$$

Then the asymptotic tail index of the h-type distribution is then given by $1/h$*. This is equivalent to the g-and-h family for* $g \neq 0$*.*

Proof. The proof of this result is found in Morgenthaler and Tukey [2000]. □

This shows that the h-type family has a Pareto heavy-tailed property, hence the restriction that moments will only exist on the order of less than $1/h$. The g-family of distributions can be shown to be sub-exponential in the tail behavior but not regularly varying. It was shown [Degen et al., 2007, theorem 2.2] that one can obtain an explicit form for the function of slow variation in the g-and-h family as detailed in Theorem 7.2.

Theorem 7.2 (Slow variation representation of g-and-h severity models). *Consider the random variable* $W \sim \text{Normal}(0,1)$ *and a random variable* ϵ*, which has distribution given by the g-and-h with parameters* $a, b, g, h \in \mathbb{R}$*, denoted* $\epsilon \sim \text{GH}(a,b,g,h)$*, with* $g > 0$ *and* $h > 0$ *and density (distribution)* $f(x)$ *(and* $F(x)$*) . Then* $\overline{F}(x) = x^{-1/h} L(x)$ *for some slowly varying function* $L(x)$ *given as* $x \to \infty$ *by*

$$L(x) = \frac{h}{\sqrt{2\pi} g^{1/h}} \frac{\left[\exp\left(\frac{g}{h}\sqrt{g^2 + 2h\ln(gx)} - \frac{g^2}{h}\right) - 1\right]^{1/h}}{\sqrt{g^2 + 2h\ln(gx)} - g} \left(1 + O\left(\frac{1}{\ln x}\right)\right). \tag{96}$$

Proof. This was proven in Degen et al. [2007]. □

From this explicit Karamata representation developed by Degen et al. [2007], it was also shown that one can obtain the second-order regular variation properties of the g-and-h family.

The implications of these findings are that the g-and-h distribution, under the parameter restrictions $g > 0$ and $h > 0$, belongs to the domain of attraction of an Extreme Value Distribution, such that $\epsilon \sim \text{GH}(a,b,g,h)$ with distribution F satisfying $F \in MDA(H_\gamma)$ where $\gamma = h > 0$. As a consequence, by the Pickands–Balkema–de Haan Theorem, discussed in detail in Embrechts et al. [2013] and recently in Cruz et al. [2015], one can state that there exists an Extreme Value Index (EVI) constant γ and a positive measurable function $\beta(\cdot)$ such that the following result between the excess distribution of the g-and-h (denoted by $F_{\epsilon,u}(x) = \Pr(\epsilon - u \leq x | \epsilon > u)$ and the generalized Pareto distribution (GPD) is satisfied in the tails

$$\lim_{u \uparrow \infty} \sup_{x \in (0,\infty)} \left| F_{\epsilon,u}(x) - G_{\gamma,\beta(u)}(x) \right| = 0. \tag{97}$$

For discussion on the rate of convergence in the tails, see Raoult and Worms [2003] and the application of this theorem to the g-and-h case by Degen et al. [2007] where it is shown that the order of convergence is given by $O\left(A \exp\left(V^{-1}(u)\right)\right)$ for functions

$$\begin{aligned} V(x) &:= \overline{F}^{-1}\left(\exp(-x)\right), \\ A(x) &:= \frac{V''(\ln x)}{V'(\ln x)} - \gamma. \end{aligned} \tag{98}$$

Hence, the conclusion from this analysis regarding the tail convergence of the excess distribution of the g-and-h family toward the GPD $G_{\gamma,\beta(u)}(x)$ is given explicitly by

$$\frac{\ln L(x)}{\ln x} \sim \sqrt{2} \frac{g}{h^{\frac{3}{2}}} \frac{1}{\sqrt{\ln(x)}} = O\left(\frac{1}{\sqrt{\ln\left(k^{-1}(x)\right)}}\right), \quad x \to \infty. \tag{99}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2018 doi:10.20944/preprints201806.0198.v1

Peer-reviewed version available at *Risks* **2018**, *6*, 97; doi:10.3390/risks6030097

process, if a goodness-of-fit test suggests that one may not reject the null hypothesis that these data came from a g-and-h distribution, then one should avoid performing estimation of the extreme quantiles, such as those used to measure the capital via the Value-at-Risk, via methods based on Peaks Over Threshold (POT) or Extreme Value Theory (EVT) based penultimate approximations.

Proposition 7.5 (Index of regular variation of the generalized g-and-h distribution)**.** *Consider the random variable $W \sim \text{Normal}(0,1)$ and a random variable ϵ, which has distribution given by the generalized g-and-h distribution with parameters $a, b, g, h, c \in \mathbb{R}$, denoted $\epsilon \sim \text{Generalized}-\text{GH}(a,b,g,h,c)$, with $g > 0$ and density (distribution) $f(x)$ (and $F(x)$) . Recall that we have, for the generalized g-and-h loss model, the function $r(x)$ with $a = 0$ and $b = 1$ given by*

$$r(x) = \left[1 + c\frac{1-\exp(-gx)}{1+\exp(-gx)}\right] x \exp\left(\frac{hx^2}{2}\right). \tag{100}$$

Using this, we can then find the index of regular variation at $x \to \infty$ given as follows

$$\lim_{x\to\infty} \frac{xf(x)}{\overline{F}(x)} = \lim_{x\to\infty} \frac{x\phi\left(r^{-1}(x)\right)}{r'\left(r^{-1}(x)\right)\left[1-\Phi\left(r^{-1}(x)\right)\right]} = \frac{1}{h} \tag{101}$$

Proof. Proof of this result is obtained in Peters et al. [2016]. □

We note that this result is not unexpected since the g transform in each case drives the skewness and not the kurtosis. We can also obtain this analysis for the g-and-k model, this yields that the g-and-k does not admit a finite limit in either sign of the parameter g, showing that such a model is not regularly varying, as we see in the case of the g-and-h models. However, even though this is the case we can still assess the relative heavy tailedness of the g-and-k models compared to the base distribution under the Tukey k-transform.

8 Alternative General Quantile Error Models: Rank Transmutation

In this section we discuss the ideas presented in Shaw and Buckley [2009] where they discuss alternative approaches to deformation maps of base random variables to obtain valid quantile functions. Again, as in previous discussions, these maps can be considered as applicable to constructing flexible families of error quantile function or treated as non-linear maps for the conditional quantile time series relationship specification.

In Shaw and Buckley [2009] they study composition maps of base quantile functions and target distribution functions that they term Rank Transmutation Maps (RTM), since the inner functional mapping of a random variable creates a random variable in $[0,1]$ which is a rank statistic. The generic idea of RTM's is presented in the following definition.

8.1 Alternative Maps that Induce Relative Skewness and Kurtosis

In this section we introduce an alternative class of mapping functions, compared to the Tukey family of transformations, that can also create relative skewness and kurtosis in the resulting transformed quantile function. These will be known as Rank Transmutation Maps (RTM's) as discussed in Shaw and Buckley [2009].

Definition 8.1 (Rank Transmutation Maps)**.** *Consider two distribution functions with a common support (domain or values of the random variable that have non-zero probability associated with their outcome), denoted by F_1, F_2. Then one can define the following pair of general RTMs as follows:*

$$\begin{aligned} G_{R_{12}}(u) &= F_2(F_1^{-1}(u)) \\ G_{R_{21}}(u) &= F_1(F_2^{-1}(u)). \end{aligned} \tag{102}$$

Remark 8.1. *It is clear that these mappings are also transformations in the same manner as described in the class of Tukey-Elongation transforms. However, they correspond to mappings given by distribution functions instead of the general Tukey class. In general to relate them consider the mapping for a base random variable $W \sim F_1(w)$ then one would have*

$$\epsilon_t = T(W) = G_{R_{12}}(U) = F_2\left(F_1^{-1}(U)\right) \tag{103}$$

assumptions, will form mutual inverses with the properties that they satisfy.

$$\begin{aligned} G_{R_{ij}}(0) &= 0, \\ G_{R_{ij}}(1) &= 1. \end{aligned} \tag{104}$$

If the resulting distributions F_1 and F_2 are are continuous or in other words the RTM maps $G_{R_{12}}$ and $G_{R_{21}}$ are continuously differentiable then the law of the resulting mapped distribution is a continuous function.

One can then define different families of RTM mappings which correspond to different classes of quantile distortion. The ones outline in Shaw and Buckley [2009] correspond to classes:

1. Quadratic class of Rank Transmutation Maps;
2. Skew-Uniform class of Rank Transmutation Maps;
3. Skew-Exponential class of Rank Transmutation Maps;
4. Symmetric-cubic class of Rank Transmutation Maps; and the
5. Skew-kurtotic class of Rank Transmutation Maps;
6. General Class of Rank Transmutation Maps.

We will outline briefly the definition of each class of maps below.

Definition 8.2. *The following functional transformations form a range of RTM function mappings for a random variable $U \sim U[0,1]$ to obtain $\epsilon_t = T(U)$ (analogously $Q_{\epsilon_t}(u) = T(Q_U(u))$. The classes of RTM maps are given by:*

1. *Family of Quadratic RTM's can be defined by a single parameter λ map given by*

$$T_Q(U) = G_{R_{12}}(u) = u + \lambda u(1-u), \quad \textit{for } |\lambda| \le 1, \tag{105}$$

This RTM map T_Q has the effect of introducing skew to a base distribution when it is symmetric. (analog of the g-transform in the Tukey elongation transforms.) If the base distribution is symmetric around the origin i.e. F_1 satisfies that $F_1(x) = 1 - F_1(-x)$, then one has that the distribution of the square of the transmuted random variable is identical to that of the distribution of the square of the original random variable.

2. *Family of Skew-Uniform RTM's can be defined by a single parameter λ map given by*

$$T_{SU}(U) = G_{R_{12}}(u) = F_2\left(F_1^{-1}(U)\right), \quad \textit{for } |\lambda| \le 1, \tag{106}$$

such that

$$\begin{aligned} F_1(x) &= x, \\ F_2(x) &= \begin{cases} 0 & x < 0, \\ (1+\lambda)x - \lambda x^2 & 0 \le x \le 1, \\ 1 & x > 1. \end{cases} \end{aligned} \tag{107}$$

Alternatively, for any $\lambda \in \mathbb{R}$ one could also consider the mapping:

$$G_{R_{12}}(u) = \min[\max[u + \lambda u(1-u), 0], 1]. \tag{108}$$

3. *Family of Skew-Exponential RTM's can be defined by two parameters λ, β map given by*

$$T_{SE}(U) = G_{R_{12}}(u) = F_2\left(F_1^{-1}(U)\right), \quad \textit{for } |\lambda| \le 1, \& \beta > 0, \tag{109}$$

such that

$$F(x;\lambda) = \begin{cases} 1 - e^{-\lambda x} & x \ge 0, \\ 0 & x < 0, \end{cases} F_2(x) = \begin{cases} 0 & x < 0 \\ (1+\lambda)x - \lambda x^2 & 0 \le x \le 1 \\ 1 & x > 1 \end{cases}. \tag{110}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2018 doi:10.20944/preprints201806.0198.v1

Peer-reviewed version available at *Risks* **2018**, *6*, 97; doi:10.3390/risks6030097

$$T_{SN}(U) = G_{R_{12}}(U) = F_2\left(F_1^{-1}(U)\right), \quad \text{for } |\lambda| \leq 1, \& \beta > 0, \tag{111}$$

such that

$$F_1(x) = \Phi(x) := \frac{1}{2}\left(1 + \text{erf}\left(\frac{z}{\sqrt{2}}\right)\right), \tag{112}$$

$$F_2(x) = \begin{cases} 0 & x < 0 \\ (1+\lambda)\,x - \lambda x^2 & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}. \tag{113}$$

5. *Family of General RTM's can be defined by a map given by*

$$T_G(U) = G_{R_{12}}(U) = U + U(1-U)P(U), \tag{114}$$

where P is a polynomial with various parameters. One could consider different order of polynomials and different parameter restrictions to generate a general family of RTMs.

Remark 8.2. *If one considers the class of T_G RTM's then a useful illustrative example can be obtained by setting:*

$$P(u) = \gamma.(u - \frac{1}{2}), \quad \gamma \in \mathbb{R}. \tag{115}$$

Such a choice would produce the practical property that $G_{R_{12}}(1-u) = 1 - G_{R_{12}}(u)$.

With all these flexible model components and the general framework provided in this manuscript practitioners should be able to understand and construct a large variety of quantile models.

9 Illustrations of Quantile Time Series Models for Mortality and Demographic Actuarial Applications

In this section we explore a range of mortality and demographic data sets via quantile time series regressions. The outputs of this analysis can be directly useful in insurance applications for instance in life insurance applications in annuities pricing and risk management as well as pension policy development. We do not go into detail on these particular areas of application in this tutorial manuscript, instead we focus on modelling and comparison of different quantile time series models on real mortality and demographic data sets obtained for England, Wales, Scotland and Northern Ireland.

The intention of these application illustrations is not to be exhaustive on all the different models explored in previous sections, instead we focus on providing examples of illustrations of these new regression techniques to show actuaries and practitioners how they may be readily applied in practical settings. To explain some of the properties of the linear vs non-linear parametric and non-parametric models.

In this manuscript we will focus on data reflecting Demographic data and health event data. In particular, the data we consider includes several different time series data sets from the following sources with the following attributes:

- the Human Mortality Database available at `http://www.mortality.org/` which provides records of annual data for aggregated births, aggregated deaths by age group with yearly stratification and population sizes. We took a specific focus on England, Wales, Scotland and Norther Ireland.

- the United Kingdom Office of National Statistics data available at `https://www.ons.gov.uk/` where we obtained weekly mortality records from 2010 to 2017. Furthermore, we also obtained decompositions of the annual death counts for England and Wales between 2001 and 2013 for avoidable mortality events and alcohol related deaths.

- the National Archives also provide the number of deaths annually by sex, age group and underlying cause from periods of 1901 to 2017, where data are available at `http://webarchive.nationalarchives.gov.uk/20160111174808/http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-215593`.

tained weekly birth and death recordings for Scotland as well as the weekly recorded deaths due to respiratory disease, from 2004 to 2017. In addition, the monthly recorded births and deaths by geographical area in Scotland was obtained from 1990 to 2017.

- the National Records of Scotland data available at `https://www.nisra.gov.uk/publications/weekly-deaths` where we obtained weekly death recordings as well as monthly death records from 2006 to 2017 for Northern Ireland. Furthermore, alcohol related deaths were also obtained monthly from 2006 to 2017 for Norther Ireland.

The illustrations of quantile time series modelling on these data sets will be undertaken by data type and regions. Since this manuscript is intentionally not focused on aspects of estimation of quantile time series models, we will utilise existing R packages to perform estimation of the models explored in this section. All the model illustrations performed in the following sections were estimated with standard quantile regression and time series packages in R based on 'rq' and 'nlrq' function outputs.

9.1 Annual Births for Males and Females: England and Wales, Scotland and Northern Ireland

We begin with analysis of the Human Mortality Data Base data sets of annual births by year. The data is presented in Figure 1.

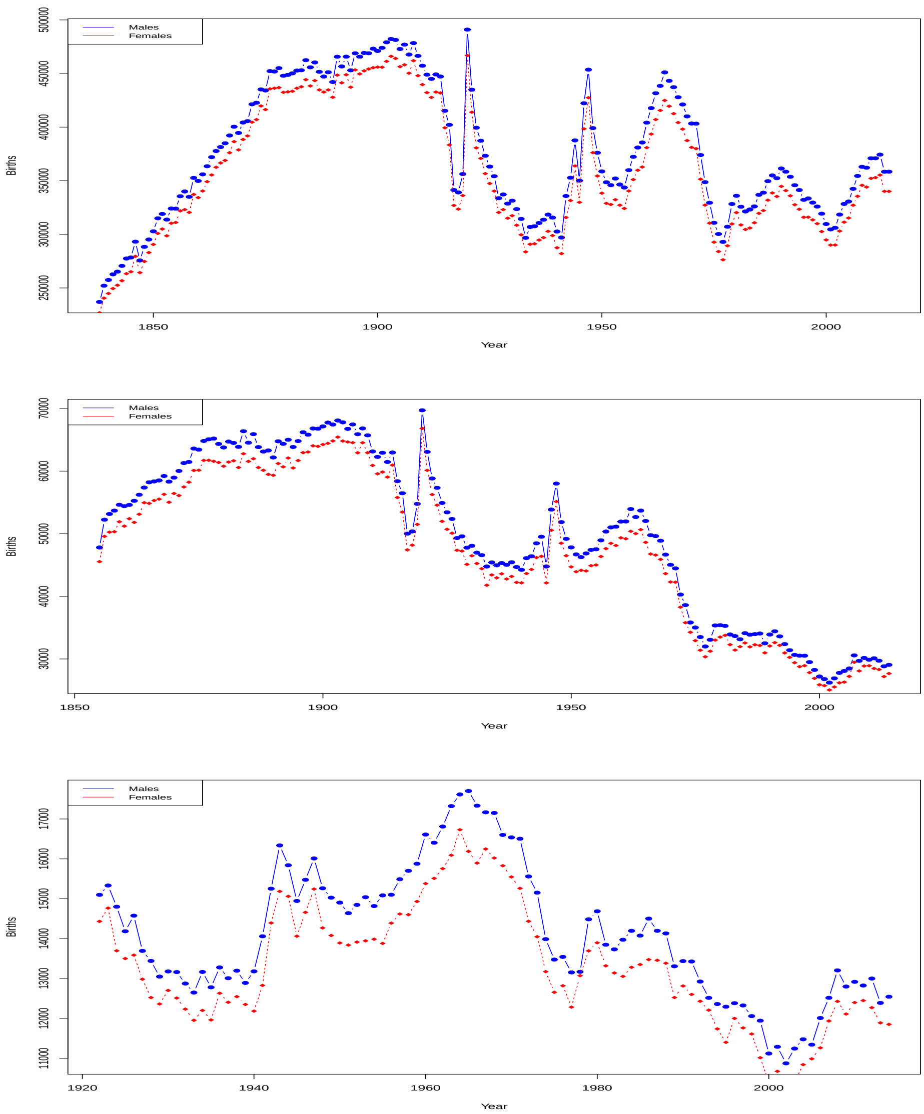

Figure 1: Births of Males and Females verus year. Top Panel: England and Wales; Middle Panel: Scotland; Bottom Panel: Northern Ireland.

We first select the order of the time series regression base on AIC criterion considering 1 through to 5 lags. In Table 9.1.1 we present the results for the AIC vs lag for Male and Female births over time.

England and Wales Male Births				
AIC 1 Lag	AIC 2 Lag	AIC 3 Lag	AIC 4 Lag	AIC 5 Lag
3945.396	3925.935	3906.828	3886.021	**3866.022**
England and Wales Female Births				
AIC 1 Lag	AIC 2 Lag	AIC 3 Lag	AIC 4 Lag	AIC 5 Lag
3945.396	3925.935	3906.828	3886.021	**3866.022**
Scotland Male Births				
AIC 1 Lag	AIC 2 Lag	AIC 3 Lag	AIC 4 Lag	AIC 5 Lag
3309.565	3294.536	3279.546	3261.000	**3245.277**
Scotland Female Births				
AIC 1 Lag	AIC 2 Lag	AIC 3 Lag	AIC 4 Lag	AIC 5 Lag
3309.565	3294.536	3279.546	3261.000	**3245.277**
Northern Ireland Male Births				
AIC 1 Lag	AIC 2 Lag	AIC 3 Lag	AIC 4 Lag	AIC 5 Lag
1398.028	1381.626	1369.940	**1359.550**	1360.158
Northern Ireland Female Births				
AIC 1 Lag	AIC 2 Lag	AIC 3 Lag	AIC 4 Lag	AIC 5 Lag
1398.028	1381.626	1369.940	1360.550	**1360.158**

We see that for all cases both male and female time series prefer lag structure, according to AIC assessment, of order 5 with the exception of Northern Ireland male births where lag of 4 was selected. Therefore, we proceed with first a linear quantile time series analysis to assess the lag 5 linear QAR model. We present the fitted quantile time series results in Figure 2. We note that the results are presented for males only as the female results in this case are very similar. The plots show the estimated AR(5) model quantile residuals for decile values $u \in \{0.1, 0.2, \ldots, 0.9\}$. First we show the lag 1 to lag 5 estimated quantile regression coefficients versus the quantile level in Figure 2

We can also assess the statistical significance of the fitted coefficients to see if they are statistically different from zero, even though AIC has suggested a model order of 5. For instance as an illustration, we will consider the case of the model for $u = 0.1$ and $u = 0.2$ the first and second deciles.

For the QAR(5) model $u = 0.1$ the coefficients corresponding to lag 2 and lag 3 are basically on the boarder of being statistically significant according to the upper bound of a 95% confidence interval of the coefficient including 0. However, the lags 4 and 5 are clearly not statistically significant as their 95% confidence intervals of the estimated coefficients clearly contain zero. In this case, one could consider to include a comparison of a QAR(3) model. Furthermore, it is evident from the estimated coefficient for the lag one coefficient that the model is very close to the boundary of a random walk type behaviour for these low quantiles. In actual fact, this type of behaviour is seen through out the entire range of decile fits in this data. In the case of the model for $u = 0.2$, one sees that all coefficients on lags 2 to 5 are clearly not statistically significant when one looks at whether their 95% confidence intervals on the coefficient estimates contain 0.

In addition to assess the quality of the fitted QAR(5) quantile models at each quantile level, we also plot the studentised residuals of each model fit as a function of the quantile level in Figure 3.

We see from these plots that the fitted QAR(5) models for each of the deciles $u = 0.1$ through to $u = 0.9$ have studentized residuals which are very well behaved. This indicates that the fitted models are doing a reasonable job at capturing the quantile time series dyanmics of the births data. To finish this aspect of the illustrations we will also fit the England and Wales data with an QAR(1) model to see how it performs relative to the QAR(5), since we found that the coefficients for lags 2 to 5 were not statistically significant. We show the fitting results for the QAR(1) model for the male births from England and Wales in Figure 4.

The top subplot in Figure 4 again confirms that the fitted model is very close to a random walk type behaviour across all decile levels, except for very low quantile levels. In this QAR(1) model analysis the bottom subplot shows the fitted lines are superimposed in gray. In the case we see in the top subplot in which the AR coefficients are basically very close to constant across quantiles one would then expect fitted lines which are parallel to each other as the only change is the quantile level fitted. there is a slight

Peer-reviewed version available at Risks 2018, 6, 97; doi:10.3390/risks6030097

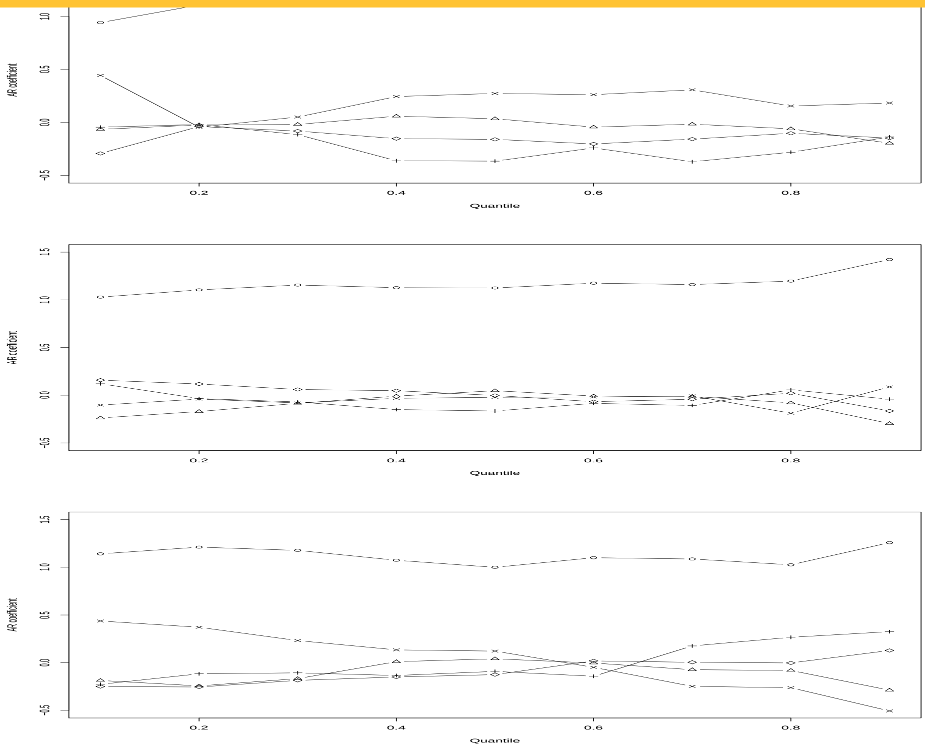

Figure 2: Plot of the estimated coefficients in the QAR(5) model verus quantile level. Top Panel: England and Wales; Middle Panel: Scotland; Bottom Panel: Northern Ireland.

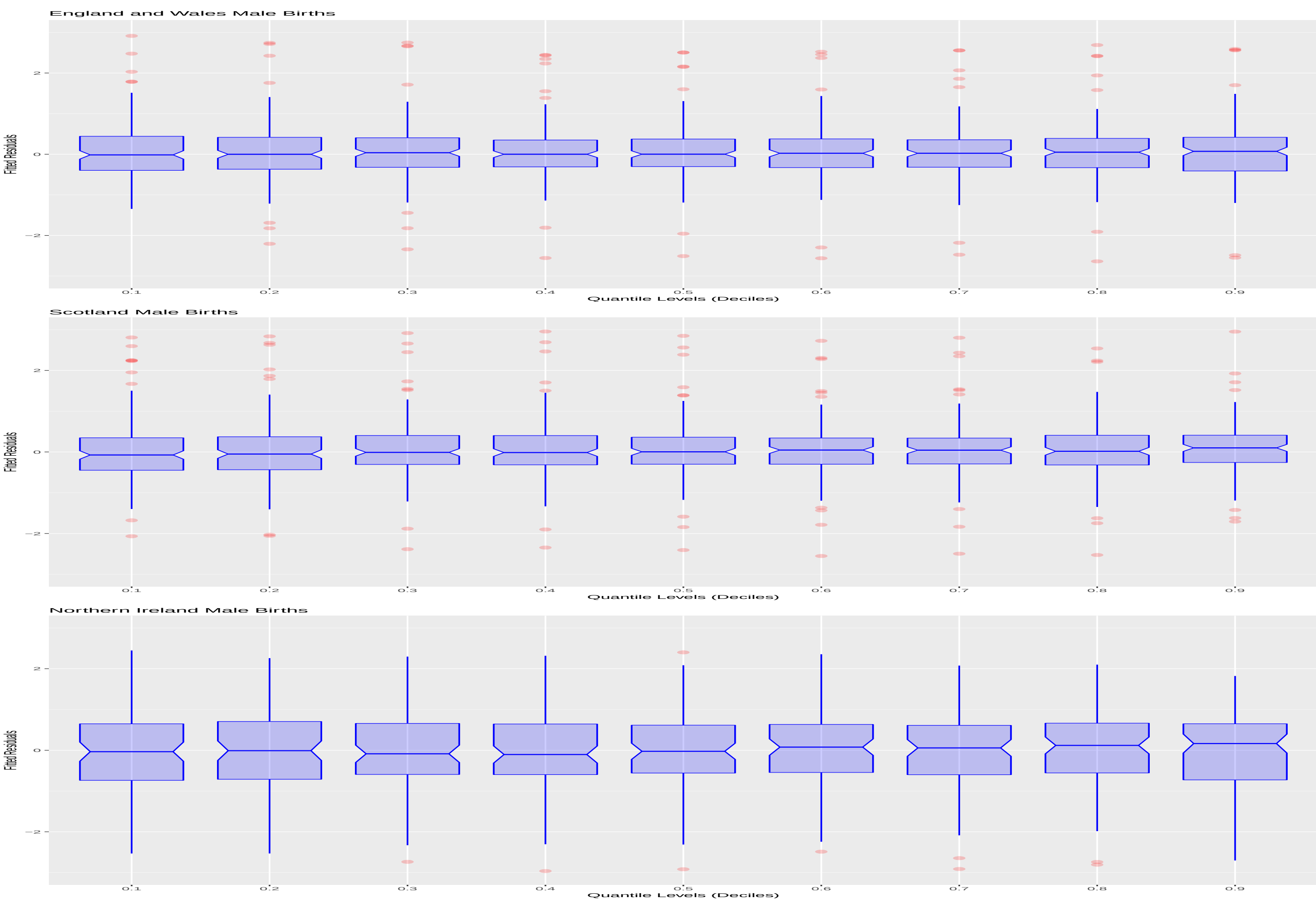

Figure 3: Top Panel: England and Wales; Middle Panel: Scotland; Bottom Panel: Northern Ireland.

fanning happening for quantile levels around the median but this is very close to uniform behaviour across all quantiles.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2018 doi:10.20944/preprints201806.0198.v1

Peer-reviewed version available at Risks 2018, 6, 97; doi:10.3390/risks6030097

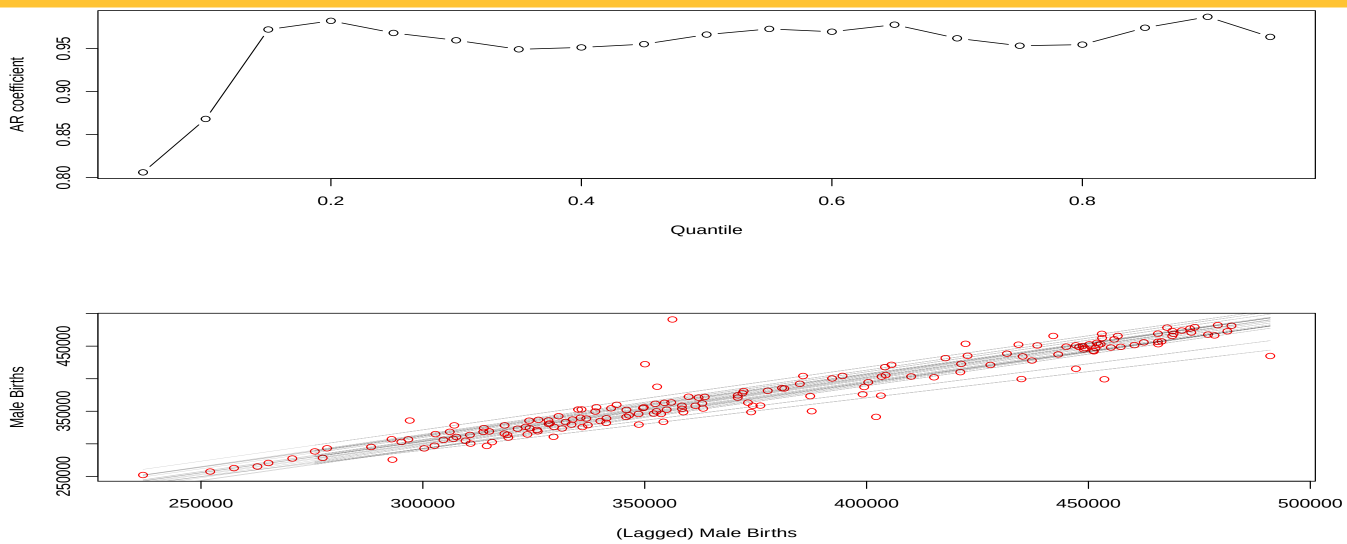

Figure 4: Top Panel: fitted AR(1) coefficient accross the quantile level.; Bottom Panel: Observation vs Lag and fitted values.

We learn from this analysis that there has been a steady decline in the birth rates of males in Scotland which is more pronounced in the last few decades than the declines seen in Northern Ireland. We will therefore focus further on the Scottish case study.

9.1.2 Examples of Non-Linear Non-Parametric QAR Modelling

We proceed next with an illustration of non-linear quantile time series models for the total weekly births for Scotland from 2004 t0 2018. We consider to fit both linear and non-linear quantile regression models to this weekly data. We will use a QAR(1) linear model as comparative reference which was suitable for the yearly aggregate data. In Figure 5 we see the fitted QAR(1) coefficients for the first lag as a function of quantile level are significantly different when looking at weekly observation patterns compared to the annual aggregate data.

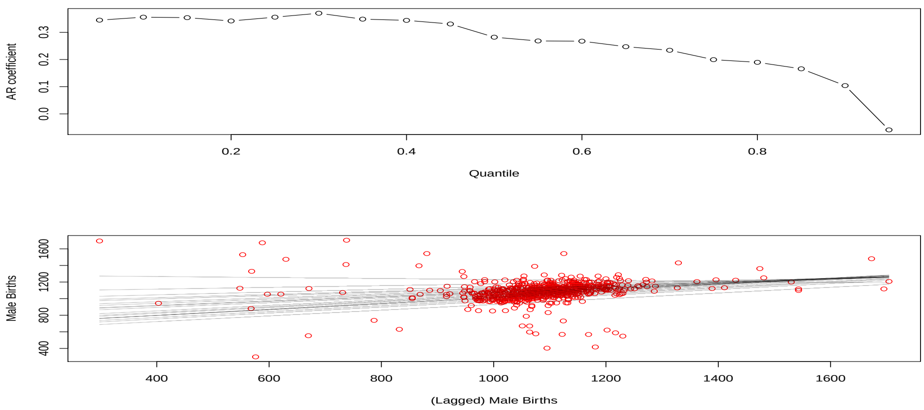

Figure 5: Scottish weekly total births from 2004 to 2017 fitted with family of QAR(1) linear models. Top Subplot: fitted coefficients vs quantile level; Bottom Subplot: observations versus lag and fitted QAR(1) models.

In particular we now see a pronounced deviation away from the random walk type behaviour observed in the annual counts models. Furthermore, we see that strength of the serial dependence present in the quantile time series of births is diminishing in strength as we move from low to high quantile levels, as reflected by the estimated magnitude of the first lag coefficient. Furthermore, we see that this change in coefficient of the lag one QAR models as a function of quantile level results in the fitted regressions fanning out much more than at the annual aggregate level where we saw almost parallel line relationships with the quantile level.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2018 doi:10.20944/preprints201806.0198.v1

Peer-reviewed version available at *Risks* **2018**, *6*, 97; doi:10.3390/risks6030097

modelling in R using the nlrq package and the lprq local polynomial quantile regression function where we explore the effect of the bandwidth parameter h. In this case, we use the weekly time as the input covariate, constructing a model with local non-linear polynomial transforms for $T(\cdot)$ and distributed lags for $\mathcal{G}_t$ of week index for the regression variable.

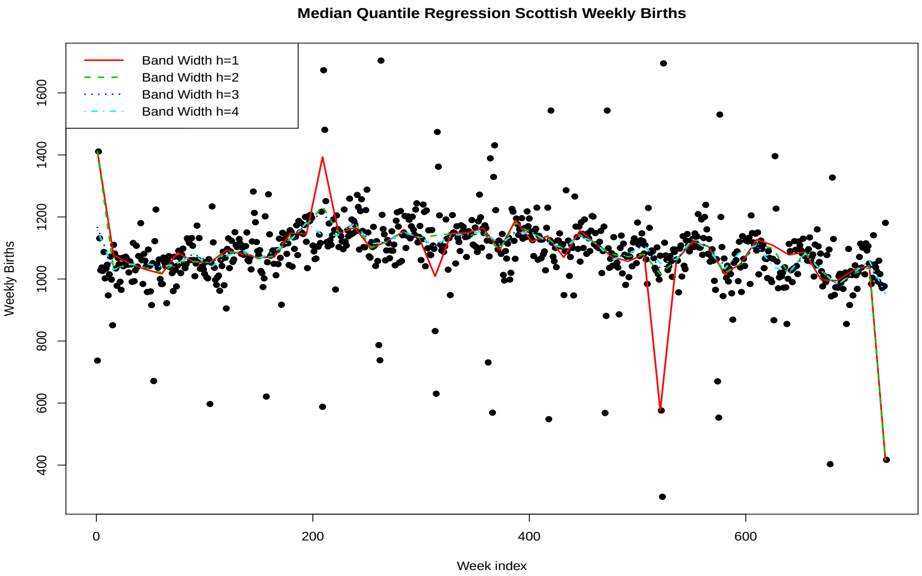

Figure 6: Non-linear local polynomial spline median regression of Scottish weekly total births from 2004 to 2017. Study of the bandwidth parameter in the local polynomial regression.

We see that the fit of the median regression is reasonable, however a stronger bandwidth is required to ensure certain points dont have too great a leverage effect on the local polynomial median quantile regression.

Next we plot in Figure 7 the fitted quantile regressions for a range of quantile levels $u \in \{0.1, 0.25, 0.5, 0.75, 0.9\}$ with bandwidth $h = 4$ and we only consider the temporal weekly covariate in the non-linear non-parameteric quantile regression.

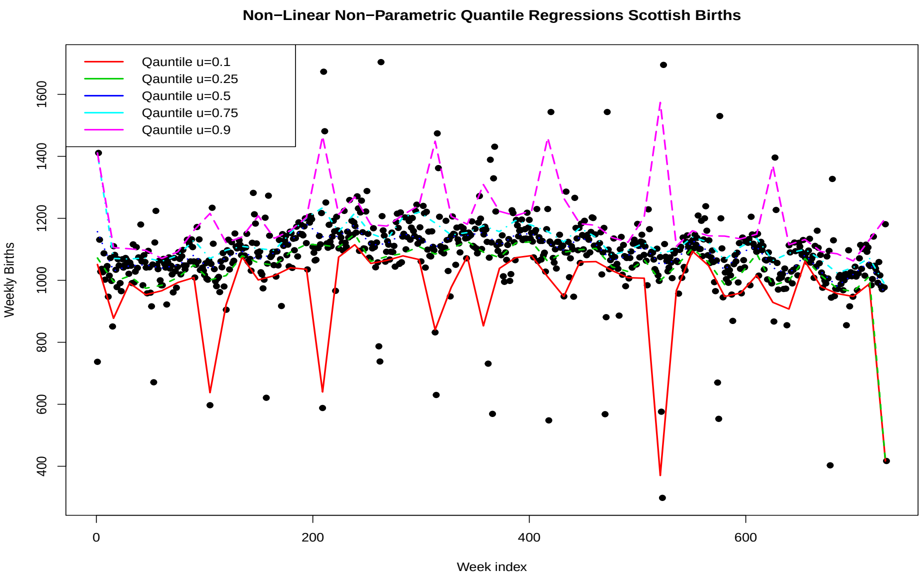

Figure 7: Non-linear local polynomial spline median regression of Scottish weekly total births from 2004 to 2017.

9.2 Annual Deaths for England and Wales, Scotland and Northern Ireland

Next we analyse the Human Mortality Data Base data sets of annual deaths by year and by age group. The data is presented in Figure 8 where we explore particular age groups $x \in \{20, 65, 75\}$.

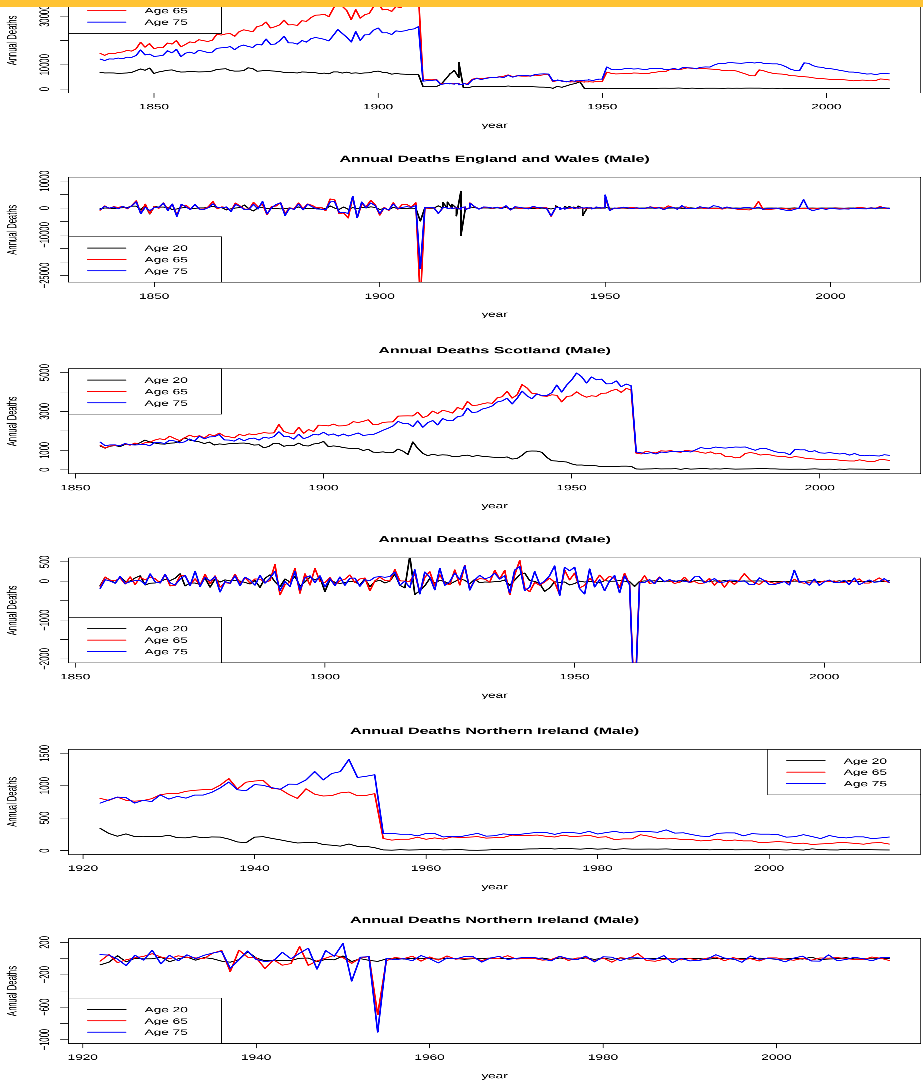

Figure 8: Deaths of Males versus year for age groups $x \in \{20, 65, 75\}$. Top subplot corresponds to annual deaths and second subplot corresponds to difference in annual deaths for each region of the UK. Top Panels: England and Wales; Middle Panels: Scotland; Bottom Panels: Northern Ireland.

9.2.1 Examples of Linear vs Non-Linear Parametric QAR Modelling

In this section we undertake a study of the first differences of the annual deaths for a range of ages for England and Wales, Scotland and Northern Ireland. We will consider a range of parameteric non-linear models of the following forms:

- Linear QARI(1,1) model (reference).
- A version of Double-QAR(1,1) model in Equation 15 which we selected accordign to the equation

$$T(x) = a + bx + c\sqrt{1 + x^2}$$

- General Non-linear QAR model as specified in Equation 22. We consider as an illustration the lagged Tukey G skewness Quantile Regression: $T(x) = a + b * (\exp(gx) - 1)/g$. The filtration $\mathcal{F}_t$ will be the natural filtration generated by the data and we will consider a single covariate of lag 1 from the observed series. $Q_\epsilon(u)$ will be selected to be standard Normal.

We present the results below for England and Wales as similar results were obtained for Scotland and Northern Ireland. The results for the fitted QARI(1,1) model are presented in Figure 9 annual deaths for ages 20, 65 and 75 years old males. We also show the fitted quantile time series models for 0.1, 0.5 and 0.9 quantiles for age group of 20 years in Figure 10.

 doi:10.20944/preprints201806.0198.v1

Peer-reviewed version available at Risks 2018, 6, 97; doi:10.3390/risks6030097

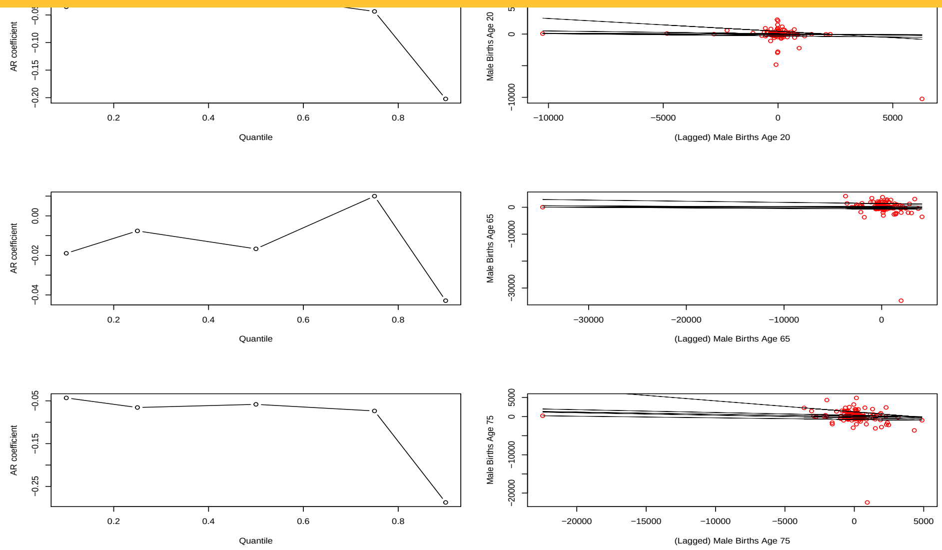

Figure 9: QARI(1,1) linear regressions at quantile levels $u \in \{0.1, 0.25, 0.5, 0.75, 0.9\}$ of English and Welsh male annual deaths.

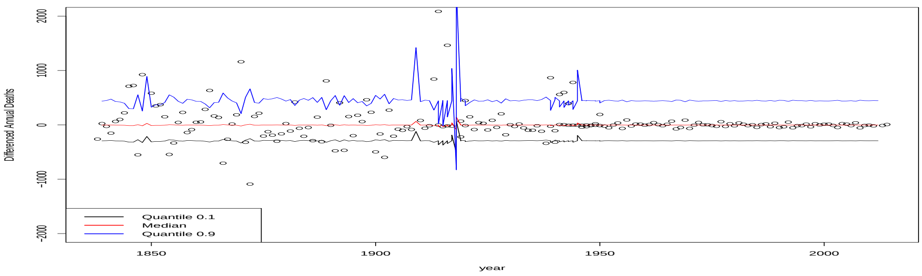

Figure 10: Fits of QARI(1,1) linear regressions at quantile levels $u \in \{0.1, 0.5, 0.9\}$ of English and Welsh male annual deaths.

The fits of the QAR(1,1) model for the 0.1 and 0.9 quantile models are clearly having difficulty in the fit quality the interval where the deaths suddenly declined in the great war periods between 1910-1940. There is also an evident asymmetry in the volatility of the fitted QARI(1,1) model for the $u = 0.9$ compared to the case of $u = 0.1$. We also observe that both the extreme quantile time series models struggle to differentiate the change that occurred in the deaths post the great wars and the behaviour before. We will therefore explore some examples of non-linear Quantile regression models.

Next we explore in Figure 11 the non-linear parametric Double QARI model described above, where we select values of $b \in \{0.1, 0.5, 0.85\}$ and we compare this model fit to the linear QAR(1) model for the median $u = 0.5$. We see that when the coefficient b is small, the model for the median time series regression is significantly skewed, however as b increases it becomes less skewed and produces a more reactive fit to the periods of large death count change relative to the results obtained from the linear QAR(1) median time series model. One would need to be careful to assess the model output in this case to ensure over fitting were not a problem, this can be achieve by looking at the residuals of the regression fits and standard model selection criterion.

Next we illustrate the difference between the QARI(1,1) linear model to the Tukey G transform model explained in Definition 7.3. We note that the behaviour of the median regression performance of this model for values of g has a highly non-linear relationship with the parameter g and the degree of responsivity that the median regression has produced when fit to the data. For instance we plot in Figure 12 the choices of $g = 0.001$ and $g = 0.00705$ where we see that for smaller values of g, one effectively recovers a model almost identical to the QARI(1,1) linear time series model with a slight deviation around 1920. However, for $g = 0.00705$ we see that the model is skewed in the median response and can be significantly more responsive to individual death yearly fluctuations than the QARI(1,1) model.

We note that the simple Tukey G model as presented can only be skewed in one direction, hence, we

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2018 doi:10.20944/preprints201806.0198.v1

Peer-reviewed version available at *Risks* **2018**, *6*, 97; doi:10.3390/risks6030097

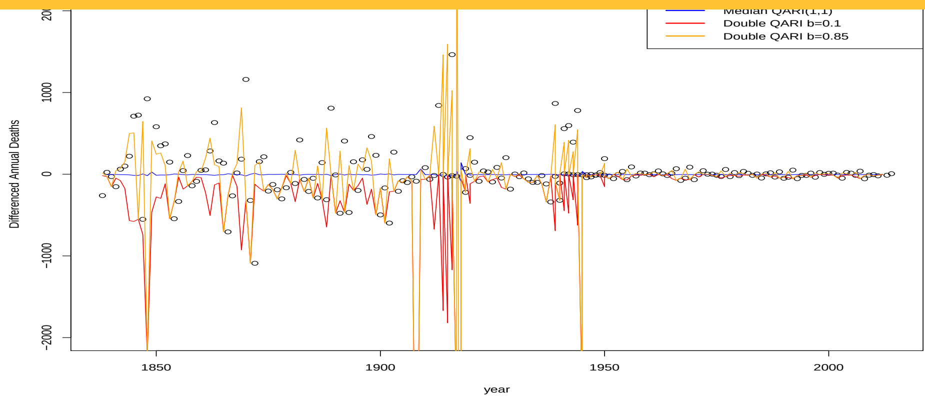

Figure 11: Comparisons between QARI(1,1) model and the Double QARI models for different parameter settings on the non-linear volatility parameter for differenced annual deaths.

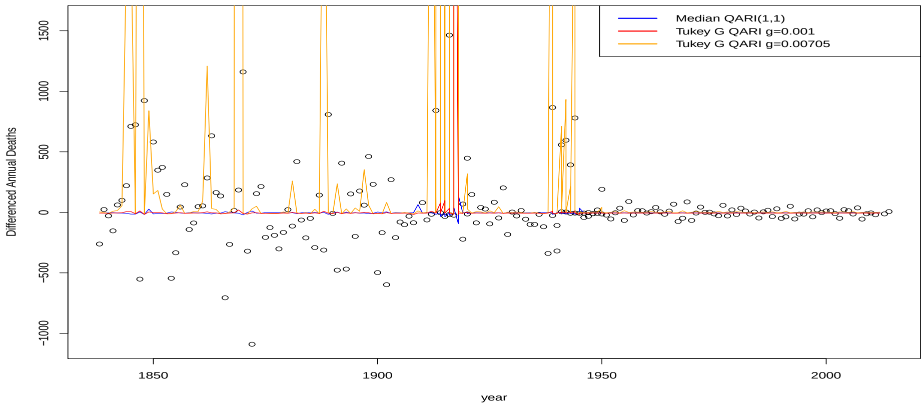

Figure 12: Comparisons between QARI(1,1) model and the Tukey G-and-H models for different parameter settings on the G skewness parameter for differenced annual deaths.

see that the values below the median have little influence on the result compared to the samples above the median. In these examples we have seen the behaviours of non-linear quantile regression models versus simple linear regressions.

9.3 Alcohol Related Age-Standardised Death Rates per 100,000 Population of UK

In this example we study the time series for alcohol related deaths in the UK for males and females between the period 1994 to 2016, see Figure 13. In this set of examples we study a model that may be interpreted according to the class of non-linear quantile time series models for the Rank Transmutation Maps (RTMs), that we demonstrate on the two real data sets. The particular example considered is a Weibul model growth curve model given by

$$f(x) = x_l - (x_l - x_u) \exp\left(-(kx)^{\delta}\right)$$

where x_l and x_u are the lower and upper asymptotes of the curve, k is the growth rate and δ controls the x co-ordinate for the point of inflection.

we then consider the fitted results for male and female alcohol annual related deaths over time under the Weibul growth curve quantile regression model, for 0.1, 0.5 and 0.9 quantile levels. The results of the non-linear quantile regression are provided in Figure 14. In both cases the Weibul RTM quantile model provides a reliable fit to at the quantile levels studied.

Peer-reviewed version available at *Risks* **2018**, *6*, 97; doi:10.3390/risks6030097

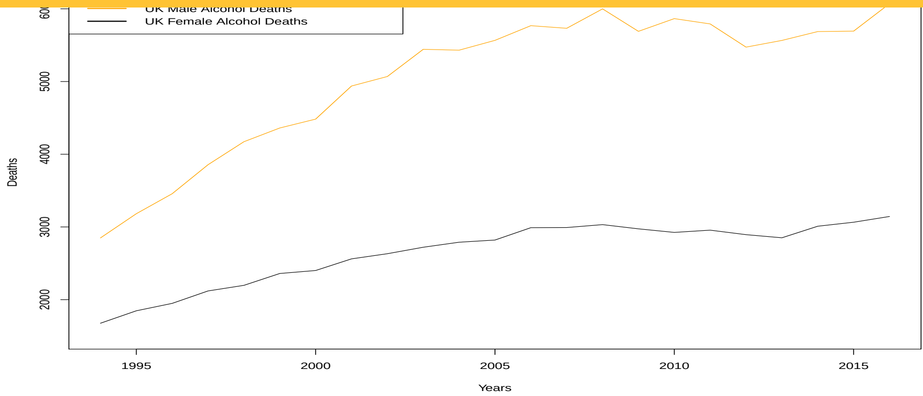

Figure 13: Alcohol Related Age-Standardised Death Rates per 100,000 Population of UK.

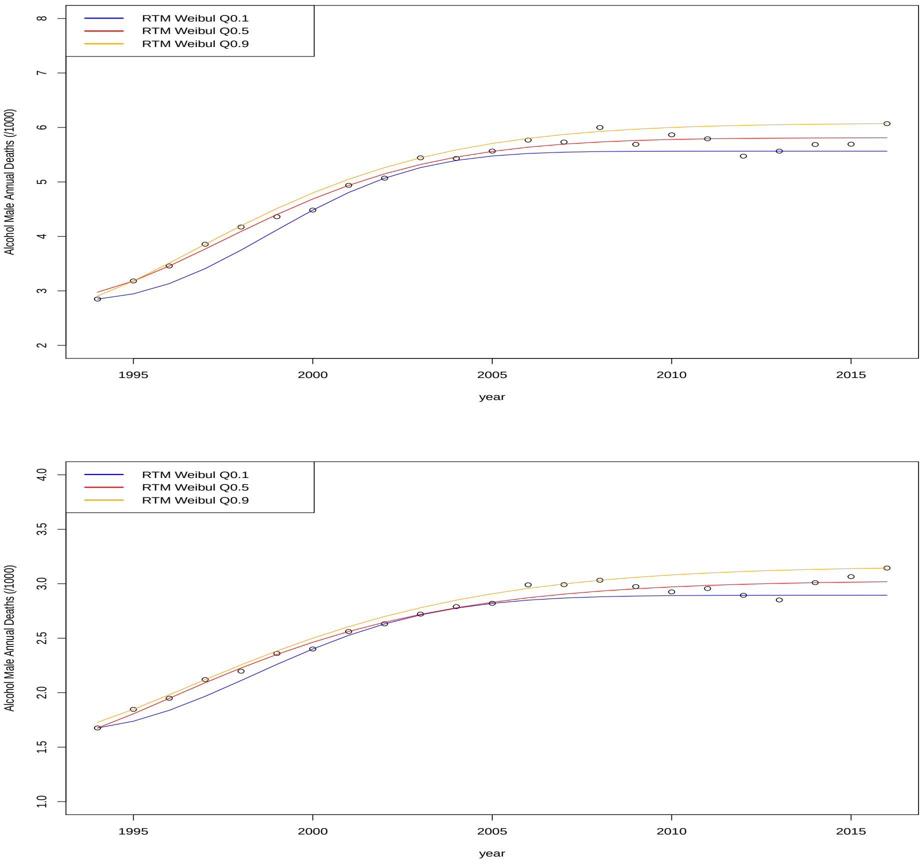

Figure 14: Alcohol Related Age-Standardised Death Rates per 100,000 Population of UK. Top Panel: Male Deaths; Bottom Panel: Female Deaths

9.4 Male Population Size 65-69: England and Wales

As a further illustration of the quantile time series models discussed in this tutorial, we will also consider fitting quantile time series models for the official recorded annual population size for males in age groups 65-69 between 1911 and the year 2000.

In this illustration a non-linear quantile regression based on a logistic growth model is considered, this is a form of RTM model with a logistic distribution. In its most basic form it is a logistic growth curve often termed the “S” shaped sigmoid curve, with equation:

$$f(x) = \frac{L}{1 + e^{-k(x - x_0)}}$$

where x_0 is the x-value of the sigmoid’s midpoint, L is the curve’s maximum value, and k is the steepness of are presented in Figure 15 for each of the three age groups.

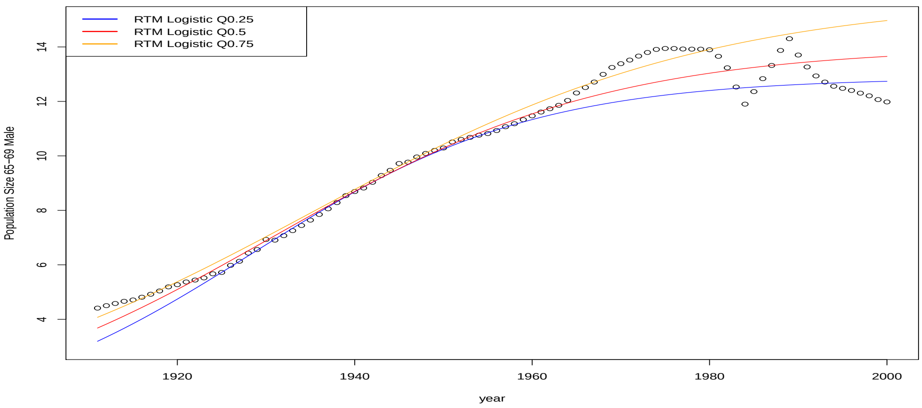

Figure 15: Quantile regressions for male population in England and Wales in the age group 65-69.

9.5 Conclusions

This manuscript has provided a detailed overview of the different strands of quantile time series regression modelling. Unifying the different approaches in a general modelling framework which allows one to treat each key component of a quantile time series model individually. In addition, several properties of special classes of quantile time series model are described and constructed in detail. The paper then concludes with detailed descriptions of different classes of quantile error family and quantile transformation maps that allow for construction of a vast array of flexible quantile time series models. The treatment of estimation is generally not addressed as the main focus of this tutorial paper is to address considerations of model construction and the properties of the resulting models.

The illustration of each of the main classes of quantile time series and regression structures explained in the tutorial are then illustrated in important mortality and demographic data sets for actuarial applications.

References

Roger Koenker and Gilbert Bassett Jr. Regression quantiles. *Econometrica: journal of the Econometric Society*, pages 33–50, 1978. 1, 2

Roger Koenker. *Quantile regression*. Number 38. Cambridge university press, 2005. 1

Warren Gilchrist. *Statistical modelling with quantile functions*. CRC Press, 2000. 1, 2, 3, 5.1, 5.2

Moshe Buchinsky. Recent advances in quantile regression models: a practical guideline for empirical research. *Journal of human resources*, pages 88–126, 1998. 1

Roger Koenker. Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, 91(1):74–89, 2004. 1

Keming Yu and MC Jones. Local linear quantile regression. *Journal of the American statistical Association*, 93(441):228–237, 1998. 1

Keming Yu and Rana A Moyeed. Bayesian quantile regression. *Statistics & Probability Letters*, 54(4):437–447, 2001. 1, 4.1

Paul Thompson, Yuzhi Cai, Rana Moyeed, Dominic Reeve, and Julian Stander. Bayesian nonparametric quantile regression using splines. *Computational Statistics & Data Analysis*, 54(4):1138–1150, 2010. 1

Keming Yu, Zudi Lu, and Julian Stander. Quantile regression: applications and current research areas. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 52(3):331–350, 2003. 1

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2018 doi:10.20944/preprints201806.0198.v1

Peer-reviewed version available at *Risks* **2018**, *6*, 97; doi:10.3390/risks6030097

arXiv:1605.05602, 2016a. 1, 6.1

Tony Lancaster and Sung Jae Jun. Bayesian quantile regression methods. *Journal of Applied Econometrics*, 25(2):287–307, 2010. 1

Qing Li, Ruibin Xi, Nan Lin, et al. Bayesian regularized quantile regression. *Bayesian Analysis*, 5(3):533–556, 2010. 1

Roger Koenker. Quantile regression: 40 years on. *Annual Review of Economics*, 9(1), 2017. 1, 2

Antonio F Galvao and Gabriel Montes-Rojas. On bootstrap inference for quantile regression panel data: A monte carlo study. *Econometrics*, 3(3):654–666, 2015. 1

Alice XD Dong, Jennifer SK Chan, and Gareth W Peters. Risk margin quantile function via parametric and non-parametric bayesian approaches. *Astin Bulletin*, 45(03):503–550, 2015. 1, 2, 6.2

Gareth W Peters, Wilson Ye Chen, and Richard H Gerlach. Estimating quantile families of loss distributions for non-life insurance modelling via l-moments. *Risks*, 4(2):14, 2016. 1, 2, 2, 5.2, 7, 7.1.1, 7.3, 7.1.2

Marcelo G Cruz, Gareth W Peters, and Pavel V Shevchenko. *Fundamental Aspects of Operational Risk and Insurance Analytics: A Handbook of Operational Risk*. John Wiley & Sons, 2015. 1, 5.2, 7, 7.1.2

Gareth W Peters and Pavel V Shevchenko. *Advances in Heavy Tailed Risk Modeling: A Handbook of Operational Risk*. John Wiley & Sons, 2015. 1, 5.2

Yuao Hu, Robert B Gramacy, and Heng Lian. Bayesian quantile regression for single-index models. *Statistics and Computing*, 23(4):437–454, 2013. 1

Mauro Bernardi, Roberto Casarin, Bertrand Maillet, and Lea Petrella. Dynamic model averaging for bayesian quantile regression. *arXiv preprint arXiv:1602.00856*, 2016b. 1

Roger Koenker. Galton, edgeworth, frisch, and prospects for quantile regression in econometrics. *Journal of Econometrics*, 95(2):347–374, 2000. 2

Roger Koenker and Zhijie Xiao. Quantile autoregression. *Journal of the American Statistical Association*, 101(475):980–990, 2006. 2, 2, 2

Bernd Fitzenberger, Roger Koenker, and José AF Machado. *Economic applications of quantile regression*. Springer Science & Business Media, 2013. 2

Guodong Li, Yang Li, and Chih-Ling Tsai. Quantile correlations and quantile autoregressive modeling. *Journal of the American Statistical Association*, 110(509):246–261, 2015. 2, 2

Heejoon Han, Oliver Linton, Tatsushi Oka, and Yoon-Jae Whang. The cross-quantilogram: Measuring quantile dependence and testing directional predictability between time series. *Journal of Econometrics*, 193(1):251–270, 2016. 2

Alexander Aue, Rex CY Cheung, Thomas CM Lee, Ming Zhong, et al. Piecewise quantile autoregressive modeling for nonstationary time series. *Bernoulli*, 23(1):1–22, 2017. 2

Michael Stephanou, Melvin Varughese, Iain Macdonald, et al. Sequential quantiles via hermite series density estimation. *Electronic Journal of Statistics*, 11(1):570–607, 2017. 2

Zongwu Cai. Regression quantiles for time series. *Econometric theory*, 18(01):169–192, 2002. 2

Zongwu Cai and Xiaoping Xu. Nonparametric quantile estimations for dynamic smooth coefficient models. *Journal of the American Statistical Association*, 103(484):1595–1608, 2008. 2

P Bloomfield and WL Steiger. Least absolute deviations. *Boston: Birkhauser. Brunner, D., Kacelnik, A., & Gibbon, J.(1992). Optimal foraging and timing processes in the starling Sturnus vulgaris: Effect of intercapture interval. Animal Behaviour*, 44:597–613, 1983. 2

Yuzhi Cai. Polynomial power-pareto quantile function models. *Extremes*, 13(3):291–314, 2010a. 2, 6.2

Peer-reviewed version available at *Risks* **2018**, *6*, 97; doi:10.3390/risks6030097

Journal of Forecasting, 32(6):551–560, 2013. 2, 2, 2, 2, 4.2

Andrew A Weiss. Estimating nonlinear dynamic models using least absolute error estimation. *Econometric Theory*, 7(01):46–68, 1991. 2

Richard A Davis and William TM Dunsmuir. Least absolute deviation estimation for regression with arma errors. *Journal of Theoretical Probability*, 10(2):481–497, 1997. 2

Roger Koenker and Quanshui Zhao. Conditional quantile estimation and inference for arch models. *Econometric Theory*, 12(05):793–813, 1996. 2

Sangyeol Lee and Jungsik Noh. Quantile regression estimator for garch models. *Scandinavian Journal of Statistics*, 40(1):2–20, 2013. 2, 2

Jungsik Noh and Sangyeol Lee. Quantile regression for location-scale time series models with conditional heteroscedasticity. *Scandinavian Journal of Statistics*, 2015. 2, 2

Robert F Engle and Simone Manganelli. Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, 22(4):367–381, 2004. 2

Yuzhi Cai. A general quantile function model for economic and financial time series. *Econometric Reviews*, 35(7):1173–1193, 2016. 2

Shiqing Ling. A double ar (p) model: structure and estimation. *Statistica Sinica*, 17(1):161, 2007. 2

Andrew A Weiss. Arma models with arch errors. *Journal of Time Series Analysis*, 5(2):129–143, 1984. 2

Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007, 1982. 2

Marshall Freimer, Georgia Kollia, Govind S Mudholkar, and C Thomas Lin. A study of the generalized tukey lambda family. *Communications in Statistics-Theory and Methods*, 17(10):3547–3567, 1988. 2

Yuzhi Cai and Julian Stander. Quantile self-exciting threshold autoregressive time series models. *Journal of Time Series Analysis*, 29(1):186–202, 2008. 2

Wilson Y Chen, Gareth William Peters, Richard H Gerlach, and Scott Sisson. Dynamic quantile function models. *Available at SSRN: https://ssrn.com/abstract=2999451*, 2017. 2, 2

Dennis Prangle. gk: An r package for the g-and-k and generalised g-and-h distributions. *arXiv preprint arXiv:1706.06889*, 2017. 2

Yuzhi Cai. Forecasting for quantile self-exciting threshold autoregressive time series models. *Biometrika*, 97 (1):199–208, 2010b. 2

Alexander J McNeil and Rüdiger Frey. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of empirical finance*, 7(3):271–300, 2000. 2

Roger Koenker and Kevin Hallock. Quantile regression: An introduction. *Journal of Economic Perspectives*, 15(4):43–56, 2001. 4.1

Roger Koenker and Jose AF Machado. Goodness of fit and related inference processes for quantile regression. *Journal of the american statistical association*, 94(448):1296–1310, 1999. 4.1

GW Peters and SA Sisson. Bayesian inference, monte carlo sampling and operational risk. *Journal of Operational Risk*, 1(3):27–50, 2006. 5.2, 7

William T Shaw and Ian RC Buckley. The alchemy of probability distributions: beyond gram-charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *arXiv preprint arXiv:0901.0434*, 2009. 5.2, 7, 8, 8.1, 8.1

Kevin P Balanda and HL MacGillivray. Kurtosis and spread. *Canadian Journal of Statistics*, 18(1):17–30, 1990. 5.2, 7.1.1

generalized g-and-h distributions. *Statistics and Computing*, 12(1):57–75, 2002. 5.2, 7.1, 7.1.1

Dongming Zhu and Victoria Zinde-Walsh. Properties and estimation of asymmetric exponential power distribution. *Journal of econometrics*, 148(1):86–99, 2009. 6.1

J Miguel Marín and Genaro Sucarrat. Modelling the skewed exponential power distribution in finance. In *Mathematical and statistical methods for actuarial sciences and finance*, pages 279–286. Springer, 2012. 6.1

Thomas J DiCiccio and Anna Clara Monti. Inferential aspects of the skew exponential power distribution. *Journal of the American Statistical Association*, 99(466):439–450, 2004. 6.1

AXD Dong and JSK Chan. Bayesian analysis of loss reserving using dynamic models with generalized beta distribution. *Insurance: Mathematics and Economics*, 53(2):355–365, 2013. 6.2, 6.2

James B McDonald. Some generalized functions for the size distribution of income. *Econometrica: journal of the Econometric Society*, pages 647–663, 1984. 6.2

Eo W Stacy. A generalization of the gamma distribution. *The Annals of mathematical statistics*, pages 1187–1192, 1962. 6.2

Fernando De Helguero. *Sulla rappresentazione analitica delle statistiche abnormali*. 1908. 7

SAMUEL KOTZ-DONATELLA VICARI. Survey of developments in the theory of continuous skewed distributions. *Metron*, 63(2):225–261, 2005. 7

Adelchi Azzalini. The skew-normal distribution and related multivariate families. *Scandinavian Journal of Statistics*, 32(2):159–188, 2005. 7

Marc G Genton. Discussion of âĂŸthe skew-normalâĂŹ. *Scandinavian Journal of Statistics*, 32(2):189–198, 2005. 7

John W Tukey. Modern techniques in data analysis. In *Proceedings of the NSF-Sponsored Regional Research Conference*, 1977. 7

David C Hoaglin. *Summarizing Shape Numerically: The g-and-h Distributions*. Wiley Online Library, 1985a. 7

Martinez Jorge and Iglewicz Boris. Some properties of the tukey g and h family of distributions. *Communications in Statistics-Theory and Methods*, 13(3):353–369, 1984. 7

Adelchi Azzalini. A class of distributions which includes the normal ones. *Scandinavian journal of statistics*, pages 171–178, 1985. 7

Matthias Fischer, Armin Horn, and Ingo Klein. Tukey-type distributions in the context of financial data. *Communications in Statisticsï£¡Theory and Methods*, 36(1):23–35, 2007. 7

Christopher Field and Marc G Genton. The multivariate g-and-h distribution. *Technometrics*, 2012. 7

Kabir Dutta and Jason Perry. A tale of tails: an empirical analysis of loss distribution models for estimating operational risk capital. 2006. 7

Matthias Degen, Paul Embrechts, and Dominik D Lambrigger. The quantitative modeling of operational risk: between g-and-h and evt. *Astin Bulletin*, 37(02):265–291, 2007. 7, 7.1.2, 7.3, 7.1.2, 7.1.2, 7.1.2

José Alfredo Jiménez and Viswanathan Arunachalam. Using tukeyï£¡s g and h family of distributions to calculate value-at-risk and conditional value-at-risk. *Journal of Risk*, 13(4):95–116, 2011. 7

Michele A Haynes, HL MacGillivray, and KL Mengersen. Robustness of ranking and selection rules using generalised g-and-k distributions. *Journal of Statistical Planning and Inference*, 65(1):45–66, 1997. 7

Md Anower Hossain and Syed Shahadat Hossain. Numerical maximum likelihood estimation for the g-and-k distribution using ranked set sample. *Journal of Statistics*, 16(1), 2009. 7

tisches Archiv, 88(1):35–50, 2004. 7.1

David C Hoaglin. Summarizing shape numerically: The g-and-h distributions. *Exploring data tables, trends, and shapes*, pages 461–513, 1985b. 7.1

Matthias Fischer. Generalized tukey-type distributions with application to financial and teletraffic data. *Statistical Papers*, 51(1):41–56, 2010. 7.1, 7.1

Kabir K Dutta, David F Babbel, et al. On measuring skewness and kurtosis in short rate distributions: The case of the us dollar london inter bank offer rates. *Center for Financial Institutions Working Papers 02*, 25, 2002. 7.1, 7.1.1, 7.2

Kevin P Balanda and HL MacGillivray. Kurtosis: a critical review. *The American Statistician*, 42(2): 111–119, 1988. 7.1.1

Stephan Morgenthaler and John W Tukey. Fitting quantiles: Doubling, hr, hq, and hhh distributions. *Journal of Computational and Graphical Statistics*, 9(1):180–195, 2000. 7.1.2, 7.1.2, 7.1.2

Ioannis Karatzas and Steven E Shreve. Brownian motion and stochastic calculus, vol. 113 of. *Graduate texts in mathematics*, 1991. 7.1.2

Nicholas H Bingham, Charles M Goldie, and Jef L Teugels. *Regular variation*, volume 27. Cambridge university press, 1989. 7.1.2, 7.1.2

Paul Embrechts, Claudia Klüppelberg, and Thomas Mikosch. *Modelling extremal events: for insurance and finance*, volume 33. Springer Science & Business Media, 2013. 7.1.2

J-P Raoult and R Worms. Rate of convergence for the generalized pareto approximation of the excesses. *Advances in Applied Probability*, pages 1007–1027, 2003. 7.1.2