12-2003

Fiducial predictive densities and econometric duration analysis

Zhenlin YANG
Singapore Management University, zlyang@smu.edu.sg

Fiducial predictive densities and econometric duration analysis

Article · January 2003

1 author:

Zhenlin Yang
Singapore Management University

Fiducial Predictive Densities and Econometric Duration Analysis

Zhenlin Yang*

School of Economics and Social Sciences, Singapore Management University
469 Kukit Timah Road, Singapore 259756

December 8, 2003

Abstract

In this article, we propose using fiducial predictive density (FPD) as a density estimate, which leads naturally to estimators of survivor and hazard functions and provides a simple way of constructing shortest prediction intervals. This approach is studied in detail in the context of two flexible duration models proposed in this paper, namely the *trans-normal* and *trans-exponential* families, by presenting the FPDs, their basic properties, their Bayesian correspondence and their applications in econometric duration analysis. Empirical evidences show that the FPD method provides better estimates of survivor and hazard functions, particularly the latter, than does the usual maximum likelihood method. It provides shortest prediction intervals for a future duration, which can be much shorter than the regular equitailed prediction intervals. The trans-normal model has an easy extension to include exogenous variables, whereas the trans-exponential model allows for the analysis of censored data. Finally, when the transformation function is indexed by unknown parameter(s), the FPD method still provides asymptotically correct inference when the transformation parameter is replaced by its estimator.

Keywords: Bayesian correspondence, Censored data, Fiducial prediction, Hazard estimate, Shortest prediction interval, Survivor function, Trans-exponential, Trans-normal.

*Corresponding author. Tel.: +65-6822-0852; fax: +65-6822-0833; *E-mail address:* zlyang@smu.edu.sg (Z. L. Yang)

1 INTRODUCTION

Prediction is an important task for the analysis of economic duration data[1], such as the strike duration, lengths of spells of unemployment, duration of marriage, lifetimes of firms, lifetimes of products, etc. Typical prediction problems include prediction of a future duration, prediction of the median or a general percentile duration, density prediction[2], survivor function prediction, hazard function prediction, etc. Parametric methods for prediction can be generally classified as non-Bayesian and Bayesian. Non-Bayesian methods include the classical method based on a predictive pivot, method based on the maximum likelihood predictive density (Lejeune and Faulkenberry, 1982), method based on the predictive likelihood (Butler, 1986, 1989), and the method based on the fiducial argument (Fisher, 1973), etc. Bayesian method for prediction is through a Bayesian predictive density (Geisser, 1993). It seems that the fiducial approach to prediction has not received much attention, especially in econometric applications, and it is the purpose of this article to explore the applicability of fiducial approach to econometric duration analysis.

In this article, we propose using the fiducial predictive density (FPD) to estimate a probability density function (pdf) which leads naturally to estimators of density function, survivor function (sf) and hazard function (hf). It also provides a simple way to construct the prediction or shortest prediction intervals for a future duration. This approach is explored in detail in the context of two flexible duration models proposed in this paper, namely the *trans-normal* and the *trans-exponential* families, by presenting the FPDs, their basic properties, their Bayesian correspondence and their various applications in econometric duration analysis. The results obtained in this paper indicate that the FPD method is very attractive: it is rather simple; it not only provides the same prediction interval as the classical approach, but also gives a shorter or the shortest possible prediction intervals; and it gives rise to density, survivor and hazard rate estimators that are less biased and

[1]Analysis of economic duration data may be one of the rapidly growing area of econometric research (Greene, 2000, p. 937). Representative works include Heckman and Singer (1984a, b), Kennan (1985), Kiefer (1988), Sider (1985), Lancaster (1972, 1979, 1985, 1990), Ryu (1993), Torelli and Trivellato (1993), Koop and Ruhm (1993), Saha and Hilton (1997), Baker and Melino (2000), and Zhang, *et al.* (2001).

[2]Density forecast may be another fast growing area of econometric research. See Tay and Wallis (2000) for a survey.

less variable than the usual maximum likelihood estimators (MLEs). The generality and extendibility of the FPD method can be seen from the facts that the trans-normal model is easily extendable to include exogenous variables, and the trans-exponential model allows for the analysis of censored data. Moreover, when the transformation function is indexed by unknown parameter(s), the FPD method still provides asymptotically correct inference when the transformation parameter is replaced by its estimator. This greatly expands the applicability of the FPD method.

The paper is organized as follows. Section 2 gives a brief introduction to the fiducial approach to prediction. Section 3 is a detailed exploration of the FPD method based on the *trans-normal* family. Section 4 studies the *trans-exponential* family, where more attention is given to the analysis of censored duration data. Section 5 extends the results of the trans-normal model to include the effect of exogenous variables. Section 6 studies the case of an unknown transformation and shows that all the results remain asymptotically valid when the transformation function is estimated. Section 7 presents some Monte Carlo simulation results to show the finite sample properties of the estimates. Section 8 presents some numerical examples to illustrate the methods. Section 9 concludes the paper.

2 FIDUCIAL PREDICTION

Let $\mathbf{Y} = \{Y_1, Y_2, ..., Y_n\}$ be a sample of past observations and Y_0 be a future observation both from a population with pdf $f(\cdot \mid \theta)$. We are interested in predicting Y_0 based on $\mathbf{Y}$. The classical method for prediction is to first find a *predictive pivot*, denoted by $q(Y_0, \mathbf{Y})$, which is a function of $\mathbf{Y}$ and Y_0 with the quantity itself and its distribution free of the unknown parameters, set up an probability inequality for the predictive pivot and then invert to give a prediction interval for Y_0.

Fiducial approach to prediction starts from a predictive pivot. This approach is best understood by taking a simple example where $\mathbf{Y}$ and Y_0 are from an exponential population with mean θ. It is well known that $Y_0/\theta \sim \chi^2_2$ and $\sum_{i=1}^n Y_i/\theta \sim \chi^2_{2n}$. Thus, $q(Y_0, \mathbf{Y}) = nY_0/\sum Y_i \sim F_{2,2n}$ with pdf $(1+q)^{-(n+1)}$. Fiducial argument then changes the gear to view only Y_0 in $q(Y_0, \mathbf{Y})$ as random and use a change of variable technique to have a fiducial distribution of Y_0 as $n(\sum y_i)^{-1}(1 + ny_0/\sum y_i)^{-(n+1)}$.

In general, if the predictive pivot $q(Y_0, \mathbf{Y})$ is a one-to-one function of Y_0 and has a pdf g, then a *fiducial predictive density* (FPD) of Y_0 given $\mathbf{Y} = \mathbf{y}$ is defined as

$$p_{\mathrm{F}}(y_0 \mid \mathbf{y}) = g[q(y_0, \mathbf{y})]q_y(y_0, \mathbf{y}), \tag{1}$$

where $q_y(y_0, \mathbf{y})$ is the derivative of q with respect to y_0 (Fisher, 1973, p117). Among many greatest contributions of R. A. Fisher to statistical foundations, fiducial argument has received the least attention. Zabell (1992) gave an excellent account on the history of the Fiducial argument[3].

Paralleled with the classical frequentist approach and the fiducial approach to prediction is the Bayesian approach based on the so-called the *Bayesian predictive density* (BPD) of Y_0 that is defined as

$$p_{\mathrm{B}}(y_0 \mid \mathbf{y}) = \int_\theta p(\theta \mid \mathbf{y}) f(y_0 \mid \theta) d\theta \tag{2}$$

where $p(\theta \mid \mathbf{y})$ is the posterior distribution of θ given $\mathbf{Y} = \mathbf{y}$ (Geisser, 1993, p49).

From the definitions, it is easy to see the connection between the frequentist and the FPD approaches for the prediction problems, but it is rather difficult to see any possible linkage between the FPD and BPD. This paper shows that a close correspondence exists between the two in the framework of a trans-normal family, a trans-exponential family, and a general trans-normal regression model. It is conjectured that such a correspondence exists in general as long as a predictive pivot with the desired property can be found.

3 FPD FOR TRANS-NORMAL FAMILY

3.1 The Trans-Normal Family

Definition 1. A family of distributions is called the *trans-normal* family if the random variable Y is such that $h(Y) \sim N(\mu, \sigma^2)$ for some $h : B \to R$, a monotonic increasing and differentiable function with range R, the whole real line, and domain B, a subset of R.

Clearly, the pdf of the trans-normal distribution has the form:

[3]More on the Fiducial arguments can be found in Lindley (1958), Seidenfeld (1979, 1992), Dawid and Stone (1982), Dawid and Wang (1993) and Barnard (1995).

$$f(y;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2\sigma^2}[h(y)-\mu]^2\right\}h'(y), \tag{3}$$

where $h'(y) = dh(y)/dy$. The domain B could be the positive half real line as in the case of duration data and a bounded interval as in the case of percentage or proportions. The trans-normal family is seen to be a very rich family. It contains popular distributions such as normal with $h(y) = y$, and lognormal with $h(y) = \log y$. It also covers several sub families such as the ξ-normal family (Saunders, 1974) where $h(y)$ satisfies $h(y) = -h(y^{-1}), y > 0$ and $\alpha^{-1}h(y/\beta) \sim N(0,1)$, and the Box and Cox (1964) power family:

$$h(y) = \begin{cases} (y^\lambda - 1)/\lambda, & \lambda \neq 0, \\ \log y, & \lambda = 0, \end{cases} \quad y > 0. \tag{4}$$

When $\lambda \neq 0$, the Box-Cox power transformation is bounded either below or above depends on whether λ is positive or negative. Thus, exact normality for $h(Y)$ is not possible. In practice, one often takes the normality as an approximation. To overcome this difficulty, Yang (2002) proposed a modified power transformation:

$$h(y) = \begin{cases} (y^\lambda - y^{-\lambda})/2\lambda, & \lambda > 0, \\ \log y, & \lambda = 0, \end{cases} \quad y > 0. \tag{5}$$

For nonnegative y, this function is one-to-one with its inverse $y = [\lambda h + (1+\lambda^2 h^2)^{1/2}]^{1/\lambda}$. Note that the distributional family generated by modified power transformation generalizes the ξ-normal family of Saunders (1974). We now give some general theoretical properties of the trans-normal family.

Theorem 1. Let $f(y) = f(y;\mu,\sigma)$ be the pdf of a trans-normal random variable Y defined in (3). Assume $h(y)$ is monotonically increasing with the first two derivatives $h'(y)$ and $h''(y)$ exist. Then $f(y)$

i) is a monotonic function of y if $m(y) = 0$ does not have a real root;

ii) is a unimodal pdf if $m(y) = 0$ has a unique real root in the interior of B;

iii) has two stationary points if $m(y) = 0$ has two real roots;

iv) is bimodal if $m(y) = 0$ has three real roots, etc.,

where $m(y) = h''(y)/h'^2(y) - \sigma^{-2}(h(y)-\mu)$.

Proof. Let $k(y) = \exp\{-[h(y)-\mu]^2/(2\sigma^2)\}$. Then, $f(y) \propto k(y)h'(y)$ and $f'(y) = k(y)h'^2(y)[h''(y)/h'^2(y) - \sigma^{-2}(h(y)-\mu)] = k(y)h'^2(y)m(y)$. Since the function $k(y)h'^2(y)$ is a positive function of y, how many times that $f'(y)$ changes its sign as y changes depends on how many real roots that $m(y) = 0$ has, which determines the behavior of $f(y)$. The results of the theorem thus follows.

Note that the case (i) in Theorem 1 rarely happens, case (ii) is the most typical case and it happens as long as $f(y)$ vanishes at both ends and $h''(y)/h'^2(y)$ is monotonic. The cases (iii) is also not common and (iv) can happen for certain special functions at certain parameter settings.

To illustrate the versatility and usefulness of the trans-normal distribution, we pick a special modified power transformation, $h(y) = y^{.5} - y^{-.5}$, and plot the pdf, the survivor function (sf) and the hazard function (hf) for serval parameter configurations. From the plots summarized in Figure 1, we see that the pdf of this trans-normal distribution has all kinds of shapes: it can be nearly symmetric, bimodal, or very skewed depending whether σ is small, medium, or large relative to the mean of Y. When σ is small relative to the mean, the pdf has one bump at the center part; as σ increases, another bump shows up at the left of the center and as σ further increases, the first bump disappeared and the distribution becomes unimodal again. Figure 1 also exhibits serval shapes of hazard function, including the interesting 'bath-tub' shape, which has a popular engineering interpretation: first bump represents the 'burn-in' period, the center flat part represents the 'stable period' and the second bump represents the 'wear-out' period. Econometricians call this the U-shaped hazard (Kiefer, 1988) and some evidence for its existence is provided by Kennan (1985) from the analysis of the strike duration data. It is interesting to note that when σ is large, the hf has a sharp increase at the very beginning and then quickly becomes flat for a long period of 'time'. This exactly reflects the failure mechanisms of certain engineering systems and electronic components which are very fragile at the very beginning, but once stabilized, can last for a very long period of time.

Figure 1 near here

3.2 Fiducial Predictive Densities

Consider the case where h is completely specified. Let $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} h(Y_i)$ and $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}[h(Y_i) - \hat{\mu}]^2$. Clearly, $\hat{\mu}$ and $\hat{\sigma}$ are the unrestricted MLEs of μ and σ, respectively.

When both parameters are unknown, we have the predictive pivot

$$q(Y_0, \mathbf{Y}) = \frac{h(Y_0) - \hat{\mu}}{\hat{\sigma}\sqrt{(n+1)/(n-1)}} \tag{6}$$

that is t-distributed with $n-1$ degrees of freedom, and the fiducial prediction density for Y_0 based on this pivot is

$$p_{\mathrm{F}}(y_0 \mid \mathbf{y}) = \frac{\Gamma(n/2)}{\sqrt{(n+1)\pi}\Gamma[(n-1)/2]} \left[1 + \frac{[h(y_0) - \hat{\mu}]^2}{(n+1)\hat{\sigma}^2}\right]^{-\frac{n}{2}} \frac{h'(y_0)}{\hat{\sigma}}. \tag{7}$$

When only σ is assumed unknown, putting $\hat{\sigma}^2(\mu) = \frac{1}{n}\sum_{i=1}^{n}[h(Y_i) - \mu]^2$, we have a predictive pivot, $[h(Y_0) - \mu]/\hat{\sigma}(\mu) \sim t_n$, and the fiducial predictive density

$$p_{\mathrm{F}}(y_0 \mid \mathbf{y}, \mu) = \frac{\Gamma[(n+1)/2]}{\sqrt{n\pi}\Gamma(n/2)} \left[1 + \frac{[h(y_0) - \mu]^2}{n\hat{\sigma}^2(\mu)}\right]^{-\frac{n+1}{2}} \frac{h'(y_0)}{\hat{\sigma}(\mu)}. \tag{8}$$

Finally, when only μ is assumed unknown, letting $\gamma = \sigma\sqrt{1 + n^{-1}}$, we have a predictive pivot, $(h(Y_0) - \bar{h})/\gamma \sim N(0, 1)$, and the fiducial predictive density

$$p_{\mathrm{F}}(y_0 \mid \mathbf{y}, \sigma) = \frac{1}{\sqrt{2\pi}\gamma} \exp\left[-\frac{1}{2\gamma^2}[h(y_0) - \bar{h}]^2\right] h'(y_0). \tag{9}$$

The FPD given by (9) is a trans-normal. The FPDs given in (7) and (8) have an identical structure, which is termed the trans-t distribution defined below.

Definition 2. A family of distributions is called the *trans-t* family with ν degrees of freedom and with parameters ξ and τ if the random variable T is such that $[h(T) - \xi]/\tau$ follows a t-distribution with ν degrees of freedom.

This definition is a generalization of the log-t distribution given in Dahiya and Guttman (1982), where they showed that a log-t pdf is either a purely decreasing function or a function with two stationary points. It can be further shown that the log-t pdf is essentially unimodal if the degrees of freedom ν is large relative to τ, which is often the case in practical applications. We give some general properties of the trans-t in the following theorem.

Theorem 2. Let $f(t; \xi, \tau)$ be the pdf of a trans-t random variable given by

$$f(t; \xi, \tau) = \frac{\Gamma[(\nu+1)/2]}{\sqrt{\nu\pi}\tau\Gamma(\nu/2)} \left[1 + \frac{[h(t) - \xi]^2}{\nu\tau^2}\right]^{-\frac{\nu+1}{2}} h'(t). \tag{10}$$

Assume that the h function satisfies the conditions of Theorem 1. Define $m(t) = h''(t)/h'^2(t) - \{(\nu - 1)[h(t) - \xi]\}/\{\nu\tau^2 + [h(t) - \xi]^2\}$. Then:

(i) $f(t;\xi,\tau)$ is a decreasing function of t if $m(t)=0$ has no real root; is unimodal if $m(t)=0$ has a unique real root; has two stationary points if $m(t)=0$ has two real roots; bimodal if $m(t)=0$ has three real roots, etc.

(ii) The trans-t pdf converges to a trans-normal pdf as $\nu\to\infty$.

Proof. Define $k(t)=\{1+[h(t)-\xi]^2/(\nu\tau^2)\}^{-(\nu+1)/2}$. Then, we have $f(t)=k(t)h'(t)$, which goes to zero as $t\to\infty$. Also,

$$f'(t)=k(t)h'^2(t)\left[\frac{h''(t)}{h'^2(t)}-\frac{(\nu-1)[h(t)-\xi]}{\nu\tau^2+[h(t)-\xi]^2}\right]=k(t)h'^2(t)m(t).$$

As both $k(t)$ and $h'^2(t)$ are positive functions, how many times that $f'(t)$ changes its sign depends on the number of real roots of the equation $m(t)=0$. Hence the part (i) of the theorem follows. The proof of part (ii) is straightforward.

From the expression of $m(t)$, it is easy to see that when $h''(t)/h'^2(t)=0$, $m(t)$ has only one root, and when $h''(t)/h'^2(t)=const$, $m(t)$ has two roots. The former corresponds to the linear transformation that gives a unimodal pdf and the latter the log transformation that gives a pdf with two stationary points. With the results of Theorems 1 and 2, the basic properties of the FPDs (7)-(9) are clear as they are either a trans-normal or a trans-t. The following theorem shows that they are consistent for estimating the true pdf of a future observation.

Theorem 3. The FPDs given by (7)-(9) are all consistent estimators of the true pdf of Y in the sense that as $n\to\infty$, each of the FPDs converges in probability to $f(y;\mu,\sigma)$ for each y.

Proof. First, for the case of both parameters unknown, it is easy to see that $\hat\mu$ and $\hat\sigma$ are root-n consistent estimators of μ and σ. Write $p_{\mathrm F}(y_0\mid\mathbf{y})=p_{\mathrm F}(y_0\mid\hat\mu,\hat\sigma)$. A first-order Taylor expansion around μ and σ gives:

$$p_{\mathrm F}(y_0\mid\hat\mu,\hat\sigma)=p_{\mathrm F}(y_0\mid\mu,\sigma)+(\partial p_{\mathrm F}/\partial\mu)(\hat\mu-\mu)+(\partial p_{\mathrm F}/\partial\sigma)(\hat\sigma-\sigma)+O_p(n^{-1})$$

As $n\to\infty$, $p_{\mathrm F}(y_0\mid\mu,\sigma)\to f(y_0;\mu,\sigma)$, and the two partial derivatives are bounded. Hence the results follows. The proofs for the other two cases are similar.

3.3 The Bayesian Correspondence

It is of interest to see how the FPD relate to the Bayesian predictive density. The results given in the following theorem shows that the FPDs for the trans-normal family corresponds to the Bayesian predictive density with flat priors.

Theorem 4. The FPD given in (7) coincides with the BPD with a flat prior $p(\mu, \sigma) \propto 1/\sigma$; the FPD given in (8) coincides with the BPD with a flat prior $p(\sigma) \propto 1/\sigma$ and the FPD given in (9) coincides with the BPD with a flat prior $p(\mu) \propto \mathit{const.}$.

Proof. To show that (7) is a Bayesian predictive density with a flat prior $p(\mu, \sigma) \propto 1/\sigma$, we first show it is true for the normal case, i.e., $h(y) = y$ and then generalize the result to the general trans-normal distribution by a change of variable technique. Let $Y_0 = Y_{n+1}$, and $\hat{\mu}_*$ and $\hat{\sigma}_*^2$ be the MLEs based on all the $n+1$ observations. We have,

$$(n+1)\hat{\mu}_* = \hat{\mu} + Y_0, \text{ and } (n+1)\hat{\sigma}_*^2 = n\hat{\sigma}^2 + n(Y_0 - \hat{\mu})^2/(n+1),$$

and Bayesian predictive density is

$$\begin{aligned}
p_{\text{B}}(y_0 \mid \mathbf{y}) &= \int_0^\infty \int_{-\infty}^\infty p(\mu, \sigma \mid \mathbf{y}) f(y_0 \mid \mu, \sigma) d\mu d\sigma \\
&\propto \int_0^\infty \int_{-\infty}^\infty \frac{1}{\sigma^{n+2}} \exp\left[-\frac{1}{2\sigma^2}\left\{\sum_{i=1}^n (y_i - \mu)^2 + (y_0 - \mu)^2\right\}\right] d\mu d\sigma \\
&\propto \int_0^\infty \int_{-\infty}^\infty \frac{1}{\sigma^{n+2}} \exp\left[-\frac{1}{2\sigma^2}\left\{(n+1)\mu^2 - 2\mu\sum_{i=1}^{n+1} y_i + \sum_{i=1}^{n+1} y_i^2\right\}\right] d\mu d\sigma \\
&\propto \int_0^\infty \frac{1}{\sigma^{n+2}} \left[\int_{-\infty}^\infty \exp\left\{-\frac{n+1}{2\sigma^2}(\mu - \hat{\mu}_*)^2\right\} d\mu\right] \exp\left\{-\frac{n+1}{2\sigma^2}\hat{\sigma}_*^2\right\} d\sigma \\
&\propto \int_0^\infty \frac{1}{\sigma^{n+1}} \exp\left\{-\frac{n+1}{2\sigma^2}\hat{\sigma}_*^2\right\} d\sigma \\
&= \frac{1}{2}\int_0^\infty \tau^{n/2-1} \exp\left\{-\frac{n+1}{2}\hat{\sigma}_*^2 \tau\right\} d\tau, \text{ where } \tau = 1/\sigma^2 \\
&\propto (\hat{\sigma}_*^2)^{-n/2} \propto \left[1 + \frac{(y_0 - \hat{\mu})^2}{(n+1)\hat{\sigma}^2}\right]^{-\frac{n}{2}}.
\end{aligned}$$

Integrating the last expression with respect to y_0 through a t density gives the normalizing constant and hence the BPD for Y_0 in the normal case. For a general h, the BPD for $h(Y_0)$ can be easily obtained by applying the change of variable technique, completing the proof.

With the definition of FPD and the conclusion of Theorem 4, it seems that the fiducial argument for prediction serves as a bridge between the classical (frequentist) and the Bayesian approaches to the prediction problems.

3.4 The Applications

We now discuss some interesting applications of the FPDs in econometric duration analysis, including the density estimation, survivor probability estimation, hazard rate estimation, constructing prediction intervals, etc. We will concentrate on the most realistic situation where both parameters are unknown. In each application, the simple maximum likelihood method is applicable, thus the FPD method is compared with the ML method.

Density estimation. Density estimation is one of the important topics in economic duration analysis. In the parametric setting, the simplest method may be the maximum likelihood where the unknown parameters in the pdf are replace by their MLEs. We now consider the FPD $p_{\mathrm{F}}(y_0 \mid \mathbf{y})$ as an estimator of the true pdf $f(y; \mu, \sigma)$, and to see it performance relative to the MLE of the true pdf. The MLE method completely ignores the effect of parameter estimation, whereas the FPD method takes the parameter estimation into account through the pivotal quantity and its distribution. In this sense, one would expect that FPD method performs better than the MLE method. Our simulation results given next show that it is indeed true.

Survivor probability. The survivor function of the trans-normal is:

$$S(t; \mu, \sigma) = 1 - \Phi\left[\frac{h(t) - \mu}{\sigma}\right]$$

where Φ denotes the CDF of the standard normal distribution. The MLE of $S(t; \mu, \sigma)$ is $\hat{S}(t) = S(t; \hat{\mu}, \hat{\sigma})$, where $\hat{\mu}$ and $\hat{\sigma}$ are the MLEs of μ and σ. Naturally, the survivor probability estimator based on the FPD should be

$$\tilde{S}(t) = 1 - \Psi_{n-1}\left[\frac{h(t) - \hat{\mu}}{\hat{\sigma}\sqrt{(n+1)/(n-1)}}\right]$$

where the last part is the CDF of $p_{\mathrm{F}}(y_0 \mid \mathbf{y})$ with Ψ_ν denoting the CDF of a t-distribution with ν degrees of freedom.

Hazard rate estimation. The hazard rate function of the trans-normal is

$$r(t; \mu, \sigma) = \frac{f(t; \mu, \sigma)}{1 - \Phi[(h(t) - \mu)/\sigma]}$$

The MLE of $r(t;\mu,\sigma)$ is $\hat{r}(t) = r(t;\hat{\mu},\hat{\sigma})$, and the corresponding FPD estimator is

$$\tilde{r}(t) = \frac{p_{\mathrm{F}}(t \mid \mathbf{y})}{1 - \Psi_{n-1}[(h(t)-\hat{\mu})/(\hat{\sigma}\sqrt{(n+1)/(n-1)})]}$$

Prediction and shortest prediction intervals. Prediction interval (PI) construction is also a important topic in life-testing and reliability studies, such as predicting the time required to perform a life-test and constructing a warranty limit, etc. As $h(Y_0)$ is normally distributed, an application of the standard method gives a prediction interval for $h(Y_0)$ and a inverse transformation gives the prediction interval for Y_0. Thus, a $100(1-\alpha)\%$ PI for Y_0 has the form:

$$h^{-1}\left\{\hat{\mu} \pm t_{n-1}(\alpha/2)\hat{\sigma}\sqrt{(n+1)/(n-1)}\right\} \tag{11}$$

Using the FPD, it is also fairly easy to construct prediction intervals for Y_0. A $100(1-\alpha)\%$ equitailed PI for Y_0 is defined as $\{L(\mathbf{Y}), \mathbf{U}(\mathbf{Y})\}$ such that

$$\Psi_{n-1}\left[\frac{h[L(\mathbf{Y})]-\hat{\mu}}{\hat{\sigma}\sqrt{(n+1)/(n-1)}}\right] = \frac{\alpha}{2} \text{ and } \Psi_{n-1}\left[\frac{h[U(\mathbf{Y})]-\hat{\mu}}{\hat{\sigma}\sqrt{(n+1)/(n-1)}}\right] = 1-\frac{\alpha}{2},$$

which, not surprisingly, gives a PI that is identical to (11). This means that the fiducial arguments gives inferences that have exact classical interpretations. An important feature of using FPD is that it allows to construct the shortest PI that is defined as follows.

Definition 3. Let Y_0 be a future observation and $\mathbf{Y}$ a sample of past observations both from a population with pdf $f(y)$. Let $\hat{f}(y)$ be in general a predictive density of the unknown $f(y)$. Then a $100(1-\alpha)\%$ shortest prediction interval for Y_0 based $\mathbf{Y}$ is defined as $\{L_s(\mathbf{Y}), U_s(\mathbf{Y})\}$ such that (a) $\int_{L_s(\mathbf{Y})}^{U_s(\mathbf{Y})} \hat{f}(t)dt = 1-\alpha$; and (b) for any $y_1 \in \{L_s(\mathbf{Y}), U_s(\mathbf{Y})\}$ and $y_2 \ni \{L_s(\mathbf{Y}), U_s(\mathbf{Y})\}$, $\hat{f}(y_2) \leq \hat{f}(y_1)$.

This definition is adapted from a definition in Dahiya and Guttman (1982). Based on the Definition 3, it is easy to see that if the predictive density is unimodal, then the condition (b) reduces to $\hat{f}(L_s(\mathbf{Y})) = \hat{f}(U_s(\mathbf{Y}))$.

4 FPD FOR TRANS-EXPONENTIAL FAMILY

Censoring is common in duration analysis. This issue is addressed in the context of the trans-exponential family defined below.

Definition 4. A family of distributions is called the *trans-exponential* family if the random variable Y is such that $h(Y) \sim \text{Exp}(\theta)$ for some $h : B \to R$, a monotonic increasing and differentiable function with range R^+, the positive half real line, and domain B, a subset of R.

With this definition, the pdf of the trans-exponential distribution has the form

$$f(y;\theta) = \frac{1}{\theta}\exp\left\{-\frac{1}{\theta}h(y)\right\}h'(y), \tag{12}$$

where h' denotes the derivative of h. This family contains the known distributions such as exponential ($h(y) = y, y > 0$), the log-exponential ($h(y) = \log(y), y > 1$), the Weibull ($h(y) = y^\beta, y > 0, \beta > 0$), etc.. Similar results as in Theorem 1 can be obtained for the trans-exponential distribution with $m(y) = h''(y)/h'^2(y) - 1/\theta$. That is, $f(y;\theta)$ is strictly decreasing if $m(y) = 0$ has no real root such as the exponential case, unimodal if it has a unique real root such as the Weibull case, etc..

Suppose now only the first r observations $Y_1 < Y_2 < \cdots < Y_r$ are available in a total sample of size n. Let again Y_0 be a future duration or lifetime. From Lawless (1982, Ch. 3), we have when h is completely specified the MLE of θ,

$$\hat{\theta} = \frac{1}{r}\left[\sum_{i=1}^{r} h(Y_i) + (n-r)h(Y_r)\right],$$

$2h(Y_0)/\theta \sim \chi^2_2$, and $2r\hat{\theta}/\theta \sim \chi^2_{2r}$. Hence, a predictive pivot is given as follow

$$q(Y_0, \mathbf{Y}) = \frac{h(Y_0)}{\hat{\theta}} \sim F_{2,2r}$$

where $F_{2,2r}$ denotes an F-distribution with 2 degrees of freedom in numerator and $2r$ degrees of freedom in the denominator. The FPD in this case has the form

$$p_F(y_0|\mathbf{y}) = \frac{1}{\hat{\theta}}\left[1 + \frac{h(y_0)}{r\hat{\theta}}\right]^{-(r+1)} h'(y_0) \tag{13}$$

It is easy to see that the FPD given in (13) is a BPD with prior $\theta \sim \frac{1}{\theta}$, and that it converges to the true pdf as n as well as r become large.

The survivor and hazard functions of the trans-exponential are, respectively, $\exp[-\frac{1}{\theta}h(y_0)]$ and $\frac{1}{\theta}h'(y_0)$; their MLEs are $\exp[-\frac{1}{\hat{\theta}}h(y_0)]$ and $\frac{1}{\hat{\theta}}h'(y_0)$; and their FPD estimators are $[1 + \frac{1}{r\hat{\theta}}h(y_0)]^{-r}$ and $\frac{1}{\hat{\theta}}h'(y_0)[1 + \frac{1}{r\hat{\theta}}h(y_0)]^{-1}$. The equitailed PI for Y_0 is $\{h^{-1}[\hat{\theta}F^{(1-\alpha/2)_{2,2r}}]$, $h^{-1}[\hat{\theta}F^{\alpha/2_{2,2r}}]\}$, and the shortest PI is also readily obtainable based on Definition 3.

The above results can be extended to the case of predicting the minimum Y_* of a future sample of size m as $mh(Y_*)/\theta \sim \chi_2^2$. The fiducial prediction based on progressive Type II censored data can also be handled in a similar way. See Balakrishnan and Basu (1995, p25) for the distributional results of the progressive Type II censored data.

5 FPD FOR TRANS-NORMAL REGRESSIONS

The FPD method for trans-normal model can easily be extended to incorporate the effect of exogenous variables. Let h again be a known generic monotonic increasing and differentiable function. The trans-normal regression model has the form

$$h(\mathbf{Y}) = \mathbf{X}\beta + \varepsilon \tag{14}$$

where $h(\mathbf{Y})$ is an $n \times 1$ vector of the transformed responses, $\mathbf{X}$ is an $n \times p$ design matrix of full column rank, β is a $p \times 1$ vector of regression coefficients, and ε is the error vector that is assumed to be $N(0_n, \sigma^2 I_n)$ with I_n being an $n \times n$ identity matrix.

The model (14) is essentially the Box-Cox transformation model (Box and Cox, 1964). It can be equivalent written as $h[y_p(\mathbf{x}_i)] = \mathbf{x}_i'\beta + \Phi(p)\sigma,\ i = 1, 2, ..., n$ where $\mathbf{x_i'}$ is the ith row of the design matrix $\mathbf{X}$, $y_p(\mathbf{x_i})$ is the pth quantile of Y at $\mathbf{x_i}$ and $\Phi(p)$ is the pth quantile of the standard normal distribution. This type of models has a nature generalization to a general location-scale family and is termed as *failure-time regression models* by Meeker and Escobar (1998), among others.

Consider the problem of predicting Y_0 at the predictor value $\mathbf{x_0}$, with h completely specified. It is well known that the following pivotal quantity

$$q(Y_0, \mathbf{Y}) = \frac{h(Y_0) - \mathbf{x_0'}\hat{\beta}}{\hat{\sigma}\sqrt{1 + \mathbf{x_0'}(\mathbf{X'X})^{-1}\mathbf{x_0}}} \cdot \sqrt{\frac{n-p}{n}}$$

is t-distributed with $n-p$ degrees of freedom, where $\hat{\beta} = (\mathbf{X'X})^{-1}\mathbf{X'}h(\mathbf{Y})$ and $\hat{\sigma}^2 = n^{-1} \parallel \mathbf{D}h(\mathbf{Y}) \parallel^{\mathbf{2}}$ are the MLEs of μ and σ, with $\mathbf{D} = I_n - \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}$ and $\parallel \cdot \parallel$ being the Euclidean norm. Using Definition 2, it is easy to see that the fiducial predictive density for Y_0 at the point $\mathbf{x_0}$ is

$$p_{\text{F}}(y_0 \mid \mathbf{y}, \mathbf{X}, \mathbf{x}_0) = \frac{\Gamma[(n-p+1)/2]}{\sqrt{n\pi}\Gamma[(n-p)/2]}\left[1 + \frac{[h(y_0) - \mathbf{x}_0'\hat{\beta}]^2}{n\hat{\sigma}^2 c_0^2}\right]^{-\frac{n-p+1}{2}} \frac{h'(y_0)}{\hat{\sigma}c_0} \tag{15}$$

where $c_0^2 = 1 + \mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x_0}$.

The FPD given above has the trans-t form, hence its properties are given in Theorems 2 and 3. Its Bayesian correspondence can be easily established in the similar manner as in Theorem 4, hence is not considered here in detail. Several applications of this FPD can be considered. The most interesting one may be the prediction of survivors probability under the existence of concomitant variables. Constructing prediction intervals may also be very interesting as the FPD does not only lead to the classical prediction interval, but also gives a shortest possible prediction interval based on the given pivotal quantity. We now explore these applications.

Density Forecasting. Under the model assumptions, the pdf of Y_0 at $\mathbf{x_0}$ is a trans-normal with parameters $\mathbf{x_0}'\beta$ and σ. Hence its MLE is $f(y_0; \mathbf{x}_0'\hat{\beta}, \hat{\sigma})$ and its FPD estimator is $p_{\text{F}}(y_0 \mid \mathbf{y}, \mathbf{X}, \mathbf{x}_0)$ given in Equation (15). It is again of interest to see which method gives a better density forecast based on a finite sample.

Predicting Survivor Probability. The true survivor function has the form

$$S(t; \beta, \sigma) = 1 - \Phi\left[\frac{h(t) - \mathbf{x_0}'\beta}{\sigma}\right]$$

The MLE of $S(t; \beta, \sigma)$ is $\hat{S}(t) = S(t; \hat{\beta}, \hat{\sigma})$. This estimator should perform poorer than the case of a single sample as more parameters are involved in the estimation. The FPD accounts for the parameter estimation, hence its estimator/predictor to the survivor function should have a better performance. The FPD estimator of $S(t; \beta, \sigma)$ has the form:

$$\tilde{S}(t) = 1 - \Psi_{n-p}\left[\frac{h(t) - \mathbf{x}_0'\hat{\beta}}{\hat{\sigma} c_0 \sqrt{n/(n-p)}}\right]$$

It is easy to see that both $\hat{S}(t)$ and $\tilde{S}(t)$ are consistent estimators of $S(t; \beta, \sigma)$. This means that when n is large, the two estimators have a similar performance.

Estimating the Hazard Function. The hazard function under this regression framework is $r(t; \beta, \sigma) = f(t; \beta, \sigma)/S(t; \beta, \sigma)$. Its MLE is $\hat{r}(t) = r(t; \hat{\beta}, \hat{\sigma})$ and its FPD estimator is $\tilde{r}(t) = p_{\text{F}}(t \mid \mathbf{y}, \mathbf{X}, \mathbf{x}_0)/\tilde{S}(t)$. Again, it is easy to see that the two estimators of hf are asymptotically equivalent.

Prediction and Shortest Prediction Intervals. The prediction interval construction can proceed in a similar way as in the one sample case. The classical and the equitailed

FPD prediction intervals have the form:

$$h^{-1}\left\{\mathbf{x}_0'\hat{\beta} \pm t_{n-p}(\alpha/2)\hat{\sigma}c_0\sqrt{n/(n-p)}\right\}. \tag{16}$$

The shortest prediction interval can be easily obtained based on the Definition 3. It should be pointed out that the issue of having the shortest PI is much more important in trans-normal regression than in trans-normal model as the regressor value can have substantial effect on the interval length (See Section 8 for numerical examples).

6 APPROXIMATE FPD WITH UNKNOWN TRANSFORMATION

The FPDs presented in Sections 3 to 5 all depend on a known transformation function. This seems restrictive as often in practice the normalizing or exponentializing transformation is unknown. We now present some results to show that use of an estimated transformation leads to fiducial inferences that are asymptotically correct.

Consider a simpler case where the transformation function is known up to an indexing parameter, i.e., $h(y) = h(y, \lambda)$. For trans-normal, it could be the Box-Cox power transformation or the modified version of it; whereas for the trans-exponential, it could be the simple power transformation that corresponds to the Weibull distribution. The former leads to a flexible three parameter duration model and the latter a two parameter one. Thus, leaving the transformation function unspecified (at least partially) is of interest in at least two aspects: it enlarges the family of duration models and ii) it extends the applicability of fiducial prediction method.

Theorem 5. Suppose that Y_0 and $\mathbf{Y} = \{Y_1, \cdots, Y_n\}$ are from a trans-normal with $h(y) = h(y, \lambda)$, λ unknown. Let $\hat{\lambda}$ be the MLE of λ. Let $(\hat{\mu}(\lambda), \hat{\sigma}(\lambda))$ and $(\hat{\mu}(\hat{\lambda}), \hat{\sigma}(\hat{\lambda}))$ be, respectively, the restricted (given λ) and unrestricted MLEs of (μ, σ). Assume $E[h_\lambda(Y_i, \lambda)]$ exists, where $h_\lambda(Y_i, \lambda) = \partial h(Y_i, \lambda)/\partial\lambda$. Then,

$$\frac{h(Y_0, \hat{\lambda}) - \hat{\mu}(\hat{\lambda})}{\hat{\sigma}(\hat{\lambda})} \sim \frac{h(Y_0, \lambda) - \hat{\mu}(\lambda)}{\hat{\sigma}(\lambda)} \tag{17}$$

Proof. Since $\hat{\lambda}$ is the MLE of λ, it can be shown to be consistent. The result follows by straightforward applications of Taylor expansions on $h(Y_0, \hat{\lambda})$, $\hat{\mu}(\hat{\lambda})$, and $1/\hat{\sigma}(\hat{\lambda})$ at $\hat{\lambda} = \lambda$, followed by an application of the law of large numbers.

Theorem 5 says that $q(Y_0, \mathbf{Y}, \lambda)$, the pivotal quantity defined in (6) with $h(y) = h(y, \lambda)$, and $q(Y_0, \mathbf{Y}, \hat{\lambda})$, the pivotal quantity obtained by replacing λ in $q(Y_0, \mathbf{Y}, \lambda)$ by $\hat{\lambda}$, are asymptotically equivalent. This result implies that when transformation is estimated all the FPDs in Section 3 and the fiducial inferences based on them remain asymptotically correct. It should be noted that the result of Theorem 5 can be easily extended to the case of trans-normal regression model discussed in Section 5. A similar result can be obtained for the trans-exponential family.

Theorem 6. Suppose that Y_0 and $\mathbf{Y} = \{Y_1, \cdots, Y_n\}$ are from a trans-exponential with $h(y) = h(y, \lambda)$, λ unknown. Let $\hat{\lambda}$ be the MLE of λ. Let $\hat{\theta}(\lambda)$ and $\hat{\theta}(\hat{\lambda})$ be, respectively, the restricted (given λ) and unrestricted MLEs of θ. Assume $E[h_\lambda(Y_i, \lambda)]$ exists, where $h_\lambda(Y_i, \lambda) = \partial h(Y_i, \lambda)/\partial\lambda$. Then

$$\frac{h(Y_0, \hat{\lambda})}{\hat{\theta}(\hat{\lambda})} \sim \frac{h(Y_0, \lambda)}{\hat{\theta}(\lambda)}.$$

Proof. Direct applications of Taylor expansion on $h(Y_0, \hat{\lambda})/\hat{\theta}(\hat{\lambda})$ at $\hat{\lambda} = \lambda$, and the law of large numbers.

With the Theorem 6, the results of Section 4 are extended to the case of an unknown transformation. The result of Theorem 6 is extendable to the cases of Type II and progressive Type-II censored data. However, the proofs are much more complicated, and hence are not discussed here.

7 MONTE CARLO SIMULATION

From the previous sections, we see that the FPD has rather easy applications in econometric duration studies. However, the properties of these applications, except for the case of prediction intervals, are not clear. For example, it is not clear whether the FPD estimator of the hazard rate performs better than the usual MLE. No doubt, it is rather difficult to compare the FPD and ML methods analytically and generally. We thus turn to Monte Carlo simulations.

The simulation process, e.g., for hf estimation, can be described as follows. For each sample generated from a given model, the MLE and FPD estimates of the hazard rate at a duration point t are calculated. Repeat this process 10,000 times to give 10,000 pairs of MLE and FPD estimates. The means and variances of these 10,000 pairs of estimates give Monte Carlo estimates of the true means and variances of the MLE and FPD estimators, which lead to estimates of bias and relative efficiency.

Extensive simulations are performed for various quantities and under various models. Almost all the results favor the FPD approach. For brevity we only report the results corresponding to the most interesting quantity, the hazard function, under two models: i) a trans-normal model of the form $h(Y, \lambda) \sim N(\mu, \sigma^2)$ with h being the Box-Cox power transformation and the transformation parameter λ assumed known or unknown; and ii) a trans-exponential model of the form $Y^\beta \sim \text{Exp}(\theta)$ with the transformation parameter β assumed known or unknown and the data generated are either complete or Type II censored.

The trans-normal model. Table 1 contains a part of simulation results for hf estimation based on a special trans-normal model, the lognormal. The simulation involves seven hf values at pth quantile duration $t_p, p = .05, .10, .25, .50, .75, .90, .95$, two sample sizes, and four σ values, to contrast the ML and FPD estimators under different function values, sample sizes, and degrees of population skewness. We see from Table 1 that the FPD estimator is always more efficient than the ML estimator. The relative efficiency of the FPDE to the MLE increases with the duration. This means that the FPD method gives much better estimate of hf at far tail area of a duration distribution. The ML method over-estimates the hf for medium to large durations, whereas the FPD method slightly under-estimates the hf. Increasing the sample size reduces significantly the discrepancy between the two estimators. Increasing the skewness of the population (i.e., increasing the value of σ) does not seem to change much on the relative performance of the two estimators, but changes the variabilities of the estimators.

Simulation is also carried out using an ML estimate of the true transformation ($\lambda = 0$). The simulation results (the right part of Table 1) show that, when λ is assumed unknown and its MLE is used, the FPD estimator preserves its excellent performance except when population is not so skewed, but the ML estimator deteriorates significantly. The over-estimation problem for MLE is worsen off and the variability of it gets larger relative to that of FPD estimator. As a result, the relative efficiency gets larger than the case of a known λ.

When σ is small (the population is not so skewed), both methods can give unsatisfactory estimates of hf at medium and large duration points. This is because when data is not so skewed, it is difficult to estimate the transformation parameter (Yang, 1999). Simulation is also carried out based a trans-normal model with covariates. Similar conclusions are drawn.

Table 1 near here

The trans-exponential model. The model considered is in fact a Weibull model with a shape parameter β and a scale parameter $\theta^{1/\beta}$. From the results in Table 2, we see that similar to the case of trans-normal model, the FPD estimator is more efficient than the ML estimator. The relative efficiency increases as duration increases. Estimation of the transformation parameter enhances the relative performance of the FPD estimator over the MLE. In addition, the effect of censoring on the hf estimation is investigated using the trans-exponential model. We found that censoring has a much greater effect on the performance of the MLE than on the performance of the FPD estimator, especially when the transformation parameter is estimated.

Table 2 near here

In summary, Monte Carlo experiments have revealed a good performance of the FPD estimator and its robustness with respect to the transformation estimation. These together with its simplicity and the flexibility of the two duration models introduced in this paper show a great tractability and applicability of the FPD method in the analysis of economic duration of event-time data.

8 EMPIRICAL EXAMPLES

We now provide some empirical examples to illustrate the various applications of the FPD method discussed above.

Example 1. A Set of Trans-normal Data. A sample of 68 observations are drawn from the model $h(Y, 0.25) \sim N(5, 1.5^2)$ with h being the power transformation: 39.00 25.38 17.08 79.04 13.65 4.80 19.22 40.55 68.56 18.16 21.53 23.08 8.80 21.05 28.34 29.09 34.49 23.67 12.15 12.02 33.83 15.43 56.59 14.93 21.50 12.69 2.01 42.66 46.74 42.42 8.23 16.51 29.75 21.54 31.69 26.25 26.28 9.74 57.97 25.68 98.70 13.71 11.50 33.82 41.69 33.97 38.65 20.34 12.70 16.52 14.03 24.64

10.15 40.15 6.05 42.09 34.45 42.93 57.06 14.01 19.00 26.11 19.69 33.25 29.12 19.26 41.83 24.06.

The MLEs of μ and σ with known $\lambda = 0.25$ are 4.8790 and 1.4079, respectively. The MLE of λ when it is assumed unknown is $\hat{\lambda} = 0.2863$, which gives the MLEs of μ and σ 5.2302 and 1.5772, respectively. Figure 2 gives plots of the estimated pdfs, sfs and hfs for the cases of λ known or unknown. The ML and FPD methods give similar estimates of the pdf and sf, but give substantially different estimates of hf. The FPD estimate of hf is much closer to the true function than the ML estimate, especially at medium and large durations. This is consistent with the simulation results given in the last section. From the plots, we do not see a clear effect of estimating transformation.

Figure 2 near here

Table 3 summarizes the prediction and shortest prediction intervals (PI) for a future duration. With $n = 68$, the shortest PIs are still 10% to 12% shorter than the regular equitailed PIs. It is interesting to note that the PIs with a known transformation and the PIs with an estimated transformation are rather similar, indicating the effect of estimating transformation is small with respect to the PI construction.

Table 3: A Summary of FPD PIs based on the Trans-normal Data

	λ known					λ known				
δ	Equitailed		Shortest		R	Equitailed		Shortest		R
.90	6.96	62.84	3.11	53.00	1.12	6.83	62.14	2.98	52.78	1.11
.95	5.15	73.97	1.95	63.65	1.12	4.97	72.74	1.76	62.97	1.11
.99	2.62	100.57	0.57	89.23	1.10	2.38	97.66	0.36	87.06	1.10

Example 2. A Set of Weibull Data. A sample of 80 observations are drawn from the model $Y^{\beta} \sim \text{Exp}(\theta)$ with $\beta = 1.5$ and $\theta^{1/\beta} = 10$. The first 60 observations are: 0.48 0.90 1.14 1.18 1.49 1.52 1.55 1.79 1.86 2.53 2.55 2.70 2.76 2.80 3.07 3.34 3.74 3.83 3.97 4.06 4.23 4.46 4.87 5.02 5.32 5.32 5.47 5.85 6.05 6.36 6.58 6.58 6.66 6.97 7.05 7.12 7.18 7.52 7.71 7.76 7.79 7.81 7.98 8.06 8.19 8.37 8.72 9.15 9.24 9.34 9.69 9.75 9.85 10.47 10.83 11.00 11.16 11.25 11.28 11.42.

Based on the first 60 observations, the MLE of α when β is assumed known is found to be 9.5055 and that when β is assumed unknown is 9.4896. The MLE of β is 1.5868. Figure 3 presents plots of the estimated pdfs, sfs and hfs using the ML and FPD methods assuming β known or unknown. Again the two methods give similar estimates of the pdf and sf, but give different estimates of the hf with the FPD estimate being closer to the true hf at the medium and large duration points. Table 4 summarizes the PIs and shortest PIs with β assumed known or unknown. We note that the shortest PIs are about 10% shorter than the corresponding equitailed PIs. The PIs with estimated β are shorter than the PIs using the known value of β.

Figure 3 near here

Table 4: A Summary of FPD PIs based on the Weibull Data

	β known					β unknown				
δ	Equitailed		Shortest		R	Equitailed		Shortest		R
.90	1.31	20.09	0.15	16.89	1.12	1.46	19.25	0.29	16.43	1.10
.95	0.82	23.17	0.05	20.12	1.11	0.94	22.03	0.11	19.31	1.10
.99	0.28	29.76	0.00	27.00	1.09	0.34	27.91	0.01	25.46	1.10

Example 3. Computer Program Execution Time Data. The data given in Table 5 is taken from Meeker and Escobar (1998, p638). It represents the amount of time it took to execute a particular computer program, on a multiuser computer system, as a function of system load (obtained with the Unix `uptime` command) at the time when execution was beginning. The data was analyzed by Meeker and Escobar using a simple log-linear regression model.

Table 5: Computer Program Execution Time Versus Load

Seconds	load	Seconds	load	Seconds	load
123	2.74	78	0.51	317	5.86
704	5.47	98	0.29	142	1.18
184	2.13	240	0.96	127	0.57
113	1.00	110	0.60	96	1.10
94	0.32	213	2.10	111	1.89
76	0.31	284	3.10		

We consider fitting a trans-normal regression model to the data by first assuming the true transformation is known to be the log transformation as did by Meeker and Escobar (1998), and then using an estimated Box-Cox power transformation. The MLEs of β_0, β_1 and σ under a log transformation of the response are, respectively, $4.4936, 0.2907$ and 0.3125. The MLE of λ is -0.4340, under which, the MLEs of β_0, β_1 and σ are, respectively, $1.9831, 0.0291$ and 0.0338.

Figures 4-6 give plots of the estimated pdfs, sfs and hfs with a chosen transformation ($\lambda = 0$) and an ML estimated transformation ($\hat{\lambda} = -0.434$) at various loading levels. We see from the plots that the ML and FPD methods give substantially different estimates of the pdf, sf and hf, in particular the hf, under both a chosen transformation and an estimated transformation. A greater discrepancy is observed at a large loading level. Thus, based on the results presented earlier it may be safer to use the FPD method than the ML method to do prediction in such a situation.

We also notice from the plots that the estimates of pdf, sf and hf based on $\hat{\lambda} = -0.434$ are quite different from the corresponding estimates based on $\lambda = 0$, especially for the pdf and hf estimation at large durations. This reflects the importance of choosing a right scale for analyzing the duration or event-time data. The equitailed and shortest PIs can also be easily calculated. For example, at load 5 with $\lambda = 0$, the two sets of 95% PIs are, respectively, {171.33, 854.59} and {133.10, 763.93}; where at the same load but with $\lambda = -0.434$, the two sets of PIs become {149.23, 1820.17} and {88.57, 1269.73}. Hence, using log transformation and an ML estimated transformation also result in substantially different PIs. The shortest PI is much shorter than the equitailed PI, showing the usefulness of the FPD method.

Figures 4-6 near here

9 CONCLUSION

The applicability of the fiducial predictive density approach in the econometric duration analysis is studied. Two flexible families of duration distributions, the trans-normal and the trans-exponential, are proposed, for detailed examinations of this approach. The former has an easy extention to the trans-normal regression and the latter allows for the analysis of censored data. It is found that the FPD approach gives simple and reliable estimates of the density, survivor, and hazard functions, and that it provides a simple way to constructing the shortest prediction intervals. The results are further extended to the case that the transformation functional form is known, but the function is indexed by an unknown parameter. The latter extension of the results greatly expands the applicability of the FPD method in the econometric duration analysis.

REFERENCES

1. Baker, M. and Melino, A. (2000) Duration dependence and nonparametric heterogeneity: A Monte Carlo study. *Journal of Econometrics* 96, 357-393.
2. Balakrishnan, N. and Basu, A. P. (1995). *The Exponential Distribution – Theorey, Methods and Applications.* Gorden and Breach, Amsterdam.
3. Barnard, G. A. (1995). Pivotal models and fiducial argument. *International Statistical Review* 63, 309-323.
4. Box, G. E. P and Cox, D. R. (1964). An analysis of transformation (with discussion). *Journal of the Royal Statistical Soceity* B 26, 211-252.
5. Butler, R. W. (1986). Predictive likelihood inference with applications (with discussion). *Journal of the Royal Statistical Soceity* B 47, 1-38.
6. Butler, R. W. (1989). Approximate predictive pivots and densities. *Biometrika* 76, 489-501.
7. Dawid, A. P. and Stone, M. (1982). The functional-model basis of fiducial inference. *The Annals Statistics* 10, 1054-1067.
8. Dawid, A. P. and Wang J. (1993). Fiducial prediction and semi-Bayesian inference. *The Annals of Statistics* 21, 1119-1138.
9. Dahiya, R. C. and Guttman, I. (1982). Shortest confidence and prediction intervals for the log-normal. *Canadian Journal Statistics* 10, 277-291.
10. Fisher, R.A. (1956-3rd Ed. 1973). *Statistical Methods and Scientific Inference.* New York: Hafner.
11. Geisser, S. (1993). *Predictive Inference: An Introduction.* Chapman & Hall, New York · London.
12. Greene, W. H. (2000) *Econometric Analysis, 4th ed.* Prentice Hall.
13. Heckman, J. and Singer, B. (1984a). Econometric duration analysis. *Journal of Econometrics* 24, 63-132.
14. Heckman, J. and Singer, B. (1984b). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* 52, 271-320.
15. Kennan, J. (1985). The duration of contract strikes in U.S. manufacturing. *Journal of Econometrics* 28, 5-28.

16. Kiefer, N. M. (1988). Economic Duration Data and Hazard Functions. *Journal of Economic Literature* 26, 646-679.
17. Koop, G. and Ruhm, C. J. (1993). Econometric estimation of proportional hazard models. *Journal of Economics and Business* 45, 421-430.
18. Lancaster, T. (1972). A stochastic model for the duration of a strike. *Journal of the Royal Statistical Society* A 135, 2, 257-271.
19. Lancaster, T. (1979). Econometric methods for the duration of unemployment. *Econometrika* 47, 939-956.
20. Lancaster, T. (1985). Generalised residuals and heterogeneous duration models - with applications to the Weibull model. *Journal of Econometrics* 28, 113-126.
21. Lancaster, T. (1990). *The Econometric Analysis of Transition Data.* Cambridge University Press.
22. Lawless (1982). *Statistical Models and Methods for Lifetime Data.* John Wiley & Sons.
23. Lejeune, M. and Faulkenberry, G. D. (1982). A simple predictive density function. *Journal of the American Statistical Association* 77, 654-659.
24. Lindley, D.V. (1958). Fiducial distribution and Bayes' theorem. *Journal of the Royal Statistical Soceity* B 20, 102-107.
25. Meeker, W. Q. and Escobar, L. A. (1998). *Statistical Methods for Reliability Data* John Wiley & Sons.
26. Ryu, K. K. (1993). Structural duration analysis of management data. *Journal of Econometrics* 57, 91-115.
27. Saha, A. and Hilton, L. (1997). Expo-power: a flexible hazard function for duration models. *Economics Letters* 54, 227-233.
28. Saunders, S. C. (1974). A family of random variables closed under reciprocation. *Journal of the American Statistical Association* 69, 533-539.
29. Seidenfeld, T. (1979). *Philosophical Problems of Statistical Inference.* Reidel, Boston.
30. Seidenfeld, T. (1992). R. A. Fisher's fiducial argument and Bayes' theorem. *Statistical Science* 7, 358-368.
31. Sider, H. (1985) Unemployment duration and incidence: 1968-82. *American Economic Review.* 75, 461-472.
32. Tay, A. S. and Wallis, K. F. (2000). Density forecasting: a survey. *Journal of Forecasting* 19, 235-254.

33. Torelli, N. and Trivellato, U. (1993). Modelling inacuracies in job-search duration data. *Journal of Econometrics* 59, 187-211.

34. Yang, Z. (1999). Estimating a transformation and its effect on Box-Cox T-ratio. *Test* 8, 167-190.

35. Yang, Z. (2002). A modified family of power transformations. *Working Paper*, School of Economics and Social Sciences, Singapore Management University.

36. Zabell, S. L. (1992). R. A. Fisher and the fiducial argument. *Statistical Science* 7, 369-387.

37. Zhang, M. Y. J., Russell J. R. and Tsay, R. S. (2001). A nonlinear autoregressive conditional duration model with applications to financial transaction data. *Journal of Econometrics* 104, 179-207.

Table 1: Simulation results for hazard rate estimation based on lognormal data: 1 = ML estimator, 2 = FPD estimator, rb = relative bias in percentage, mse = mean squared error, $ref_{2,1}$ = relative efficiency of FPDE over MLE.

$\mu = 1.0$				log transformation				estimated transformation			
n	σ	p	hf	rb_1	rb_2	mse_1	$ref_{2,1}$	rb_1	rb_2	mse_1	$ref_{2,1}$
30	2.0	.05	.536	-3.54	3.45	.04634	1.10	-6.28	0.88	.04880	1.11
		.10	.465	-2.51	-0.32	.02066	1.13	-3.83	-1.48	.02236	1.13
		.25	.300	1.24	-0.68	.00457	1.12	2.21	0.29	.00625	1.11
		.50	.147	4.87	0.43	.00113	1.20	8.36	3.68	.00210	1.21
		.75	.061	7.56	-0.34	.00026	1.42	11.76	3.18	.00051	1.47
		.90	.025	8.78	-4.01	.00005	1.71	14.09	-0.18	.00010	1.93
		.95	.014	9.36	-7.19	.00002	1.89	14.44	-3.97	.00003	2.27
	1.5	.05	.314	-3.39	3.62	.01573	1.10	-6.63	0.52	.01649	1.12
		.10	.327	-2.69	-0.48	.00951	1.13	-3.87	-1.50	.01098	1.13
		.25	.286	1.03	-0.89	.00399	1.11	2.74	0.80	.00630	1.11
		.50	.196	4.75	0.31	.00194	1.20	8.90	4.16	.00454	1.21
		.75	.113	7.63	-0.28	.00090	1.43	12.69	3.88	.00236	1.47
		.90	.063	8.46	-4.24	.00031	1.69	15.19	0.38	.00086	1.97
		.95	.043	9.16	-7.33	.00015	1.85	15.47	-3.59	.00041	2.36
	1.0	.05	.207	-3.32	3.67	.00672	1.10	-7.23	-0.14	.00730	1.12
		.10	.258	-2.15	0.02	.00608	1.13	-4.50	-2.08	.00743	1.12
		.25	.306	1.41	-0.53	.00466	1.12	3.52	1.57	.01005	1.10
		.50	.294	5.17	0.70	.00455	1.21	12.48	7.41	.01882	1.21
		.75	.238	7.40	-0.47	.00386	1.42	17.20	7.34	.01893	1.51
		.90	.179	8.40	-4.29	.00246	1.69	20.75	3.63	.01377	2.22
		.95	.146	9.35	-7.17	.00187	1.89	–	–	–	–
	0.5	.05	.182	-2.45	4.49	.00510	1.09	-8.86	-1.42	.00634	1.09
		.10	.272	-2.35	-0.15	.00673	1.13	1.34	3.86	.01654	1.03
		.25	.437	1.38	-0.55	.00949	1.12	26.26	23.00	.24132	1.15
		.50	.587	4.94	0.49	.01775	1.21	42.63	31.89	.86173	1.61
		.75	.668	7.30	-0.55	.02969	1.42	–	–	–	–
		.90	.680	8.86	-3.93	.03793	1.72	–	–	–	–
		.95	.667	8.70	-7.68	.03729	1.81	–	–	–	–
60	2.0	.05	.536	-2.03	1.63	.02289	1.05	-3.54	0.18	.02414	1.07
		.10	.465	-1.37	-0.26	.00983	1.07	-1.97	-0.82	.01063	1.07
		.25	.300	0.85	-0.13	.00202	1.06	1.20	0.24	.00286	1.05
		.50	.147	2.54	0.38	.00048	1.10	3.70	1.50	.00080	1.10
		.75	.061	3.16	-0.56	.00010	1.17	5.41	1.53	.00018	1.20
		.90	.025	4.25	-1.85	.00002	1.30	5.96	-0.39	.00003	1.35
		.95	.014	4.62	-3.37	.00001	1.37	6.72	-1.68	.00001	1.48
	1.5	.05	.314	-1.80	1.85	.00782	1.05	-3.57	0.14	.00840	1.07
		.10	.327	-1.02	0.07	.00476	1.07	-2.17	-1.00	.00527	1.07
		.25	.286	0.43	-0.53	.00186	1.05	0.99	0.03	.00276	1.05
		.50	.196	2.49	0.33	.00085	1.10	4.16	1.95	.00172	1.10
		.75	.113	3.65	-0.10	.00036	1.19	5.38	1.47	.00076	1.19
		.90	.063	3.74	-2.30	.00012	1.28	6.61	0.13	.00028	1.36
		.95	.043	4.14	-3.79	.00006	1.34	6.76	-1.71	.00013	1.48
	1.0	.05	.207	-1.94	1.70	.00347	1.05	-4.76	-1.02	.00367	1.08
		.10	.258	-0.90	0.20	.00297	1.07	-2.21	-1.04	.00332	1.06
		.25	.306	0.58	-0.38	.00211	1.06	1.85	0.87	.00432	1.05
		.50	.294	2.64	0.48	.00192	1.10	5.92	3.64	.00675	1.10
		.75	.238	3.53	-0.21	.00153	1.19	7.59	3.46	.00595	1.20
		.90	.179	4.10	-1.98	.00103	1.29	8.55	1.65	.00395	1.38
		.95	.146	4.61	-3.38	.00073	1.37	8.11	-0.82	.00262	1.51
	0.5	.05	.182	-2.01	1.64	.00265	1.05	-3.67	0.10	.00304	1.05
		.10	.272	-1.10	0.00	.00331	1.07	0.13	1.32	.00691	1.02
		.25	.437	0.58	-0.39	.00427	1.06	10.19	9.00	.05374	1.05
		.50	.587	2.40	0.25	.00765	1.10	17.36	14.33	.17077	1.14
		.75	.668	3.73	-0.02	.01248	1.19	19.51	13.52	.24324	1.30
		.90	.680	4.04	-2.05	.01481	1.29	–	10.25	–	–
		.95	.667	4.36	-3.59	.01507	1.36	–	–	–	–

Table 2: Simulation results for hazard rate estimation based on complete or censored Weibull data: 1 = ML estimator, 2 = FPD estimator, rb = relative bias in percentage, mse = mean squared error, $ref_{2,1}$ = relative efficiency of FPDE over MLE.

$n = 60, \theta^{1/\beta} = 10$				known transformation				estimated transformation			
r	β	p	hf	rb_1	rb_2	mse_1	$ref_{2,1}$	rb_1	rb_2	mse_1	$ref_{2,1}$
20	0.5	.05	0.97	5.52	5.22	.06182	1.02	-3.40	-3.66	0.09731	1.01
		.10	0.47	5.59	4.97	.01474	1.03	0.54	0.27	0.01688	1.02
		.25	0.17	5.25	3.60	.00194	1.09	2.37	2.05	0.00570	1.09
		.50	0.07	5.44	1.54	.00033	1.22	2.47	2.02	0.00455	1.30
		.75	0.04	5.24	-2.25	.00008	1.41	2.07	1.42	0.00330	1.93
		.90	0.02	5.22	-6.63	.00003	1.53	2.02	1.02	0.00506	5.51
		.95	0.02	4.86	-9.92	.00002	1.48	1.79	0.70	0.00558	11.48
	2.0	.05	0.05	4.94	4.64	.00013	1.02	-0.15	-0.16	0.00021	1.01
		.10	0.06	4.95	4.34	.00026	1.03	0.05	0.01	0.00031	1.02
		.25	0.11	5.33	3.67	.00073	1.09	1.44	1.24	0.00211	1.08
		.50	0.17	5.58	1.66	.00183	1.23	5.84	4.78	0.02382	1.27
		.75	0.24	5.74	-1.83	.00371	1.43	13.76	9.40	0.15687	2.08
		.90	0.30	5.24	-6.62	.00592	1.55	27.63	13.88	0.90760	5.14
		.95	0.35	5.36	-9.58	.00779	1.55	38.86	15.02	3.55742	16.48
30	0.5	.05	0.97	3.74	3.55	.03755	1.01	-2.91	-3.09	0.08619	1.01
		.10	0.47	3.84	3.45	.00893	1.02	-0.53	-0.70	0.01183	1.01
		.25	0.17	3.05	2.01	.00113	1.06	0.82	0.64	0.00163	1.05
		.50	0.07	3.51	1.01	.00021	1.14	0.99	0.79	0.00079	1.14
		.75	0.04	3.59	-1.30	.00005	1.26	0.84	0.59	0.00052	1.32
		.90	0.02	3.37	-4.45	.00002	1.33	0.72	0.42	0.00041	1.73
		.95	0.02	3.51	-6.46	.00001	1.34	0.64	0.33	0.00030	2.03
	2.0	.05	0.05	3.65	3.46	.00008	1.01	-0.14	-0.15	0.00019	1.01
		.10	0.06	3.64	3.25	.00016	1.02	-0.08	-0.10	0.00023	1.01
		.25	0.11	3.41	2.36	.00045	1.06	0.50	0.39	0.00062	1.05
		.50	0.17	3.61	1.10	.00111	1.14	2.30	1.82	0.00445	1.14
		.75	0.24	3.31	-1.54	.00211	1.25	5.43	3.82	0.02401	1.34
		.90	0.30	3.68	-4.19	.00359	1.35	10.37	6.30	0.07396	1.66
		.95	0.35	3.76	-6.26	.00477	1.37	14.00	7.40	0.13944	2.07
40	0.5	.05	0.97	2.34	2.20	.02606	1.01	-2.19	-2.32	0.08288	1.01
		.10	0.47	2.23	1.95	.00634	1.02	-0.60	-0.73	0.01124	1.01
		.25	0.17	2.79	2.01	.00088	1.04	0.29	0.16	0.00092	1.04
		.50	0.07	2.73	0.89	.00015	1.10	0.46	0.33	0.00031	1.10
		.75	0.04	2.19	-1.39	.00004	1.17	0.43	0.28	0.00017	1.21
		.90	0.02	2.61	-3.25	.00001	1.24	0.38	0.21	0.00012	1.38
		.95	0.02	2.71	-4.79	.00001	1.26	0.33	0.16	0.00009	1.54
	2.0	.05	0.05	2.45	2.32	.00006	1.01	-0.11	-0.12	0.00018	1.01
		.10	0.06	2.37	2.08	.00012	1.02	-0.11	-0.13	0.00021	1.01
		.25	0.11	2.53	1.76	.00032	1.04	0.18	0.10	0.00036	1.04
		.50	0.17	2.48	0.65	.00078	1.10	1.08	0.77	0.00164	1.10
		.75	0.24	2.79	-0.83	.00158	1.19	2.93	1.94	0.00751	1.21
		.90	0.30	2.72	-3.15	.00265	1.25	5.18	2.89	0.02350	1.41
		.95	0.35	2.46	-5.01	.00341	1.24	7.00	3.46	0.03871	1.53
60	0.5	.05	0.97	1.50	1.41	.01696	1.01	-0.97	-1.06	0.07048	1.00
		.10	0.47	1.68	1.49	.00407	1.01	-0.43	-0.52	0.01035	1.01
		.25	0.17	1.94	1.43	.00055	1.03	0.01	-0.07	0.00068	1.02
		.50	0.07	1.64	0.44	.00009	1.06	0.14	0.06	0.00010	1.06
		.75	0.04	1.78	-0.60	.00002	1.12	0.16	0.07	0.00005	1.13
		.90	0.02	1.54	-2.33	.00001	1.15	0.14	0.05	0.00003	1.22
		.95	0.02	1.57	-3.41	.00000	1.15	0.12	0.03	0.00002	1.29
	2.0	.05	0.05	2.06	1.97	.00004	1.01	-0.04	-0.05	0.00015	1.00
		.10	0.06	1.61	1.43	.00008	1.01	-0.08	-0.09	0.00019	1.01
		.25	0.11	1.30	0.80	.00021	1.03	0.02	-0.03	0.00026	1.02
		.50	0.17	1.66	0.46	.00051	1.06	0.34	0.15	0.00056	1.06
		.75	0.24	1.63	-0.74	.00097	1.11	0.98	0.41	0.00187	1.13
		.90	0.30	1.81	-2.08	.00170	1.16	1.95	0.68	0.00497	1.22
		.95	0.35	1.89	-3.12	.00221	1.17	2.74	0.83	0.00866	1.30

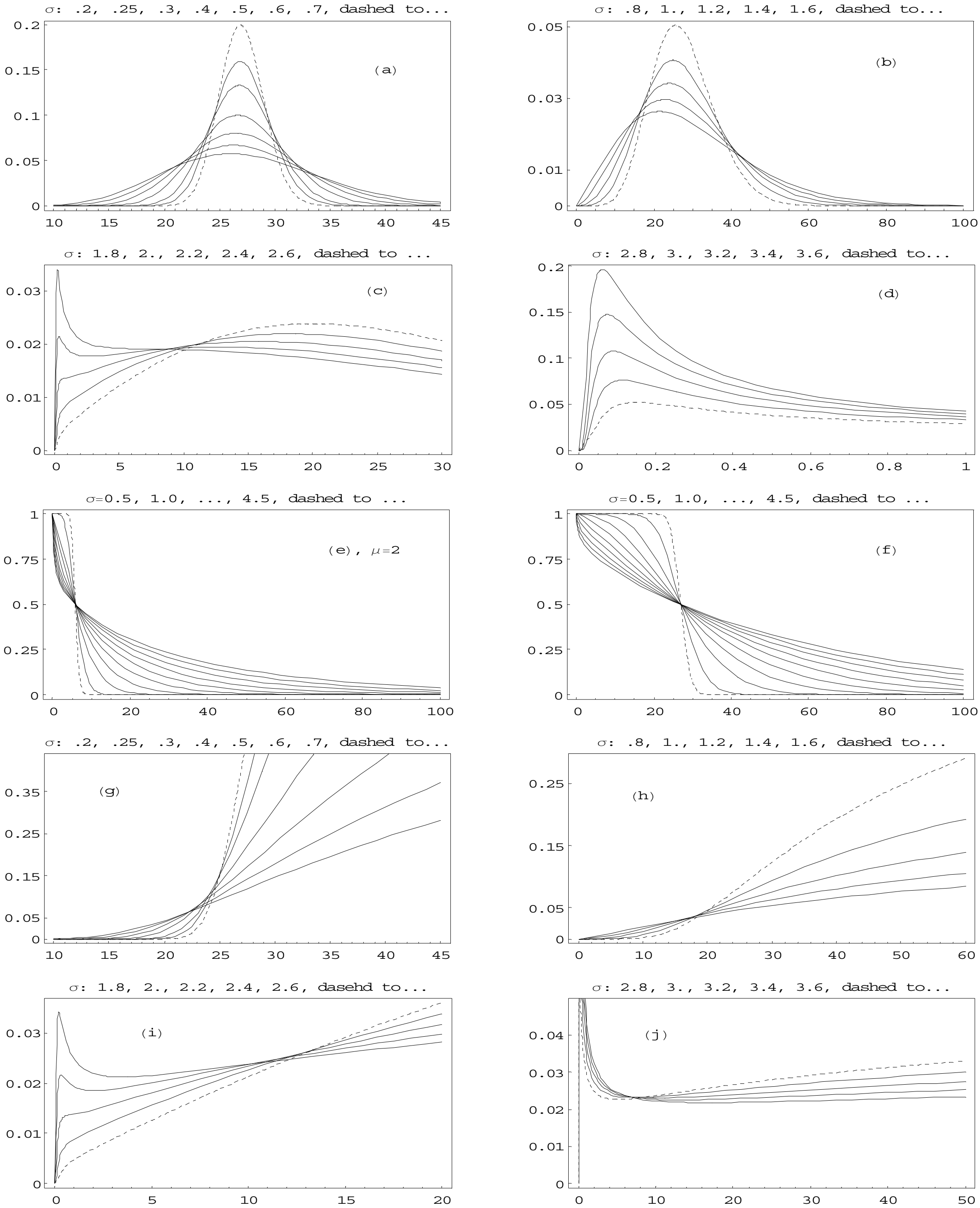

Figure 1. Plots of the trans-normal pdfs (a)-(d), sfs (e)-(f), and hfs (g)-(j), $h(y) = y^{.5} - y^{-.5}$, $\mu = 5$ except (e).

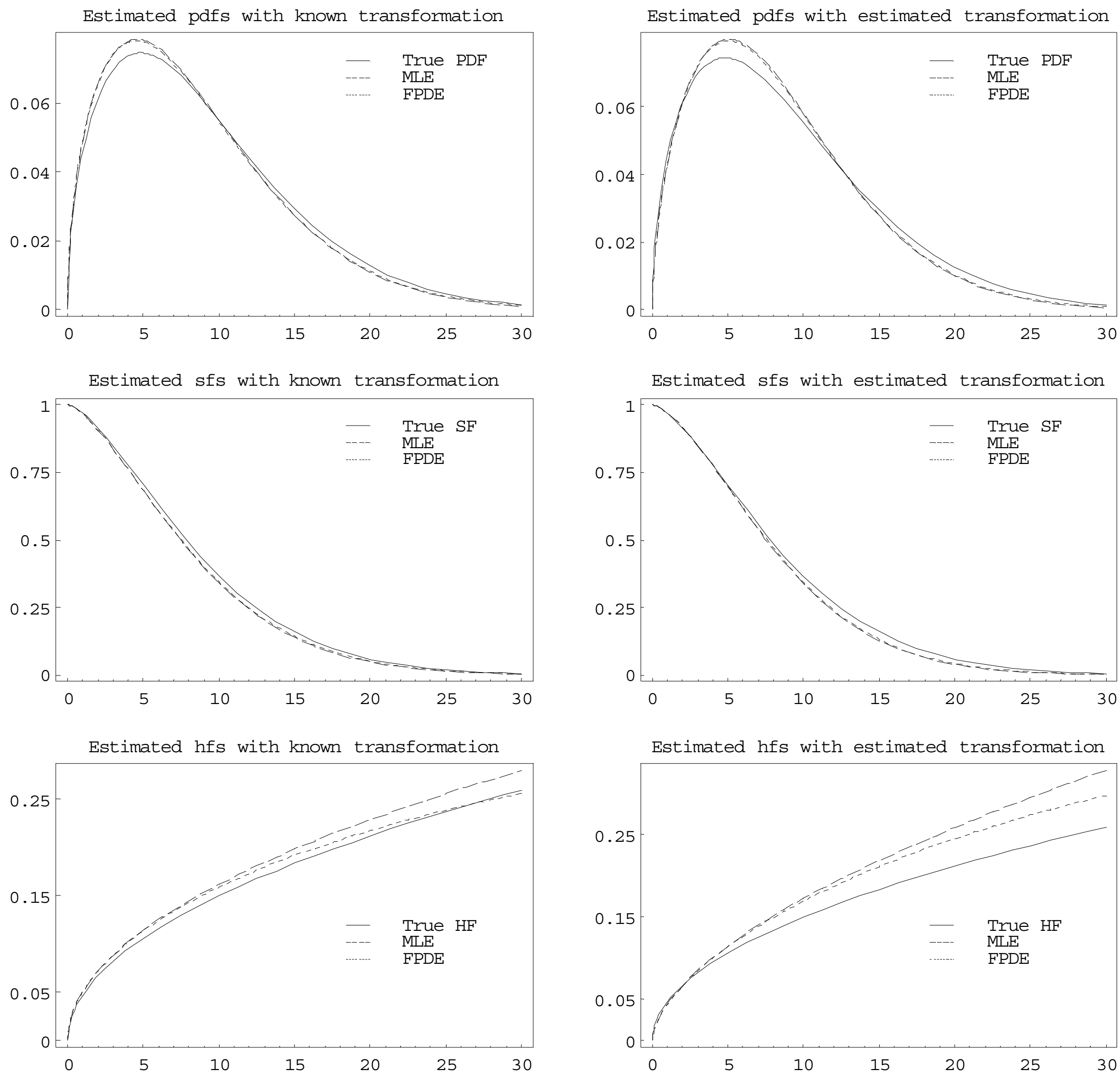

Figure 2. Plots of the estimated pdfs, sfs and hfs for the trans-normal data

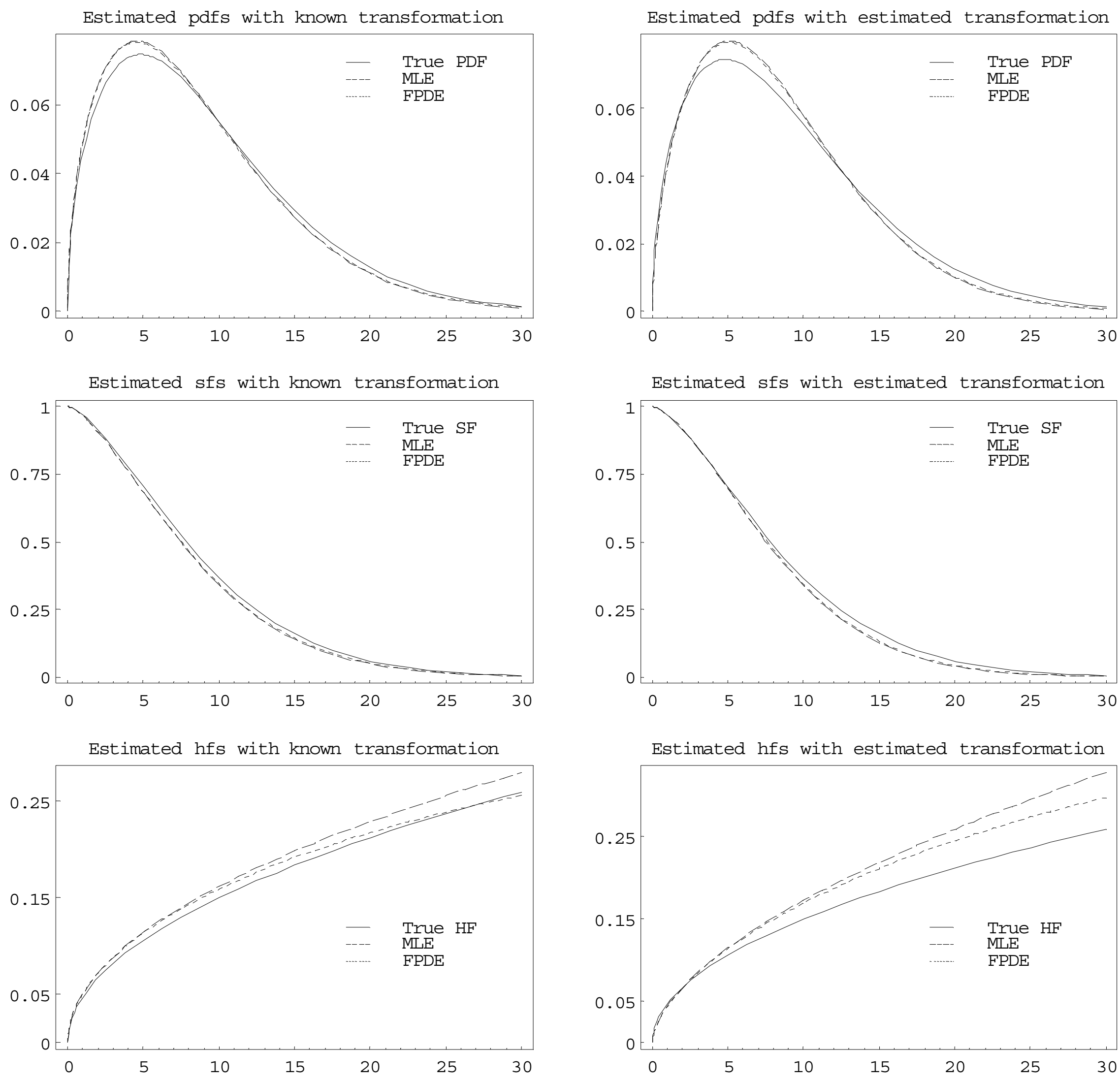

Figure 3. Plots of the estimated pdfs, sfs and hfs for the Weibull data

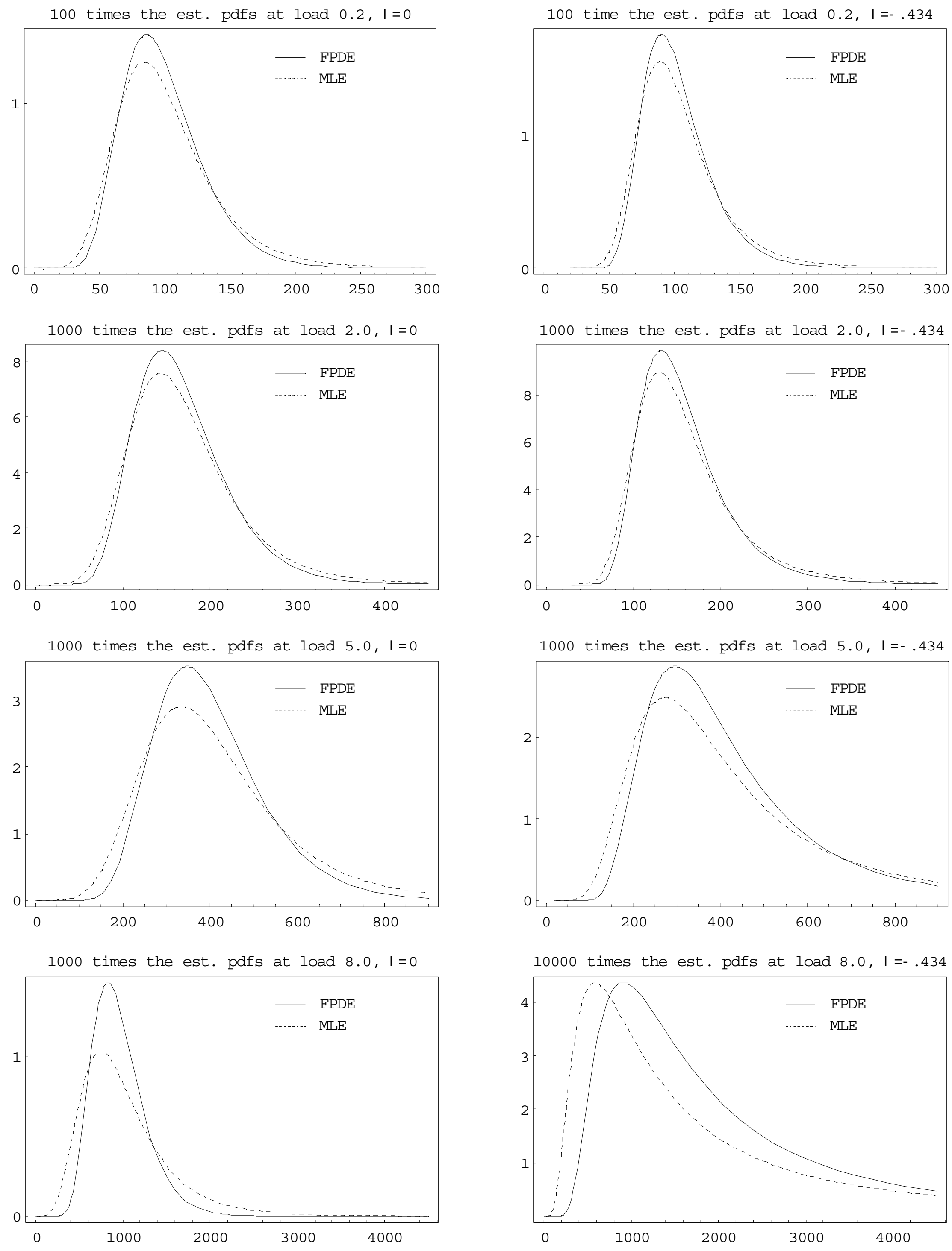

Figure 4. Plots of the estimated pdfs for the Computer Execution Time Example

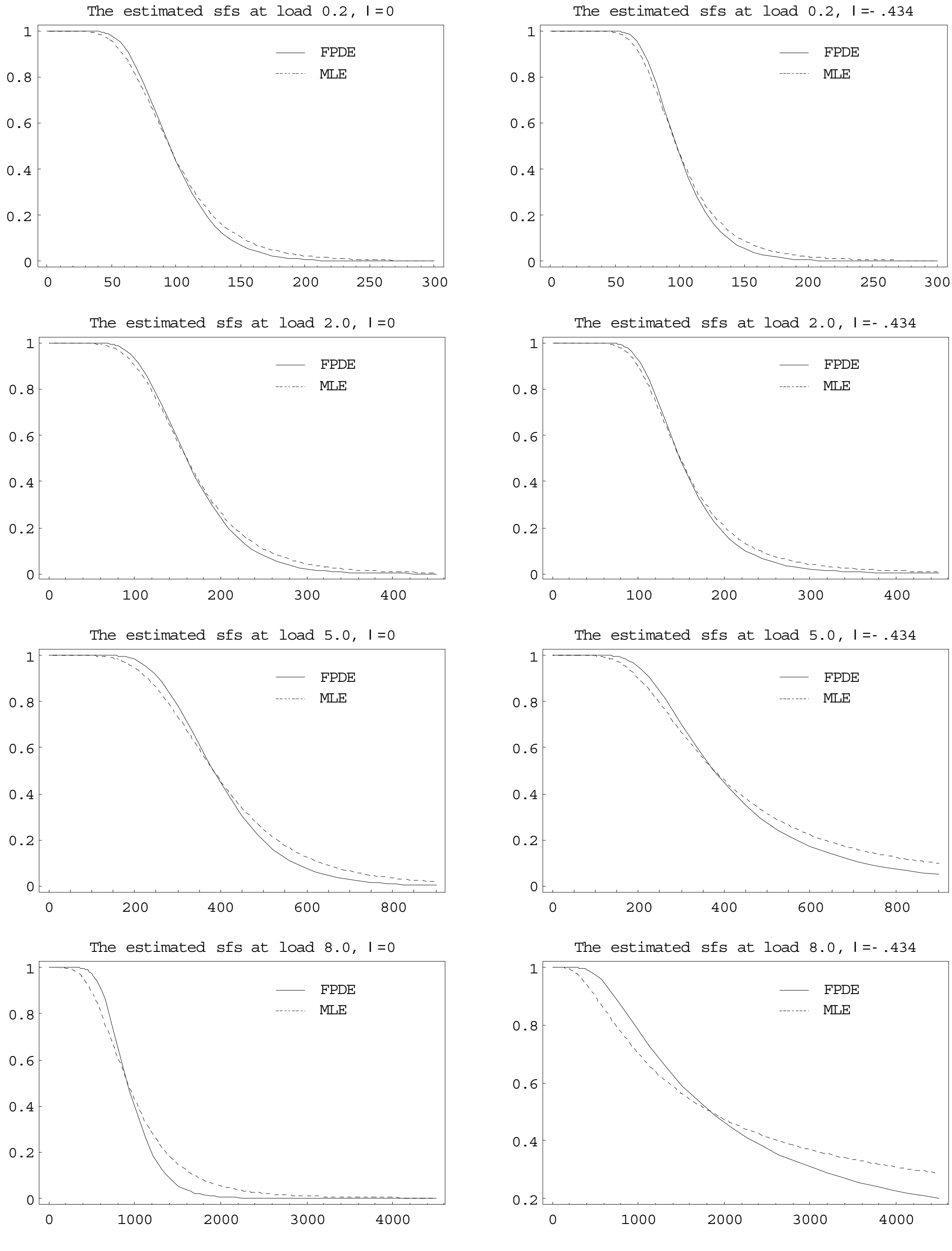

Figure 5. Plots of the estimated sfs for the Computer Execution Time Example

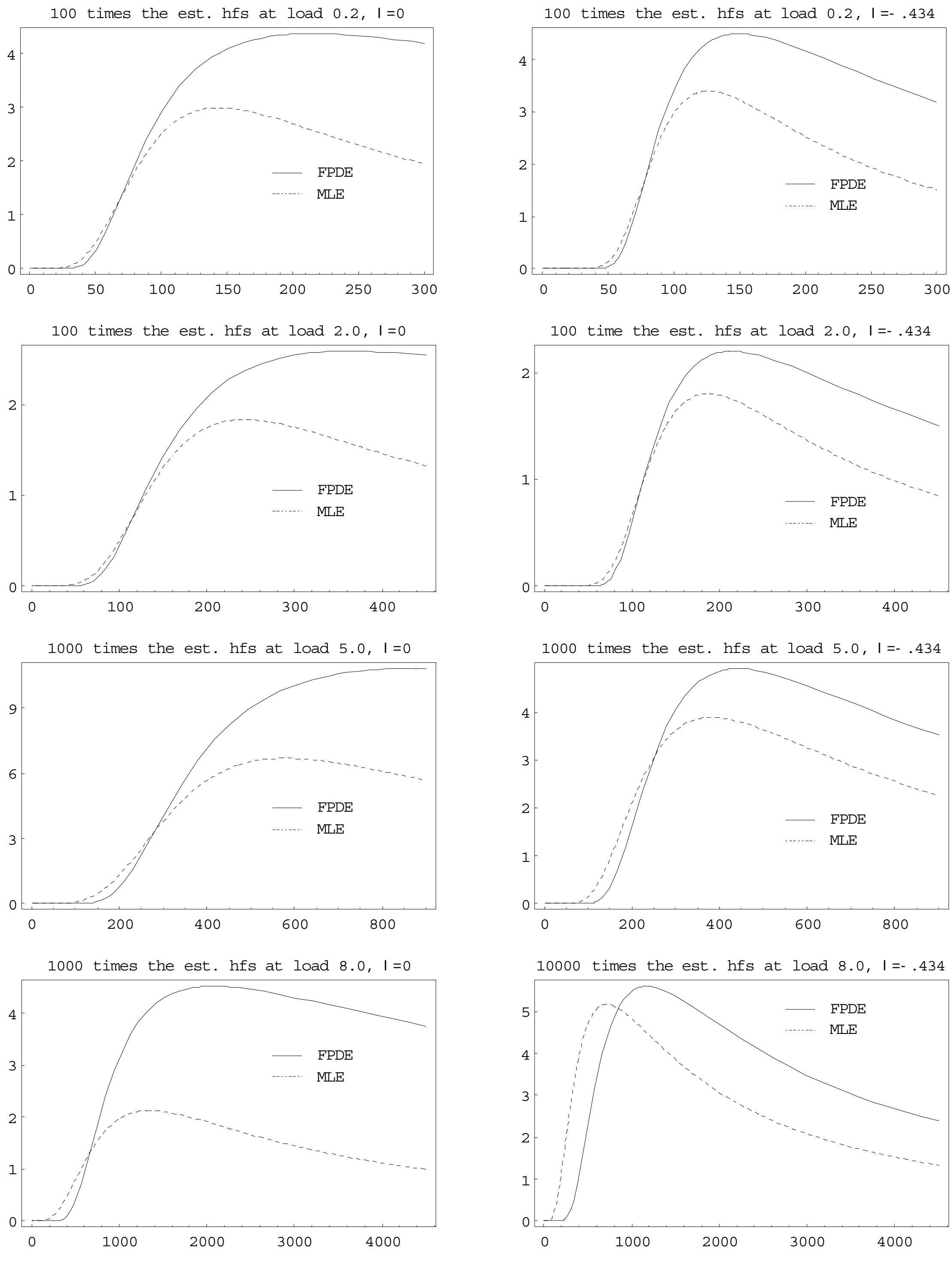

Figure 6. Plots of the estimated hfs for the Computer Execution Time Example