Generalized gamma linear transformation model with application to limit order execution

Yue Jiang* Xianhua Peng† Kani Chen‡

Abstract

With accurate forecasting as a goal in security market, adequate statistical modeling of time to execution of limit orders is much desired. In survival analysis, Cox's proportional hazards model and accelerated failure time model are two distinct and popular regression models. It has been reported in the literature, Cox's model is generally a poorer fit of time to execution of limit orders than accelerated failure time model. Cox's model is a special case of linear transformation model by assuming the error term follows the extreme value distribution, which may be too restrictive in practice. In this paper, we build a novel generalized gamma linear transformation model and apply to the analysis of execution time of limit orders. This model assumes the error term belongs to a generalized gamma distribution family and allows more flexibility than Cox's model. We provide theoretical justification of the statistical properties of maximum likelihood estimation and derive a fast algorithm to calculate the maximum likelihood estimates. Simulation studies are carried out to evaluate the finite sample performance. Analysis of real data of limit order execution times show that the proposed model provides better fit and forecast than accelerated failure time model. We believe the proposed model, along with the provided computation algorithm, may have broad applications.

*Department of Mathematics, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. Email: yjiangan@connect.ust.hk.

†Department of Mathematics, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. Email: maxhpeng@ust.hk.

‡Department of Mathematics, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. Email: makchen@ust.hk.

Keywords: survival analysis, maximum likelihood estimates, nonparametric estimation, generalized gamma distribution, execution time, limit orders, asymptotic property

JEL classification: C13, C14, C46, C53, C61

1 Introduction

In analysis of time to event data, with or without censoring, Cox's proportional hazard model Cox (1972) and accelerated failure time (AFT) model are two distinct classes of regression models. The Cox's model assumes

$$h(t|X) = \lambda(t)\exp(\beta^T X), \tag{1}$$

where $h(t|X)$ is the hazard rate function given covariate X, $\lambda(t)$ is an unknown baseline hazard function and β is a p-dimensional regression parameter. The accelerated failure time model has the form

$$T = e^{-\beta^T X} T_0, \tag{2}$$

where T is the failure time, X is a covariate vector, β is a parameter vector, and T_0 is the baseline failure time and its distribution is called the baseline distribution. The above two models are applied extensively in survival analysis.

Time to event data with the presence of censoring also occur in limit order execution in security market and there exist extensive research on analysis of limit order books. Order cancelation after the placement of a limit order is naturally treated as censored observations. Lo et al. (2002) is among the first to apply survival models to develop and estimate econometric models of limit order execution times. By treating order cancelation as censoring, they estimate separate models for time-to-first-fill and time-to-completion for both buy and sell limit orders. For each model, they estimate the cumulative distribution function of the execution time of limit order conditioning on the market and order information at the time of submission for the limit order. As a major finding, they reported that Cox's model does not fit time to execution data. Alternatively, they applied a generalized gamma accelerated failure time (AFT-GG) model and a better fit was found. With the error distribution in the family of generalized gamma distribution, their model is a parametric model rather than semiparametric. We note that, for the sake of prediction, an entirely unknown error distribution is generally not preferred.

The AFT-GG model was also used in Gava (2005) to analyze the effect of variables on the speed of order execution and their quantitative impact in the Spanish stock exchange. In particular, they considered covariates for order aggressiveness, relative inside spread, volatility, trading activity, priority and market depth, among others. Moreover, to see if the time of order placement affects time to execution, dummy variables for the day of the week and the time of the trading session were inserted. Several findings were reported. Time to execution is generally shorter for limit orders priced at the quotes or within the quotes or when the asset is more volatile and active. The time of the day affects the expected execution time; for orders placed during the first 30 minutes and the last two hours of the trading session, the expected execution time is shorter. A negative relationship is found between expected time of execution and a variable which captures the interaction between the bid ask spread and the trading activity.

Wen (2009) replicated the AFT-GG models in Lo et al. (2002) on UK market data and confirmed the finding of Lo et al. (2002) that Cox's model does not fit time to execution data. Evidence were reported by showing Q-Q plots that deviate markedly from the 45-degree line. However, the goodness-of-fit test of the AFT-GG model shows that there may be room for it to be further improved. They also identified potential multicollinearity effect on the estimation of model parameters, which were dealt with by introducing new covariates, resulting in the correlations among the covariates no higher than 40%.

In survival analysis, linear transformation (LT) model is extensively studied and well understood; see Chen et al. (2002), Zeng and Lin (2006, 2007, 2009), among many others. It assumes

$$H(T) = -\beta^T X + \epsilon, \tag{3}$$

where H is an unknown monotone transformation function, ϵ is a random variable with a known distribution and is independent of X, and β is an unknown p-dimensional regression parameter of interest. Cox's model becomes a special case of (3) with ϵ following the extreme-value distribution. Proportional odds model, another popular model, is also a special case with error following the logistic distribution. The AFT model is also formally a special case of (3) when H is taken to be the log function.

Apart from understanding, forecasting is a primary goal is financial market. For this reason, accurate modeling that allows adequate flexibility becomes vital. Both Cox's model and AFT-GG model have their own pros and cons. The AFT-GG model, despite its better fit of time to execution data, is a parametric model. In general, parametric model is more restrictive, though not necessarily a drawback, than semiparametric. Cox's model is a semiparametric model with the unknown baseline hazard function. Viewed as a special case of linear transformation model, it is flexible in allowing unknown transformation function. On the other hand, it is also restrictive in the given error distribution. which may cause the poor fit of Cox's model to time to execution data.

In view of the poor fit of Cox's model and the "room to improve" of the AFT-GG model, we propose a novel generalized gamma linear transformation (LT-GG) model, a linear transformation with error term in the family of generalized gamma distribution. The proposed model combines the strengths of Cox's model and the AFT-GG model. We provide theoretical justification of the maximum likelihood estimation and the inference, with the aid of the results of Zeng and Lin (2006, 2007, 2009). Moreover, we derive an efficient computation algorithm, tailor made for this model. Extensive simulations are carried to examine the finite sample performance. We then apply the LT-GG model to a real data time to execution of limit orders. It is found that the proposed model provides better fit and forecast than AFT-GG model. We conclude that generalized gamma family is a proper modeling of the error term, and better fit and forecast is achieved by allowing an unknown transformation function, rather than fixing it to be the log function as in the AFT model. It is likely that the proposed model may have broader applications than analysis of time to execution data.

The rest of the paper is organized as follows. In Section 2, we introduce the generalized gamma linear transformation model and clarify the model identifiability problem. We calculate the likelihood function and propose an algorithm to get maximum likelihood estimates of the linear transformation model in Section 3. In Section 4, we conduct simulation studies to evaluate the finite sample performance of linear transformation model. In Section 5, we apply our model to a real dataset. Cross validation is done in Section 6 to illustrate the advantage of our linear transformation model. A brief summary is given in Section 7.

Before moving to the next section, we note an alternative formulation of the linear transformation model in terms of conditional survival function, as given in Zeng and Lin (2006):

$$S(t|X) = \Psi \left\{ \exp(\beta^T X)\Lambda(t) \right\}, \tag{4}$$

where $\Lambda(t)$ is an unknown increasing function, $\Lambda(t) = \int_0^t \lambda(u)du$, and Ψ is a known link function. (4) is equivalent to (3) by taking $\Lambda(t) = e^{H(t)}$ and $\Psi(z) = P(\epsilon > \log z)$. Yin and Zeng (2012) provides an MLE approach to fit the linear transformation model.

2 Model Specification

2.1 The LT-GG Model

To conform with the presentation in the relevant literature, we formulate the LT model (3) by adding a log transformation:

$$\log H(T) = -\beta^T X + \epsilon, \tag{5}$$

where $H(\cdot)$ is an unknown monotone transformation function defined on $[0, \infty)$, the range of the failure time T. Throughout this paper, $e^\epsilon = \xi$ follows generalized gamma distribution, $GG(\mu, \sigma, Q)$, $\sigma > 0$ with density function:

$$f(x, \mu, \sigma, Q) = \begin{cases} \frac{|Q|\left(Q^{-2}\right)^{Q^{-2}}}{\sigma x \Gamma(Q^{-2})} e^{Q^{-2}\left(Qw - e^{Qw}\right)}, \text{ with } w = \frac{\log x - \mu}{\sigma}, \text{ if } Q \neq 0, \\ \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\log x - \mu}{\sigma}\right)^2}, \text{ if } Q = 0. \end{cases} \tag{6}$$

The survival function of ξ is:

$$\begin{aligned} S(x, \mu, \sigma, Q) &= 1 - F(x, \mu, \sigma, Q) \\ &= \begin{cases} 1 - \Gamma_I\left(\frac{1}{Q^2}, \frac{e^{Q\left(\frac{\log x - \mu}{\sigma}\right)}}{Q^2}\right), \text{ if } Q > 0, \\ 1 - \Phi\left(\frac{\log x - \mu}{\sigma}\right), \text{ if } Q = 0, \\ \Gamma_I\left(\frac{1}{Q^2}, \frac{e^{Q\left(\frac{\log x - \mu}{\sigma}\right)}}{Q^2}\right), \text{ if } Q < 0, \end{cases} \end{aligned} \tag{7}$$

where $\Gamma_I(a, x) = \frac{1}{\Gamma(a)} \int_0^x s^{a-1} e^{-s} ds$ is the lower incomplete gamma function, μ is location parameter, σ is scale parameter and Q is shape parameter.

The family of generalized gamma distribution is enough rich to include many commonly used distributions. In the special cases of $Q = 1$, $Q = 0$, $Q = \sigma = 1$ and $Q = \sigma$, the generalized gamma distribution reduces to a Weibull distribution a lognormal distribution, an exponential distribution and a Gamma distribution, respectively. The density function of ϵ is

$$f_\epsilon(y, \mu, \sigma, Q) = \begin{cases} \frac{|Q|\left(Q^{-2}\right)^{Q^{-2}}}{\sigma\Gamma(Q^{-2})} e^{Q^{-2}\left(Q\frac{y-\mu}{\sigma} - e^{Q\frac{y-\mu}{\sigma}}\right)}, \text{ if } Q \neq 0, \\ \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}, \text{ if } Q = 0. \end{cases} \tag{8}$$

The survival function of ϵ is:

$$\begin{aligned} S_\epsilon(y, \mu, \sigma, Q) &= 1 - F_\epsilon(y, \mu, \sigma, Q) \\ &= \begin{cases} 1 - \Gamma_I\left(\frac{1}{Q^2}, \frac{e^{Q\left(\frac{y-\mu}{\sigma}\right)}}{Q^2}\right), \text{ if } Q > 0, \\ 1 - \Phi\left(\frac{y-\mu}{\sigma}\right), \text{ if } Q = 0, \\ \Gamma_I\left(\frac{1}{Q^2}, \frac{e^{Q\left(\frac{y-\mu}{\sigma}\right)}}{Q^2}\right), \text{ if } Q < 0. \end{cases} \end{aligned} \tag{9}$$

2.2 Model Identifiability

If $\xi \sim GG(\mu, \sigma, Q)$, then $e^a\xi^b \sim GG(a + b\mu, |b|\sigma, Q\,\mathrm{sgn}(b))$. So (5) continues to hold if $H(T)$ is replaced by $e^a H(T)^b$, β is replaced by $b\beta$ and ϵ is replaced by $a + b\epsilon$ for any constant a and any positive constant b. In order to make the model identifiable, we set $\mu = 0$ and $\sigma = 1$. Then we fit the model

$$\log H(T) = -\beta^T X + \epsilon, \tag{10}$$

where $e^\epsilon \sim GG(0, 1, Q)$.

3 THE MLE and a Computation Algorithm

3.1 Observed Likelihood Function

In the presence of censoring, we assume conditional independence of T and the censoring variable C given X. Let $Y = \min(T, C)$ be the event time and $\delta = I(T \leq C)$ be the failure/censoring index. The observations are independent copies of (Y, δ, X), denoted as $(Y_i, \delta_i, X_i), i = 1, 2, \dots, n$. The observed likelihood function is

$$L(\beta, Q, H) = \prod_{i=1}^{n} \left\{ f_\epsilon \left(\log H(Y_i) + \beta^T X_i \right) \frac{H^{'}(Y_i)}{H(Y_i)} \right\}^{\delta_i} \times \\ \left\{ 1 - F_\epsilon \left(\log H(Y_i) + \beta^T X_i \right) \right\}^{1-\delta_i}, \tag{11}$$

where $f_\epsilon(\cdot)$ and $F_\epsilon(\cdot)$ are the density function and cumulative distribution function of ϵ, and $H^{'}(\cdot)$ is the first derivative of $H(\cdot)$.

Throughout the paper, we assume the observed failure times are ordered from the smallest to the largest and $m = \sum_{i=1}^{n} \delta_i$ is the number of all the noncensored observations.

The maximum of (11) does not exist because there always exists some function $H(\cdot)$ such that $H^{'}(Y_i)$ is close to infinity for some Y_i with $\delta_i = 1$. Thus, the nonparametric maximum likelihood estimation has been proposed for deriving the maximum likelihood estimator (MLE) in such problems. In the maximization, $H(\cdot)$ is assumed to be a nondecreasing step function only with jumps at the observed failure times and $H(0) = 0$ (Chen et al.; 2002), and the objective function (11) becomes

$$L(\beta, Q, h_1, h_2, \dots, h_m) = \prod_{i=1}^{n} \left\{ f_\epsilon \Big(\log \big(\sum_{Y_j \le Y_i, \delta_j = 1} h_j \big) + \beta^T X_i \Big) \frac{h_i}{\sum_{j=1}^{i} h_j} \right\}^{\delta_i} \times \\ \left\{ 1 - F_\epsilon \Big(\log \big(\sum_{Y_j \le Y_i, \delta_j = 1} h_j \big) + \beta^T X_i \Big) \right\}^{1-\delta_i}, \tag{12}$$

where h_i denotes the jump size of $H(\cdot)$ at Y_i. Taking log on (12), the loglikelihood function is

$$\begin{aligned} & l(\beta, Q, h_1, h_2, \dots, h_m) \\ & = \sum_{i=1}^{n} \Bigg\{ \delta_i \Big[\log f_\epsilon \Big(\log\big(\sum_{Y_j \le Y_i, \delta_j = 1} h_j \big) + \beta^T X_i \Big) + \log h_i - \log \sum_{j=1}^{i} h_j \Big] \\ & \quad + (1 - \delta_i) \log \Big[1 - F_\epsilon \Big(\log\big(\sum_{Y_j \le Y_i, \delta_j = 1} h_j \big) + \beta^T X_i \Big) \Big] \Bigg\}. \end{aligned} \tag{13}$$

The MLE of (β, Q, H) is denoted $(\hat{\beta}, \hat{Q}, \hat{H})$ and $(\hat{h}_1, ..., \hat{h}_m)$ are sizes of jumps of $\hat{H}$.

3.2 Computation Algorithm

Solving for the MLE of (13) is equivalent to

$$\max l(\beta, Q, h_1, h_2, \ldots, h_m), \text{ s.t. } h_i > 0, i = 1, 2, \ldots, m. \tag{14}$$

Note that we require constraints $h_i > 0, i = 1, 2, \ldots, m$ to be satisfied as they denote the jump sizes of $H(\cdot)$. We let $h_i = e^{d_i}, i = 1, 2, \ldots, m$, $\vec{h} = (h_1, h_2, \ldots, h_m)$ and $\vec{d} = (d_1, d_2, \ldots, d_m)$, then (14) becomes an unconstrained optimization problem:

$$\max g(\beta, Q, \vec{d}), \tag{15}$$

where

$$g(\beta, Q, \vec{d}) = l(\beta, Q, e^{d_1}, e^{d_2}, \ldots, e^{d_m}). \tag{16}$$

To solve the optimization problem (15), we use conjugate gradient method. The computation algorithm is detailed as follows.

Step 1: Choose initial values for β, Q and $d_i, i = 1, 2, \ldots, m$, denoted by $\beta^{(0)}, Q^{(0)}$ and $d_i^{(0)}, i = 1, 2, \ldots, m$.

Step 2: Calculate gradient vector at the current iteration point and search for new β, Q and $d_i, i = 1, 2, \ldots, m$.

Step 3: Repeat Step 2 until the prescribed convergence criteria are satisfied.

Instead of using finite difference approximations for first derivatives of $g(\beta, Q, \vec{d})$, we provide explicit formulae for them to accelerate the search process. The first derivatives of $g(\beta, Q, \vec{d})$ are calculated as follows.

For $k = 1, 2, \ldots, p$,

$$\begin{aligned}\frac{\partial g}{\partial \beta_k} = \sum_{i=1}^{n} \Bigg\{ & \delta_i \frac{f'_\epsilon\Big(\log(\sum_{Y_j \le Y_i, \delta_j = 1} e^{d_j}) + \beta^T X_i\Big)}{f_\epsilon\Big(\log(\sum_{Y_j \le Y_i, \delta_j = 1} e^{d_j}) + \beta^T X_i\Big)} X_{ik} \\ & + (1 - \delta_i) \frac{-f_\epsilon\Big(\log(\sum_{Y_j \le Y_i, \delta_j = 1} e^{d_j}) + \beta^T X_i\Big)}{1 - F_\epsilon\Big(\log(\sum_{Y_j \le Y_i, \delta_j = 1} e^{d_j}) + \beta^T X_i\Big)} X_{ik} \Bigg\},\end{aligned} \tag{17}$$

where $f^{'}(\cdot)$ is the first derivative of $f(\cdot)$. For $k = 1, 2, \ldots, m$,

$$\begin{aligned}
\frac{\partial g}{\partial d_k} &= 1 - \sum_{i=k}^{m} \frac{e^{d_k}}{\sum_{j=1}^{i} e^{d_j}} \\
&+ \sum_{Y_i \geq Y_k} \left\{ \delta_i \frac{f_\epsilon^{'} \left(\log(\sum_{Y_j \leq Y_i, \delta_j = 1} e^{d_j} + \beta^T X_i \right)}{f_\epsilon \left(\log(\sum_{Y_j \leq Y_i, \delta_j = 1} e^{d_j}) + \beta^T X_i \right)} \frac{e^{d_k}}{\sum_{Y_j \leq Y_i, \delta_j = 1} e^{d_j}} \right. \\
&+ \left. (1 - \delta_i) \frac{-f_\epsilon \left(\log(\sum_{Y_j \leq Y_i, \delta_j = 1} e^{d_j}) + \beta^T X_i \right)}{1 - F_\epsilon \left(\log(\sum_{Y_j \leq Y_i, \delta_j = 1} e^{d_j}) + \beta^T X_i \right)} \frac{e^{d_k}}{\sum_{Y_j \leq Y_i, \delta_j = 1} e^{d_j}} \right\}. \qquad (18)
\end{aligned}$$

And

$$\begin{aligned}
\frac{\partial g}{\partial Q} &= \sum_{i=1}^{n} \left\{ \delta_i \frac{\partial f_\epsilon \left(\log(\sum_{Y_j \leq Y_i, \delta_j = 1} e^{d_j}) + \beta^T X_i \right) / \partial Q}{f_\epsilon \left(\log(\sum_{Y_j \leq Y_i, \delta_j = 1} e^{d_j}) + \beta^T X_i \right)} \right. \\
&+ \left. (1 - \delta_i) \frac{-\partial F_\epsilon \left(\log(\sum_{Y_j \leq Y_i, \delta_j = 1} e^{d_j}) + \beta^T X_i \right) / \partial Q}{1 - F_\epsilon \left(\log(\sum_{Y_j \leq Y_i, \delta_j = 1} e^{d_j}) + \beta^T X_i \right)} \right\}. \qquad (19)
\end{aligned}$$

3.3 Variance Estimation

The variance estimation is based on the loglikelihood function of β, Q and $h_i, i = 1, 2, \ldots, m$. In particular, the covariance matrix of the first derivatives vector at the MLE is used to approximate the inverted covariance matrix of β, Q and $h_i, i = 1, 2, \ldots, m$. The theoretical justification of such a method can be found in Greene (2003, Chapter 14, p. 521-522). Denote as $l_i(\beta, Q, \vec{h})$ the loglikelihood function at observation $(Y_i, \delta_i, X_i), i = 1, 2, \ldots, n$. Then the first derivatives vector is

$$\vec{v_i} = \left(\frac{\partial l_i(\beta, Q, \vec{h})}{\partial \beta_1}, \ldots, \frac{\partial l_i(\beta, Q, \vec{h})}{\partial \beta_p}, \frac{\partial l_i(\beta, Q, \vec{h})}{\partial Q}, \frac{\partial l_i(\beta, Q, \vec{h})}{\partial h_1}, \ldots, \frac{\partial l_i(\beta, Q, \vec{h})}{\partial h_m} \right)^T. \qquad (20)$$

Let $\hat{G} = \left[\hat{\vec{v}}_1, \hat{\vec{v}}_2, \ldots, \hat{\vec{v}}_n \right]$, where $\hat{\vec{v}}_i$ is the value of $\vec{v_i}$ at the MLE. Then

$$\text{cov}(\beta, Q, \vec{h}) \approx \left[\sum_{i=1}^{n} \hat{\vec{v}}_i \hat{\vec{v}}_i' \right]^{-1} = \left[\hat{G} \hat{G}^T \right]^{-1}. \qquad (21)$$

As a result, the estimates of variance of β, Q can be found from the relevant diagonal elements of matrix $\text{cov}(\beta, Q, \vec{h})$.

3.4 Theoretical Justification

We list the following regularity assumptions (Zeng et al.; 2006):
(1) The covariate X takes value in a bounded set and does not concentrate in any $p-1$ dimension hyperplane.
(2) Conditional on X, the right-censoring time C is independent of T, and $P(C = \infty|X) > 0$.
(3) The true value of β, denoted as β_0, belong to the interior of a known compact set $\mathfrak{B}_0$, and the true function H_0 is differentiable with $H_0^{'}(x) > 0$ for all $x > 0$.

We first state the identifiability of our model.

Theorem 1. *Under Condition (1), the parameters (β_0, Q_0, H_0) in the LT-GG model are identifiable.*

Proof. Suppose that two sets of parameters (β, Q, H) and $(\tilde{\beta}, \tilde{Q}, \tilde{H})$ give the same likelihood function for the observed data. Then,

$$\left\{ f_\epsilon \left(\log H(Y) + \beta^T X \right) \frac{H^{'}(Y)}{H(Y)} \right\}^\delta \times \left\{ 1 - F_\epsilon \left(\log H(Y) + \beta^T X \right) \right\}^{1-\delta}$$
$$= \left\{ f_{\tilde{\epsilon}} \left(\log \tilde{H}(Y) + \tilde{\beta}^T X \right) \frac{\tilde{H}^{'}(Y)}{\tilde{H}(Y)} \right\}^\delta \times \left\{ 1 - F_{\tilde{\epsilon}} \left(\log \tilde{H}(Y) + \tilde{\beta}^T X \right) \right\}^{1-\delta}. \quad (22)$$

By letting $\delta = 0$, we have

$$F_\epsilon \left(\log H(Y) + \beta^T X \right) = F_{\tilde{\epsilon}} \left(\log \tilde{H}(Y) + \tilde{\beta}^T X \right). \quad (23)$$

Differentiate the above equation with respect to β and cancel X on both sides, we get

$$f_\epsilon \left(\log H(Y) + \beta^T X \right) = f_{\tilde{\epsilon}} \left(\log \tilde{H}(Y) + \tilde{\beta}^T X \right). \quad (24)$$

If $Q \neq 0, \tilde{Q} \neq 0$, we have

$$\frac{|Q| \left(Q^{-2}\right)^{Q^{-2}}}{\Gamma\left(Q^{-2}\right)} e^{Q^{-2}\left(Qw - e^{Qw}\right)} = \frac{|\tilde{Q}| \left(\tilde{Q}^{-2}\right)^{\tilde{Q}^{-2}}}{\Gamma\left(\tilde{Q}^{-2}\right)} e^{\tilde{Q}^{-2}\left(\tilde{Q}\tilde{w} - e^{\tilde{Q}\tilde{w}}\right)}, \tag{25}$$

where $w = \log H(Y) + \beta^T X, \tilde{w} = \log \tilde{H}(Y) + \tilde{\beta}^T X$. Hence,

$$\frac{|Q| \left(Q^{-2}\right)^{Q^{-2}} \Gamma\left(\tilde{Q}^{-2}\right)}{|\tilde{Q}| \left(\tilde{Q}^{-2}\right)^{\tilde{Q}^{-2}} \Gamma\left(Q^{-2}\right)} = \frac{e^{\tilde{Q}^{-2}\left(\tilde{Q}\tilde{w} - e^{\tilde{Q}\tilde{w}}\right)}}{e^{Q^{-2}\left(Qw - e^{Qw}\right)}} = c_1, \forall X, Y, \tag{26}$$

where c_1 is a constant. As a result,

$$\tilde{Q}^{-2}\left(\tilde{Q}\tilde{w} - e^{\tilde{Q}\tilde{w}}\right) - Q^{-2}\left(Qw - e^{Qw}\right) = \log c_1, \forall X, Y. \tag{27}$$

So we have $Q = \tilde{Q}, w = \tilde{w}$. i.e.,

$$\log H(Y) + \beta^T X = \log \tilde{H}(Y) + \tilde{\beta}^T X, \forall X, Y. \tag{28}$$

Hence,

$$\log \frac{H(Y)}{\tilde{H}(Y)} = \tilde{\beta}^T X - \beta^T X = c_2, \forall X, Y, \tag{29}$$

where c_2 is a constant. Condition (1) ensures $\beta = \tilde{\beta}$ and $c_2 = 0$. Then $H(Y) = \tilde{H}(Y), \forall Y$.
If $Q = 0, \tilde{Q} \neq 0$, or if $Q \neq 0, \tilde{Q} = 0$, it's easy to prove that (24) does not hold. If $Q = 0, \tilde{Q} = 0$, similar proof can be carried out. The identifiability holds in the LT-GG model. □

Next we look into the asymptotic properties of the MLE. Define function $G_Q(x) = P(e^\epsilon > x)$. Throughout the paper, G'_Q, G''_Q and $G^{(3)}_Q$ denote first, second and third deriatives of G_Q with respect to x and $\dot{G}_Q$ denotes the derivative of G_Q with respect to Q. The following theorem establishes the consistency of $\hat{\beta}, \hat{Q}$ and $\hat{H}$.

Theorem 2. *Under Conditions (1)-(3), with probability one,*

$$|\hat{\beta} - \beta_0| \to 0, |\hat{Q} - Q_0| \to 0, \text{ and } \sup_{t>0} |\hat{H}(t) - H_0(t)| \to 0,$$

i.e. $\hat{\beta}, \hat{Q}$ *and* $\hat{H}$ *are strongly consistent.*

Proof. According to Theorem 1, the parameters (β_0, Q_0, H_0) in our model are identifiable. We next show the following statement (4') hold.
(4'): $G_Q(x)$ is three time differentiable with respect to Q and x and all the derivatives are uniformly bounded with $G_Q^{'}(x) < 0$.
We briefly show the verification steps of (4') in the case of first order derivatives as an example.

$$G_Q^{'}(x) = \begin{cases} -\frac{|Q|\left(Q^{-2}\right)^{Q^{-2}}}{\Gamma(Q^{-2})} e^{Q^{-2}\left(Q\log x - e^{Q\log x}\right)} \frac{1}{x}, \text{ if } Q \neq 0, \\ -\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\log x)^2} \frac{1}{x}, \text{ if } Q = 0. \end{cases} \tag{30}$$

$G_Q^{'}(x) < 0$ holds. $\lim_{x\to\infty} G_Q^{'}(x) = 0, \lim_{x\to 0+} G_Q^{'}(x) = 0$ for any Q. Hence $G_Q^{'}(x)$ is uniformly bounded.

$$G_Q^{''}(x) = \begin{cases} \frac{|Q|\left(Q^{-2}\right)^{Q^{-2}}}{\Gamma(Q^{-2})} e^{Q^{-2}\left(Q\log x - e^{Q\log x}\right)} \frac{1}{x^2}(1 - Q^{-1} + Q^{-1}e^{Q\log x}), \text{ if } Q \neq 0, \\ \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\log x)^2} \frac{1}{x^2}(1 + \log x), \text{ if } Q = 0. \end{cases} \tag{31}$$

$$G_Q^{(3)}(x) = \begin{cases} \frac{|Q|\left(Q^{-2}\right)^{Q^{-2}}}{\Gamma(Q^{-2})} e^{Q^{-2}\left(Q\log x - e^{Q\log x}\right)} \frac{1}{x^3} \times \\ \left[\frac{3}{Q}(1 - e^{Q\log x})(1 - \frac{1}{3Q}(1 - e^{Q\log x}) + e^{Q\log x} - 2\right], \text{ if } Q \neq 0, \\ -\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\log x)^2} \frac{(\log x)^2 + 3\log x - 1}{x^3}, \text{ if } Q = 0. \end{cases} \tag{32}$$

The uniformly boundedness of $G_Q^{''}(x)$ and $G_Q^{(3)}(x)$ can be proved similarly and $G_Q^{(3)}(x)$ is continuous.

$$\dot{G}_Q(x) = \begin{cases} -\frac{\partial\Gamma_I\left(\frac{1}{Q^2}, \frac{e^{Q\log x}}{Q^2}\right)}{\partial Q}, \text{ if } Q > 0, \\ 0, \text{ if } Q = 0, \\ \frac{\partial\Gamma_I\left(\frac{1}{Q^2}, \frac{e^{Q\log x}}{Q^2}\right)}{\partial Q}, \text{ if } Q < 0, \end{cases} \tag{33}$$

where

$$\frac{\partial\Gamma_I\left(\frac{1}{Q^2}, \frac{e^{Q\log x}}{Q^2}\right)}{\partial Q}$$

$$=\frac{1}{\Gamma\left(\frac{1}{Q^2}\right)}\left(\frac{e^{Q\log x}}{Q^2}\right)^{\frac{1}{Q^2}-1}e^{-\frac{e^{Q\log x}}{Q^2}}\frac{e^{Q\log x}(\log x)Q^2-2e^{Q\log x}Q}{Q^4}$$
$$+\frac{2}{Q^3}\left[\left(Q\log x-\log Q^2\right)\Gamma\left(\frac{1}{Q^2},\frac{e^{Q\log x}}{Q^2}\right)+\frac{e^{Q\log x}}{Q^2}T\left(3,\frac{1}{Q^2},\frac{e^{Q\log x}}{Q^2}\right)\right]\frac{1}{\Gamma\left(\frac{1}{Q^2}\right)}$$
$$-\frac{2}{Q^3}\Gamma\left(\frac{1}{Q^2},\frac{e^{Q\log x}}{Q^2}\right)\Psi^{(0)}\left(\frac{1}{Q^2}\right)\frac{1}{\Gamma\left(\frac{1}{Q^2}\right)}, \tag{34}$$

function $\Gamma(a,x)=\int_x^\infty s^{a-1}e^{-s}ds$, the function $T(m,s,x)$ is a special case of the Meijer G-function

$$T(m,s,x)=G^{m,0}_{m-1,m}\left(\begin{smallmatrix}0,0,\ldots,0\\ s-1,-1,\ldots,-1\end{smallmatrix}\,\middle|\,x\right), \tag{35}$$

and function $\Psi^{(n)}(x)$ is the polygamma function of order n,

$$\Psi^{(n)}(x)=\frac{d^{n+1}}{dx^{n+1}}\log\Gamma(x). \tag{36}$$

$\lim_{Q\to 0}\dot{G}_Q(x)=0$ so $\dot{G}_Q(x)$ is continuous. Using the property that

$$G^{m,0}_{m-1,m}\left(\begin{smallmatrix}0,0,\ldots,0\\ s-1,-1,\ldots,-1\end{smallmatrix}\,\middle|\,x\right)=0,\ \text{ if } |x|>1, \tag{37}$$

we get $\lim_{x\to\infty}\dot{G}_Q(x)=0, \lim_{x\to 0+}\dot{G}_Q(x)=0$ for any Q. Hence $\dot{G}_Q(x)$ is uniformly bounded.

Using the following properties of gamma function, incomplete gamma function, polygamma function and Meijer G-function,

$$\Gamma^{'}(x)=\Gamma(x)\Psi^{(0)}(x), \tag{38}$$
$$\frac{\partial\Gamma(a,x)}{\partial x}=-x^{a-1}e^{-x}, \tag{39}$$
$$\frac{\partial\Gamma(a,x)}{\partial a}=\log x\Gamma(a,x)+xT(3,a,x), \tag{40}$$
$$\frac{\partial T(m,s,x)}{\partial s}=\log xT(m,s,x)+(m-1)T(m+1,s,x), \tag{41}$$
$$\frac{\partial T(m,s,x)}{\partial x}=-\frac{1}{x}\left[T(m-1,s,x)+T(m,s,x)\right], \tag{42}$$

the second and third derivatives of $S_Q(x)$ with respect to Q can be calculated tediously but easily. And the existence and uniformly boundedness of them can be verified

similarly, taking advantage of (37). Hence the statement (4') holds.
By applying Theorem 1 and Remark A.1 in Zeng et al. (2006), we reach the conclusion that $\hat{\beta}, \hat{Q}$ and $\hat{H}$ are strongly consistent. $\square$

The following theorem establishes the asymptotic normality of $\hat{\beta}, \hat{Q}$ and $\hat{H}$.

Theorem 3. *Assume Q_0 belongs to the interior of a compact set $\mathfrak{Q}$ in Euclidean space. Under Conditions (1)-(3), $\sqrt{n}(\hat{\beta}-\beta_0, \hat{H}-H_0)$ converges weakly to a zero-mean Gaussian process. Furthermore, $\hat{\beta}$ is efficient and its asymptotic variance attains the semiparametric efficiency bound for β_0.*

Proof. We have shown that statement (4') holds in our model in the proof of above theorem. Next, we will show the following statements (4) and (5) hold.
(4): G_Q satisfies

$$G_Q(0) = 1, G_Q(x) > 0, G_Q'(x) < 0, G_Q^{(3)}(x) \text{ exists and is continuous,}$$

where $G_Q^{(3)}(x)$ is the third derivative of $G_Q(x)$.
(5): If with probability one,

$$G_Q'(e^{\beta^T X})e^{\beta^T X}X^T\boldsymbol{a_1} + \dot{G}_Q(e^{\beta^T X})a_2 = 0,$$

where $\boldsymbol{a_1}$ is a constant vector, a_2 is a constant and $\dot{G}_Q$ denotes the derivative with respect to Q, then $\boldsymbol{a_1} = \mathbf{0}$ and $a_2 = 0$.
Statement (4) can be verified easily. For statement (5),

$$\begin{aligned} &G_Q'(e^{\beta^T X})e^{\beta^T X}X^T\boldsymbol{a_1} + \dot{G}_Q(e^{\beta^T X})a_2 \\ &= \begin{cases} -\frac{|Q|\left(Q^{-2}\right)^{Q^{-2}}}{\Gamma(Q^{-2})}e^{Q^{-2}\left(Q\beta^T X - e^{Q\beta^T X}\right)}X^T\boldsymbol{a_1} + \frac{\partial\Gamma_I\left(\frac{1}{Q^2}, \frac{e^{Q\beta^T X}}{Q^2}\right)}{\partial Q}a_2, \text{ if } Q < 0 \\ -\frac{|Q|\left(Q^{-2}\right)^{Q^{-2}}}{\Gamma(Q^{-2})}e^{Q^{-2}\left(Q\beta^T X - e^{Q\beta^T X}\right)}X^T\boldsymbol{a_1} - \frac{\partial\Gamma_I\left(\frac{1}{Q^2}, \frac{e^{Q\beta^T X}}{Q^2}\right)}{\partial Q}a_2, \text{ if } Q > 0. \end{cases} \end{aligned} \tag{43}$$

We have

$$-\frac{|Q|\left(Q^{-2}\right)^{Q^{-2}}}{\Gamma\left(Q^{-2}\right)}e^{Q^{-2}\left(Q\beta^T X - e^{Q\beta^T X}\right)}X^T \neq 0, \frac{\partial\Gamma_I\left(\frac{1}{Q^2}, \frac{e^{Q\beta^T X}}{Q^2}\right)}{\partial Q} \neq 0 \text{ a.e..} \tag{44}$$

Hence statement (5) holds. By applying Theorem 2 and Remark A.2.1 in Zeng et al.

(2006), the asymptotic normality of $\hat{\beta}, \hat{Q}$ and $\hat{H}$ holds. $\square$

In summary, Theorem 2 and Theorem 3 ensure the MLE of β in our model is consistent and asymptotically efficient; the MLE of Q and H is consistent and asymptotically normal.

4 Simulation Studies

We carried out simulation studies to examine the finite sample performance of our method. In the simulation, the underlying regression model was

$$\log T = -\beta_1 X_1 - \beta_2 X_2 + \epsilon, \tag{45}$$

where X_1 was a uniformly distributed random variable in $[0, 1]$, X_2 was a Bernoulli(0.5) random variable and e^{ϵ} followed a $GG(0, 1, Q)$ distribution. The true parameters were $\beta_1 = -0.5$, $\beta_2 = -1$ and $Q \in \{0.5, 1\}$. When $Q = 1$, the true model reduces to proportional hazard model. The censoring times were independently generated from a uniform distribution in order to have approximately 15% and 30% censoring rates. The sample size was chosen to be $n = 100$, 200 and 500. For each configuration, we simulated $1,000$ datasets. Then we applied the above method to obtain the maximum likelihood estimates of β_1, β_2 and Q.
For each dataset, we chose initial values $\beta_1 \in \{-1, 0, 1\}, \beta_2 \in \{-1, 0, 1\}, Q \in \{-1, 0.5, 1\}$ and then picked out the estimation results with largest likelihood values. The estimates we got were robust with respect to the initial values. The estimation results are summarized in Table 1 and Table 2 , where the column labeled as "estimate" denotes the average values of the parameter estimates over 1,000 replications; "SE" is the standard deviation of the estimates; and "CP(%)" is the coverage proportion of 95% confidence intervals constructed based on the asymptotic normal approximation.

Table 1: Simulation Results under $\log T = -\beta_1 X_1 - \beta_2 X_2 + \epsilon$, where $e^{\epsilon} \sim GG(0, 1, Q)$

censoring	n	true value	estimate	SE	CP(%)
15%	100	$\beta_1 = -0.5$	-0.517	0.395	94.3
		$\beta_2 = -1$	-1.052	0.249	94.7
		$Q = 1$	0.609	0.426	89.0
15%	200	$\beta_1 = -0.5$	-0.522	0.270	94.9
		$\beta_2 = -1$	-1.042	0.189	96.7
		$Q = 1$	0.744	0.339	96.5
15%	500	$\beta_1 = -0.5$	-0.521	0.169	95.1
		$\beta_2 = -1$	-0.999	0.084	94.3
		$Q = 1$	0.991	0.223	95.5
30%	100	$\beta_1 = -0.5$	-0.512	0.455	95.2
		$\beta_2 = -1$	-1.004	0.269	95.0
		$Q = 1$	0.509	0.500	95.0
30%	200	$\beta_1 = -0.5$	-0.498	0.302	95.3
		$\beta_2 = -1$	-0.998	0.190	95.8
		$Q = 1$	0.635	0.375	95.1
30%	500	$\beta_1 = -0.5$	-0.532	0.198	94.7
		$\beta_2 = -1$	-1.068	0.098	95.1
		$Q = 1$	0.993	0.241	96.2

Table 2: Simulation Results under $\log T = -\beta_1 X_1 - \beta_2 X_2 + \epsilon$, where $e^\epsilon \sim GG(0, 1, Q)$

censoring	n	true value	estimate	SE	CP(%)
15%	100	$\beta_1 = -0.5$	-0.517	0.407	95.0
		$\beta_2 = -1$	-1.029	0.236	94.8
		$Q = 0.5$	0.251	0.410	94.1
15%	200	$\beta_1 = -0.5$	-0.500	0.264	95.0
		$\beta_2 = -1$	-1.023	0.163	94.3
		$Q = 0.5$	0.365	0.318	94.0
15%	500	$\beta_1 = -0.5$	-0.493	0.184	95.2
		$\beta_2 = -1$	-1.098	0.163	94.9
		$Q = 0.5$	0.483	0.099	95.1
30%	100	$\beta_1 = -0.5$	-0.527	0.426	95.3
		$\beta_2 = -1$	-1.050	0.268	95.0
		$Q = 0.5$	0.179	0.463	95.4
30%	200	$\beta_1 = -0.5$	-0.499	0.297	95.5
		$\beta_2 = -1$	-1.027	0.188	96.3
		$Q = 0.5$	0.320	0.347	95.3
30%	500	$\beta_1 = -0.5$	-0.513	0.198	95.0
		$\beta_2 = -1$	-1.015	0.120	94.9
		$Q = 0.5$	0.498	0.214	95.2

The simulation results indicate that the LT-GG model performs well with sample sizes $n = 100, 200$ and 500, $Q = 0.5$ and 1 under censoring rates of 15% and 30%. The biases are small. The standard error increases when censoring rate increases and decreases when sample size increases.

5 Real Example

We applied the LT-GG model to a real data set of time to execution of limit buy orders of Microsoft from 9:30 a.m. to 10:30 a.m. on June 21, 2012. The original data sets were downloaded from lobsterdata.com and contained the evolution of the limit order book, as well as indicators for the type of event causing an update of the limit order book. The order reference numbers in the original data sets were used to calculate accurate time-to-execution of each order.

The data set contains 33912 observations, among which 2742 were non-censored. The covariates included in the analysis are the same as Lo et al. (2002). Let P_l[1] denote the limit-order price, S_l the limit-order size, P_b the best bid price, S_b the best bid size, P_o the best offer price, S_o the best offer size, P_q the mid-quote price and P the market price of the most recent transaction. The covariates are defined as follows:

$$
\begin{aligned}
MQLP &= P_q - P_l, \\
BSID &= \begin{cases} 1 & \text{if prior trade occurred above } P_q, \\ 0 & \text{if prior trade occurred at } P_q, \\ -1 & \text{if prior trade occurred below } P_q, \end{cases} \\
MKD1^{2} &= \begin{cases} (1 + P_b - P_l)\log S_b & \text{if } P_l \leq P_b, \\ 0 & \text{if } P_l > P_b, \end{cases} \\
MKD1X &= \begin{cases} (P - P_l)MKD1 & \text{if } P \geq P_l, \\ 0 & \text{if } P < P_l, \end{cases} \\
MKD2 &= \begin{cases} \log S_o/(1 + P_o - P_l) & \text{if } P_o \geq P_l, \\ \log S_o & \text{if } P_o < P_l, \end{cases} \\
SZSD &= \begin{cases} \log S_l/(1 + P_o - P_l) & \text{if } P_o > P_l, \\ \log(S_l - S_o) & \text{if } P_o = P_l \text{ and } S_l > S_o, \\ 0 & \text{otherwise}, \end{cases} \\
STKV &= \#\text{trades last half hour}/\#\text{trades last one hour}, \\
TURN &= \log(\#\text{trades last one hour}),
\end{aligned}
$$

We first fitted the AFT-GG model to the data set. The estimation results are summarized in Table 3 and Table 4.

[1]The price in original data sets downloaded from lobsterdata.com is the actual price multiplied by 1e+04. Tick size of Microsoft stock is \$0.01. We divided the price by 100 to get P_l such that the minimum change in P_l is \$1. Similar preprocessing was done for P_b, P_o, P_q and P.

[2]We add $P_b - P_l$ by 1 to make MKD1 different from 0 when $P_l \leq P_b$. Similar technique is used in covariate MKD2 and SZSD.

Table 3: Maximum Likelihood Estimates Under the AFT-GG Model for the Stock Data

Covariate	Estimate	Standard Error	p value
Intercept	-8.4529	0.5791	<.0001
mqlp	-71.7222	48.5170	0.1393
bsid	-0.0369	0.0506	0.4666
mkd1	-17.7018	7.5432	0.0189
mkd1x	52.1361	9.9011	<.0001
mkd2	18.9035	0.7731	<.0001
szsd	-3.8614	0.4813	<.0001
stkv	4.8704	0.2525	<.0001
turn	-6.9832	0.5282	<.0001
Scale	2.6364	0.0856	
Shape	0.6499	0.0470	

Table 4: Fit Statistics Under the AFT-GG Model for the Stock Data

Fit Statistics	Value
-2 Log Likelihood	27276.34
AIC	27298.34
BIC	27391.09

Next, we fitted the LT-GG model to the stock data. Since we fixed location parameter $\mu = 0$ and shape parameter $\sigma = 1$ in the underlying distribution, we used estimates divided by estimated scale parameter under the AFT-GG model as initial values for computation. The estimation results are summarized in Table 5 and Table 6.

Table 5: Maximum Likelihood Estimates Under the LT-GG Model for the Stock Data

Covariate	Estimate	Standard Error	p value
mqlp	-250.001	70.3890	0.0004
bsid	-11.511	0.0634	$<$ .0001
mkd1	-49.989	12.5740	$<$ .0001
mkd1x	10.001	449.7666	0.9823
mkd2	70.147	0.5247	$<$ .0001
szsd	-1.983	0.4365	$<$ .0001
stkv	-0.167	0.1655	0.3129
turn	-4.965	0.2501	$<$ .0001
Q	0.872	0.4738	
h1	0.0009	0.0004	

Table 6: Fit Statistics Under the LT-GG Model for the Stock Data

Fit Statistics	Value
-2 Log Likelihood	60840.17
AIC	66342.17
BIC	89537.23

Table 3 and Table 5 indicate that the sign and scale of parameters under the LT-GG and AFT-GG models are generally close. However, there is no significant evidence to show that the covariate stkv, which is a proxy for high frequency changes in volatility (Lo et al.; 2002), is an important predictor of time-to-execution of limit buy orders in the LT-GG model. A more appropriate definition of volatility may be explored in further research. The covariate mkd1x is also not significant in the LT-GG model, which may be due to multicollinearity between covariates. Besides, covariates mqlp, bsid and mkd1 are significant under the LT-GG model, which indicates that the proposed LT-GG model may be better at identifying important features than the AFT-GG model. In addition, Lo et al. (2002) and Wen (2009) reported that the AFT-GG model performs better than Cox model. Here, Table 5 shows that the estimated value of parameter Q is marginally different from 1.

6 Cross Validation

We performed 5-fold cross validation on the LT-GG and AFT-GG models using the stock data. Censoring rate of each fold were the same. The covariates included in the analysis were the same as previous chapter.

We calculated forecasting residuals of each folder. For each folder $\mathcal{F}$ the residual is defined to be

$$\text{res} = \sum_{i \in \mathcal{F}} \delta_i \|\phi(T_i) - \widehat{\phi(T_i)}\|^2, \tag{46}$$

where $\phi(T_i) = \log(T_i)$ in AFT-GG model and $\phi(T_i) = \log \hat{H}(T_i)$ in the LT-GG model with $\hat{H}$ computed based on data in $\mathcal{F}$, $\widehat{\phi(T_i)} = -\hat{\beta}^T X_i$ with $\hat{\beta}$ computed based on data not in the folder $\mathcal{F}$.

The forecasting residuals of each folder are summarized in Table 7.

Table 7: Cross Validation Results Under the AFT-GG and LT-GG models for the Stock Data

Forecasting Residual	AFT-GG Model	LT-GG Model
Fold 1	13575.58	3027.89
Fold 2	12622.99	2159.75
Fold 3	10970.14	5104.64
Fold 4	14225.38	4585.61
Fold 5	13505.88	5485.31

The cross validation results indicate that the LT-GG model performs much better on the stock data than the AFT-GG model, as the forecasting residuals are reduced by 2-6 times.

To conclude, compared with the the AFT-GG model, the LT-GG model is the best in forecasting execution time of limit orders.

7 Summary

We have introduced a new LT-GG model and provided an algorithm to calculate the MLE using CFSQP method. The asymptotic properties are justified under several regularity conditions. In our real data study, we first fitted both AFT-GG and LT-GG models and compared their results. Cross validation were carried out and we reached

the conclusion that the LT-GG model is comparatively more accurate in predictions of time to execution of limit orders. The forecasting residuals are reduced by 4-5 times under the by the proposed LT-GG model.
The new LT-GG model may have important implications when applied to the analysis of trading processes. It is expected to provide more accurate predictions on execution time of limit orders and identify the key features that affect the execution time, which can help market participants design better trading strategies. It may also offer a more detailed look into the price discovery process and provide more accurate justification of market efficiency.

References

Chen, K., Jin, Z. and Ying, Z. (2002). Semiparametric analysis of transformation models with censored data, *Biometrika* **89**(3): 659–668.

Cox, D. R. (1972). Regression models and life tables (with discussion), *Journal of the Royal Statistical Society* **34**: 187–220.

Gava, L. (2005). The speed of limit order execution in the spanish stock exchange.

Greene, W. H. (2003). *Econometric analysis*, Pearson Education India.

Lo, A. W., MacKinlay, A. C. and Zhang, J. (2002). Econometric models of limit-order executions, *Journal of Financial Economics* **65**(1): 31–71.

Wen, Q. (2009). *Limit-order completion time in the London stock market*, PhD thesis, Heriot-Watt University.

Yin, G. and Zeng, D. (2012). Efficient algorithm for computing maximum likelihood estimates in linear transformation models, *Journal of Computational and Graphical Statistics* .

Zeng, D. and Lin, D. (2006). Efficient estimation of semiparametric transformation models for counting processes, *Biometrika* **93**(3): 627–640.

Zeng, D. and Lin, D. (2007). Maximum likelihood estimation in semiparametric models with censored data (with discussion), *Journal of the Royal Statistical Society B* .

Zeng, D. and Lin, D. (2009). A generalized asymptotic theory for maximum likelihood estimation in semiparametric regression models with censored data, *Statistica Sinica* .

Zeng, D., Yin, G. and Ibrahim, J. G. (2006). Semiparametric transformation models for survival data with a cure fraction, *Journal of the American Statistical Association* **101**(474): 670–684.