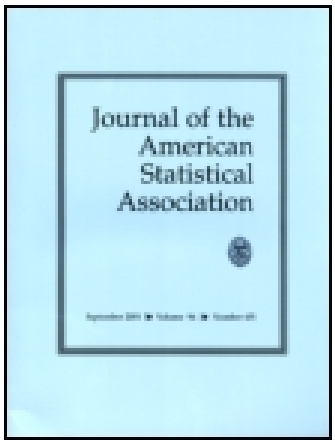

Journal of the American Statistical Association
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/uasa20

An Approach to Multivariate Covariate-Dependent Quantile Contours With Application to Bivariate Conditional Growth Charts

Ying Wei[a]

[a] YingWei is Assistant Professor, Department of Biostatistics, Columbia University, New York, NY 10032 . This work was supported in part by National Science Foundation grant DMS-0504972. The author thanks the associate editor and two referees for their helpful comments and suggestions.
Published online: 01 Jan 2012.

To cite this article: Ying Wei (2008) An Approach to Multivariate Covariate-Dependent Quantile Contours With Application to Bivariate Conditional Growth Charts, Journal of the American Statistical Association, 103:481, 397-409, DOI: 10.1198/016214507000001472

To link to this article: http://dx.doi.org/10.1198/016214507000001472

An Approach to Multivariate Covariate-Dependent Quantile Contours With Application to Bivariate Conditional Growth Charts

Ying WEI

Multivariate quantile contours are useful in numerous applications and have been studied in different contexts. However, no easy solutions exist when dynamic and conditional quantile contours are needed without strong distributional assumptions. In this article we propose a new form of bivariate quantile contours and a two-stage estimation procedure to take time effect into account. The proposed procedure relies on quantile regression for longitudinal data and is flexible to include potentially important covariates as necessary. In addition, we propose a visual model assessment tool and discuss a practical guideline for model selection. The performance of the proposed methodology is demonstrated by a simulation study, as well as an application to joint height–weight screening of young children in the United States. We construct bivariate growth charts by a nested sequence of age-dependent and covariate-varying quantile contours of height and weight, and use it to locate an individual subject's percentile rank with respect to a reference population. Our work shows that the proposed method is valuable for pediatric growth monitoring and provides more informative readings than the conventional approach based on univariate growth charts.

KEY WORDS: Bivariate quantile contours; Longitudinal; Quantile regression.

1. INTRODUCTION

The quantile function plays an important role in univariate data analysis. The univariate quantile function $Q_\tau(X)$ of a random variable X is the inverse of its cumulative distribution function (cdf), that is, $Q_\tau(X) = F_X^{-1}(\tau) = \inf\{x : F(x) \geq \tau\}$, where $\tau \in (0, 1)$ is the percentile rank and F is the cdf of X. Due to the lack of natural ordering in a multidimensional space, there is no universally preferred definition of multivariate quantiles, but various ideas have been developed in the literature; for example, Liu, Parelius, and Singh (1999) and Zuo and Serfling (2000) used multivariate quantile functions based on the half-space depth function. Other approaches have been given by Parzen (1979), Abdous and Theodorescu (1992), Hettmansperger, Nyblom, and Oja (1992), Chaudhuri (1996), Koltchinskii (1997), Chakraborty (2003), and McDermott and Lin (2007), to name just a few. Serfling (2002) presented a nice survey of multivariate quantile functions and outlined the probabilistic properties that a multivariate quantile function should have.

Because life is never static, the distributions of the random variables often change over time. For example, children's height and weight grow with age, stock prices and trading volumes vary constantly, CD4 cell counts and viral loads of patients positive for human immunodeficiency virus change with the duration of therapy, and so on. It is often desirable to obtain dynamic quantile functions that take into account the effect of time and/or other covariate effects. Quantile regression (Koenker and Bassett 1978) provides a direct means of estimating $Q_\tau(Y|X)$, the conditional quantiles of Y given X, where Y is a real-valued variable. Under this framework, various dynamic quantile models have been proposed by Kim (2007), Cai and Xu (2007), and Wei and He (2006). But, it remains challenging to incorporate the time/covariate effect into multivariate quantiles, that is, when Y is multivariate. As mentioned earlier, existing methods for constructing multivariate quantile/depth functions have difficulty incorporating the effects of covariates without strict distributional assumptions.

The present article proposes a new form of dynamic bivariate quantile function that is probabilistically interpretable and is sufficiently flexible to capture a wide range of distributions. It also proposes a two-stage method to estimate such dynamic (time-dependent and covariate-varying) bivariate quantile functions without parametric distributional assumptions and without requiring large sample sizes at all time points, and applies the proposed methodology to constructing bivariate age-dependent growth charts to give pediatricians a more comprehensive assessment of children's growth/health status.

The proposed bivariate quantile function is a natural generalization to the univariate reference centiles (Wei, Pere, Koenker, and He 2006). To differentiate these from the other bivariate quantile functions, we call them *reference quantile contours*. Let $\mathbf{Y}_t = (Y_{t,1}, Y_{t,2}) \in \mathbb{R}^2$ be a two-dimensional random vector measured at time t and let $\mathbf{F}(y|t, X)$ be its conditional joint distribution given time t and covariate X. We then denote the reference conditional quantile contour at time t and a given covariate value X as $C_\tau(\cdot|t, X)$. The formal definition of C_τ and its probabilistic properties are provided in Section 2. To estimate the time-varying conditional C_τ, we propose a two-stage approach. We first construct a set of *stratified quantile models*, a marginal quantile model of $Y_{t,1}$ as a function of time t and covariate X, followed by a conditional quantile model of $Y_{t,2}$ as a function of t, X, and $Y_{t,1}$. We show that these estimated marginal and conditional quantiles provide consistent estimation of the time-dependent conditional joint distribution $\mathbf{F}(y|t, X)$. Moreover, fixing t and X, there is a simple way to simulate a pair of random variables sequentially from the estimated marginal and conditional quantiles. The simulated pair follows the estimated joint distribution of $\mathbf{Y}_t$ conditional on t and X, denoted as $\hat{\mathbf{F}}(y|t, X)$. Given that $\hat{\mathbf{F}}(y|t, X)$ is a good approximation of $\mathbf{F}(y|t, X)$, this simulation can be viewed as a random draw from the underlying joint distribution $\mathbf{F}(y|t, X)$. Therefore, we could repeat the *model-based simulation* in the

Ying Wei is Assistant Professor, Department of Biostatistics, Columbia University, New York, NY 10032 (E-mail: *ying.wei@columbia.edu*). This work was supported in part by National Science Foundation grant DMS-0504972. The author thanks the associate editor and two referees for their helpful comments and suggestions.

Journal of the American Statistical Association
March 2008, Vol. 103, No. 481, Theory and Methods
DOI 10.1198/016214507000001472

second stage to obtain a sufficiently large sample. Once we obtained simulated data, we could estimate the reference quantile contour at each time point t and each given covariate value using the nonparametric algorithm proposed in Section 2.2. This modeling and simulation approach avoids the difficulty arising from the insufficient sample size at individual time points and also is sufficiently flexible to include the other covariate effect when necessary.

Noting that the performance of the proposed method relies on how well the stratified quantile models fit the data, we also propose a visual goodness-of-fit assessment comparing the model-estimated distribution function and empirical distribution on a grid of time points. Finally, it is easy to see that switching the order of $Y_{t,1}$ and $Y_{t,2}$ will produce another distinct set of stratified quantile models. To answer the questions of which order to use and whether the two orders should be combined, we discussed a practical guideline on model selection and combination based partly on the proposed model fitting assessment.

The rest of the article is organized as follows. Section 2 describes our general methodology, with a simulation study presented in Section 2.4. Section 3 considers a height–weight growth chart for young children in the United States and provides some illustrative examples of bivariate screening versus marginal screening. Section 4 provides brief discussion on its general applications and possible extensions.

2. DYNAMIC BIVARIATE QUANTILE FUNCTION

2.1 General Setting and Notation

Suppose that $\mathbf{Y}_t = (Y_{t_1}, Y_{t_2}) \in \mathbb{R}^2$ is a pair of measurements of interest at time t and that $\mathbf{X}$ is q-dimensional associated covariates. We assume that at any given time t and covariate $\mathbf{X}$, the measurement $\mathbf{Y}_t$ follows a continuous joint density $f(y|t, X)$ defined on $\mathbb{R}^2$. Moreover, we denote a random sample consisting of n subjects as $\{(\mathbf{Y}_{i,j}, \mathbf{X}_{i,j}, t_{i,j}); i = 1, \ldots, n; j = 1, \ldots, m_i\}$, where $\mathbf{Y}_{i,j} = (Y_{i,j,1}, Y_{i,j,2})^\top$ is the jth paired measurements of the ith subject at time $t_{i,j}$ with $\mathbf{X}_{i,j} = (X_{i,j,1}, X_{i,j,2}, \ldots, X_{i,j,q})^\top$. We assume that the measurement times $t_{i,j}$ have finite support $[0, T]$ for some constant T but are not necessary evenly spaced for each i. The covariates $\mathbf{X}_{i,j}$ could be the characteristics of the ith subject, another measurement at time $t_{i,j}$, or his or her previous records, and thus they may or may not be time-dependent. The main goal of our proposed methodology is to define and estimate dynamic quantile function of $\mathbf{Y}_t$ based on the random sample $(\mathbf{Y}_{i,j}, \mathbf{X}_{i,j}, t_{i,j})$. The following notations are frequently used throughout the article: for a random variable X, $Q_\tau(X)$ denotes its τth quantile; for vectors, $\|\cdot\|$ denotes the Euclidean norm and $\langle \mathbf{x}, \mathbf{y} \rangle$ denotes for the inner product of $\mathbf{x}$ and $\mathbf{y}$.

2.2 Bivariate Reference Quantile Contours

In this section we first describe the proposed multivariate reference quantile contours assuming that there is no time or covariate effect, and then propose a nonparametric algorithm to estimate such quantile contours based on a random sample $\{\boldsymbol{y}_i = (y_{i,1}, y_{i,2})^\top, i = 1, \ldots, n\}$. We use this algorithm in later sections at selected values of time t and covariate X. In Section 2.3 we discuss how to include time t and covariates X into the reference quantile contours.

We designate $\mathbf{Y}$ as a p-dimensional measurement vector, even though the rest of the article we restrict ourselves to $p = 2$. Let $\mathbb{S}^{(p-1)}$ be the unit sphere around a center in $\mathbb{R}^p$; then the spatial direction can be given by a line connecting a pair of opposite points $(\mathbf{u}, -\mathbf{u})$ on $\mathbb{S}^{(p-1)}$. Our definition starts by defining a central interval as $[F_{\mathbf{u}}^{-1}(\frac{1-\tau}{2}), F_{\mathbf{u}}^{-1}(\frac{1+\tau}{2})]$ along each spatial direction, where $F_{\mathbf{u}}$ is the distribution of condition $\mathbf{Y}$ on the line connecting $\mathbf{u}$ and $-\mathbf{u}$. The union of these intervals forms a compact central set with $\tau 100\%$ coverage probability, the boundary of which serves naturally as the corresponding τth reference quantile contour. By construction, the reference quantile contours adapt to the underlying shape of the distribution of interest and are not constrained to a particular family of a priori shapes (e.g., ellipses). They essentially measure the spatial "outlyingness" of an observation relative to a center. To facilitate the formal definition, we first introduce the following notations: For any two points $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^p$, we write $\mathbf{y} \overset{(\mathbf{u})}{=} \mathbf{x}$ if both of them stay on the $(\mathbf{u}, -\mathbf{u})$ axis—that is, they follow the same spatial direction. Moreover, we say that $\mathbf{y}$ is more *outlying* than $\mathbf{x}$ toward $\mathbf{u}$, denoted as $\mathbf{y} \geq_{(\mathbf{u})} \mathbf{x}$, iff $\langle \mathbf{y}, \mathbf{u} \rangle \geq \langle \mathbf{x}, \mathbf{u} \rangle$. With these notations, we have the following definition.

Definition 1 (p-dimensional reference quantile contour). For a p-dimensional random vector $\mathbf{Y}$, we choose $\boldsymbol{\mu}$ as a componentwise location parameter of $\mathbf{Y}$ and $\mathbf{S} = \operatorname{diag}(s_i)_{i=1,\ldots,p}$ as the $p \times p$ diagonal matrix with componentwise scale parameters of $\mathbf{Y}$. Let $\mathbf{Z} = \mathbf{S}^{-1}(\mathbf{Y} - \boldsymbol{\mu})$ be the standardized measurement. Then for any $\tau \in (0, 1)$, the following conditions hold:

a. The $\tau 100\%$ directional reference interval of $\mathbf{Z}$ is the closed interval $[l_\tau(\mathbf{u}), u_\tau(\mathbf{u})]$ along the axis $(\mathbf{u}, -\mathbf{u})$, which contains all of the points $\mathbf{x} \overset{(\mathbf{u})}{=} \mathbf{u}$ satisfying

$$\mathrm{P}\big(\mathbf{Z} \leq_{(\mathbf{u})} \mathbf{x} \big| \mathbf{Z} \overset{(\mathbf{u})}{=} \mathbf{x}\big) \leq \frac{1+\tau}{2}$$

and

$$\mathrm{P}\big(\mathbf{Z} \leq_{(\mathbf{u})} \mathbf{x} \big| \mathbf{Z} \overset{(\mathbf{u})}{=} \mathbf{x}\big) \geq \frac{1-\tau}{2}.$$

b. The τth reference quantile contour C_τ is the boundary of the central set

$$R_\tau = \left\{ S\mathbf{x} + \boldsymbol{\mu} : \mathbf{x} \in \bigcup_{\mathbf{u} \in \mathbb{S}^{(p-1)}} [l_\tau(\mathbf{u}), u_\tau(\mathbf{u})] \right\}, \tag{1}$$

that is, $C_\tau = \partial R_\tau$.

Intuitively speaking, the central set R_τ consists of a family of directional reference intervals, each of which reflects the outlyingness of measurements from $\boldsymbol{\mu}$ along one spatial direction. When $p = 1$, R_τ is the interval $[F^{-1}(\frac{1-\tau}{2}), F^{-1}(\frac{1+\tau}{2})]$, and C_τ is its two endpoints. Moreover, the spatial direction is equivalent to the constant ratio between the standardized measurements, which may reflect certain characteristics of the measurements. Again using human height and weight again as an example, the ratio between standardized weight and height characterizes the body composition; consequently, the directional reference interval is the range of normal height and weight values among those with similar body composition. In the rest of the article, we have $p = 2$, and choose $\boldsymbol{\mu}$ to be the componentwise median and take the scale parameter in $\mathbf{S}$ as

the componentwise median absolute deviation (MAD) for robust purposes. Because each component in $\mathbf{Y}$ usually has its own implications, we stick to componentwise standardizations. The defined bivariate contours have the following properties. Appendix A provides a sketch of the proof.

Proposition 1. If $\mathbf{Y}$ has positive probability distribution function everywhere in $\mathbb{R}^p$, then (a) for any $\tau \in (0,1)$, R_τ is uniquely defined with coverage probability τ; (b) C_τ is a continuous contour; (c) the C_τ's are nested, that is, for any $\tau_1 < \tau_2$, the normal regions satisfy $R_{\tau_1} \subseteq R_{\tau_2}$; and (d) C_τ is equivariant to the location and scale transformation on $\mathbf{Y}$.

Estimation of Reference Quantile Contours. Suppose that $\{\mathbf{y}_i = (y_{i,1}, y_{i,2})^\top, i = 1, \ldots, n\}$ is a random sample following the joint density of a random vector $\mathbf{Y}$. Then we could estimate the reference quantile contours C_τ of $\mathbf{Y}$ using the following nonparametric estimation algorithm:

Step 1. We standardize that $\mathbf{y}_i$'s by $\mathbf{z}_i = \hat{\mathbf{S}}^{-1}(\mathbf{y}_i - \hat{\boldsymbol{\mu}})$, where $\hat{\boldsymbol{\mu}}$ is the vector of sample medians and $\hat{\mathbf{S}}$ is the diagonal matrix of sample MADs.

Step 2. First we note that the τth reference quantile contour of $\mathbf{z}$ can be written as a function of the polar angle $g_\tau(\theta) : [0, 2\pi] \to \mathbb{R}$ subject to the constraint $g_\tau(0) = g_\tau(2\pi)$. This $g_\tau(\theta)$ consists of two conditional quantiles (the upper and lower) along each direction $(-\mathbf{u}, \mathbf{u})$. Using dual-polar representations for each $\mathbf{z}_i$, that is,

$$\theta_i^{(1)} = \arctan(z_{i,2}/z_{i,1}) \cap [0, \pi),$$
$$r_i^{(1)} = \left\langle \mathbf{z}_i, \left(\cos\theta_i^{(1)}, \sin\theta_i^{(1)}\right)^\top \right\rangle$$

and

$$\theta_i^{(2)} = \arctan(z_{i,2}/z_{i,1}) \cap [\pi, 2\pi),$$
$$r_i^{(2)} = \left\langle \mathbf{z}_i, \left(\cos\theta_i^{(2)}, \sin\theta_i^{(2)}\right)^\top \right\rangle = -r_i^{(1)},$$

where $\theta_i^{(1)}$ and $\theta_i^{(2)}$ differ exactly by π, we can estimate $g_\tau(\theta)$ by one-step optimization minimizing

$$\sum_{k=1}^{2}\sum_{i=1}^{n} \rho_{(\tau+1)/2}\left(r_i^{(k)} - g\left(\theta_i^{(k)}\right)\right), \tag{2}$$

where $\rho_\tau(\xi) = \xi(\tau - \mathbf{I}_{\{\xi<0\}})$, over a family of constrained smooth functions,

$$\mathcal{G} = \{g(\theta) : [0, 2\pi] \to \mathbb{R} \text{ s.t. } g(0) = g(2\pi)\}.$$

The two summands of (2) correspond to the objective functions for the two conditional quantiles; one yields a half contour on $\theta \in [0, \pi)$, and the other gives the second half on $[\pi, 2\pi)$. These can be estimated simultaneously due to the fact that $Q_\tau(-Y) = -Q_{1-\tau}(Y)$. In our proposal we approximate g by a linear combination of B-spline basis functions using the constrained algorithm of He and Ng (1999).

Step 3. We let $\hat{g}_\tau$ be the estimated function in Step 2. Then the rectangular coordinates of the τth reference quantile contour of Y is

$$\hat{C}_\tau = \hat{\mathbf{S}}\hat{g}(\theta)\begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix} + \hat{\boldsymbol{\mu}}, \qquad \theta \in [0, 2\pi),$$

which is a consistent estimate of the true reference quantile contour C_τ.

Remark 1. We use the constrained smoothing algorithm of He and Ng (1999) for convenience in estimating g, but other nonparametric estimators may be equally effective, such as kernel smoothing or locally constant(linear) estimates (Yu and Jones 1998).

Remark 2. The proposed quantile contours work well for unimodal distributions, although unimodality is not required. If we view the interior of each contour as a prediction region, then in general the prediction region does not have the minimum volume property among all possible prediction regions with the same coverage probability. The choice of the centering parameter μ may influence the shape of the contours. Alternative quantile contours, such as density-based quantile contours (e.g., Fraiman and Meloehe 1999), may be used, and the methodology presented in the next section for adjusting for the effect of covariates would apply just as well.

Remark 3. Our estimation algorithm assumes that the $g_\tau(\cdot)$ in (2) is a smooth function, which implicitly implies that the center μ is contained in all of the directional reference intervals $[l_\tau(\mathbf{u}), u_\tau(\mathbf{u})], \forall \mathbf{u}$. When this assumption fails, the proposed estimation procedure could fail to produce a closed contour. In Appendix B we illustrate this phenomenon with an artificially created banana-shaped data suggested by a referee.

As we discussed in Section 1, our main interest in this work is to estimate the conditional quantile contour C_τ given time t and covariate X. If we have a sufficient sample size at any time t and any given covariate value X, then we can well estimate the conditional $C_\tau(y|t, X)$'s following the our algorithm. But, the sample size is often inadequate at individual time points, as is common in the presence of irregular measurement times. This can be even more problematic when the covariates are included. For this reason, we propose a quantile-based modeling strategy to "borrow" information from the neighbors and from the entire data set, which in turn allows us to include the time/covariate effects without requiring a large reference sample everywhere. This modeling step connects to the estimation algorithm proposed in this section by model-based simulation, as we describe in the next section.

2.3 A Two-Stage Method for Estimating Dynamic Bivariate Quantile Contour

In this section we propose a two-stage modeling and estimation procedure, that enables estimation of the time-dependent and/or covariate-adjusted reference quantile contours. Quantile regression plays a major role in the proposed methodology.

Stratified Quantile Regression and Model-Based Simulation. Recall that $(\mathbf{Y}_{i,j}, t_{i,j}, X_{i,j})$ is the random sample as defined in Section 2.1. We model the distribution of $\mathbf{Y}$ given t and X by constructing two stratified quantile models: a marginal model of $Y_{t,1}$ and a conditional model of $Y_{t,2}$ given $Y_{t,1}$; that is, for any $\tau \in (0, 1)$, we specify

$$Q_\tau(Y_{i,j,1}) = g_1(\tau; t_{i,j}, \mathbf{X}_{i,j}) \tag{3}$$

and

$$Q_\tau(Y_{i,j,2}) = g_2(\tau; t_{i,j}, \mathbf{X}_{i,j}, Y_{i,j,1}), \tag{4}$$

where we assume that $g_1(\tau; t, \mathbf{X})$ is the τth quantile function of $Y_{t,1}$ given time t and covariate $\mathbf{X}$, and $g_2(\tau; t, \mathbf{X}, y_1)$ is the τth conditional quantile function of $Y_{t,2}$ at time t and covariate $\mathbf{X}$, and given the value of $Y_{t,1} = y_1$. The estimates of g_1 and g_2 can be obtained by minimizing the quantile regression objective functions

$$\hat{g}_{n1}(\tau) = \operatorname*{argmin}_{g_1 \in \mathcal{G}_1} \sum_{i,j} \rho_\tau(Y_{i,j,1} - g_1(t_{i,j}, \mathbf{X}_{i,j}))$$

and

$$\hat{g}_{n2}(\tau) = \operatorname*{argmin}_{g_2 \in \mathcal{G}_2} \sum_{i,j} \rho_\tau(Y_{i,j,1} - g_2(t_{i,j}, \mathbf{X}_{i,j}, Y_{i,j,1})),$$

where $\mathcal{G}_1$ is a certain family of measurable functions of t and $\mathbf{X}$, and $\mathcal{G}_2$ is that of $t, \mathbf{X}$, and y_1. This marginal-conditional structure also was used by Ma and Koenker (2006) to construct structural equation models. Based on the stratified quantile regression models, there is a nice, easy way to generate a random sample that follows the joint distribution of $\mathbf{Y}$ at any given time t and covariate X. We elaborate this model-based simulation procedure by first introducing the following lemma.

Lemma 1. Suppose that (Y_1, Y_2) is a pair of random variables with absolutely continuous joint distribution F_{Y_1,Y_2}, and let U_1 and U_2 be two independent random variables uniformly distributed on (0, 1); then

$$\left(F_{Y_1}^{-1}(U_1), F_{Y_2|Y_1}^{-1}(U_2|U_1)\right) \sim F_{Y_1,Y_2},$$

where F_{Y_1} is the marginal distribution of Y_1 and $F_{Y_2|Y_1}$ is the conditional distribution of Y_2 given Y_1.

Lemma 1 provides an one-to-one continuous mapping between the sample space of (Y_1, Y_2) and a unit square $(0, 1) \times (0, 1)$. The corresponding inverse mapping is known as Rosenblatt's transformation (Mardia, Kent, and Bibby 1979). Using this idea, we could generate a random sample of (Y_1, Y_2) sequentially from the stratified quantile models (3) and (4). Specifically, we first draw U_1 from a uniform distribution on (0, 1), and then generate Y_1 as the U_1th quantile from the (estimated) marginal model given t and X. We then draw U_2 from the uniform (0, 1) distribution and generate Y_2 as the U_2th quantile from the (estimated) conditional model at the same values of t, X, and the Y_1 generated in the first step. According to Lemma 1, we have simulated the pair (Y_1, Y_2) that follows the joint distribution of Y_t given time t and covariate X. The simulated data based on the regression quantile estimates $\hat{g}_{n1}$ and $\hat{g}_{n2}$, that is,

$$Y_{n_1}^* = \hat{g}_{n_1}(U_1; t, X)$$

and

$$Y_{n_2}^* = \hat{g}_{n_2}\left(U_2; t, X, Y_{n_1}^*\right)$$

follow the model-estimated joint distribution $\mathbf{Y}$ given t and $\mathbf{X}$, denoted as F_n^*. Let $F_0(y_1, y_2)$ be the true distribution; then, under certain regulation conditions that (3) and (4) are consistently estimated, we have, for any $(y_1, y_2) \in \Omega$,

$$|F_n^*(y_1, y_2) - F_0(y_1, y_2)| = o_p(1), \tag{5}$$

as n goes to infinity. The mathematical form of $F_n^*(y_1, y_2)$ is provided in (9) of Section 2.4. Therefore, with sufficient sample size n, the simulated $(Y_{n_1}^*, Y_{n_2}^*)$ can be taken as a draw from the underlying distribution at the given time t and covariate X, and we use those draws to construct the desired reference quantile contours. Moreover, once the models are estimated, the sample generating procedure is computationally simple, and thus we can generate a sufficiently large sample size for quantile estimates in the following step. Here the model-based simulation nicely bridges the stratified quantile regression modeling with the estimation procedure outlined in Section 2.2, which, consequently, allows us to incorporate t and X into the bivariate quantile function.

To sum up, the two-stage approach works as follows. The first stage is model estimation, in which we fit the stratified quantile regression model using the entire data set, and the second stage is the contour estimation, in which we use model-based simulation to obtain a sufficiently large sample at given t and X and use it to estimate the desired reference quantile contours.

Remark 4. Models (3) and (4) may be estimated by likelihood methods if the error distributions can be fully parameterized. The estimation based on quantile regression avoids the distributional assumptions on the measurements but may not be practical if the reference data set is very small.

Remark 5. In the modeling stage, we decompose the bivariate joint distribution of (Y_1, Y_2) by a marginal distribution and a conditional distribution. Although not explored in this article, the joint distribution also can be decomposed by two marginal functions and a copula function. A link between the copula function and the conditional quantile function has been elucidated by Bouyé and Salmon (2003).

Time-Varying Coefficient Models. We now consider the problem of estimating the quantile functions g_1 and g_2. One natural approach is to assume a linear and additive relationship among $\mathbf{Y}$ and $\mathbf{X}$ but allow the coefficients to vary over time t. In other words, we write models (3) and (4) as

$$Q_\tau(Y_{i,j,1}) = \alpha_1(\tau, t_{i,j}) + \sum_{k=1}^{q} \gamma_{1k}(\tau, t_{i,j}) X_{i,j,k} \tag{6}$$

and

$$Q_\tau(Y_{i,j,2}) = \alpha_2(\tau, t_{i,j}) + \beta_2(\tau, t_{i,j}) Y_{i,j,1} + \sum_{k=1}^{q} \gamma_{2k}(\tau, t_{i,j}) X_{i,j,k}, \tag{7}$$

where α_1 and α_2 are the intercepts in the quantiles of Y_1 and Y_2, β_2 is the slope of the conditional quantile of Y_2 given Y_1, and γ_{1k} and γ_{2k} are the coefficients associated with X_k in the two models. Due to the dynamic biological process, these coefficients are set to be the functions of both time t and quantile level τ. We can view models (6) and (7) as the linear approximations of the general quantile functions g_1 and g_2. If $\mathbf{Y}_t$ and $\mathbf{X}_t$ are jointly Gaussian given t, then linearity holds immediately.

If we approximate all of the coefficient functions by splines, then models (6) and (7) are known as varying-coefficient quantile regression models; these have been studied by Kim (2007) and Cai and Xu (2007). Here we follow this approach, which

offers flexibility in two directions: The nonparametric approximations to the time-varying coefficients capture the underlying global dependence structure among the measurements, whereas the regression quantiles, estimated separately at individual quantile levels, reflect the underlying data distribution without restriction to a narrow family of distributions. If we estimate the intercept functions α_1 and α_2 nonparametrically, while using polynomial functions to model the rest of the coefficients, then the models are reduced to the semiparametric (or partly linear) quantile regression models, as discussed by He, Zhu, and Fung (2002) and Wei and He (2006). Moreover, if we have sufficient prior knowledge to fully parameterize the coefficient functions, then models (6) and (7) are simply linear or nonlinear quantile regression models.

Linear conditional quantile functions is the major assumption in (6) and (7). If this assumption is not satisfied, then we may find a monotone transformation on the measurements toward linearity (see Mu and He 2007 on power transformation) or adopt a more general model to relax the linearity assumption.

2.4 Model Selection and Combination

In the preceding sections, we estimated the joint distribution of $(Y_{t,1}, Y_{t,2})$ by modeling the marginal distribution of $Y_{t,1}$ and then the conditional distribution of $Y_{t,1}|Y_{t,2}$. Switching the order of $Y_{t,1}$ and $Y_{t,2}$, we have another estimate of the joint distribution of $(Y_{t,1}, Y_{t,2})$. One naturally asks which order is to be used. In some applications, there is a natural ordering between the two measurements such that one of them can be viewed as demonstrating the primary character of the growth process. Using human growth, for example, although body size should be characterized by both height and weight, growth in height is more closely associated with the underlying biological growth mechanism and is generally stable and monotone. Problems related to growth in height might be directly due to the deficiency of growth hormones. Growth of weight, in contrast, can be more influenced by external factors, such as environment and nutrition. However, the conditional quantile rank of weight given height measures the degree of body fatness, another important indictor of growth/health status. Therefore, to model the age-dependent joint distribution of height and weight, a natural approach is to construct the marginal model of height first, followed by the conditional model of weight given height. We call this the *primary order*. Using weight first is considered the *secondary order*. We use this approach in a later application. In the applications where such natural ordering is ambiguous, model goodness of fit may be used; whichever approach fits the data better can be considered the primary order.

To assess the goodness of fit, we may compare the empirical distribution and the estimated distribution at a series of time points and combinations of covariate values. Specifically, let $(y_{i,1}, y_{i,2})$, $i = 1, \ldots, m$, be a subsample retrieved at (near) the given time t and X; then the empirical distribution of $\mathbf{Y}$ given t and X is approximately

$$\hat{F}(y_1, y_2) = \frac{1}{m}\sum_{j=1}^{m} \mathbf{I}_{\{y_{j,1}\le y_1; y_{j,2}\le y_2\}}. \tag{8}$$

On the other hand, let $\hat{g}_{n1}(\tau_1; t, X)$ be the estimated τ_1th regression quantile of Y_1 from model (3) and let $\hat{g}_{n2}(\tau_2|\tau_1; t, X)$ be the estimated τ_2th quantile of Y_2 from model (4) given that $Y_1 = \hat{g}_{n1}(\tau_1)$. Then the estimated distribution of $\mathbf{Y}$ given t and X is

$$F_n^*(y_1, y_2) = \int_0^{\hat{g}_{n1}^-(y_1)} \hat{g}_{n2}^-(y_2, v)\, dv, \tag{9}$$

where $\hat{g}_{n1}^-(y_1) = \inf_\tau\{\tau \in (0,1) : \hat{g}_{n1}(\tau; t, X) \ge y_1\}$ and $\hat{g}_{n2}^-(y_2, v) = \inf_\tau\{\tau \in (0,1) : \hat{g}_{n2}(\tau|v; t, X) \ge y_2\}$. Equation (9) can be verified by an application of Lemma 1. If the model fit the data well, then we would expect a good agreement between the empirical and model-estimated distributions at any t and X. Let $\hat{F}_{i,j}$ and $F_{i,j}^*$ be the empirical and estimated distributions evaluated on a prespecified grid $\Lambda = \{(y_1^{(i)}, y_2^{(j)}) : i, j = 1, \ldots, K\}$. We can plot $F_{i,j}^*$ against $\hat{F}_{i,j}$, as in a P–P plot, to get a quick idea of how well the two distributions match. Here Λ can be the subsample itself if it is of small size, or a set of representative points if the sample size is large.

Sometimes both primary and secondary orders yield good approximations to the underlying distribution, in which case it is advantageous to simulate data in both orders and take the combined data for contour estimation. The combined data provide a more complete coverage of the sample space, which in turn yields more stable contour estimation. If one order provides a much better goodness of fit, then it should be used to generate data for contour estimation. In summary, we have three options: (a) Use the primary order if there exists the natural ordering or if one order fits the data much better than the other, (b) combine the data generated from both orders when either one provides decent fit to the data, or (c) refine the models if neither order leads to a clear lack of fit to the data.

2.5 Simulation Study

In this section we present a simulation study to demonstrate the performance of the proposed methodology. Two models are considered:

Model S1:

$$Y_1 = 40t/(1+2t) + e_1, \qquad e_1 \sim \mathrm{N}(0,1);$$
$$Y_2 = \ln(1+t) + (1+.2t)Y_1 + e_2, \qquad e_2 \sim \mathrm{N}(0,1)$$

Model S2:

$$Y_1 = 40t/(1+2t) + (1+.2t)e_1,$$
$$e_1 \sim \mathrm{SN}(0,1,5);$$
$$Y_2 = \ln(1+t) + (1+.2t)Y_1 + (1+0.1Y_1)e_2,$$
$$e_2 \sim \mathrm{N}(0,1).$$

Here SN is the skew-normal distribution (Azzalini and Capitanio 2003). The two models have the same structure but very different error distributions. In both settings the marginal mean curves of Y_1 and Y_2 are $40t/(1+2t)$ and $\ln(1+t) + 40t(1+.2t)/(1+2t)$, which resemble the paths of human growth trajectories on the log scale. In both models we also assume that the quantiles of Y_2 are linear functions of Y_1 with time-varying intercept and slope functions. But the errors e_1 and e_2 in model S1 are iid standard normal, so that (Y_1, Y_2) is jointly Gaussian

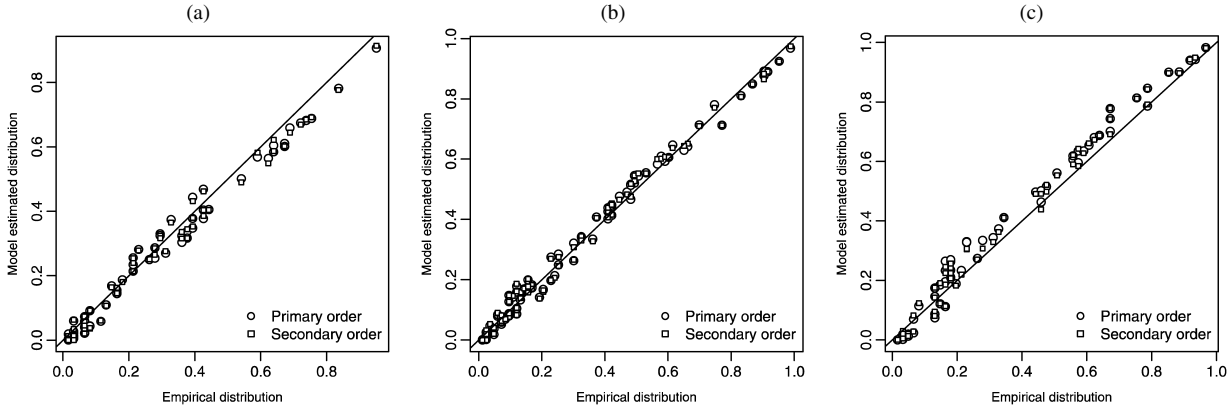

Figure 1. Diagnostic P–P plot for the goodness of fit of model S2 at times .5 (a), 1 (b), and 2 (c). The reference line is $y = x$.

with mean $(40t/(1+2t), \ln(1+t) + (40t(1+.2t))/(1+2t))$ and variance–covariance matrix

$$\begin{pmatrix} 1 & 1+.2t \\ 1+.2t & (1+.2t)^2+1 \end{pmatrix}.$$

More "sophisticated" error distributions are used in model S2, where e_2 remains standard normal but e_1 follows a skewed normal distribution with mean 0, variance 1, and shape parameter 5 (skewness = .85). Moreover, the inclusion of the multipliers $(1 + .2t)$ and $(1 + .1Y_1)$ to the errors forces the standard deviation of Y_1 to increase linearly with time t and that of Y_2 to increase linearly with Y_1. The skewed distribution and heteroscedastic errors are common characteristics of the growth data.

To generate a Monte Carlo sample from models S1 and S2, we first draw a random t from the uniform distribution on the interval (0, 3), and then generate Y_1 and Y_2 sequentially according to the given models. The sample size used is 5,000. In the first stage of model estimation, we fit the entire data set with the stratified coefficient-varying quantile regression models $Q_\tau(Y_1(t)) = a_\tau(t)$ and $Q_\tau(Y_2(t)) = b_\tau(t) + c_\tau(t)Y_1(t)$, where $a_\tau(t)$, $b_\tau(t)$, and $c_\tau(t)$ are estimated by cubic regression B-splines with evenly spaced internal knots (.5, 1, 1.5, 2, 2.5). Instead of estimating the entire quantile process on $\tau \in (0, 1)$, we estimate the model on a evenly spaced grid of (0, 1), denoted as Γ. The number of grids is 300. This model estimation procedure is used on both orders of modeling Y_1 and Y_2.

To assess the goodness of fit at various time points, we retrieve a subsample by drawing narrow windows ($\pm$ .02) around the selected times .5, 1, and 2. The subsample size is around 60 at each time point. We evaluate at each selected time point the empirical distribution and model-estimated distribution based on equations (8) and (9). To save space, Figure 1 presents only the corresponding P–P plots under model S2. The points are all distributed along the reference line $y = x$, suggesting a good match between the empirical and model-based distributions. Because both orders provide good fit to the data, we combine them in the second stage for estimation of the reference quantile contours; that is, we generate random samples of (Y_1^*, Y_2^*) at times $t = .5$, 1, and 2, from the estimated models under both orders (Y_1 first or Y_2 first), and combine them at each time point. With the simulated sample size of the combined data at 10,000, we construct the .5th, .75th, and .95th reference quantile contours at each time point following the algorithm proposed in Section 2.2. Specifically, we use the constrained quadratic smoothing B-spline with 10 evenly spaced internal knots over $(0, 2\pi)$. The resulting reference quantiles contours are displayed in Figures 2 and 3, overlaid with the underlying

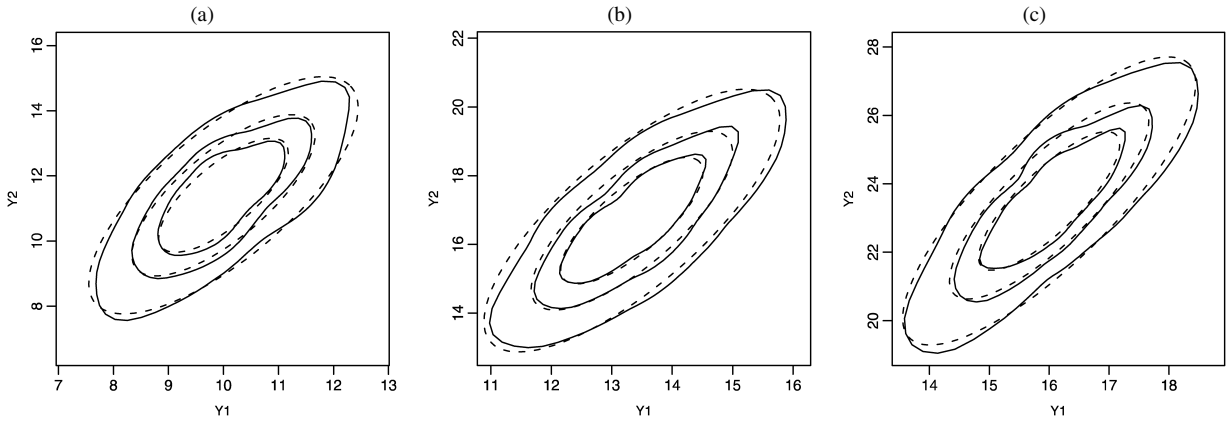

Figure 2. Estimated reference quantile contours under model S1 at times .5 (a), 1 (b), and 2 (c) and quantile levels .5, .75, and .95. The dotted contours are the true τ 100% confidence ellipse.

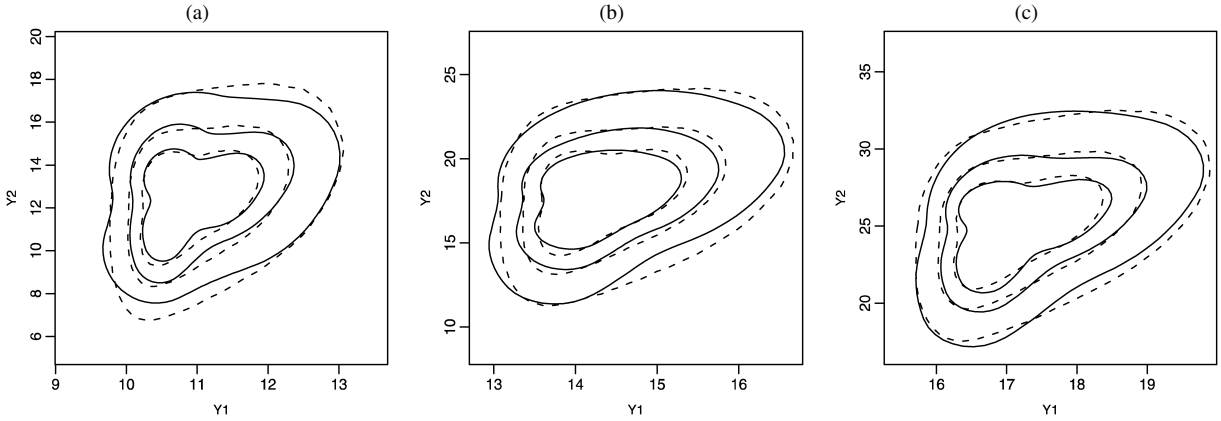

Figure 3. Estimated reference quantile contours under model S2 at times .5 (a), 1 (b), and 2 (c) and quantile levels .5, .75, and .95. The dotted contours are the true reference quantile contour under Definition 1.

theoretical contours. We can see that the estimated quantile contours agree with the true ones fairly well under both models at different time points and different quantile levels.

In model S1, Y_1 and Y_2 are jointly normal, and thus the linear time-varying coefficient models hold for both primary and secondary orders. However, in model S2, due to the skewness of e_1, the quantiles of Y_1 are no longer linear in Y_2, but we still find that the linear model closely approximates the quantile contours. In other words, the conditional model in the secondary order is misspecified, but this does not automatically mean that the secondary order leads to poor contour estimates. To further study the impact of misspecification, we altered the distribution of e_1 to increase its skewness, only to find that the quality of the secondary order in contour estimation decreases with the skewness at a slow rate. Figure 4 shows the difference in goodness of fit at $t = 2$ between the primary order (a correctly specified model here) and the secondary order (a misspecified model) when $e_1 = Z^4$, where Z is standard normal. In this case, the skewness of e_1 reaches 10, and the diagnostic plots are now able to differentiate between the two orders; the primary order maintains goodness of fit, but the secondary order begins to falter. Our empirical experience has shown that the estimated contours from combined data (generated under both orders) tend to be preferable unless the secondary order provides a clearly inferior fit.

3. APPLICATION: BIVARIATE LONGITUDINAL GROWTH CHARTS

To demonstrate the performance and practical value of the proposed methodology, in this section we present an application to constructing age-dependent, bivariate pediatric growth charts of height and weight. Univariate growth charts of height and weight are widely used to monitor the growth status of infants and children. The basic idea is to plot a series of reference quantile curves that depict the time-varying distribution of a certain growth measurement within a reference population and thus can be used to locate an individual subject's quantile rank among the population. A quantile rank appearing to be extreme (either too high or too low) may indicate a certain abnormality in growth status. Early work on reference growth charts relied heavily on parametric distributional assumptions, but recent research has emphasized the need for more flexible models. Various methods have been proposed for constructing growth charts (see Cole and Green 1992; Thompson and Fatti 1997; Wei et al. 2006; and references therein), but these methods generally are aimed at screening one measurement at a time.

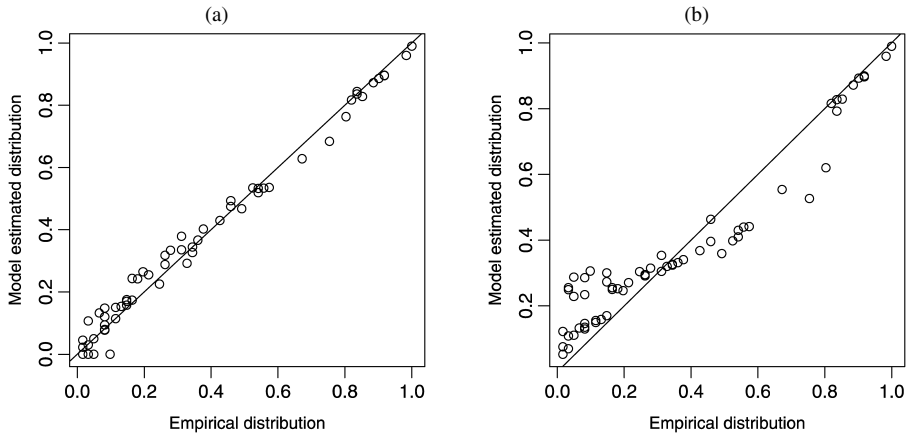

Figure 4. Comparison of goodness of fit between the primary (a) and secondary (b) modeling orders under model misspecification ($t = 2$).

Because growth processes are usually dynamic and sophisticated, a single parameter is rarely adequate to describe their entire evolution. With each parameter revealing some information, overall growth/health status should be assessed by considering all of the measurements together. Because univariate charts fail to take measurement correlations into account, abnormality in growth may be manifested jointly but be invisible marginally. Pediatricians have expressed their preference for screening height and weight simultaneously (Pere 2006). One natural way to do this is to construct a nested sequence of reference quantile contours of height and weight, based on which unusual and extreme cases can be ranked and abnormal growth detected. On the other hand, because the distributions of the growth parameters, as well as their associations, are generally time-dependent, screening is unlikely to be effective without incorporating the time effect into the charts. Often we wish to incorporate other potentially important covariates, such as cohort effect, parental information, or an individual's prior growth path, to provide more informative screening. In other words, in principles, a bivariate growth chart should consist of dynamic age-dependent and covariate-varying quantile contours.

When $p = 1$, the defined reference quantile function C_τ reduces to an interval $[F^{-1}(\frac{1-\tau}{2}), F^{-1}(\frac{1+\tau}{2})]$, which is considered the normal range for univariate screening. A univariate growth chart can be viewed as a nested sequence of normal ranges. Being the multivariate analog of the normal range, the defined reference quantile contour is particularly applicable for constructing bivariate growth charts. Moreover, biological variables, including height and weight, rarely follow the multivariate normal distributions. It is known that height and weight are not normally distributed, and their joint distribution is even less likely to be normal. Hence using a Gaussian model may lead to nonignorable bias, and joint transformation is usually difficult (Wei and He 2006). But distributional assumptions play an important role in growth chart construction, because the extreme quantiles are of interest here; thus relaxing the distributional assumptions for accurate estimation is desirable. Finally, we note that the measurement times in most growth data typically are irregular. Consequently, the number of observations at a fixed time may be small even in a large reference data set. This makes it difficult to estimate bivariate quantile functions at a single time t, even more so when we wish to adjust for potentially important covariates.

Based on the foregoing discussion, we believe that the proposed methodology can be naturally applied to construct bivariate conditional growth charts, because it does not rely on distributional assumptions and has nice adaptivity to unequally spaced time series for longitudinal data. The concept of constructing bivariate growth charts based on quantile contours has been explored by Fatti, Senaoana, and Thompson (1998) and Petersen (2003); however, these methods either assume multivariate normal distribution nor assume homogeneous association along the time.

We use the data from the 1988 National Maternal and Infant Health Survey (NMIHS) and its 1991 Longitudinal Follow-Up (Sanderson, Placek, and Keppel 1991), which selected a nationally representative sample of subjects born during calendar year 1988 in the United States. There are 5,065 subjects in total (2,555 boys and 2,510 girls). Their body size measurements were obtained at birth and on each visit to the hospital or to other pediatric care provider up to age 4 years. The median number of visits for each subject is six; however, the actual measurement time intervals, due to different individual needs, are irregular across the subjects. In addition to the body size measurements, the data also include parental information, such as their heights and socioeconomic status, and the child's health status on each visit. In this article, we call this data set the NMIHS data and focus on boys only.

3.1 Bivariate Charts of Height and Weight

The simplest bivariate joint charts consist of the reference quantile contours that depend solely on time (age). In this section we construct such a joint chart of height and weight from age 0 to 2 years based on the NMIHS data. Let (W_t, H_t) be a pair of height and weight at age t; we model them by the following stratified quantile regression models, a marginal model of height and a conditional model of weight given height, the primary order described in Section 2.4:

$$\begin{aligned} Q_\tau(\ln(H_t)) &= a_\tau(t), \\ Q_\tau(\ln(W_t)) &= b_\tau(t) + c_\tau(t)\ln(H_t). \end{aligned} \tag{10}$$

Model (10) uses height H_t and weight W_t on the log scale, which is more consistent with the general belief that weight tends to change linearly with a power of height. Because the reference growth charts must be constructed based on normal subjects, we excluded from model estimation any subject who was clinically designated as low birth weight ($\leq$ 2,500 g), had an overall health status rated as "poor," or was considered as abnormal based on the developmental screening. The remaining boys serve as the reference population. The sample size (the number of height–weight measurements) for estimation was 12,122. We used cubic B-splines with internal knots (.25, .5, 1.0, 1.5, 2.0, 3.0) to approximate the coefficient functions at a given τ, and estimated the model on a grid of 200 evenly spaced $\tau \in (0, 1)$.

Because the success of the joint charts depends on how well model (10) captures the underlying age-varying distribution of height and weight, we assessed the goodness of fit of the model before proceeding to the second stage. The diagnostic plots comparing the empirical distributions with the model-estimated distributions at age .5, 1, and 2 years are shown in Figure 5. The subsamples used at each time point were within ($\pm$0.02 year) of the target age. The subsample sizes were 677, 560, and 348. Although not perfect, the overall goodness of fit at those time points indicates that model (10) is reasonable.

In the second stage of contour estimation, we simulated 10,000 pairs of height and weight from the estimated models (10) at age .5, 1, and 2 years and used these to construct the reference quantile contours at quantile levels .5, .8, and .9. (A quantile level of .9 in the bivariate chart corresponds to the 5th and the 95th quantiles in a univariate chart.) Following the algorithm in Section 2.2 and using constrained quadratic B-splines with eight evenly spaced internal knots over $(0, 2\pi)$, we obtained the quantile contours as displayed in Figure 6, overlaid with the subsample around each time point. Based on Figure 6, we say that the distribution of height and weight indeed changes over age in both location and scale. To further quantify the quality of the contours, Table 1 gives the sample coverage $\hat{p}$ of each

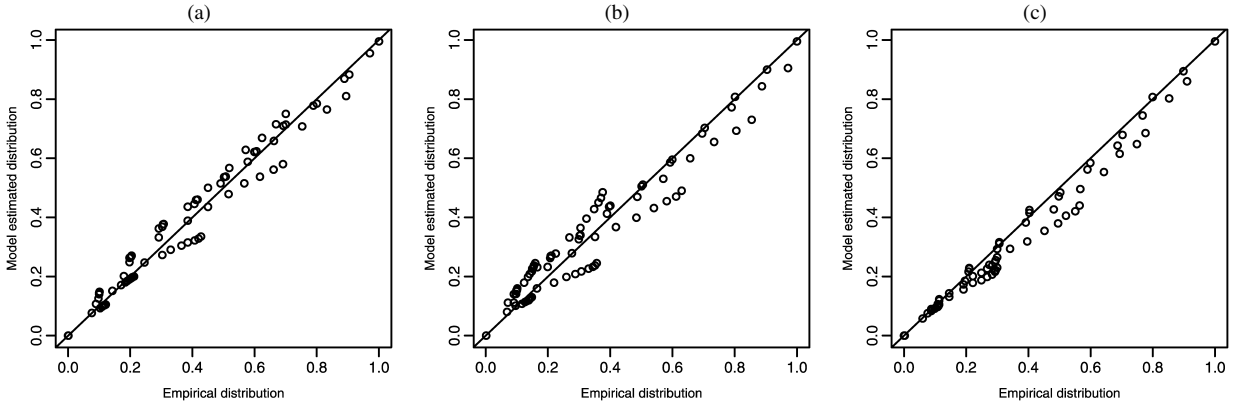

Figure 5. Diagnostic P–P plot for the goodness of fit of model (10) at age .5 (a), 1 (b), and 2 (c) years. The reference line is $y = x$. The two distribution functions are evaluated at each pair of the $i/10$th and $j/10$th sample quantiles of height and weight; $i, j = 0, 1, \ldots, 10$.

contour based on the proportion of the subsample constrained within each contour; the 95% confidence interval for the true coverage is $\hat{p} \pm 1.96\sqrt{\hat{p}(1-\hat{p})/n_t}$, where n_t is the subsample size at time t. Based on the findings in Table 1, we see that all of the estimated reference quantile contours have the desired coverage around τ.

Remark 6. We also estimated the models of the secondary order (marginal model of weight plus the conditional model of height given weight), and constructed the reference quantile contours based on the combined data. The resulting contours (not presented herein) are nearly identical to the results presented here.

To demonstrate the screening function of the bivariate growth charts, we selected five illustrative subjects from the NMIHS data, designating them subjects A, B, C1, C2, and D. With .5 year as the target age, we plotted their height–weight measurements against the estimated reference quantile contours to locate their percentile ranks, as shown Figure 7(a). Their height–weight measurements remained outside the .9th reference quantile contour, indicating a is somewhat unusual current body size. Moreover, the directions of their outlyingness provide good information. Subject D had an unusually small body size with both height and weight in the lower quantiles. In contrast, subject B had a comparatively larger body size than most of his peers. Subjects A, C1, and C2 had normal height and weight, but subject A was overweight given his height, whereas subjects C1 and C2 appeared to be underweight given their heights.

Although the abnormality in height and weight cannot be related to any specific health hazard, further examination of these subjects' medical profiles revealed good information. Subject D suffered an acute gastroduodenal perforation at age .25, followed by inguinal hernia at age .33. His weak digestive condition may partly explain his unusually small body size. Subject B was diagnosed as obese at age 2 months; subject C1 had chronic endocarditis at age 2 months. Subject C2 had few very few entries of diagnosis records, and thus we could not evaluate his health status. Finally, subject A had hepatomegaly at age .5. These records confirm the clinical value of pediatric height–weight screening.

Comparison With Univariate Growth Charts. In the United States, the National Center for Health Statistics (NCHS) publishes univariate pediatric growth charts periodically, including weight-for-age, height-for-age, and weight-for-height charts. These charts are routinely used by pedestrians for screening. To compare the proposed bivariate growth charts to the conventional univariate growth charts, we estimate the reference quantile curves using the method of Wei et al. (2006). The results, shown in Figure 7, are similar to the NCHS charts. Figure 7(b) and (c) provide marginal charts for height and weight,

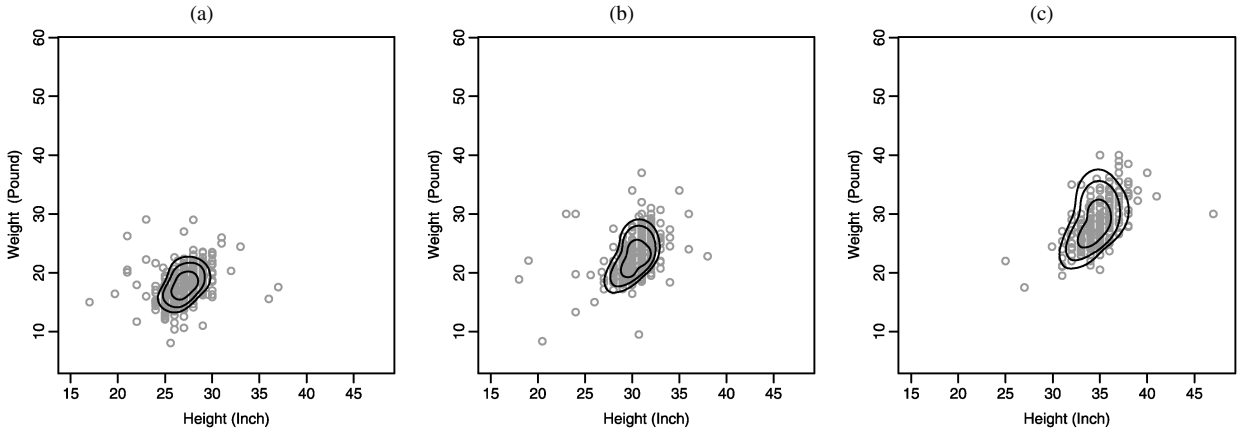

Figure 6. Estimated height–weight reference quantile contours at age .5 (a), 1 (b), and 2 (c) years and quantile levels $\tau = .5, .8$, and .9.

Table 1. Sample coverages of the estimated reference quantile contours and the 95% confidence intervals of the coverage probability at nominal quantile levels .5, .8, and .9 and age .5, 1, and 2 years

Age (year)	Nominal quantile level $\tau = .50$	$\tau = .80$	$\tau = .90$
.5 ($n = 677$)	.528 (.491, .567)	.799 (.769, .829)	.886 (.862, .910)
1 ($n = 560$)	.507 (.466, .548)	.834 (.803, .865)	.893 (.867, .919)
2 ($n = 348$)	.557 (.505, .609)	.819 (.779, .859)	.899 (.867, .931)

and 7(d) is a conditional chart of weight given height. We plotted subjects A–D against those reference curves. Clear, no individual univariate growth chart was as informative as the bivariate chart. For example, subjects A and C2 were within the 90% normal ranges of both height and weight at age .5 year, and their abnormality is visible only through the conditional analysis of weight for height. The bivariate growth chart effectively identifies the type of abnormality. Even more importantly, the conventional univariate growth charts do not adjust for covariates and so may be misleading as shown in the example of the next section.

3.2 Bivariate Conditional Charts of Height and Weight

Parental heights carry genetic information and have a nonignorable influence on their children's body size; for example, a child with short stature may not be unusual if both of his parents are short. In this section we take a further step to incorporate an individual subject's average parental height P into the joint height–weight growth charts, with the new model

$$Q_\tau(\ln(H_t)) = a_\tau(t) + d_{1\tau}(t)\ln(P), \tag{11}$$

$$Q_\tau(\ln(W_t)) = b_\tau(t) + c_\tau(t)\ln(H_t) + d_{2\tau}(t)\ln(P). \tag{12}$$

Parental height is incorporated into the charts by assuming the age-varying multiplicative contributions to the child's height and weight. Following the same procedure, we can construct the reference quantile contours given each individual subject's age and average parental height. Using age 2 years as an example, Figure 8 displays the reference quantile contours for children with tall parents (average height, 72 inches) and children with short parents (average height, 63 inches). Both are compared with the height–weight contours without adjusting for parental height. According to Figure 8(a), the reference quantile contours of the children of tall parents shift toward the higher end, with the obvious implication that the children of tall parents generally are of larger body size than the normal cohort. In contrast, the children of short parents are close to the

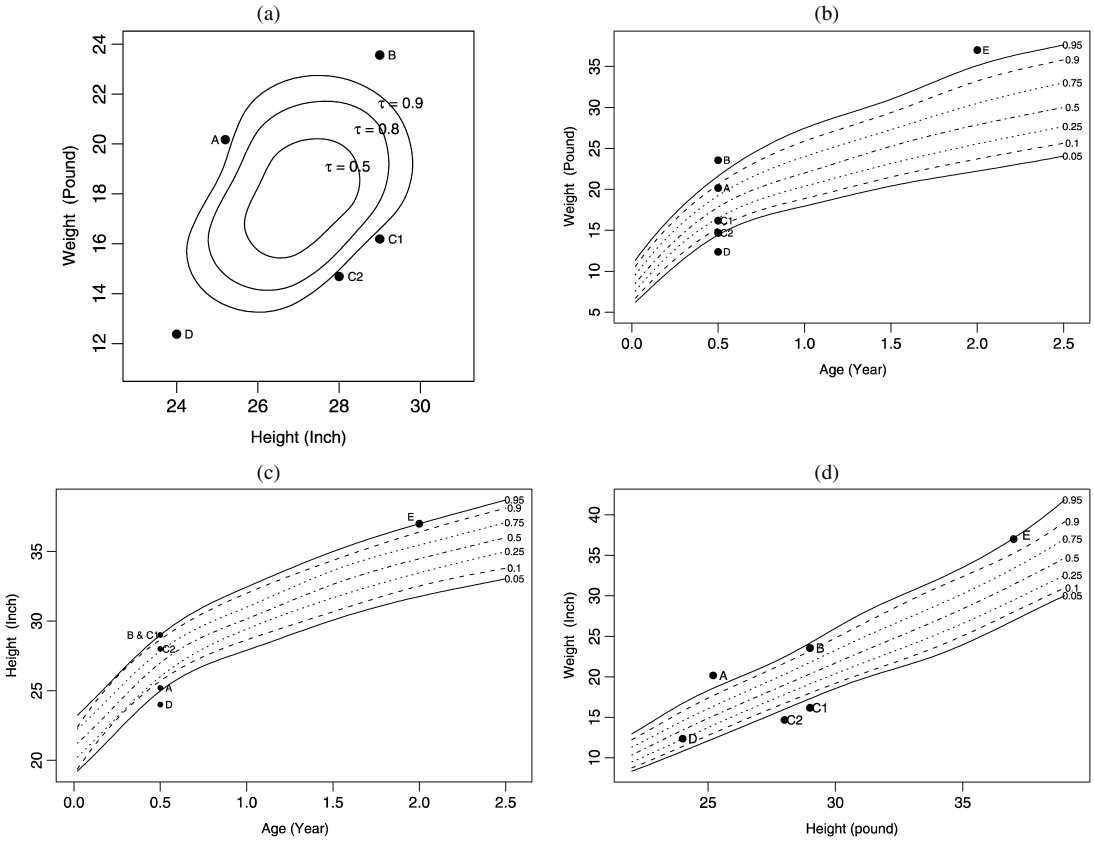

Figure 7. Screening subjects A–E based on bivariate screening and conventional univariate growth charts. (a) Bivariate reference contours of height and weight at age .5. From inside to outside, the quantile levels are .5, .8, and .9. Marginal height (b) and weight (c) charts for the age range 0–2.5 years. (d) Conditional weight-for-height chart for the height range 22–39 inches. In each chart, the reference quantile levels are .05, .10, .25, .50, .75, .90, and .95.

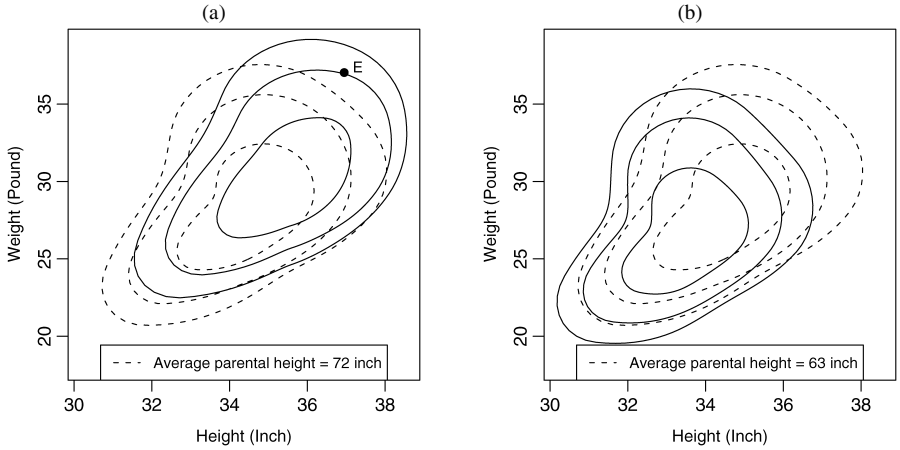

Figure 8. Comparison of reference quantile contours with and without considering parental heights. (a) Tall parents. (b) Short parents. The dotted contours in both plots are the reference quantile contours at age 2 without considering parental heights, whereas the solid ones are their conditional version with the inclusion of parental heights. The solid dot in (a) is subject E selected from NMIHS data, with average parental height of 72 inches.

marginal contours at the lower end but differ more markedly at the higher end. This finding is interesting, because it suggests that the lower extremes of the body sizes of the children of short parents are actually comparable to those of the general population, but that children with short parents are much less likely to be tall and heavy. For example, subject E, chosen from the NMIHS data for illustration, had a weight of 37 pound and a height 37 inches at age 2. Based on the bivariate chart of height and weight (the dotted contours in Fig. 8), his body size appears to be overly large, clearly lying outside the 90% contour. However, subject E has tall parents with an average parental height of 72 inches. Based on the conditional joint charts [the solid contours in Fig. 8(a)], his percentile rank given his parental heights is about 80%. Therefore, we have reason to believe that this subject's unusual body size is likely due to the "tall" gene(s) that he inherited from his parents. To complete the comparison, we also plot the height and weight of subject E against the three univariate growth charts in Figure 7. Because those conventional univariate charts do not include the effect of parental height, subject E would be considered overweight, clearly heavier than 95% of his peers.

To address the natural question of how well the two sets of conditional contours in Figure 8 fit the data, we also access the goodness of fit of model (11) at age 2 and parental heights 72 and 63 inches using the same diagnostic tool; that is, we estimate the empirical distribution of a subsample that is similar to the target case and compare it with the corresponding model-estimated distribution. The resulting P–P plots are presented in Figure 9. The empirical distribution in Figure 9(a) is calculated based on heights and weights at age 2 years from 48 subjects whose average parental height is around 72 inches ($\pm$ 2 inches), whereas the empirical distribution in Figure 9(b) is from 39 2-year-old subjects whose average parental height is around 63 inches ($\pm$ 2 inches). Based on Figure 9, we can conclude that model (11) fits the data reasonably well in both cases.

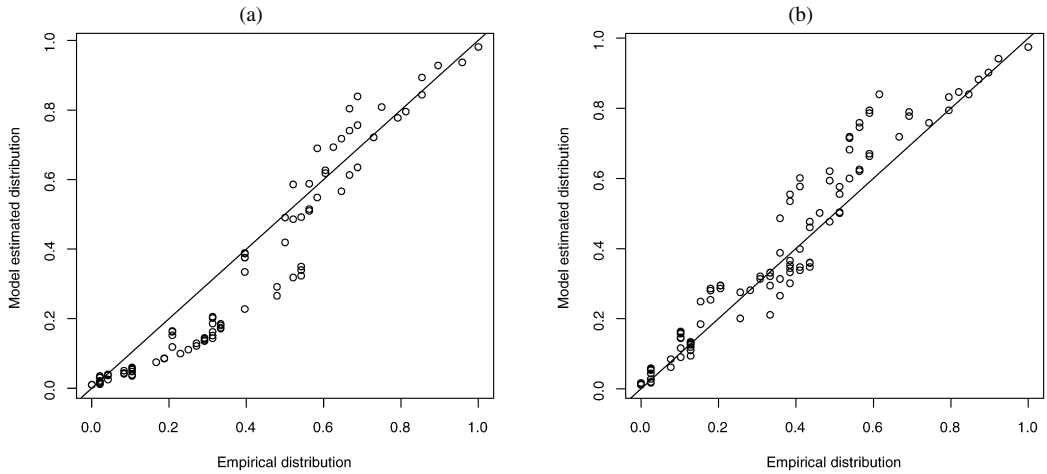

Figure 9. Diagnostic P–P plot for the goodness of fit of model (11) at age 2 and average parental heights of 72 and 63 inches. (a) Tall parents. (b) Short parents. The reference line is $y = x$. The two distribution functions are evaluated at each pair of the $i/10$th and $j/10$th sample quantiles of height and weight; $i, j = 0, 1, \ldots, 10$.

4. CONCLUSION AND DISCUSSION

In this article we have proposed methodologies for constructing conditional bivariate quantile functions and have applied them for height–weight screening of young children in the United States based on the NMIHS data. Although there are many proposals regarding bivariate quantile contours, incorporating covariate effects is usually difficult. The common assumption of multivariate normality is unsound, as pointed out by Wei and He (2006) or demonstrated when the individual measurements are taken irregularly to allow convenient modeling of a multivariate vector that may include previous measurements of the same subject. The strategy of stratified quantile modeling and estimation of contours based on simulated data from such models has proven useful. This method is sufficiently flexible to incorporate dynamic time and covariate structure without making strong distributional assumptions.

General Applications. Growth charts (also known as reference centile charts) have been used in a wide range of applications, including monitoring CD4 and CD8 cell counts in children in human immunodeficiency virus studies (Aldhous et al. 1994), monitoring obesity for cancer prevention, and monitoring tumor markers in cancer treatment (Oesterling et al. 1993). Although illustrated by pediatric growth charts, the proposed methodology is readily applicable to other reference contours, such as in multivariate tolerance regions used in quality control charts. The linear quantile models that work well in height–weight charts can be modified to suit other applications. The proposed diagnosis tool may be used to guide the selection of alternative models.

Extension to Multivariate Quantiles. The two-stage modeling and estimation procedure can be extended to multivariate contexts. We start with a marginal model for $Y_1(t)$, followed by a conditional model for $Y_2(t)$ given $Y_1(t)$ and t, then a conditional model for $Y_3(t)$ given $Y_1(t)$, $Y_2(t)$, and t, and so on. The optimum order of this conditional sequence awaits further research, but subject matter knowledge may provide a primary order, and the diagnostic tools introduced in this article can be developed further. We do note that when the dimension exceeds three, the visualization of reference quantile contours becomes difficult, and the computational burden increases, and so a dimension-reduction technique may be helpful. Bivariate growth charts of the first two principal components of $\mathbf{Y}_t$ could be used. Additional research is warranted.

APPENDIX A: PROOF OF PROPOSITION 1

Properties (1) and (3) follow directly from the definition of the contours, so we show only (2) and (4) here. Without loss of generality, we assume in the proof that $\boldsymbol{\mu} = (0, 0)$ and $S = I_2$. Let $\mathbf{u}(\theta) = (\cos\theta, \sin\theta)$ and $l(\theta) = \{\mathbf{x} : \mathbf{x} \overset{(\mathbf{u}(\theta))}{=} \mathbf{u}(\theta)\}$ be the subset containing all of the points on the line connecting $\mathbf{u}(\theta)$ and $-\mathbf{u}(\theta)$. Clearly, we have $\mathbb{R}^2 = \bigcup_{\theta\in[0,2\pi)} l(\theta)$. For any $\mathbf{y} \in \mathbb{R}^2$, we denote by $\mathbf{d}(\boldsymbol{y}, \theta) = \langle \mathbf{y}, \mathbf{u}(\theta)\rangle$ the distance of $\mathbf{y}$ to the origin with respect to $\mathbf{u}(\theta)$. Let $F_\theta(d) = \mathrm{P}\{\mathbf{d}(\mathbf{Y}, \theta) \le d | \mathbf{Y} \in l(\theta)\}$ be the conditional distribution function of $\mathbf{d}(\mathbf{Y}, \theta)$ given $\mathbf{Y}$ on the line of $l(\theta)$. Because $\mathbf{Y}$ has a continuous and positive density over $\mathbb{R}^2$, $\mathbf{d}(\mathbf{y}, \theta)$ is a continuous transformation, and $l(\theta)$ is a connected support, $F_\theta(d)$ clearly is a continuous function of d. Therefore, along the direction indexed by θ, the τth quantile contour of $\mathbf{Y}$ can be written as

$$C_\tau = g_\tau(\theta) = F_\theta^{-1}\left(\frac{1+\tau}{2}\right), \qquad \theta \in [0, 2\pi),$$

where F_θ^{-1} is the inverse function of F_θ, or the conditional quantile of $\mathbf{d}(\mathbf{Y}, \theta)$ given $\mathbf{Y} \in l(\theta)$. Following lemma 8.3.1 of Resnick (1998), a sufficient condition for the continuity of $g_\tau(\theta)$ is that for any $\theta_0 \in [0, 2\pi]$ and $d \in \mathbb{R}$, $F_{\theta_n}(d) \to F_{\theta_0}(d)$ as $|\theta_n - \theta_0| \to 0$, which holds under our assumption that $\mathbf{Y}$ has a continuous and positive density over $\mathbb{R}^2$. Thus the proof of property (2) is hence completed.

Property (4) follows easily from the facts that the median is equivariant to location-scale transformations and that MAD is invariant to location shift and equivariant to scale transformation.

APPENDIX B: AN ADDITIONAL EXAMPLE

In this appendix we demonstrate by example that the quantile contours given in Section 2.2 should be applied with caution if the support of the variable Y is restricted to certain subset of R^2. For this purpose, we consider an artificially generated banana-shaped dataset. In particular, we generated 200 sample points (x, y) from the following model:

$$x = r * \cos\left(\frac{3\pi}{4}s\right) \qquad \text{and} \qquad y = r * \sin\left(\frac{3\pi}{4}s\right),$$
$$\text{where } r \sim \mathrm{N}(3, .2) \text{ and } s \sim \mathrm{Unif}[0, 1]. \quad \text{(B.1)}$$

As illustrated in Figure B.1, the data are banana-shaped. Following the algorithm of Section 2.2, we converted the data into dual-polar coordinates using (0, 2) as the center of the data. We then estimated the .9th reference quantile contour using (COBS) with 22 evenly spaced internal knots and a periodicity constraint. The resulting quantile contour, shown in Figure B.1(b), exhibits an interesting crossing pattern at the center (0, 2). To understand why this crossing pattern occurs, Figure B.1(c) plots the dual-polar coordinates of the data, $(r_i^{(k)}, \theta_i^{(l)})_{k=1,2}$, overlaid with the estimated quantile curve from COBS. We notice that the true $(1 + .9)/2$th conditional quantile function of $r_i^{(k)}$ given $\theta_i^{(k)}$ is discontinuous at around π and 1.8π, but that COBS enforces a smooth function. The discontinuity is due to the fact that at a certain range of θ, the center stays out of the directional reference intervals. If we remove part the of curves near the discontinuity points, then we find that the remaining two segments of the curve make sense—one segment tracking the upper tail of the banana-shaped data and the other tracking the lower tail, as illustrated in Figure B.1(d). Connecting these two segments by the dotted lines Figure B.1(d), we obtain a reasonably good contour, capturing the shape of the data. Although such banana-shaped data are unlikely to occur in growth chart applications, this experiment suggests that further modification of our algorithm given in Section 2.2 may be necessary for general applications in which a natural center of the distribution is subject to debate and the quantile contour curves have discontinuities as a function of θ.

[Received May 2007. Revised September 2007.]

REFERENCES

Abdous, B., and Theodorescu, R. (1992), "Note on the Spatial Quantile of a Random Vector," *Statistics and Probability Letters*, 13, 333–336.

Aldhous, M. C., Raab, G. M., Doherty, K. V., Mok, J. Y. Q., Bird, A. G., and Froebel, K. S. (1994), "Age-Related Ranges of Memory, Activation, and Cytotoxic Markers on CD4 and CD8 Cells in Children," *Journal of Clinical Immunology*, 14, 289–298.

Azzalini, A., and Capitanio, A. (2003), "Distributions Generated by Perturbation of Symmetry With Emphasis on a Multivariate Skew-t Distribution," *Journal of the Royal Statistical Society*, Ser. B, 65, 367–389.

Bouyé, E., and Salmon, M. (2003), "Dynamic Copula Quantile Regressions and Tail Area Dynamic Dependence in Forex Markets," preprint.

Cai, Z., and Xu, X. (2007), "Nonparametric Quantile Regression for Dynamic Smooth Coefficient Models," preprint.

Chakraborty, B. (2003), "On Multivariate Quantile Regression," *Journal of Statistical Planning and Inference*, 110, 109–132.

Chaudhuri, P. (1996), "On a Geometric Notion of Quantiles for Multivariate Data," *Journal of the American Statistical Association*, 91, 862–972.

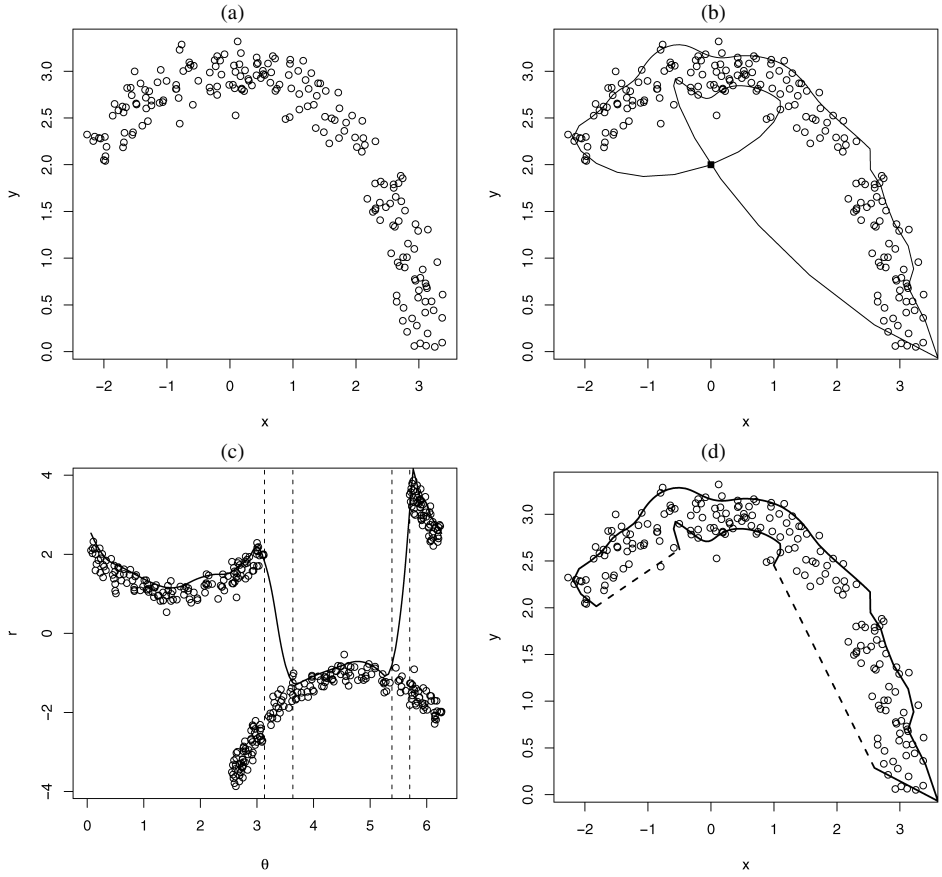

Figure B.1. An additional experiment on banana-shaped data. (a) Scatterplot of the banana-shaped data. (b) The estimated .9th reference quantile contour using the algorithm in Section 2.2. The gray square point is the designated "center" at (0, 2). (c) The dual polar coordinates of the banana-shaped data, overlayed with the estimated $(1 + .9)/2$th quantile curve using COBS. The two sets of dotted lines define the areas where the estimated quantile curves do not fit the data. (d) The remaining contour after removal of the poorly fitted segments due to discontinuity.

Cole, T. J., and Green, P. J. (1992), "Smoothing Reference Centile Curves: The LMS Method and Penalized Likelihood," *Statistics in Medicine*, 11, 1305–1319.

Fatti, L. P., Senaoana, E. M., and Thompson, M. L. (1998), "Bayesian Updating in Reference Centile Charts," *Journal of the Royal Statistical Society*, Ser. A, 161, 103–115.

Fraiman, R., and Meloehe, J. (1999), "Multivariate L–Estimation" (with discussion), *Test*, 8, 255–317.

He, X., and Ng, P. (1999), "COBS: Qualitatively Constrained Smoothing via Linear Programming," *Computational Statistics*, 14, 315–337.

He, X., Zhu, Z., and Fung, W. (2002), "Estimation in a Semiparametric Model for Longitudinal Data With Unspecified Dependence Structure," *Biometrika*, 89, 579–590.

Hettmansperger, T. P., Nyblom, J., and Oja, H. (1992), "On Multivariate Notions of Sign and Rank," in L_1-*Statistical Analysis and Related Methods*, ed. Y. Dodge, Amsterdam: North-Holland, pp. 267–278.

Kim, M. (2007), "Quantile Regression With Varying Coefficients," *The Annals of Statistics*, 35, 92–108.

Koenker, R., and Bassett, G. J. (1978), "Regression Quantiles," *Econometrica*, 46, 33–50.

Koltchinskii, V. (1997), "M–Estimation, Convexity and Quantiles," *The Annals of Statistics*, 25, 435–477.

Liu, R. Y., Parelius, J. M., and Singh, K. (1999), "Multivariate Analysis by Data Depth: Descriptive Statistics, Graphics and Inference" (with discussion), *The Annals of Statistics*, 27, 783–858.

Ma, L., and Koenker, R. (2006), "Quantile Regression Methods for Recursive Structural Equation Models," *Journal of Econometrics*, 134, 471–506.

Mardia, K. V., Kent, J. T., and Bibby, J. M. (1979), *Multivariate Analysis*, New York: Academic Press.

McDermott, J. P., and Lin, D. K. J. (2007), "Quantile Contours and Multivariate Density Estimation for Massive Datasets via Sequantile Convex Hull Peeling," *IIE Transactions*, 39, 581–591.

Mu, Y., and He, X. (2007), "Power Transformation Towards a Linear Regression Quantile," *Journal of the American Statistical Association*, 102, 269–279.

Oesterling, J., Jacobsen, S. J., Chute, C. G., Guess, H. A., Girman, C. J., Panser, L. A., and Lieber, M. M. (1993), "Serum Prostate-Specific Antigen in a Community-Based Population of Healthy Men; Establishment of Age-Specific Reference Ranges," *Journal of the American Medical Association*, 270, 860–864.

Parzen, E. (1979), "Nonparametric Statistical Data Modeling," *Journal of the American Statistical Association*, 74, 105–131.

Pere, A. (2006), Comment on "Conditional Growth Charts," by Y. Wei and X. He, *The Annals of Statistics*, 34, 2105–2112.

Petersen, J. H. (2003), "Two Bivariate Geometrically Defined Reference Regions With Application to Male Reproductive Hormones and Human Growth," *Statistics in Medicine*, 22, 2603–2618.

Resnick, S. I. (1998), *A Probability Path*, Boston: Birkhauser.

Sanderson, M., Placek, P. J., and Keppel, K. G. (1991), "The 1988 National Maternal and Infant Health Survey: Design, Content, and Data Availability," *Birth*, 18, 26–32.

Serfling, R. (2002), "Quantile Functions for Multivariate Analysis: Approaches and Applications," *Statistica Neerlandica*, 56, 214–232.

Thompson, M. L., and Fatti, L. P. (1997), "Construction of Multivariate Centile Charts for Longitudinal Measurements," *Statistics in Medicine*, 16, 333–345.

Wei, Y., and He, X. (2006), "Conditional Growth Charts" (with discussion), *The Annals of Statistics*, 34, 2069–2097.

Wei, Y., Pere, A., Koenker, R., and He, X. (2006), "Quantile Regression Methods for Reference Growth Charts," *Statistics in Medicine*, 25, 1369–1382.

Yu, K., and Jones, M. C. (1998), "Local Linear Quantile Regression," *Journal of the American Statistical Association*, 93, 228–238.

Zuo, Y., and Serfling, R. (2000), "General Notations of Statistical Depth Function," *The Annals of Statistics*, 28, 461–482.