Online Supplement for the Paper Titled “Monotonic Quantile Regression with Bernstein Polynomials for Stochastic Simulation”

Matthias H.Y. Tan

Systems Engineering and Engineering Management Department, City University of Hong Kong

Appendix A: Proofs of Mathematical Results

Proof of Equation (10)

Denote $b_{j_i}^i = b(z_i, j_i, N_i)$ and $B_{j_i}^i = B(z_i, j_i, N_i)$. We split $\sum_{j_1=0}^{N_1} \cdots \sum_{j_d=0}^{N_d} \alpha_{(j_1, \dots, j_d)} \prod_{i=1}^d b_{j_i}^i$ into 2^d sums, where each sum is obtained by fixing $j_p = 0$ for $p \in \mathcal{A} \subset \{1, \dots, d\}$:

$$\begin{aligned}
 & \sum_{j_1=0}^{N_1} \cdots \sum_{j_d=0}^{N_d} \alpha_{(j_1, \dots, j_d)} \prod_{i=1}^d b_{j_i}^i \\
 &= \sum_{\mathcal{A} \cup \mathcal{B} = \{1, \dots, d\}, \mathcal{A} \cap \mathcal{B} = \emptyset} \sum_{i \in \mathcal{B}} \sum_{j_i=1}^{N_i} \alpha_{(j_1, \dots, j_d)} \prod_{p \in \mathcal{A}} 1 \prod_{k \in \mathcal{B}} b_{j_k}^k \\
 &= \sum_{\mathcal{A} \cup \mathcal{B} = \{1, \dots, d\}, \mathcal{A} \cap \mathcal{B} = \emptyset} \sum_{i \in \mathcal{B}} \sum_{j_i=1}^{N_i} \alpha_{(j_1, \dots, j_d)} \prod_{p \in \mathcal{A}} \left(\sum_{l_p=0}^{N_p} B_{l_p}^p \right) \prod_{k \in \mathcal{B}} B_{j_k}^k \\
 &= \sum_{\mathcal{A} \cup \mathcal{B} = \{1, \dots, d\}, \mathcal{A} \cap \mathcal{B} = \emptyset} \sum_{i \in \mathcal{B}} \sum_{j_i=1}^{N_i} \sum_{q \in \mathcal{A}} \sum_{l_q=0}^{N_q} \alpha_{(j_1, \dots, j_d)} \left(\prod_{p \in \mathcal{A}} B_{l_p}^p \prod_{k \in \mathcal{B}} B_{j_k}^k \right) \quad (A1) \\
 &= \sum_{j_1=0}^{N_1} \cdots \sum_{j_d=0}^{N_d} \beta_{(j_1, \dots, j_d)} \prod_{i=1}^d B_{j_i}^i.
 \end{aligned}$$

If $j_p = 0$, terms involving $\prod_{i=1}^d B_{j_i}^i$ can only be found with $p \in \mathcal{A}$ in (A1). On the other hand, if $j_p > 0$, terms involving $\prod_{i=1}^d B_{j_i}^i$ can be found with $p \in \mathcal{A}$ or $p \in \mathcal{B}$. This implies that

$$\beta_{(l_1, \dots, l_d)} = \sum_{e=1}^d \sum_{j_e \in \{0, l_e\}} \alpha_{(j_1, \dots, j_d)}.$$

Theorem 2: Let $F: [0,1]^d \rightarrow \mathbb{R}$ be a function. Then, F has a unique FANOVA decomposition:

$$F(\mathbf{z}) = F_0 + \sum_{i=1}^d F_{(i)}(z_i) + \sum_{i=1}^{d-1} \sum_{j=i+1}^d F_{(i,j)}(z_i, z_j) + \cdots + F_{(1,2,\dots,d)}(\mathbf{z}) \quad (A2)$$

if $F_{(i_1, \dots, i_e)}(x_{i_1}, \dots, x_{i_e}) = 0 \forall (x_{i_1}, \dots, x_{i_e}) \in \{\zeta \in [0,1]^e: \prod_{i=1}^e \zeta_i = 0\}, e = 1, \dots, d$.

Proof: We first show that such a decomposition exists. Define

$$\zeta_{(i)}(a) = (z_1 = 0, \dots, z_{i-1} = 0, z_i = a, z_{i+1} = 0, z_d = 0),$$

$$\zeta_{(i,j)}(a, b) = (z_1 = 0, \dots, z_{i-1} = 0, z_i = a, z_{i+1} = 0, \dots, z_{j-1} = 0, z_j = b, z_{j+1} = 0, \dots, z_d = 0),$$

and similarly for $\zeta_{(i,j,k)}(a, b, c)$. We set $F_0 = F(\mathbf{0})$,

$$F_{(i)}(z_i) = F(\zeta_{(i)}(z_i)) - F_0,$$

$$F_{(i,j)}(z_i, z_j) = F(\zeta_{(i,j)}(z_i, z_j)) - F(\zeta_{(i)}(z_i)) - F(\zeta_{(j)}(z_j)) + F_0 = F(\zeta_{(i,j)}(z_i, z_j)) -$$

$$F_{(i)}(z_i) - F_{(j)}(z_j) - F_0,$$

$$F_{(i,j,k)}(z_i, z_j, z_k) = F(\zeta_{(i,j,k)}(z_i, z_j, z_k)) - F(\zeta_{(i,j)}(z_i, z_j)) - F(\zeta_{(i,k)}(z_i, z_k)) -$$

$$F(\zeta_{(j,k)}(z_j, z_k)) + F(\zeta_{(i)}(z_i)) + F(\zeta_{(j)}(z_j)) + F(\zeta_{(k)}(z_k)) - F_0$$

$$= F(\zeta_{(i,j,k)}(z_i, z_j, z_k)) - F_{(i,j)}(z_i, z_j) - F_{(i,k)}(z_i, z_k) - F_{(j,k)}(z_j, z_k) - F_{(i)}(z_i) - F_{(j)}(z_j) -$$

$$F_{(k)}(z_k) - F_0,$$

and similarly for all higher order interactions. It is easy to see that $F_{(i)}(0) = F_{(i,j)}(0, z_j) =$

$F_{(i,j)}(z_i, 0) = F_{(i,j,k)}(0, z_j, z_k) = F_{(i,j,k)}(z_i, 0, z_k) = F_{(i,j,k)}(z_i, z_j, 0) = 0$ and similarly for all

higher order interactions.

To show that the decomposition is unique, suppose that there is another decomposition:

$$F(\mathbf{z}) = G_0 + \sum_{i=1}^d G_{(i)}(z_i) + \sum_{i=1}^{d-1} \sum_{j=i+1}^d G_{(i,j)}(z_i, z_j) + \dots + G_{(1,2,\dots,d)}(\mathbf{z}), \text{ where } G_{(i)}(0) =$$

$$0, G_{(i,j)}(0, z_j) = G_{(i,j)}(z_i, 0) = 0, \dots, G_{(1,2,\dots,d)}(\mathbf{z}) = 0 \forall \mathbf{z} \in \{\zeta \in [0,1]^d : \prod_{i=1}^d \zeta_i = 0\}.$$

Then, we must have $G_0 = F(\mathbf{0}) = F_0$. Setting $\mathbf{z} = \zeta_{(i)}(z_i)$ gives $G_{(i)}(z_i) = F(\zeta_{(i)}(z_i)) - G_0 =$

$$F(\zeta_{(i)}(z_i)) - F_0 = F_{(i)}(z_i). \text{ Setting } \mathbf{z} = \zeta_{(i,j)}(z_i, z_j) \text{ yields } G_{(i,j)}(z_i, z_j) = F(\zeta_{(i,j)}(z_i, z_j)) -$$

$$G_{(i)}(z_i) - G_{(j)}(z_j) - G_0 = F(\zeta_{(i,j)}(z_i, z_j)) - F_{(i)}(z_i) - F_{(j)}(z_j) - F_0 = F_{(i,j)}(z_i, z_j). \text{ The proof}$$

that higher order interactions of both decompositions are equal can be obtained by induction.

Theorem 4: Let $F: [0,1]^d \rightarrow \mathbb{R}$ be a continuous function that is increasing in $z_k, k \in \Omega \subset$

$\{1, \dots, d\}$ and w be a density on $[0,1]^d$. Let $Y_{\mathbf{z}} = F(\mathbf{z}) + \varepsilon$, where $P(\varepsilon \leq 0 | \mathbf{z}) = P(Y_{\mathbf{z}} \leq$

$F(\mathbf{z}) | \mathbf{z}) = \tau \in (0,1)$. Assume that $h_{\mathbf{z}}(y)w(\mathbf{z})$ is the probability density function of $(Y_{\mathbf{z}}, \mathbf{z})$.

Denote by $\bar{\mathcal{F}}_{\mathbf{N}}(\mathbf{z})$ the function $\mathcal{F}(\mathbf{z}) = f_{\tau}(\mathbf{z}) | (\boldsymbol{\delta} = \mathbf{1}, \mathbf{N})$ (see (18)) with $\boldsymbol{\alpha}$ set to $\bar{\boldsymbol{\alpha}} =$

$$\operatorname{argmin}_{\alpha} \left\{ \int_{[0,1]^d} E\{L_{\tau}[Y_{\mathbf{z}} - \mathcal{F}(\mathbf{z})] - L_{\tau}[Y_{\mathbf{z}} - F(\mathbf{z})]\} w(\mathbf{z}) d\mathbf{z} : (15) \text{ holds for all } k \in \Omega, \alpha \in \mathcal{S} \right\},$$

where $\mathcal{S}$ is a compact set that contains the coefficients of $\tilde{f}_N$. Then,

$$\lim_{\min\{N_1, \dots, N_d\} \rightarrow \infty} \int_{[0,1]^d} E\{L_{\tau}[Y_{\mathbf{z}} - \bar{\mathcal{F}}_N(\mathbf{z})] - L_{\tau}[Y_{\mathbf{z}} - F(\mathbf{z})]\} w(\mathbf{z}) d\mathbf{z} = 0.$$

Proof: Note that

$$L_{\tau}[Y_{\mathbf{z}} - g(\mathbf{z})] - L_{\tau}[Y_{\mathbf{z}} - F(\mathbf{z})] = \begin{cases} \tau[F(\mathbf{z}) - g(\mathbf{z})] \text{ if } Y_{\mathbf{z}} - g(\mathbf{z}) > 0, Y_{\mathbf{z}} - F(\mathbf{z}) > 0 \\ (1 - \tau)[g(\mathbf{z}) - F(\mathbf{z})] \text{ if } Y_{\mathbf{z}} - g(\mathbf{z}) < 0, Y_{\mathbf{z}} - F(\mathbf{z}) < 0 \\ \tau[F(\mathbf{z}) - g(\mathbf{z})] - [F(\mathbf{z}) - Y_{\mathbf{z}}] \text{ if } g(\mathbf{z}) < Y_{\mathbf{z}} < F(\mathbf{z}) \\ -\tau[g(\mathbf{z}) - F(\mathbf{z})] + [g(\mathbf{z}) - Y_{\mathbf{z}}] \text{ if } F(\mathbf{z}) < Y_{\mathbf{z}} < g(\mathbf{z}) \end{cases}.$$

If $g(\mathbf{z}) < F(\mathbf{z})$, then,

$$\begin{aligned} & E\{L_{\tau}[Y_{\mathbf{z}} - g(\mathbf{z})] - L_{\tau}[Y_{\mathbf{z}} - F(\mathbf{z})]\} \\ &= \int_{F(\mathbf{z})}^{\infty} \tau[F(\mathbf{z}) - g(\mathbf{z})] h_{\mathbf{z}}(y) dy + \int_{-\infty}^{g(\mathbf{z})} (1 - \tau)[g(\mathbf{z}) - F(\mathbf{z})] h_{\mathbf{z}}(y) dy + \int_{g(\mathbf{z})}^{F(\mathbf{z})} \{\tau[F(\mathbf{z}) - g(\mathbf{z})] - [F(\mathbf{z}) - y]\} h_{\mathbf{z}}(y) dy \\ &= (1 - \tau)\tau[F(\mathbf{z}) - g(\mathbf{z})] + (1 - \tau)[g(\mathbf{z}) - F(\mathbf{z})]P(Y_{\mathbf{z}} \leq g(\mathbf{z})) + \tau[F(\mathbf{z}) - g(\mathbf{z})]P(g(\mathbf{z}) \leq Y_{\mathbf{z}} \leq F(\mathbf{z})) - \int_{-\infty}^{\infty} [F(\mathbf{z}) - y]1(g(\mathbf{z}) \leq y \leq F(\mathbf{z})) h_{\mathbf{z}}(y) dy. \end{aligned} \quad (\text{A3})$$

On the other hand, if $g(\mathbf{z}) > F(\mathbf{z})$, then,

$$\begin{aligned} & E\{L_{\tau}[Y_{\mathbf{z}} - g(\mathbf{z})] - L_{\tau}[Y_{\mathbf{z}} - F(\mathbf{z})]\} \\ &= \int_{g(\mathbf{z})}^{\infty} \tau[F(\mathbf{z}) - g(\mathbf{z})] h_{\mathbf{z}}(y) dy + \int_{-\infty}^{F(\mathbf{z})} (1 - \tau)[g(\mathbf{z}) - F(\mathbf{z})] h_{\mathbf{z}}(y) dy + \int_{F(\mathbf{z})}^{g(\mathbf{z})} \{-\tau[g(\mathbf{z}) - F(\mathbf{z})] + [g(\mathbf{z}) - y]\} h_{\mathbf{z}}(y) dy \\ &= \tau[F(\mathbf{z}) - g(\mathbf{z})]P(Y_{\mathbf{z}} \geq g(\mathbf{z})) + (1 - \tau)\tau[g(\mathbf{z}) - F(\mathbf{z})] - \tau[g(\mathbf{z}) - F(\mathbf{z})]P(F(\mathbf{z}) \leq Y_{\mathbf{z}} \leq g(\mathbf{z})) + \int_{-\infty}^{\infty} [g(\mathbf{z}) - y]1(F(\mathbf{z}) \leq y \leq g(\mathbf{z})) h_{\mathbf{z}}(y) dy. \end{aligned} \quad (\text{A4})$$

It is well known that $g(\mathbf{z}) = F(\mathbf{z})$ minimizes $E\{L_{\tau}[Y_{\mathbf{z}} - g(\mathbf{z})] - L_{\tau}[Y_{\mathbf{z}} - F(\mathbf{z})]\}$ because (A3) and (A4) are bounded from below by 0. Suppose that $g(\mathbf{z}) = \tilde{f}_N(\mathbf{z})$, where $\tilde{f}_N(\mathbf{z})$ is as defined in Corollary 2. Because the Bernstein bases $B(z_i, l, N_i)$, $l = 1, \dots, N_i$ are nonnegative and sum to less than or equal to one, $|\tilde{f}_{(i)}(z_i)| = |\sum_{l=1}^{N_i} F_{(i)}(l/N_i) B(z_i, l, N_i)| \leq \sum_{l=1}^{N_i} |F_{(i)}(l/N_i)|$

$N_i) |B(z_i, l, N_i) \leq \max_{z_i \in [0,1]} |F_{(i)}(z_i)|$. Similarly, $|\tilde{f}_{(i,j)}(z_i, z_j)| \leq \max_{(z_i, z_j) \in [0,1]^2} |F_{(i,j)}(z_i, z_j)|$.

Thus,

$$|\tilde{f}_N(\mathbf{z})| \leq |F_0| + \sum_{i=1}^d \max_{z_i \in [0,1]} |F_{(i)}(z_i)| + \sum_{i=1}^{d-1} \sum_{j=i+1}^d \max_{(z_i, z_j) \in [0,1]^2} |F_{(i,j)}(z_i, z_j)| = H.$$

Consequently, we have

$$\begin{aligned} & |L_\tau[Y_{\mathbf{z}} - \tilde{f}_N(\mathbf{z})] - L_\tau[Y_{\mathbf{z}} - F(\mathbf{z})]| \\ & \leq \begin{cases} \tau(H + |F(\mathbf{z})|) \text{ if } Y_{\mathbf{z}} - \tilde{f}_N(\mathbf{z}) > 0, Y_{\mathbf{z}} - F(\mathbf{z}) > 0 \\ (1 - \tau)(H + |F(\mathbf{z})|) \text{ if } Y_{\mathbf{z}} - \tilde{f}_N(\mathbf{z}) < 0, Y_{\mathbf{z}} - F(\mathbf{z}) < 0 \\ \tau(H + |F(\mathbf{z})|) + (H + |F(\mathbf{z})|) \text{ if } \tilde{f}_N(\mathbf{z}) < Y_{\mathbf{z}} < F(\mathbf{z}) \\ \tau(H + |F(\mathbf{z})|) + (H + |F(\mathbf{z})|) \text{ if } F(\mathbf{z}) < Y_{\mathbf{z}} < \tilde{f}_N(\mathbf{z}) \end{cases} \\ & \leq \tau(H + |F(\mathbf{z})|)1(Y_{\mathbf{z}} + H > 0, Y_{\mathbf{z}} - F(\mathbf{z}) > 0) + (1 - \tau)(H + |F(\mathbf{z})|)1(Y_{\mathbf{z}} - H < 0, Y_{\mathbf{z}} - \\ & F(\mathbf{z}) < 0) + (1 + \tau)(H + |F(\mathbf{z})|)1(-H < Y_{\mathbf{z}} < F(\mathbf{z})) + (1 + \tau)(H + |F(\mathbf{z})|)1(F(\mathbf{z}) < Y_{\mathbf{z}} < \\ & H) = H_2. \end{aligned} \tag{A5}$$

Since

$$H_2 \leq \tau(H + \|F\|_\infty) + (1 - \tau)(H + \|F\|_\infty) + 2(1 + \tau)(H + \|F\|_\infty) < \infty, \tag{A6}$$

it follows from the dominated convergence theorem (page 133 of Resnick (1999)) that

$$\begin{aligned} & \lim_{\min\{N_1, \dots, N_d\} \rightarrow \infty} \int_{[0,1]^d} E\{L_\tau[Y_{\mathbf{z}} - \tilde{f}_N(\mathbf{z})] - L_\tau[Y_{\mathbf{z}} - F(\mathbf{z})]\} w(\mathbf{z}) d\mathbf{z} \\ & = \int_{[0,1]^d} E\left(\lim_{\min\{N_1, \dots, N_d\} \rightarrow \infty} \{L_\tau[Y_{\mathbf{z}} - \tilde{f}_N(\mathbf{z})]\} - L_\tau[Y_{\mathbf{z}} - F(\mathbf{z})]\right) w(\mathbf{z}) d\mathbf{z} \\ & = \int_{[0,1]^d} E\{L_\tau[Y_{\mathbf{z}} - F(\mathbf{z})] - L_\tau[Y_{\mathbf{z}} - F(\mathbf{z})]\} w(\mathbf{z}) d\mathbf{z} = 0. \end{aligned} \tag{A7}$$

Note that the integral $\int_{[0,1]^d} E\{L_\tau[Y_{\mathbf{z}} - g(\mathbf{z})] - L_\tau[Y_{\mathbf{z}} - F(\mathbf{z})]\} w(\mathbf{z}) d\mathbf{z}$ is well defined for any continuous g because $h_{\mathbf{z}}(y)w(\mathbf{z})$ is a probability density function with arguments $(y, \mathbf{z})$ and $|L_\tau[Y_{\mathbf{z}} - g(\mathbf{z})] - L_\tau[Y_{\mathbf{z}} - F(\mathbf{z})]|$ is bounded (a bound like (A6) holds).

Because $F(\mathbf{z})$ is the τ th quantile of the distribution of $Y_{\mathbf{z}}$, we must have $\int_{[0,1]^d} E\{L_\tau[Y_{\mathbf{z}} - g(\mathbf{z})] - L_\tau[Y_{\mathbf{z}} - F(\mathbf{z})]\} w(\mathbf{z}) d\mathbf{z} \geq 0$ for all continuous functions g . Existence of a minimizer $\bar{\alpha}$ of $\varphi(\alpha) = \int_{[0,1]^d} E\{L_\tau[Y_{\mathbf{z}} - \mathcal{F}(\mathbf{z})] - L_\tau[Y_{\mathbf{z}} - F(\mathbf{z})]\} w(\mathbf{z}) d\mathbf{z}$ is guaranteed due to the fact that

$\varphi(\alpha)$ is continuous in α and α is restricted to a compact set: Since the supremum of $|\mathcal{F}(\mathbf{z})| = |\mathbf{x}(\mathbf{z})\alpha|$ over all $\mathbf{z} \in [0,1]^d$ and all α in a compact set is a finite constant C , we have

$$\begin{aligned} & |L_\tau[Y_{\mathbf{z}} - \mathbf{x}(\mathbf{z})\alpha] - L_\tau[Y_{\mathbf{z}} - F(\mathbf{z})]| \\ & \leq \tau(C + \|F\|_\infty) + (1 - \tau)(C + \|F\|_\infty) + 2(1 + \tau)(C + \|F\|_\infty). \end{aligned}$$

By the dominated convergence theorem we conclude that $\varphi(\alpha) = \int_{[0,1]^d} E\{L_\tau[Y_{\mathbf{z}} - \mathcal{F}(\mathbf{z})] - L_\tau[Y_{\mathbf{z}} - F(\mathbf{z})]\}w(\mathbf{z})d\mathbf{z}$ is continuous in α . Thus, $\varphi(\alpha)$ must achieve a minimum $\varphi(\bar{\alpha})$.

Note that $\int_{[0,1]^d} E\{L_\tau[Y_{\mathbf{z}} - \bar{\mathcal{F}}_N(\mathbf{z})] - L_\tau[Y_{\mathbf{z}} - F(\mathbf{z})]\}w(\mathbf{z})d\mathbf{z} \leq \int_{[0,1]^d} E\{L_\tau[Y_{\mathbf{z}} - \mathcal{F}(\mathbf{z})] - L_\tau[Y_{\mathbf{z}} - F(\mathbf{z})]\}w(\mathbf{z})d\mathbf{z}$ for all $\mathcal{F}(\mathbf{z})$ with α that satisfies (15) for each $k \in \Omega$ and belongs to a compact set $\mathcal{S}$ containing the coefficients of $\tilde{f}_N$. This includes $\mathcal{F} = \tilde{f}_N$. Note that the compact set $\mathcal{S}$ can be taken as a Cartesian product set such that each of the coefficients $\{\alpha_l^{(i)} : l = 1, \dots, N_i\}$ are restricted to $[\min_{z_i \in [0,1]} F_{(i)}(z_i), \max_{z_i \in [0,1]} F_{(i)}(z_i)]$ and each of the coefficients $\{\alpha_{(l,m)}^{(i,j)} : l = 1, \dots, N_i, m = 1, \dots, N_j\}$ are restricted to

$$\begin{aligned} & \left[\min_{(z_i, z_j) \in [0,1]^2} F_{(i,j)}(z_i, z_j), \max_{(z_i, z_j) \in [0,1]^2} F_{(i,j)}(z_i, z_j) \right]. \text{ Thus,} \\ & 0 \leq \int_{[0,1]^d} E\{L_\tau[Y_{\mathbf{z}} - \bar{\mathcal{F}}_N(\mathbf{z})] - L_\tau[Y_{\mathbf{z}} - F(\mathbf{z})]\}w(\mathbf{z})d\mathbf{z} \leq \int_{[0,1]^d} E\{L_\tau[Y_{\mathbf{z}} - \tilde{f}_N(\mathbf{z})] - L_\tau[Y_{\mathbf{z}} - F(\mathbf{z})]\}w(\mathbf{z})d\mathbf{z}. \end{aligned} \quad (\text{A8})$$

It follows from (A7) and (A8) that

$$\lim_{\min\{N_1, \dots, N_d\} \rightarrow \infty} \int_{[0,1]^d} E\{L_\tau[Y_{\mathbf{z}} - \bar{\mathcal{F}}_N(\mathbf{z})] - L_\tau[Y_{\mathbf{z}} - F(\mathbf{z})]\}w(\mathbf{z})d\mathbf{z} = 0. \quad (\text{A9})$$

References

1. Resnick, S. I. (1999). *A probability path*. Boston: Birkhäuser.

Appendix B: Example 3: Two-indenture Repairable Item System

In this example, we demonstrate the utility of the proposed method for modeling a two-indenture repairable item system slightly more elaborate than the one described in Daryanto et al. (2003). The system is illustrated in Figure B1. There are a total of 20 machines in a production system at time 0. Each of the machines fails with independently and identically distributed (iid) time between failures $2\rho_T T$, where ρ_T has a beta distribution with density $p(x) \propto x(1-x)1_{[0,1]}(x)$. Each failed machine is sent to the disassembly station, which is a single server queuing system with iid disassembly time $2\rho_D D$, where ρ_D has density $p(x)$. A failed machine is replaced by a reserved machine if available (there are two at time 0) and there cannot be more than 20 machines in the production system. After disassembly, the failed component from the machine is sent to the repair station, which is a single server queuing system with iid repair time $2\rho_R R$, where ρ_R has density $p(x)$. The machine itself is sent to a queue awaiting a component in good working condition. If there is such a component available (there are two at time 0) and the assembly server is idle, the machine and component are assembled with iid assembly time $2\rho_A A$, where ρ_A has density $p(x)$. The repaired machine is then sent to the reserved machine queue if the production system has a total of 20 working machines. The repaired machine is sent to the production system if the system has less than 20 working machines.

Figure B1: Diagram of two-indenture repairable item system

Table B1: Minimum and maximum levels of inputs

Input	Minimum	Maximum
T	100	200
D	0.5	1
A	0.5	1
R	5	30

The output Y of interest is the average number of machines in the production system from time 0 to time 300. We are interested to model the quantiles of Y as a function of T, D, A , and R , which are the mean-time between failures of a machine, the mean disassembly time, the mean assembly time, and the mean repair time respectively. It is clear from our knowledge of the system that the quantiles of Y are increasing in T , and decreasing with D, A , and R . An increase in mean time between failures of a machine will increase the quantiles of the average number of machines in production. On the other hand, a decrease in mean disassembly/assembly/repair time will increase the quantiles of the average number of machines in production. We shall model the 0.05, 0.5, and 0.95 quantiles over the hypercube region bounded by the minimum and maximum levels of T, D, A , and R given in Table B1. The 0.05 quantile may be of interest to a decision maker who wants to ensure that there is a 95% chance that the average number of machines in production (which is proportional to the average number of products produced) is above a given threshold. For modeling purposes, the inputs are transformed to $z_1 = (T - 100)/100$, $z_2 = (-D + 1)/0.5$, $z_3 = (-A + 1)/0.5$, and $z_4 = (-R + 30)/25$. Thus, $\mathbf{z} \in [0,1]^4$ and f_{τ} is an increasing function in each z_i .

We obtain 1000 independent replicates of the simulation model at each point in the grid $\{0,0.25,0.5,0.75,1\}^4$. Because of the large sample size, the data give quite accurate estimates of the true conditional quantile functions and it will be the *validation data set*. Figure B2 presents various contour plots of the 0.05 sample quantile of Y . We see from the contour plots of the 0.05 quantile versus (z_1, z_4) with $(z_2, z_3) = (0,0), (0.5,0.5), (1,1)$ (the left column of Figure B2) that z_1 and z_4 have large joint effect on $f_{0.05}$ and there seems to be no interaction between (z_1, z_4)

and (z_2, z_3) . The contour plots of the 0.05 quantile versus (z_2, z_3) with $(z_1, z_4) = (0,0)$, $(0.5,0.5)$, $(1,1)$ (the right column of Figure B2) appear noisy and indicate that (z_2, z_3) is inert.

Figure B2: Contour plots of 0.05 quantile of simulation output estimated with 1000 simulations

For illustration purposes, we plot the true values (accurate estimates based on 5000 simulations) of the 0.05, 0.5, and 0.95 quantiles of Y versus z_4 conditional on $(z_1, z_2, z_3) =$

(1,1,1) in Figure B3 as solid lines. We see that the distribution of Y changes significantly with z_4 and is highly skewed when $z_4 > 0.8$. Estimates of the quantiles given by the FANOVA Bernstein and rearranged quantile GP models for one dataset are also plotted in Figure B3. The design employed is a 30-run cosine MLHD replicated 20 times. The figure shows that the rearranged quantile GP can have an abrupt slope change, and rise sharply and linearly near $z_4 = 1$. This is observed in a number of randomly generated dataset. In contrast, the FANOVA Bernstein model seems to give more stable and accurate estimates of the quantiles.

Figure B3: 0.05, 0.5, and 0.95 quantiles of simulation output versus z_4 conditioned on $(z_1, z_2, z_3) = (1, 1, 1)$ (solid lines), and point estimates of the quantiles given by the FANOVA Bernstein model (dashed lines) and rearranged quantile GP model (dotted lines).

We perform 100 simulations to compare the performance of the proposed approach to the multivariate Bernstein (with fixed $\mathbf{N} = (1, 1, 1, 1)$ and monotonicity constraints), quantile GP, and rearrange quantile models. In the simulations, we employ a cosine MLHD with 30 runs replicated 20 times. It is not possible to fit a multivariate Bernstein model with $\mathbf{N} = (2, 2, 2, 2)$ with the design. In each simulation, the design is generated and the output from the computer model is obtained. Then, quantile regression is performed using the multivariate Bernstein model

(4) with the monotonicity constraints (5), the FANOVA Bernstein model, the quantile GP model, and the rearranged quantile GP model. The value of $N_{\max}$ for the FANOVA Bernstein model is fixed at seven. The MSE and MAE of the estimates over the validation data points are computed.

Table B2: Quantile estimation performance of FANOVA Bernstein model and comparison with multivariate Bernstein, quantile GP, and rearranged quantile GP models for Example 3. The table gives the simulation mean and standard deviation of the MSE and MAE for the FANOVA Bernstein model, mean difference in MSE and MAE between a competing model and the FANOVA Bernstein model, and fraction of simulations in which the difference is positive.

$\tau = 0.95$						
	FANOVA Bernstein			Multivariate Bernstein - FANOVA Bernstein	Quantile GP - FANOVA Bernstein	Rearranged Quantile GP - FANOVA Bernstein
MSE	Mean	0.110	Mean	0.625	0.056	0.036
	Std	0.050	Fraction >0	1.000	0.770	0.700
MAE	Mean	0.242	Mean	0.339	0.044	0.031
	Std	0.046	Fraction >0	1.000	0.760	0.690
$\tau = 0.05$						
	FANOVA Bernstein			Multivariate Bernstein - FANOVA Bernstein	Quantile GP - FANOVA Bernstein	Rearranged Quantile GP - FANOVA Bernstein
MSE	Mean	0.158	Mean	1.423	0.068	0.043
	Std	0.089	Fraction >0	1.000	0.680	0.650
MAE	Mean	0.293	Mean	0.555	0.047	0.027
	Std	0.069	Fraction >0	1.000	0.670	0.610
$\tau = 0.5$						
	FANOVA Bernstein			Multivariate Bernstein - FANOVA Bernstein	Quantile GP - FANOVA Bernstein	Rearranged Quantile GP - FANOVA Bernstein
MSE	Mean	0.052	Mean	0.813	0.069	0.059
	Std	0.024	Fraction >0	1.000	0.940	0.900
MAE	Mean	0.167	Mean	0.496	0.082	0.071
	Std	0.035	Fraction >0	1.000	0.920	0.870

Predictions of the quantile GP on the grid $\{0,0.25,0.5,0.75,1\}^4$ obtained in the simulations indicate that it is almost never monotonic. Table B2 presents the simulation mean and standard deviation of the MSE and MAE for the FANOVA Bernstein model. The mean difference in MSE and MAE between the multivariate Bernstein/quantile GP/rearranged quantile GP and the FANOVA model is also given in Table B2. In all cases the mean difference is positive, i.e., the FANOVA Bernstein model performs better in the simulations. We see that the

FANOVA Bernstein model performs much better than the multivariate Bernstein model in all cases. Table B2 also suggests that the FANOVA Bernstein model is better than the quantile GP and rearranged quantile GP models. The fraction of simulation runs in which the difference in MSE/MAE between a competing model and the FANOVA Bernstein model is positive, which are presented in Table B2, can be used to test the hypothesis that the FANOVA Bernstein model gives more accurate estimates more often than the competing model. Since $P(B/100 \geq 0.62) = 0.0105$, where B is a binomial random variable with 100 trials and success probability 0.5, we can say that the FANOVA Bernstein model is significantly better whenever the observed fraction is larger than or equal to 0.62. Thus, we see that FANOVA Bernstein model is significantly better (with respect to both MSE and MAE) in all cases except when compared based on the MAE to the rearranged quantile GP in estimating the 0.05 quantile.

Table B3: Number of times FANOVA component is selected out of 100 simulations

Component	Number of times selected		
	$\tau = 0.95$	$\tau = 0.05$	$\tau = 0.5$
(1)	100	100	100
(2)	19	23	24
(3)	21	23	25
(4)	100	100	100
(1,2)	0	3	1
(1,3)	0	1	1
(1,4)	100	100	100
(2,3)	1	2	1
(2,4)	1	4	0
(3,4)	3	4	0

Finally, we examine the model selection results for the FANOVA Bernstein model. Table B3 presents the frequencies with which each FANOVA component is selected. We see that in all simulations, $f_{(1)}$, $f_{(4)}$, and $f_{(1,4)}$ are selected for all three values of τ . The significance of these components is confirmed by Figure B2 for $\tau = 0.05$. Figures similar to Figure B2 are obtained for $\tau = 0.5$ and $\tau = 0.95$. We also see that $f_{(2)}$ and $f_{(3)}$ are quite often selected to enter the

model even though Figure B2 suggests that z_2 and z_3 are inert for $\tau = 0.05$. However, physical knowledge and some analysis of the validation data suggest that z_2 and z_3 do have an effect on Y . Reducing the mean disassembly/ assembly time should reduce the time it takes to fix a machine that breaks down, and this should cause an increase in average number of machines in the production system. The average values of N_1, N_2, N_3 , and N_4 are presented in Table B4. We see that the average value of N_1 is around two for all three values of τ and the average value of N_4 is around three for $\tau = 0.95$ and $\tau = 0.5$ and is near four for $\tau = 0.05$.

Table B4: Average N_i values

	$\tau = 0.95$	$\tau = 0.05$	$\tau = 0.5$
N_1	2.42	2.15	2.05
N_2	0.25	0.38	0.39
N_3	0.25	0.43	0.33
N_4	2.9	3.78	3.17

References

1. Daryanto, A., van Ommeren, J. C. W., and Zijm, W. H. M. (2003). "A closed-loop two-indenture repairable item system." In *Proceedings of the Fourth Aegean International Conference on Analysis of Manufacturing Systems*, 1-4.

Appendix C: Simulation Results for Example 1

We perform 200 simulation runs to compare the prediction accuracy and confidence/credible interval coverage of the FANOVA Bernstein, cubic Bernstein, quantile GP, and rearranged quantile GP models for estimating the 0.1, 0.5, and 0.9 quantiles of Y as a function of z_1 . The design has four distinct points $\{0.02, 0.34, 0.66, 0.98\}$ and ten replicates at each point. We assess the prediction accuracy by computing the mean square error (MSE) and mean absolute error (MAE) on the test set $\{0, 0.01, \dots, 1\}$. Quantile estimates based on 50,000 samples are taken as the true values. The simulation mean and standard deviation of the MSE and MAE for the FANOVA Bernstein model are given in Table C1. The mean differences in MSE and MAE between the cubic Bernstein/quantile GP/rearranged quantile GP and the FANOVA Bernstein

Table C1: Quantile estimation performance of FANOVA Bernstein model and comparison with cubic Bernstein, quantile GP, and rearranged quantile GP models for Example 1. The table gives the simulation mean and standard deviation of the MSE and MAE for the FANOVA Bernstein model, mean difference in MSE and MAE between a competing model and the FANOVA Bernstein model, and fraction of simulations in which the difference is positive.

$\tau = 0.9$						
	FANOVA Bernstein			Cubic Bernstein - FANOVA Bernstein	Quantile GP - FANOVA Bernstein	Rearranged Quantile GP - FANOVA Bernstein
MSE	Mean	0.34	Mean	0.532	0.110	0.097
	Std	0.79	Fraction >0	0.775	0.635	0.610
MAE	Mean	0.35	Mean	0.238	0.073	0.065
	Std	0.28	Fraction >0	0.795	0.625	0.615
$\tau = 0.1$						
	FANOVA Bernstein			Cubic Bernstein - FANOVA Bernstein	Quantile GP - FANOVA Bernstein	Rearranged Quantile GP - FANOVA Bernstein
MSE	Mean	0.011	Mean	0.0049	0.0060	0.0048
	Std	0.015	Fraction >0	0.640	0.640	0.630
MAE	Mean	0.070	Mean	0.0161	0.0172	0.0151
	Std	0.040	Fraction >0	0.645	0.690	0.680
$\tau = 0.5$						
	FANOVA Bernstein			Cubic Bernstein - FANOVA Bernstein	Quantile GP - FANOVA Bernstein	Rearranged Quantile GP - FANOVA Bernstein
MSE	Mean	0.049	Mean	0.0070	0.0072	0.0046
	Std	0.078	Fraction >0	0.510	0.525	0.510
MAE	Mean	0.138	Mean	0.0099	0.0100	0.0070
	Std	0.074	Fraction >0	0.525	0.560	0.545

model are also given in Table C1. In all cases the mean difference is positive, i.e., the FANOVA Bernstein model performs better in the simulations. The cubic Bernstein model performs poorly for estimating the 0.9 quantile because the mean differences are large. Since the same data is used to fit the four models in each simulation, the fraction of simulations in which the difference in MSE/MAE is positive can be used to assess whether the FANOVA Bernstein model is significantly better. These fractions are given in Table C1. Since $P(B/200 \geq 0.585) = 0.0097$, where B is a binomial random variable with 200 trials and success probability 0.5, we can say that the FANOVA Bernstein model is significantly better if the observed fraction is larger than or equal to 0.585. We see that FANOVA Bernstein model is significantly better (with respect to

both MSE and MAE) than all other models for the 0.1 and 0.9 quantile but not the 0.5 quantile.

Table C2: Mean and standard deviation of coverage of 95% confidence/credible intervals given by FANOVA Bernstein, quantile GP, and rearranged quantile GP models.

		FANOVA Bernstein	Quantile GP	Rearranged Quantile GP
$\tau = 0.9$	Mean	0.86	0.47	0.49
	Std	0.20	0.40	0.39
$\tau = 0.1$	Mean	0.87	0.41	0.42
	Std	0.21	0.40	0.39
$\tau = 0.5$	Mean	0.96	0.47	0.48
	Std	0.09	0.42	0.42

We also assess the confidence/credible interval coverage of the FANOVA Bernstein (200 bootstrap samples), quantile GP, and rearranged quantile GP models. The fraction of 95% confidence/credible interval that contains the true values at $\{0, 0.01, \dots, 1\}$ (coverage) is recorded in each simulation. Table C2 presents the mean and standard deviation of the coverage. It is seen that the FANOVA Bernstein model gives coverage closer to the nominal coverage than the quantile GP or rearranged quantile GP. The poor performance of the latter two models is likely due to the erroneous assumption of constant error variance and the inadequacy of the design (GP models tend to perform better when the design has many distinct runs). Plumlee and Tuo (2014) do not indicate how the quantile GP can be modified when the error variance is heterogeneous. Lastly, we point out that the rearrangement procedure only yields small improvements to the estimation accuracy and credible interval coverage, as can be seen from Table C1 and Table C2.

Appendix D: Illustrations and Model Selection Results for Example 2

For illustration, we plot in Figure D1a estimates of the 0.1, 0.5, and 0.9 quantiles of Y versus z_2 conditional on $(z_1, z_3, z_4, z_5) = (0.5, 0.5, 0.5, 0.5)$ given by the FANOVA Bernstein and rearranged quantile GP for one dataset. The design is a 40-run cosine MLHD replicated 10 times. The figure shows that the rearranged quantile GP can have sharp corners. In contrast, the FANOVA Bernstein model is smooth. Figure D1a (solid lines) and (22) suggest that the true conditional quantile functions are smooth. Thus, the FANOVA Bernstein model appears to be

more suitable than the rearranged quantile GP for this problem. In Figure D1b, we plot point estimates and 95% confidence/credible intervals of the 0.1 and 0.9 quantiles given by the FANOVA Bernstein and rearranged quantile GP models for one dataset. Again, a 40-run cosine MLHD replicated 10 times is used. We see that the confidence/credible intervals are narrow.

Figure D1: (a) 0.1, 0.5, and 0.9 quantiles (estimated with 5000 simulations) of temperature versus time conditioned on $(z_1, z_3, z_4, z_5) = (0.5, 0.5, 0.5, 0.5)$ (solid lines), and point estimates of the quantiles given by the FANOVA Bernstein (dashed lines) and rearranged quantile GP (dotted lines) models for one dataset; (b) 0.1 and 0.9 quantile of temperature versus time (solid lines), point estimates and 95% confidence/credible intervals of the quantiles given by the FANOVA Bernstein (dashed lines), and quantile GP (dotted lines) models for one dataset.

We examine the model selection results for the FANOVA Bernstein model. Note that a model expressed in terms of a handful of basis functions like the FANOVA Bernstein model can be very useful for understanding and interpreting the computer simulation model. Table D1 presents the frequencies with which each FANOVA component is selected. We see that in all simulations, $f_{(2)}$, $f_{(3)}$, $f_{(4)}$, and $f_{(2,4)}$ are selected. The significance of these components is confirmed by Figure D1 and Equation (22) (observe from the equation that $Y = u_3 + u_4 h(u_1, u_2, u_5, \dots, u_9)$). We also see from Table D1 that $f_{(1)}$ and $f_{(5)}$ are sometimes selected, which suggests that inputs u_1 and u_5 have only small effects. The average values of $N_1, \dots, N_5$ are given in Table D2. The average value of N_2 is the largest for each τ and range from 3.1 for $\tau = 0.1$ to 5.6 for $\tau = 0.9$, which indicate that Bernstein bases of high degrees are needed to

model the asymptotes seen in Figure D1 accurately. The average values of N_3 and N_4 are close to one, which is the right value to use as (22) implies that $Y = u_3 + u_4 h(u_1, u_2, u_5, \dots, u_9)$.

Table D1: Number of times FANOVA component is selected out of 50 simulations

Component		(1)	(2)	(3)	(4)	(5)	(1,2)	(1,3)	(2,3)	(1,4)	(2,4)	(3,4)	(1,5)	(2,5)	(3,5)	(4,5)
Number of times selected	$\tau = 0.9$	4	50	50	50	8	0	0	0	0	50	7	0	0	2	0
	$\tau = 0.1$	17	50	50	50	15	4	3	15	3	50	13	1	1	2	2
	$\tau = 0.5$	12	50	50	50	14	2	1	3	3	50	10	0	0	5	2

Table D2: Average N_i values

i	1	2	3	4	5
$\tau = 0.9$	0.1	5.56	1.18	1.02	0.18
$\tau = 0.1$	0.48	3.08	2	1.54	0.56
$\tau = 0.5$	0.28	3.62	1.26	1.06	0.32