GENERALIZED GAMMA FAMILY REGRESSION MODELS FOR LONG-DISTANCE TELEPHONE CALL DURATIONS

by

T. A. Cameron
University of California, Los Angeles

and

K. J. White
University of British Columbia

Discussion Paper #363
January 1985

The authors wish to thank D. Gorham, S. Haun, D. Smeaton, and T. Wales for helpful discussions and assistance.

ABSTRACT

In regression applications where the dependent variable cannot logically take on negative values, the usual normal conditional probability density function is often inappropriate. In these circumstances, the Generalized Gamma (GG) distribution can be adopted as a flexible distribution for regression error terms. This distribution contains the simple Gamma (G), Weibull (W), Exponential (E), Lognormal (LN), and other distributions as special cases. These models can be specified as linear or log-linear. Discrimination among the special cases is achieved by examination of the appropriate Wald, Lagrange Multiplier, or Likelihood Ratio test statistics. We show how the GG family can be used in linear and log-linear regression to model the duration of long distance telephone calls. Previous regression models of call duration have assumed normal errors (usually on aggregated data); Weibull and Gamma distributions have been used only in a non-regression context. Our results imply that experimentation with this model may be indicated for other economic analyses of duration.

1. Introduction

In regression applications where the dependent variable cannot logically take on negative values, the usual normal conditional probability density function is often inappropriate. In these circumstances, the Generalized Gamma (GG) distribution can be adopted as a flexible distribution for regression error terms. This distribution contains the simple Gamma (G), Weibull (W), Exponential (E), Lognormal (LN), and other distributions as special cases. Discrimination among these special cases is achieved by examination of the appropriate Wald, Lagrange Multiplier, or Likelihood Ratio test statistics. McDonald [1984] has used the simple GG distribution to estimate the parameters of an income distribution. In this paper, we show how the GG family can be extended to regression models of the duration of long distance telephone calls. The existing literature observes that call durations vary with a number of characteristics of the call, such as marginal price per minute, distance, time-of-day, type of call and origin of call. Previous regression models of call duration have assumed normal errors (usually on aggregated data); Weibull and Gamma distributions have been used only in a non-regression context.

In section 2, a brief description of related research on the demand for telephone services is provided. Section 3 focusses on call duration as an important aspect of telephone demand and develops a regression model with generalized gamma errors for analyzing the determinants of call duration. The theoretical regression models described in section 3 are rendered operational in section 4 with the selection of specific explanatory variables. The estimation results are discussed and compared to the findings of previous analyses. The important methodological issue of discrimination among various models in the generalized gamma family is addressed in section 5. Finally, section 6 describes a policy experiment using the estimated equations.

2. Related Research on Telephone Demand

In the United States, the recent AT&T divestiture decision has aroused considerable interest in the determinants of demand for a variety of telephone services. Park, Wetzel and Mitchell [1983] use aggregated data on local telephone calls from the General Telephone and Electronics (GTE) pricing experiment in central Illinois. Their objective is the measurement of price elasticities both for the total number of calls and total minutes of conversation. Pacey [1983] concentrates specifically on the estimation of "point-to-point" price elasticities for intercity long distance service. One conclusion was that "more disaggregated data needs to be made available in order to estimate the mean duration of a call". The present study responds to this point, but carries the analysis even further by allowing mean duration in the disaggregated model to be a function of a whole range of explanatory variables, while still constraining the distribution of durations to be strictly non-negative. Rea and Lage [1978] utilize time-series/cross-sectional data on communications originating in the United States and directed to 37 foreign countries. Regarding the pitfalls of aggregation, the authors state that "...since demand for telegraph and telephone services arises from both household and business sectors, it would be desirable to estimate disaggregated functions. However, the data are not available, and the assumption must be made in estimating the demand equations that meaningful aggregate relationships exist".

Gale [1971,1974] made several studies of call duration and reported a number of basic results. Among these were that (1) the mean duration of a toll call is longer, the more distant the call. (2) Evening calls are longer than daytime calls. However, since night rates are lower than daytime rates, the difference incorporates an unknown price effect. (3) Person-to-person calls are longer than station-to-station calls, and (4) collect calls and calls billed to third parties are longer than "paid" calls. Finally, (5) day-of-week makes a large difference in the distribution of calls by time-of-day for business traffic, but makes only a moderate difference for residence traffic.

The simple, unconditional distribution of telephone call durations has been explored by Wong [1981] using the Weibull distribution, by Pavarini [1979] using the Powernormal distribution, and by Curien [1981] using the exponential and Erlang distributions. De Fontenay, Gorham, Manning and Lee [1983] estimate separate Weibull distributions for the average lengths of calls with destinations in each of seven mileage bands to obtain mean duration values used in a subsequent ordinary least squares regression to derive price and income elasticities of demand.

In this study, we attempt to consolidate the methodology used in the previous studies to show how the determinants of duration can be explored using maximum likelihood estimation of a regression model with non-normal errors. We examine a sample of disaggregated data for long-distance phone calls over a twenty-four hour period for the Canadian province of British Columbia. As deregulation of long-distance telephone markets progresses, further work on the demand for such services is warranted. Our data permits separate analyses of particular types of calls. For example, we have chosen (a) residential calls to Canada and United States destinations and (b) both business and residential calls to overseas destinations.

3. Distributional Assumptions for Call Duration

In the linear regression model with normal errors, a non-zero probability is associated with negative values of the dependent variable, even though the regression line might be strictly positive, thus permitting prediction intervals to include values which may be theoretically (and empirically) impossible in some contexts. Lawless [1982] provides a comprehensive analysis of "Lifetime" models for product testing (often called "accelerated failure time" (AFT) models) which fall into this category. More recently, Heckman and Singer [1984] have explored the estimation of these models in studies of unemployment durations where the data are censored. They found that estimation becomes complicated when typical durations are long, (both relative to the sample period and in real time), and the sample will often contain a substantial proportion of incomplete spells. Further complexity is introduced when durations span an interval which must be acknowledged as "time inhomogeneous", and when the distribution exhibits "duration dependence".[1] In these contexts, the presence of left- or right-censoring due to a relatively narrow sampling window means that the hazard function, rather than the probability density function, is a more useful statistical concept upon which to base an econometric analysis.

In contrast to unemployment duration data, the telephone call durations modelled here are very short relative to the sample period. Since the data consist of all calls initiated during a twenty-four hour sample period, and complete durations are recorded, censoring is not a problem. Furthermore, since individual call durations are so very short, ceteris paribus may more readily be assumed to hold during each call. Most importantly, the rates applicable to each call are usually determined solely by the time of initiation of the call, even if the call itself spans two rate periods.

A. "Linear" Models

We focus first on the class of "linear" regression models where duration (t) for a particular call is assumed to be a linear function of a set of exogenous variables (x) so that $E(t|x)=x'\beta$. While this allows the possibility of

[1] For example, in job search models of unemployment, the existence of a "declining reservation wage" can result in positive duration dependence.

negative fitted values for the dependent variable, it does permit a simple interpretation of the coefficients, which are analogous to those in linear OLS regressions.

The simple GG probability density function (see Johnson and Kotz [1970,p.197] for a single observation (t) on the random variable T is given by:

$$(1) \qquad f(t) = c \, \frac{t^{ck-1}}{b^{ck}} \; \frac{1}{\Gamma(k)} \; \exp\left[-\left(\frac{t}{b}\right)^{c} \right] , \; t \geq 0$$

where b is a "scale" parameter and c and k are "shape" parameters and Γ is the mathematical Gamma function. The simple Gamma (G) distribution is obtained when $c = 1$, while the Weibull (W) imposes $k = 1$. For the Exponential (E) model, $c = k = 1$, while as $k \to \infty$ the lognormal (LN) distribution results.

The mean of the GG distribution is:

$$(2) \qquad E(t) = b\Gamma(k+\frac{1}{c})/\Gamma(k)$$

If we wish to make this density conditional upon a $(p \times 1)$ vector of explanatory variables, x, we can adopt the usual assumption that the scale parameter varies with x while the shape parameters are constant.[2] Thus for a "linear" regression model, $t = x'\beta + \epsilon$, with GG errors:

$$(3) \qquad E(t|x) = b(x)\Gamma(k+\frac{1}{c})/\Gamma(k) = x'\beta$$

[2] To be comparable with ordinary normal regression models, the mean of the conditional distribution should be $x'\beta$. However, the mean of the G distribution, as a simple example, is the product: bc. The least restrictive assumption would allow both b=b(x)= and c=c(x). However, the fitted "regression line" would then be given by the quadratic function b(x)c(x). Holding the shape parameters constant is an assumption no stricter than that of homoscedasticity in the normal regression model.

This means that we must substitute b(x) for b wherever the latter appears in the density function (1).[3] We find that:

$$(4) \qquad b(x) = x'\beta \frac{\Gamma(k)}{\Gamma(k+\frac{1}{c})} = \frac{x'\beta}{G}$$

$$\text{where} \quad G = \frac{\Gamma(k+\frac{1}{c})}{\Gamma(k)}$$

The conditional density function for t is now:

$$(5) \qquad f(t|x) = c \frac{t^{(ck-1)}}{(x'\beta/G)^{ck}} \cdot \frac{1}{\Gamma(k)} \exp\left[-(\frac{t}{x'\beta/G})^{c}\right], \ t \geq 0$$

The joint density of n independent observations on the variable $T = (t_1, t_2 \ldots)$ yields a log-likelihood function to be maximized by an appropriate choice of the parameters (β, k, c).

In order to simplify the exposition, we define the ψ function as the derivative of $\log(\Gamma)$. First and second derivatives of the Γ and ψ functions are denoted as Γ', Γ'', ψ', ψ''. In addition, the following abbreviations will be adopted:

$$(6) \qquad P = \psi(k)$$

$$D = \psi(k+\frac{1}{c}) \qquad D_k = \psi'(k+\frac{1}{c}) \qquad D_c = \frac{-D_k}{c^2}$$

[3] Straightforward substitution of $x'\beta$ for b(x) would mean that the fitted value of t would be a non-linear function of the estimated parameters. This complicates the process of inference. It is desirable to preserve the fitted value of t as a linear function of the estimated parameters.

$$G_k = G(D-P) \qquad G_c = \frac{-DG}{c^2} \qquad G_{kk} = \frac{\Gamma''(k+\frac{1}{c})}{\Gamma(k)} - G\left[\frac{\Gamma''(k)}{\Gamma(k)} + 2P(D-P)\right]$$

$$H = (P-D)/G \qquad H_k = \frac{\Gamma''(k)}{\Gamma(k+\frac{1}{c})} - \frac{1}{G}\left[\frac{\Gamma''(k+\frac{1}{c})}{\Gamma(k+\frac{1}{c})} + 2D(P-D)\right]$$

$$t_i^* = \frac{t_i}{x_i'\beta} \qquad x_{ij}^* = \frac{x_{ij}}{x_i'\beta}$$

The GG log-likelihood function can then be expressed as:

$$(7) \qquad \ell = n \log c - n \log \Gamma(k) + nck \log G + ck \sum_{i=1}^{n} \log t_i^* - \sum_{i=1}^{n} \log t_i - \sum_{i=1}^{n} (t_i^* G)^c$$

The first derivatives, $\ell^{(1)}$, of this log-likelihood with respect to the unknown parameters are:

$$(8) \qquad \frac{\partial \ell}{\partial \beta_r} = - ck \sum_{i=1}^{n} x_{ir}^* + cG^c \sum_{i=1}^{n} t_i^{*c} x_{ir}^* \qquad r=1,\ldots,p$$

$$\frac{\partial \ell}{\partial k} = -n(P+ckGH) + c \sum_{i=1}^{n} \log(t_i^* G) - cG_k G^{c-1} \sum_{i=1}^{n} t_i^{*c} \quad G$$

$$\frac{\partial \ell}{\partial c} = \frac{n}{c}(1-kD) + k \sum_{i=1}^{n} \log(t_i^* G) - \sum_{i=1}^{n} (t_i^* G)^c \log t_i^* + \left(\frac{D}{c} - \log G\right) \sum_{i=1}^{n} (t_i^* G)^c$$

The second derivatives, $\ell^{(2)}$, are:

$$(9) \qquad \frac{\partial^2 \ell}{\partial \beta_r \partial \beta_s} = ck \sum_{i=1}^{n} x_{ir}^* x_{is}^* - c(1+c)G^c \sum_{i=1}^{n} t_i^{*c} x_{ir}^* x_{is}^* \qquad r,s=1,\ldots,p$$

$$\frac{\partial^2 \ell}{\partial \beta_r \partial k} = c^2 G_k G^{c-1} \sum_{i=1}^{n} (t_i^*)^c x_{ir}^* - c \sum_{i=1}^{n} x_{ir}^* \qquad r=1,\ldots,p$$

$$\frac{\partial^2 \ell}{\partial \beta_r \partial c} = -k \sum_{i=1}^{n} x_{ir}^* + c \sum_{i=1}^{n} (t_i^* G)^c x_{ir}^* \log t_i^* + (1-D+c \log G) \sum_{i=1}^{n} (t_i^* G)^c x_{ir}^*$$

$$r=1,\ldots,p$$

$$\frac{\partial^2 \ell}{\partial k^2} = - n\, \psi'(k) - nc \left[GH+k(G_k H+GH_k) \right] + \frac{ncG_k}{G}$$

$$- c \left[(c-1)G^{(c-2)} G_k^2 + G^{(c-1)} G_{kk} \right] \sum_{i=1}^{n} t_i^{*c}$$

$$\frac{\partial^2 \ell}{\partial k \partial c} = -\frac{n}{c}(D+kD_k) + \frac{nkG_k}{G} + \sum_{i=1}^{n} \log(t_i^* G) - c \sum_{i=1}^{n} \log t_i^* (t_i^* G)^{c-1} (t_i^* G_k)$$

$$+ \left[\frac{D_k}{c} - \frac{G_k}{G} \right] \sum_{i=1}^{n} (t_i^* G)^c + c \left[\frac{D}{c} - \log G \right] \sum_{i=1}^{n} (t_i^* G)^{c-1} (t_i^* G_k)$$

$$\frac{\partial^2 \ell}{\partial c^2} = -\frac{n}{c^2} (1-kD) - \frac{n}{c} k\, D_c + nk \frac{G_c}{G} - c \sum_{i=1}^{n} \log t_i^* (t_i^* G)^{c-1} (t_i^* G_c)$$

$$- \sum_{i=1}^{n} \log t_i^* \log(t_i^* G)(t_i^* G)^c + \left[\left(\frac{cD_c - D}{c^2}\right) - \frac{G_c}{G} \right] \sum_{i=1}^{n} (t_i^* G)^c$$

$$+ c \left[\frac{D}{c} - \log G \right] \sum_{i=1}^{n} (t_i^* G)^{c-1} (t_i^* G_c)$$

$$+ \left[\frac{D}{c} - \log G \right] \sum_{i=1}^{n} \log(t_i^* G)(t_i^* G)^c$$

Given the maximum likelihood estimates of the parameters (β,k,c) the negative of the inverse of the matrix of second derivatives can be used to estimate the asymptotic covariance matrix of parameter estimates. These estimates can be used for Wald tests involving hypotheses about the coefficients. In particular, specific tests concerning the values of the shape parameters can be used to aid in model selection.

Table I details the analogous formulas used to generate each of the special cases of the linear GG model: the E, G, W, and LN models. Derivation of these formulas is quite tedious and since most have not been previously published it is important to display them here.

B. "Log-linear" Models

Accelerated failure time models such as those discussed in Lawless [1982], generally assume that $E(t|x)=\exp(x'\beta)$. The error term then enters additively into the exponent. These models have the advantage of forcing $E(t|x)>0$, regardless of the sign of the inner product, $x'\beta$. For these models, the transformation $y=\log(t)$ renders the right-hand side linear-in-parameters, and it is clear how to derive the corresponding density functions for the transformed variable. However, this procedure makes the model "log-linear" rather than "linear". Table IIa provides a summary of the likelihood function calculations for "log-linear" versions of the E, G, and W models. The GG model appears in Table IIb. (Of course, the "log-linear" LN model is simply ordinary least squares (OLS) with $y=\log(t)$ as the dependent variable.)

4. Estimation and Results

Data were obtained on a stratified sample of approximately 65,000 long distance telephone calls originating in the Canadian province of British Columbia on July 13, 1983. From this sample, a subsample of 21,738 residential calls to Canada and the U.S. (excluding Alaska) were found. A second subsample consists of both business and residential calls to all overseas destinations, yielding a total of 4934 calls. Table III provides variable names, definitions, sample means, and standard deviations for the variables in the sample. The Appendix provides a more complete description of the data.

TABLE I

DERIVATION OF NON-NORMAL LINEAR REGRESSION MODELS[a]

		Gamma		Lognormal	
		$f(t) = \frac{t^{k-1}}{b^k} \cdot \frac{1}{\Gamma(k)} \cdot \exp \frac{-t}{b}$		$f(t) = \frac{1}{\sqrt{2\pi}\,\sigma\, t} \exp\left(-\frac{1}{2\sigma^2} (\log t - \mu)^2\right)$	
		$E(t) = bk$		$E(t) = \exp\left(\mu + \frac{\sigma^2}{2}\right)$	
		$E(t\|x) = b(x) \cdot k = x'\beta$		$E(t\|x) = \left(\exp b(x) + \frac{\sigma^2}{2}\right) = x'\beta$	
		$b(x) = x'\beta/k$		$b(x) = \log x'\beta - \frac{\sigma^2}{2}$	
		$f(t\|x) = \frac{kt}{x'\beta}^{k} \cdot \frac{1}{t\Gamma(k)} \exp - \frac{kt}{x'\beta}$		$f(t\|x) = \frac{1}{\sqrt{2\pi}\,\sigma t} \exp\left[\frac{-1}{2\sigma^2} (\log \frac{t}{x'\beta} + \frac{\sigma^2}{2}^{2}\right]$	
$t_i^* - \sum_{i=1}^{n} \left[\Gamma(c^*) t_i^*\right]^c$		$\ell = n(k \log k - \log \Gamma(k)) - \sum_{i=1}^{n} \log t_i + k \sum_{i=1}^{n} \log t_i^* - k \sum_{i=1}^{n} t_i^*$		$\ell = - n \log\left(\sqrt{2\pi}\,\sigma\right) - \sum_{i=1}^{n} \log t_i - \frac{1}{2\sigma^2} \sum_{i=1}^{n} U_i^2$	
	$r = 1,\ldots,p$	$\frac{\partial \ell}{\partial \beta_r} = k \sum_{i=1}^{n} x_{ir}^* (t_i^* - 1)$	$r = 1,\ldots,p$	$\frac{\partial \ell}{\partial \beta_r} = \frac{1}{\sigma^2} \sum_{i=1}^{n} x_{ir}^* U_i$	$r = 1,\ldots,p$
		$\frac{\partial \ell}{\partial k} = n\left[\log k + 1 - \frac{\Gamma'(k)}{\Gamma(k)}\right] - \sum_{i=1}^{n} t_i^* + \sum_{i=1}^{n} \log t_i^*$		$\frac{\partial \ell}{\partial \sigma} = - \frac{n}{\sigma} - \frac{1}{\sigma^3} \sum_{i=1}^{n} U_i^2 - \frac{1}{\sigma} \sum_{i=1}^{n} U_i$	
	$r,s = 1,\ldots,p$	$\frac{\partial^2 \ell}{\partial \beta_r \partial \beta_s} = k \sum_{i=1}^{n} x_{ir}^* x_{is}^* - 2k \sum_{i=1}^{n} x_{ir}^* x_{is}^* t_i^*$	$r,s = 1,\ldots p$	$\frac{\partial^2 \ell}{\partial \beta_r \partial \beta_s} = - \frac{1}{\sigma^2} \sum_{i=1}^{n} x_{ir}^* x_{is}^* (U_i + 1)$	$r,s = 1,\ldots,p$
	$r = 1,\ldots p$	$\frac{\partial^2 \ell}{\partial \beta_r \partial k} = \sum_{i=1}^{n} x_{ir}^* (t_i^* - 1)$	$r = 1,\ldots,p$	$\frac{\partial^2 \ell}{\partial \beta_r \partial \sigma} = - \frac{2}{\sigma^3} \sum_{i=1}^{n} x_{ir}^* U_i + \frac{1}{\sigma} \sum_{i=1}^{n} x_{ir}^*$	$r = 1,\ldots,p$
R_i		$\frac{\partial^2 \ell}{\partial k^2} = n\left[\frac{1}{k} - \psi'(k)\right]$		$\frac{\partial^2 \ell}{\partial c^2} = - n + \frac{n}{\sigma^2} - \frac{3}{\sigma^4} \sum_{i=1}^{n} U_i^2 + \frac{3}{\sigma^2} \sum_{i=1}^{n} U_i$	

1)/c; $R_i = \left(\Gamma(c^*) t_i^*\right)^c$, $i=1,\ldots,n$; $S = \Psi'(c)$; $V = \frac{1}{c}\left(\frac{\Gamma'(c^*)}{\Gamma(c^*)}\right) - \log \Gamma(c^*)$; $U_i = \log\left(\frac{t_i}{x'\beta}\right) + \frac{\sigma^2}{2}$

TABLE II .A

DERIVATION OF NON-NORMAL NON-LINEAR REGRESSION MODELS

	Exponential	Gamma	Weibull
Conditional density:	$f(t\|x) = \frac{1}{b(x)} \exp\left(\frac{-t}{b(x)}\right)$	$f(t\|x) = \frac{t^{k-1}}{b(x)^k} \cdot \frac{1}{\Gamma(k)} \exp\left[-\left(\frac{t}{b(x)}\right)\right]$	$f(t\|x) = c\frac{t^{c-1}}{[b(x)]^c} \exp\left[-\left(\frac{t}{b(x)}\right)^c\right]$
Redefined parameters:	$b(x) = \exp(x'\beta)$	$b(x) = \exp(x'\beta)$	$b(x) = \exp(x'\beta)$; $\sigma = 1/c$
Let Y = log T:	$f(y\|x) = \exp\left[(y-x'\beta) - \exp(y-x'\beta)\right]$	$f(y\|x) = \frac{1}{\Gamma(k)} \exp\left[(k\varepsilon) - \exp(\varepsilon)\right]$	$f(y\|x) = \frac{1}{\sigma} \exp\left[w - \exp(w)\right]$
Log-likelihood:	$\ell = \sum_{i=1}^{n} \varepsilon_i - \sum_{i=1}^{n} \exp(\varepsilon_i)$	$\ell = -n\log \Gamma(k) + k \sum_{i=1}^{n} \varepsilon_i - \sum_{i=1}^{n} \exp(\varepsilon_i)$	$\ell = -n\log\sigma + \sum_{i=1}^{n} w_i - \sum_{i=1}^{n} \exp(w_i)$
Gradient vector:	$\frac{\partial \ell}{\partial \beta_r} = -\sum_{i=1}^{n} x_{ir} + \sum_{i=1}^{n} x_{ir}\exp(\varepsilon_i) \quad r=1,\ldots,p$	$\frac{\partial \ell}{\partial \beta_r} = \sum_{i=1}^{n} x_{ir}\left[\exp(\varepsilon_i) - k\right] \quad r=1,\ldots,p$	$\frac{\partial \ell}{\partial \beta_r} = \frac{-1}{\sigma} \sum_{i=1}^{n} x_{ir} + \frac{1}{\sigma} \sum_{i=1}^{n} x_{ir}e^{w_i} \quad r=1,\ldots,p$
		$\frac{\partial \ell}{\partial k} = -n\psi(k) + \sum_{i=1}^{n} \varepsilon_i$	$\frac{\partial \ell}{\partial \sigma} = \frac{-n}{\sigma} - \frac{1}{\sigma} \sum_{i=1}^{n} w_i + \frac{1}{\sigma} \sum_{i=1}^{n} w_i e^{w_i}$
Hessian matrix:	$\frac{\partial^2 \ell}{\partial \beta_r \partial \beta_s} = -\sum_{i=1}^{n} x_{ir}x_{is}\exp(\varepsilon_i) \quad r,s=1\ldots,p$	$\frac{\partial^2 \ell}{\partial \beta_r \partial \beta_s} = -\sum_{i=1}^{n} x_{ir}x_{is}\exp(\varepsilon_i) \quad r,s=1,\ldots,p$	$\frac{\partial^2 \ell}{\partial \beta_r \partial \beta_s} = \frac{-1}{\sigma^2} \sum_{i=1}^{n} x_{ir}x_{is}e^{w_i} \quad r,s=1,\ldots,p$
		$\frac{\partial^2 \ell}{\partial \beta_r \partial k} = -\sum_{i=1}^{n} x_{ir} \quad r=1,\ldots,p$	$\frac{\partial^2 \ell}{\partial \beta_r \partial \beta_s} = \frac{1}{\sigma^2} \sum_{i=1}^{n} x_{ir} - \frac{1}{\sigma^2} \sum_{i=1}^{n} x_{ir}e^{w_i} - \frac{1}{\sigma^2} \sum_{i=1}^{n} x_{ir}w_i e^{w_i} \quad r=1,\ldots,p$
		$\frac{\partial^2 \ell}{\partial k^2} = -n\psi'(k)$	$\frac{\partial^2 \ell}{\partial \sigma^2} = \frac{n}{\sigma^2} + \frac{2}{\sigma^2} \sum_{i=1}^{n} w_i - \frac{2}{\sigma^2} \sum_{i=1}^{n} w_i e^{w_i} - \frac{1}{\sigma^2} \sum_{i=1}^{n} w_i^2 e^{w_i}$

$\varepsilon_i = y_i - x_i'\beta$ $\quad$ $w_i = (y_i - x_i'\beta)/\sigma$

- TABLE II .B

Log Generalized Gamma

Conditional density: $$f(t|x) = c\,\frac{t^{ck-1}}{(b(x))^{ck}} \cdot \frac{1}{\Gamma(k)} \exp\left[\left(\frac{-t}{b(x)}\right)^{c} \right]$$

Redefined parameters: $b(x) = \exp(x'\beta)$; $\sigma = 1/c$

Let Y = log T: $$f(y|x) = \frac{1}{\sigma\Gamma(k)} \exp\left[(wk) - \exp(w) \right]$$

Log-likelihood: $$\ell = -n \log\sigma - n \log\Gamma(k) + k \sum_{i=1}^{n} (w_i) - \sum_{i=1}^{n} \exp(w_i)$$

Gradient vector: $$\frac{\partial\ell}{\partial\beta_r} = -\frac{k}{\sigma} \sum_{i=1}^{n} x_{ir} + \frac{1}{\sigma} \sum_{i=1}^{n} x_{ir} e^{w_i}$$

$$\frac{\partial\ell}{\partial k} = -n\psi(k) + \sum_{i=1}^{n} w_i$$

$$\frac{\partial\ell}{\partial\sigma} = -\frac{n}{\sigma} - \frac{k}{\sigma} \sum_{i=1}^{n} w_i + \frac{1}{\sigma} \sum_{i=1}^{n} w_i e^{w_i}$$

Hessian Matrix: $$\frac{\partial^2\ell}{\partial\beta_r\partial\beta_s} = -\frac{1}{\sigma^2} \sum_{i=1}^{n} x_{ir} x_{is} e^{w_i} \qquad r,s = 1,\ldots,p$$

$$\frac{\partial^2\ell}{\partial\beta_r\partial k} = -\frac{1}{\sigma} \sum_{i=1}^{n} x_{ir} \qquad r = 1,\ldots,p$$

$$\frac{\partial^2\ell}{\partial\beta_r\partial\sigma} = \frac{k}{\sigma^2} \sum_{i=1}^{n} x_{ir} - \frac{1}{\sigma^2} \sum_{i=1}^{n} x_{ir} e^{w_i} - \frac{1}{\sigma^2} \sum_{i=1}^{n} x_{ir} w_i e^{w_i}$$

$r=1,\ldots,p$

$$\frac{\partial^2\ell}{\partial k^2} = -n\psi'(k)$$

$$\frac{\partial^2\ell}{\partial k\partial\sigma} = -\frac{1}{\sigma} \sum_{i=1}^{n} w_i$$

$$\frac{\partial^2\ell}{\partial\sigma^2} = \frac{n}{\sigma^2} + \frac{2k}{\sigma^2} \sum_{i=1}^{n} w_i - \frac{2}{\sigma^2} \sum_{i=1}^{n} w_i e^{w_i} - \frac{1}{\sigma^2} \sum_{i=1}^{n} w_i^2 e^{w_i}$$

$w_i = (y_i - x_i'\beta)/\sigma$

TABLE III

MEANS (STANDARD DEVIATIONS) OF DATA: LONG DISTANCE TELEPHONE CALLS

	Description	Canada and U.S.	Overseas
	(Dichotomous variables	Destinations	Destinations
		(weighted)	
Variable	= 0 otherwise)	(n = 21738)	(n = 4934)
DUR	duration in minutes	6.5655 (7.9394)	7.7383 (7.5121)
LOG(DUR)	log of duration	1.3682 (1.0065)	1.6349 (0.9484)
RATE	marginal rate in dollars	0.3738 (0.1879)	2.1105 (0.4675)
LOG(DIST)	log of distance in miles	4.5151 (1.4242)	8.5669 (0.1931)
EVENING	= 1 if evening rate period	0.4617	-
NIGHT	= 1 if night rate period	0.0834	0.2475
BUSINESS	= 1 if business service	-	0.3032
COLLECT	= 1 if call collect	0.0645	-
CARD	= 1 if credit or third-party	0.0555	0.0359
PERSON	= 1 if person-to-person	0.0082	0.0387
NORTHEAST	= 1 if originating in Northeast	0.3788	0.1092

A. Canada and United States Destinations

1. "Linear" Models

The dependent variable in all cases is the duration of each long distance telephone call (in minutes). For each alternative model, the estimated coefficients, shape parameters, and regression statistics for this sample appear in Table IV.A.[4] Among the linear specifications, the LN model achieved the highest value of the log-likelihood function. The GG model should have had a higher log-likelihood function than the LN, but final convergence was not possible due to a very flat likelihood surface at large values of k and small values of c. This is not too surprising given the multiplicative term involving these two parameters which appears twice in the likelihood function making estimation somewhat difficult. However, it seems that the β-vector for the GG would have been very close to the LN estimates. Since the GG model becomes LN as $k \rightarrow \infty$, an intermediate k value for this model of 697.0 suggests that the algorithm is moving the model in the direction of the LN parameters.

The fitted coefficients of the "linear" model suggest that a ten cent increase in marginal rates should decrease call duration by .34 minutes or about 20 seconds. A one percent increase in the distance between call origin and call destination is found to increase expected duration by 1.3 minutes. This supports previous findings summarized by the phrase "the longer the haul, the longer the call" (see Taylor [1980]). Duration is usually expected to increase with distance because the frequency of calls declines.

The dummy variables for the time-of-day rate periods indicate that evening

[4] The first and second derivatives derived in Section 3 have been incorporated in the Quasi-Newton algorithms utilized by the MLE command in Version 5 of White's [1978] SHAZAM Econometrics Computer Program. All analytical derivative formulas were verified by numerical differencing of the log-likelihood function. Convergence was generally attained rapidly except in the GG case.

TABLE IV.A

MAXIMUM LIKELIHOOD ESTIMATES, CANADA AND U.S. DESTINATIONS

(LINEAR REGRESSION)

Dep. Var. = DUR	N	E	G	W	GG[b]	LN
RATE	-1.626 (.7135)	-3.770 (.7581)	-3.770 (.6774)	-3.677 (.7117)	-3.5066 (.7129)	-3.422 (.7122)
LOG(DIST)	1.238 (.0885)	1.428 (.1001)	1.428 (.0894)	1.418 (.0939)	1.3431 (.0938)	1.330 (.0936)
EVENING	2.716 (.1578)	2.057 (.1459)	2.057 (.1304)	2.089 (.1372)	1.8116 (.1345)	1.815 (.1347)
NIGHT	2.867 (.3043)	1.947 (.3211)	1.947 (.2869)	2.021 (.3026)	1.3664 (.2852)	1.370 (.2851)
COLLECT	.4210 (.2120)	.4956 (.1832)	.4956 (.1637)	.4725 (.1715)	1.0027 (.1891)	1.031 (.1899)
CARD	-.9992 (.2267)	-.6006 (.1351)	-.6006 (.1207)	-.6274 (.1261)	-.3084 (.1421)	-.3007 (.1432)
PERSON	-.4133 (.5792)	-.1044 (.4668)	-.1045 (.4171)	-.1547 (.4337)	.4614 (.5048)	.4783 (.5087)
NORTHEAST	-.3780 (.1064)	-.4520 (.0737)	-.4520 (.6586)	-.4496 (.0692)	-.4449 (.0724)	-.4445 (.0728)
intercept	.2644 (.1821)	.5508 (.1396)	.5508 (.1247)	.5694 (.1311)	.8769 (.1377)	.9237 (.1362)
Gamma Shape, k	-	1	1.252 (.0108)	1	696.42 (755.60)	-
Weibull Shape, c	-	1	1	1.068 (.0052)	.0398 (.0216)	-
Lognormal Shape, σ	-	-	-	-	-	.9524 (.0046)
Log L	-80661.4	-61253.28	-60931.76	-61166.5	-59542.34	-59525.91

[a] asymptotic standard error estimates in parentheses

[b] convergence not achieved

calls are about 1.8 minutes longer, and night calls are about 1.4 minutes longer, on the average, than day calls. This time-of-day effect is distinct from the influence of the lower marginal rates charged during these off-peak periods. Person-to-person calls are a half a minute longer than station calls but the difference is not statistically significant. Collect calls are longer than paid calls by about one minute and statistically significant. This may be explained by the fact that the initiator of the call might not be paying for the call. Credit card calls, on the other hand, are about 20 seconds shorter, on average. This probably reflects the fact that the caller is often using someone else's residence telephone. Courtesy may require that such calls be kept "short".

The coefficient on the NORTHEAST dummy variable indicates that calls originating in this generally less-populated area of B.C. tend to be about a half-minute shorter than those in other regions. The greater geographical dispersion of the population in this region may mean that a greater proportion of calls made by these subscribers must be long distance calls. Hence we may be observing a substitution effect between frequency of calls (or number of different destinations called) and duration of each individual call, subject to the subscriber's overall budget constraint.

It is important to note that the estimated coefficients of the GG and LN models are similar to each other, and noticably different from those of the E, G, and W models. The simple OLS coefficients are markedly different from any of the other models and the calculated OLS log likelihood function is roughly 20,000 lower than the other models.

2. "Log-linear" Models

Table IV.B gives the results for the Canada/U.S. sample when the same family of conditional distributions is assumed for duration itself, but it is assumed that the log of duration is linearly related to the explanatory

TABLE IV.B

MAXIMUM LIKELIHOOD ESTIMATES, CANADA AND U.S. DESTINATIONS

(LOG-LINEAR REGRESSION)

Dep.Var=LOG(DUR)	E	G	W	GG	LN
RATE	-.0010	-.0010	.0046	.0274	.0280
	(.1051)	(.0939)	(.0988)	(.0898)	(0.894)
LOG(DIST)	.1645	.1645	.1633	.1485	.1482
	(.0132)	(.0118)	(.0124)	(.0111)	(.0111)
EVENING	.4773	.4773	.4787	.4326	.4318
	(.0223)	(.0199)	(.0209)	(.0198)	(.0198)
NIGHT	.5026	.5026	.5082	.4181	.4171
	(.0432)	(.0385)	(.0405	(.0383)	(.0381)
COLLECT	.0955	.0955	.0914	.1733	.1740
	(.0279)	(.0249)	(.0261)	(.0266)	(.0266)
CARD	-.1399	-.1398	-.1452	-.0714	-.0709
	(.0298)	(.0266)	(.0279)	(.0285)	(.0284)
PERSON	-.0600	-.0600	-.0674	.0276	.0284
	(.0765)	(.0683)	(.0716)	(.0728)	(.0726)
NORTHEAST	-.0789	-.0789	-.0780	-.0733	-.0731
	(.0140)	(.0125)	(.0131)	(.0134)	(.0133)
intercept	.8443	.6172	.8760	-1063.7	.4747
	(.0247)	(.0237)	(.0233)	(123.13)	(.0228)
Gamma Parameter, k	1	1.2549	1	111.71	-
		(.0108)		(10.730)	
Weibull Parameter,	1	1	.9359	13713.	-
σ=1/c			(.0046)	(2570.6)	
Lognormal	-	-	-	-	.9511
Parameter, σ					
Log L	-61232.01	-60905.09	-61144.33	-59495.48	-59492.0

variables. Optimization proceeds using the associated conditional distributions for the logarithmically transformed variable. Point estimates of the slope parameters thus represent the expected percentage change in duration as a result of a one unit change in each explanatory variable.

Again, the E,G, and W models yield very similar parameter estimates, but the LN and GG specifications result in a considerably higher maximized value of the likelihood function and (for some variables) different parameter estimates. All computed log-likelihood functions in the log-linear models were transformed to be made comparable to those in linear models by using the appropriate Jacobian transformation as discussed in Box and Cox [1964].

B. Overseas Destinations

1. "Linear" Models

Table V.A exhibits the estimated results for the subsample of overseas calls. The linear GG regression model achieved a substantially higher value for the log-likelihood function than all the other linear models. In addition, the linear models proved to be marginally better than the log-linear models, in contrast to our results for the Canada/U.S. sample. The estimated coefficients imply that a ten cent increase in the marginal rate would be expected to decrease average duration by only .08 minutes, in contrast to the .34 minutes observed for the Canada/U.S. sample. Distance, however, has no significant impact upon call duration, in sharp contrast to our results for the previous sample. The coefficients on the dummy variable for the off-peak nighttime period suggests that such calls are marginally longer by .73 minutes, an amount which is statistically significantly different from zero. Business calls are shorter by about a half a minute and also statistically significant. Credit card calls and calls billed to a third number are longer by about 1.4 minutes, also a

TABLE V.A

MAXIMUM LIKELIHOOD ESTIMATES, OVERSEAS DESTINATIONS

(LINEAR REGRESSION)

Dep. Var. = DUR	N	E	G	W	GG	LN
RATE	-.0999	-.1845	-.1856	-.0697	-.8250	-1.0130
	(.4431)	(.4680)	(.3998)	(.4149)	(.4097)	(.4162)
LOG(DIST)	-.7857	-.6523	-.6508	-.7841	-.1674	-.1368
	(.7591)	(.7742)	(.6614)	(.6829)	(.6888)	(.7036)
NIGHT	1.012	.9656	.9651	1.009	.7343	.6730
	(.3896)	(.4197)	(.3586)	(.3681)	(.3869)	(.3998)
BUSINESS	-.1490	-.1925	-.1929	-.1045	-.5642	-.6425
	(.2396)	(.2429)	(.2075)	(.2141)	(.2206)	(.2267)
CARD	1.315	1.445	1.445	1.440	1.424	1.419
	(.5816)	(.6997)	(.5978)	(.6123)	(.6490)	(.6709)
PERSON	1.890	2.017	2.017	1.850	3.394	4.048
	(.5621)	(.7007)	(.5987)	(.6032)	(.7437)	(.7966)
NORTHEAST	-.0409	-.0253	-.0252	-.0631	.2036	.2924
	(.3447)	(.3569)	(.3049)	(.3104)	(.3435)	(.3589)
intercept	14.36	13.41	13.40	14.30	10.79	11.14
	(5.934)	(5.982)	(5.111)	(5.269)	(5.365)	(5.492)
Gamma Shape, k	-	1	1.370	1	21.66	-
			(.0249)		(8.338)	
Weibull Shape, c	-	1	1	1.150	.2311	-
				(.0121)	(.0451)	
Lognormal Shape, σ	-	-	-	-	-	.9394
						(.0095)
Log L	-16929.4	-15011.0	-14873.6	-14930.1	-14746.4	-14759.4

[a] asymptotic standard error estimates in parentheses

significant difference. The largest increase in duration is observed for person-to-person calls, which are longer by almost three and a half minutes and statistically significant. Finally, in contrast to the results for the Canada/U.S. sample, we find that for the overseas calls the region of origin (NORTHEAST) has no significant influence on call duration.

2. "Log-linear" Models

In Table V.B the estimated parameters are displayed for the log-linear model using the overseas data. Once again, the estimated parameters for the GG and LN models are more similar to each other than they are to the estimates for the other special cases of the the GG model. None of the estimated slope parameters are more than just marginally significantly different from zero. It seems that the explanatory variables available on billing records are not particularly reliable predictors of the expected duration of an overseas call.

It is important to draw attention to the possible consequences of using parameter estimates from a regression model with misspecified errors. Tables IV and V show that, relative to the GG estimates, the implied influence of marginal rates is sometimes substantially distorted by the E, G, or W shape parameter restrictions (and especially by the Normal assumption in the linear model). In the log-linear models, the effect of rate becomes insignificant. Thus, not only an inappropriate functional form, but also the choice of an inappropriate error distribution could have significant implications for rate policy decisions. Bear in mind that previous studies of duration have been inclined to use the W or the E distribution. For the overseas sample, both the linear and the log-linear E, G, and W models imply that business calls are not significantly shorter. The GG and LN specifications suggest that they are.

Section 2 described a number of results from previous studies. The positive effect of distance on duration is strongly supported by our models for

TABLE V.B

MAXIMUM LIKELIHOOD ESTIMATES, OVERSEAS DESTINATIONS

(LOG-LINEAR REGRESSION)

Dep.Var=LOG(DUR)	E	G	W	GG	LN
RATE	-.0190	-.0189	-.0054	-.0973	-.1184
	(.0611)	(.0522)	(.0534)	(.0556)	(.0557)
LOG(DIST)	-.0973	-.0974	-.1116	-.0466	-.0453
	(.1034)	(.0883)	(.0902)	(.0948)	(.0954)
NIGHT	.1225	.1226	.1272	.0907	0794
	(.0529)	(.0452)	(.0461)	(.0486)	(.0489)
BUSINESS	-.0225	-.0225	-.0110	-.0699	-.0780
	(.0321)	(.0275)	(.0280	(.0298)	(.0301)
CARD	.1655	.1655	.1649	.1609	.1572
	(.0778)	(.0665)	(.0677)	(.0722)	(.0731)
PERSON	.2334	.2335	.2140	.3726	.4245
	(.0752)	(.0642)	(.0654)	(.0705)	(.0706)
NORTHEAST	-.0040	-.0040	-.0090	-.0245	-.0344
	(.0461)	(.0394)	(.0401)	(.0429)	(.0433)
intercept	2.878	2.564	3.022	-10.80	2.251
	(.8040)	(.6872)	(.7011)	(4.234)	(.7454)
Gamma Parameter, k	1	1.3697	1	21.27	-
		(.0249)		(8.074)	
Weibull Parameter,	1	1	.8696	4.289	-
σ=1/c			(.0092)	(.8257)	
Lognormal	-	-	-	-	.9404
Parameter, σ					
Log L	-15011.32	-14874.05	-14930.43	-14747.48	-14760.7

Canada/U.S. data, but we do not find the same result for the overseas calls. Perhaps the incremental effect of distance diminishes with distance, and the overseas calls are at distances beyond where this effect vanishes. For previous studies, the result that night calls are longer than day calls was derived without controlling for differing marginal rates. With regression, we have been able to isolate the separate effect of time-of-day on duration. We find little support in the Canada/U.S. sample for the result that person-to-person calls are longer than station-to-station calls, but this is strongly corroborated by the overseas destination sample. Collect calls and calls billed to third parties have previously been found to be longer than sent-paid calls. We find in the Canada/U.S. sample that collect calls are longer, but third party (credit card) calls are shorter. There were no collect calls in the overseas sample, but third-party calls are, somewhat surprisingly, significantly longer by more than a minute.

5. Model Discrimination

In Figures 1 and 2 we present some diagrams to illustrate the differences in the fitted models. Since the shape of the conditional distribution for t depends upon the x vector, Figure 1 depicts the fitted conditional distributions for t, in the linear models, for a "representative" telephone call from Vancouver, British Columbia, to Toronto, Ontario (RATE=1.05, DIST=2098, all dummy variables=0). The domain of the conditional density functions for the log-linear models is y=log t. Figure 2 compares the transformed densities for the GG model with those for each of its special cases, along with the symmetric normal density appropriate to the log-linear normal model. It is interesting to observe that the transformed densities in the log-linear models appear much more alike than the densities of the linear model.

While the fitted densities appear to differ considerably across models, it

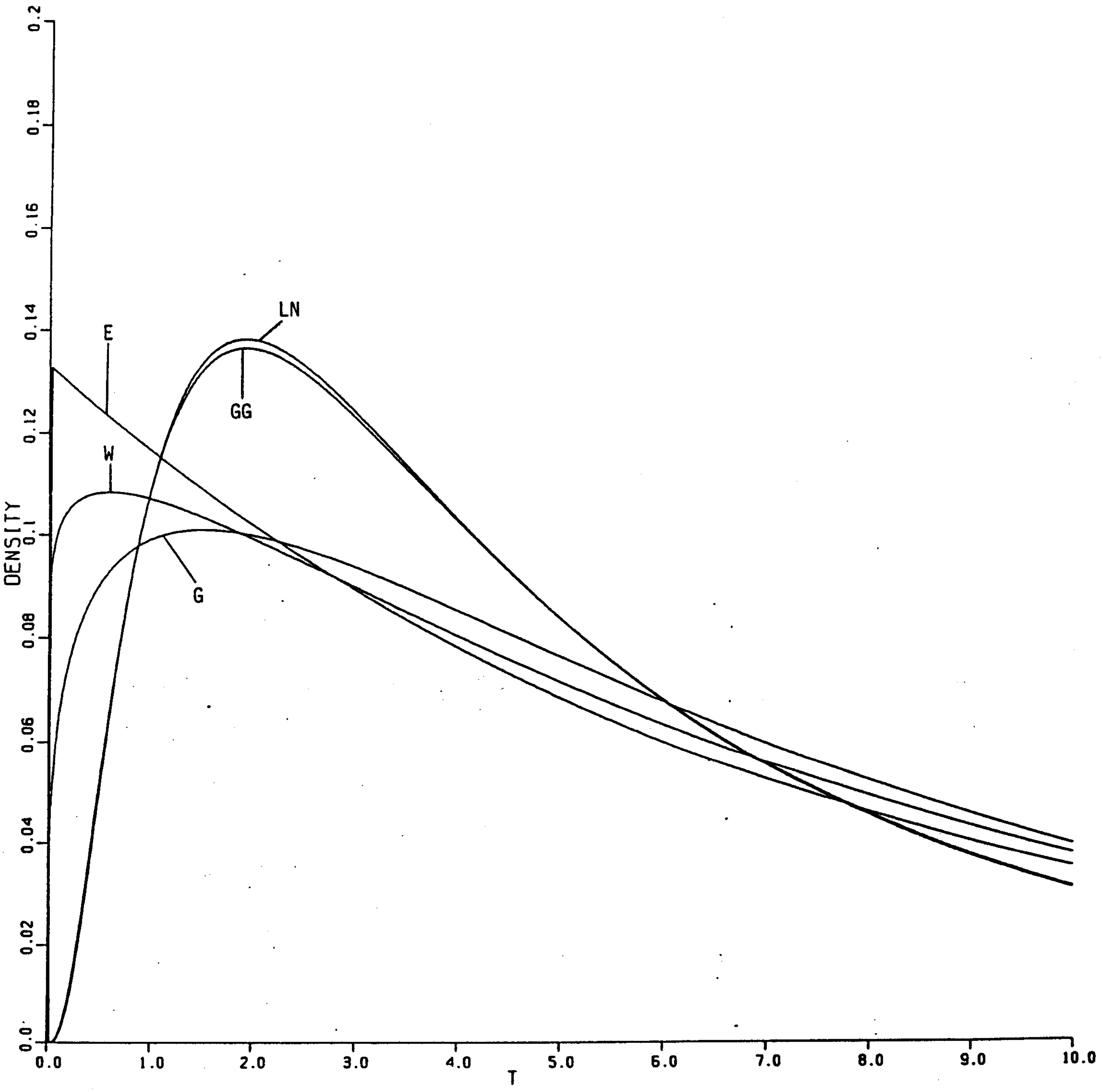
DENSITY
0.0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16
0.18
0.2
T
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
LN
E
GG
W
G

FIGURE 1

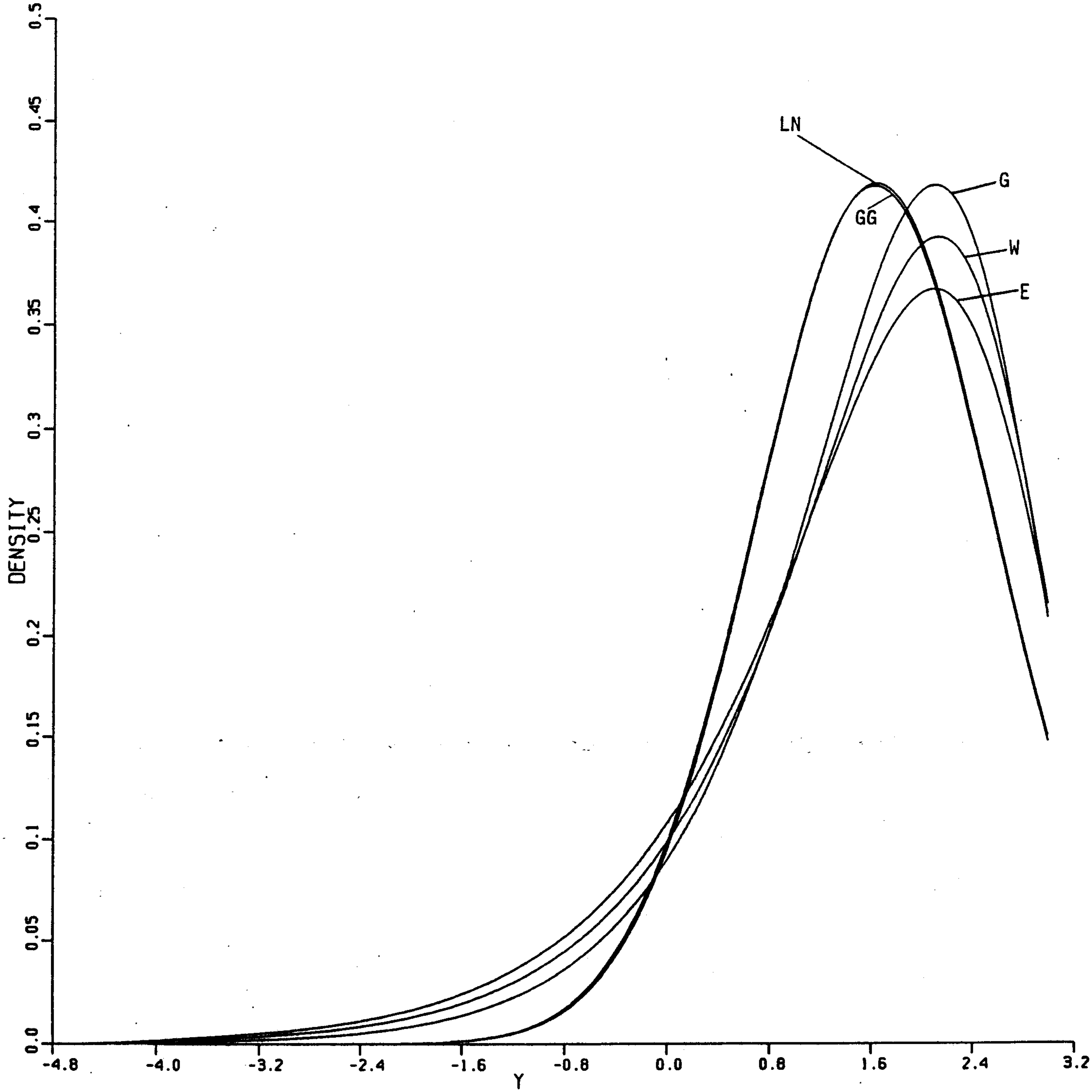
LN
GG
G
W
E
DENSITY
0.0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
-4.8
-4.0
-3.2
-2.4
-1.6
-0.8
0.0
0.8
1.6
2.4
3.2
Y

FIGURE 2

is important to assess whether the alternative formulations yield results which are statistically significantly different. We must appeal to the asymptotic distributions of the appropriate test statistics to evaluate the null hypotheses of model equivalence. Lee [1984] proposes Lagrange Multiplier (LM) tests for specific accelerated failure time models. This procedure can be used here to test estimates of a restricted model (for example: linear E) against a more general model (for example: linear W). Alternative tests are the Wald and the Likelihood Ratio (LR) tests. It is computationally convenient to employ the Hessian formulas from Section 3 in computing the Wald and LM tests. Following Amemiya [1983,p.350] the appropriate test statistics are:

$$(10)\ \mathrm{LR} = 2\ [\ell(\hat{\gamma})-\ell(\tilde{\gamma})]$$

$$(11)\ \mathrm{Wald} = (\hat{\gamma}-\gamma)'[V(\hat{\gamma})]^{-1}(\hat{\gamma}-\gamma)$$

$$(12)\ \mathrm{LM} = [\ell^{(1)}(\tilde{\gamma})]'\{\ell^{(2)}(\tilde{\gamma})\}^{-1}[\ell^{(1)}(\tilde{\gamma})]$$

where $\hat{\gamma}$ is the vector of unrestricted estimated parameters and $\tilde{\gamma}$ is the vector of parameters subject to a set of distribution restrictions. In our case $\hat{\gamma}$ includes the parameters (β,k,c) and $V(\hat{\gamma})$ is the estimated covariance matrix of parameters.

Comparison of the maximized value of the log-likelihood function under alternative distributional assumptions is a simple method for choosing among the GG, G, W, LN, and E models. Hypotheses concerning the adequacy of special cases such as the G, W, or LN models can be formally tested using the appropriate LR, LM, or Wald tests against the more-general GG distribution. Similarly, these formal tests can be used to compare the E model against either the G or W models. Tables VI and VII give the values of the test statistics for each

TABLE VI

SPECIFICATION TEST STATISTICS: CANADA AND U.S. DESTINATIONS

LINEAR REGRESSION

Comparison	LR	LM	Wald
E:G	643.04	548.53	547.95
E:W	173.58	169.05	169.18
G:GG	2778.84+	4581.80	-
W:GG	3248.30+	771.29	-

+using highest value of likelihood achieved for GG

LOG-LINEAR REGRESSION

Comparison	LR	LM	Wald
E:G	653.84	926.19	556.38
E:W	175.36	197.0	195.47
G:GG	2819.22	7820.5	-
W:GG	3297.70	778.87	-

*5% critical value for $\chi^2(1) = 3.84$

TABLE VII

SPECIFICATION TEST STATISTICS: OVERSEAS DESTINATIONS

LINEAR REGRESSION

Comparison	LR	LM	Wald
E:G	274.8	220.54	220.15
E:W	161.8	152.07	152.56
G:GG	254.4	349.34	290.50
W:GG	367.4	112.26	6.1381
LN:GG	26.0	-	-

LOG-LINEAR REGRESSION

Comparison	LR	LM	Wald
E:G	274.54	465.16	219.94
E:W	161.78	209.16	201.66
G:GG	253.14	104.66	15.865
W:GG	365.90	112.04	6.304
LN:GG	26.44	-	-

5% critical value for $\chi^2(1) = 3.84$

pairwise comparison between nested models (within the linear and log-linear families), for the Canada/U.S. and overseas samples respectively. As we might expect in very large samples the null hypotheses of model equivalence are soundly rejected in all cases. Although the computed χ^2 statistics for the three tests are often quite different numerically, there is apparently no conflict in the test results. In all cases the tail probability values associated with the observed values of the test statistics are extremely small. Nevertheless, it is interesting to note the relative magnitudes of the statistics. The Berndt-Savin [1977] inequality (Wald ≥ LR ≥ LM) for linear normal models is not expected to hold here and it does not. We find the LM and Wald statistics to be quite close in comparing E to either G or W models, while the corresponding LR statistic is higher. Tests against the GG model show much less agreement. In this case the LR or LM statistic is always the largest. Magee [1984] has shown that conflict among the test statistics can often be explained by examination of the third and fourth derivatives of the log-likelihood function. His results indicate that when the third and fourth derivatives are negative, the LR statistic will be the smallest. In cases where the third derivative is not negative the LR statistic will be the largest. As an illustration of Magee's result in the present case we can easily compute the third and fourth derivatives of the likelihood function for the G distribution and examine the ranking of the test statistics of the E distribution against the G distribution in the linear model. The third and fourth derivatives of the linear G log-likelihood function with respect to shape parameter k are:

(13) $\ell^{(3)}(k) = n\ [-k^{-1} - \psi^{(2)}(k)]$

(14) $\ell^{(4)}(k) = n\ [2k^{-3} - \psi^{(3)}(k)]$,

Tables VI and VII show that the LR statistics in this test always exceed the corresponding LM and Wald statistics. For the linear G model, evaluation of the third and fourth derivatives at the maximum likelihood estimates of k resulted in $\ell^{(3)}=14874$ and $\ell^{(4)}=-37962$ for the Canada-U.S. sample and $\ell^{(3)}=2532.1$ and $\ell^{(4)}=-5898.7$ for the Overseas sample. Under these circumstances Magee reports that we should obtain $LR \geq \max(W,LM)$ which is exactly what happened. While computation of the third and fourth derivatives is often quite difficult we have shown that in a few cases it is relatively simple and can be used to explain conflicts among the test statistics. Examination of Tables VI and VII shows that the LR statistic was the largest in all tests except G versus GG, in which case it was the smallest.

Model discrimination among the G, W, and LN models is more difficult since these are non-nested. However, the distribution of the likelihood ratio test statistics for some non-nested tests have been tabulated. For example, Dumonceaux and Antle [1973] provide small sample tables (derived by Monte Carlo methods) for the log-linear W and LN models. Unfortunately these tables are not sufficiently detailed for use here.

Note that it is not possible to compare the linear and the log-linear families of models using the techniques appropriate for nested models. While the familiar Box-Cox transformation is often invoked to compare linear and log-linear models, it is not appropriate here. The density function, in this case, is modified to accommodate the log transformation of the dependent variable. The Box-Cox approach assumes the same error distribution for both the linear and log-linear specifications.

6. Policy Implications

One outcome of deregulation in long-distance markets has been decreases in marginal rates. This is likely to increase both the number of calls and the mean duration of calls. The fitted models obtained above can be employed to shed some light on current rate-setting issues facing telephone utilities. Although the model is not designed to predict the number of new calls induced by a change in rates, one can predict the influence of hypothetical long-distance rate reductions upon the durations of existing calls.

The estimated coefficients reported in Tables IV and V can be used with the summary statistics reported in Table III to generate estimates of the "rate elasticity of duration" at the mean values of the observed data. In the Canada - U.S. sample these estimated rate elasticities are nearly zero for the log-linear regressions. While the rate elasticity for the Normal Linear model was also small (-.0925), the non-normal linear model rate elasticities were over twice as large, ranging from -.1948 to -.2146. This indicates that substantial losses in revenue on existing calls are to be expected as a result of any decreases in marginal rates. For example, a 25% reduction in marginal rates as a result of deregulation (reductions of this magnitude are likely) could increase duration at most by 5% according to these results. The expected 20-25% loss in revenue will need to be offset by an increase in the number of calls or increases in local service charges. As new services such as MCI and SPRINT in the U.S. and CNCP in Canada offer these rate reductions we are likely to observe substantial increases in their business at the expense of traditional companies such as AT&T and Bell Canada. However, our results indicate that total industry revenue will decline unless it is offset from other sources.

Although deregulation has had little impact on overseas rates the results indicate similar rate elasticities of duration to those in the Canada - U.S. sample. In the Overseas sample the calculated rate elasticities at the means ranged from -.2054 to -.2763 for the GG and LN models and from -.0114 to -.0506 for the remaining models. In this case, use of a model with the wrong distributional assumptions could lead to an understatement of the magnitude of the rate elastiticity.

7. Conclusions

This paper has demonstrated that in regression applications where the dependent variable is strictly non-negative the usual normal conditional probability density function may be inappropriate. Whereas a log-linear regression model with normal errors may sometimes suffice, we have experimented with a wide range of linear and log-linear models with error distributions in the GG family of distributions.

In our application of these alternative models to long-distance telephone call durations, we have utilized two samples, two assumptions about functional form, and five major assumptions about error distributions (the GG and its four special cases). We have found that different error assumptions can result in considerably different point estimates for the regression coefficients. Lagrange Multiplier, Wald, and Likelihood Ratio tests have been employed to distinguish between nested pairs of models. Unfortunately, the two classes of models, linear and log-linear, are not formally comparably by these methods because they are non-nested.

Overall, the simple log-linear model with normal errors seems to provide a good fit for our sample of calls with Canada/U.S. destinations. In contrast, the linear model with GG errors fits better for our sample of overseas calls. The regression technique improves upon previous correlation studies between duration and other variables; it also formalizes the fitting of non-normal distributions for different categories of calls. Qualitatively, our results regarding the determinants of duration are generally consistent with previous findings, but the regression model represents a more systematic mode of analysis and should therefore yield more reliable quantitative estimates. The estimated rate elasticities indicate that rate reductions are likely to result in a small increase in call duration. As a result, the revenue loss on existing calls must be offset by a substantially increased number of calls or higher charges for local service.

In general, we conclude that researchers seeking to explain the determinants of a strictly non-negative dependent variable should not limit themselves to a log-linear, normal-error model. There are more-flexible distributional assumptions which can be made, and linear regression models can quite readily be adapted to non-negative distributions. The usual assumption of log-linear functional relationship between the dependent and the explanatory variables is not mandatory. At the very least, one should perform a Lagrange Multiplier specification test to assess whether generalization to the log-linear GG model is warranted.

REFERENCES

[1] Amemiya, T.: "Nonlinear Regression Models", Chapter 6, in Griliches and Intriligator, Handbook of Econometrics, North-Holland, 1983.

[2] Berndt, E. and N.E. Savin: "Conflict Among Criteria for Testing Hypotheses in the Multivariate Linear Regression Model", Econometrica, 45, (1977), pp. 1263-1278.

]3] Box, G., and D. Cox: "An Analysis of Transformations", Journal of the Royal Statistical Society, Series B, (1964), 211-252.

[4] Brandon, Belinda B. (ED.): The Effect of the Demographics of Individual Households on their Telephone Usage. Cambridge, Massachusetts: Ballinger, 1981.

[5] Curien, N.: "Modelisation de l'effet des tarifs sur la consommation et le trafic; Application a l'etude de la taxation des communications locales a la duree", Note DGT, 1981.

[6] de Fontenay, Alain, Debra Gorham, J.T. Marshall Lee, and George Manning: "Stochastic Demand for a Continuum of Goods and Services: The Demand for Long Distance Telephone Services", Paper presented to the Transportation and Public Utilities Group Session of the American Economic Association Meetings. New York, New York: Dec. 29, 1981.

[7] Dumonceaux R., and C. Antle: "Discrimination Between the Lognormal and Weibull Distributions", Technometrics, 15, pp. 923-926.

[8] Gale, W.A.: "Duration of Interstate Calls, March 1969," unpublished Bell Laboratories Memorandum, December 1971.

[9] Gale, W.A.: "Elasticity of Duration for Intrastate Calls," unpublished Bell Laboratories Memorandum, October 1974.

[10] Heckman, James J., and Burton Singer: "Econometric Duration Analysis", Journal of Econometrics, 24, (1984),pp. 63-132.

[11] Jensik, John M.: "Dynamics of Consumer Usage," in Perspectives on Local Measured Service, ed. by J.A. Baude et al. Kansas City, Mo.: Telecommunications Industry Workshop Organizing Committee, 1979.

[12] Johnson, N.L., and S. Kotz: Continuous Univariate Distributions (1), New York, New York: John Wiley and Sons, 1970.

[13] Lawless, J.F.: Statistical Models and Methods for Lifetime Data. New York, New York: John Wiley and Sons, 1982.

[14] Lee, Lung-fei: "Maximum Likelihood Estimation and a Specification Test for Non-normal Distributional Assumptions for the Accelerated Failure Time Models", Journal of Econometrics, 24 (1984), 159-179.

[15] Magee, L.: "Some Inequalities for LR, W, and LM Tests", McMaster University, 1984. Presented at First Annual Meeting of Canadian Econometrics Study Group, Kingston, Ontario, September 1984.

[16] McDonald, James B.: "Some Generalized Functions for the Size Distribution

of Income", Econometrics, 52 (1984), 647-663.

[17] Pacey, Patricia L.: "Long Distance Demand: A Point-to-Point Model", Southern Economic Journal, 49 (1983), pp. 1094-1107.

[18] Park, Rolla Edward, Bruce M. Wetzel, and Bridger M. Mitchell: "Price Elasticities for Local Telephone Calls", Econometrica, 51 (1983),

[19] Pavarini, Carl: "The Effect of Flat-to-Measured Rate Conversions on Local Telephone Usage," in Pricing in Regulated Industries: Theory and Application, ed. by J. Wenders. Denver: Mountain States Telephone and Telegraph Co., 1979.

[20] Rea, John D. and Lage, Gerald M.: "Estimates of Demand Elasticities for International Telecommunications Services", Journal of Industrial Economics, 26 (1978), 363-381.

[21] Taylor, Lester D.: Telecommunications Demand: A Survey and Critique. Cambridge, Massachusetts: Ballinger, 1980.

[22] White, Kenneth J.: "A General Computer Program for Econometric Models - SHAZAM", Econometrica, 46 (1978), 239-240.

[23] Wong, T.F.: "Identifying Tariff-Induced Shifts in the Subscriber Distribution of Local Telephone Usage", Bell Laboratories, 1981.

APPENDIX A: DATA

Two specific subsets from a stratified sample of approximately 65,000 long distance calls were obtained for the Canadian Province of British Columbia. The first subsample consists of calls originated by residential service customers with destinations either in Canada or the U.S. (except Alaska). These represent about 41 percent of the full sample. Because we suspect that there may be systematic differences between calling patterns from residences and businesses (i.e. different perceived budget constraints, differing flexibility, different availability of substitutes), we chose not to incorporate both types of calls in this first model. The second subsample consists of both residential and business calls with destinations overseas.

Most models of telephone demand include a variable for marginal per-minute rates (RATE). Here, we can be much more accurate than studies using aggregated data because we know the exact price of an extra minute for every individual call (computed directly from the rate schedule in effect when our sample was collected). Deducing the effective marginal rate requires a distinction between calls with one-minute and three-minute minimum charges. The data records only whole numbers of minutes and callers must always pay for at least one minute. For direct distance dialed (DDD) calls, with their one-minute minimum, the telephone company rate schedules were used to determine the price for "each additional minute". For operator-handled and person-to-person calls, however, a three-minute minimum usually applies. If the observed call duration is three minutes or longer, the marginal price is again drawn from the rate schedule as the price for "each additional minute". However, if the observed duration is less than three minutes, the marginal price is zero, since the full three minutes must be paid for regardless of whether they are used.

As Gale [1971] has found, distance is another factor likely to affect duration of a long distance call. A crude plot of the results derived by de Fontenay, Gorham, Mitchell and Lee suggests that the relationship is highly non-linear. The logarithm of distance, LOG(DIST) is much more appropriate (and, indeed, yields a much higher value of the likelihood function in any specification). More remote calls may display longer durations because such calls are made less frequently. An overall "budget constraint" for calls at greater distances may encourage subscribers to substitute duration for frequency of calls.

For the Canada/U.S. subsample, the data provide exact mileages between each origin and destination, except for the 433 calls to Hawaii. Distances are irrelevant to the official billing formula for these calls, since rates to Hawaii are identical from anywhere in British Columbia. We opted to construct an approximate distance for each of these calls, based on the air-miles from Vancouver to Hawaii (2704) plus or minus the distance from Vancouver to the city where the call originates. For the overseas subsample, no mileage data were provided at all. Distances had to be approximated by using air-miles from the city of Vancouver to either the capital city or the largest city of the destination country.

The time-of-day at which the call is made also affects duration, even beyond the influence of differing marginal prices. All calls in the Canada/U.S. sample were subject to three different rate periods. "Daytime" calls, defined by the time period with the highest rates, (either 8:00-17:00 or 8:00-18:00), are considered to be the base period. Dummy variables were then defined for the "evening" time period (EVENING, either 17:00-23:00 or 18:00-24:00) and for the night time period (NIGHT). For overseas calls, only two rate periods are defined. This difference in billing practices explains the need for separate analyses for the two samples. The delimiting hours vary considerably by call destination in the overseas sample. We have arbitrarily termed the off-peak discount period NIGHT, but the actual time period involved could easily overlap daytime hours in either location.

Type of service has also been recognized to influence call lengths. We undertook some grouping of values of the categorical variables provided in the raw data. Most calls are designated as "sent paid". We defined one dummy variable, COLLECT, to indicate either "collect" or "special collect" calls, and a second dummy variable, CARD, signifying credit card or third number calls. Among call classes, we grouped direct-distance-dialed calls (DDD), DDD-equivalent calls, and operator-handled station-to-station calls. A dummy variable, PERSON, was then defined for person-to-person and "person call back" classes, since we suspect that the nature or importance of these messages may have a systematic influence on call duration. The model for overseas calls of course includes an additional dummy variable, BUSINESS, to discriminate between business and residence calls.

Studies using aggregated data must select as dependent variables either the total number of calls of each duration (regardless of destination) or the total number of minutes for all calls. By using completely disaggregated data, we can control for each individual factor suspected to contribute to the explanation of call length. Further, by focussing on such a narrow "window" in time, we have effectively controlled for a host of other factors which might confuse the estimation process if they were not entered explicitly. Day-of-week effects are certainly present, as are seasonal effects. A single day's sample, which controls for such factors, is particularly expedient given the complexity and expense of estimating regression models with non-normal errors. Summer data were chosen to minimize the discrepancies in temperature between different regions of the province.[5]

One last dummy variable responds to the observation that the population is

[5] July 13 was chosen simply because it was the "middle" Wednesday in the month which can be considered the middle of the summer.

heavily concentrated in the southwest corner of the province, where ocean currents result in a moderate climate throughout the year. The interior and northern regions of the province, however, experience much more extreme weather conditions, especially in winter. Since it is widely held that telephone calls are longer when the weather is bad, we opted to avoid these effects as much as possible by focussing on summer conditions. Still, different patterns of settlement may affect calling behavior in the Interior and in the North even in the summer months. Lower population densities may mean that a larger population of a household's calls are long distance calls. The NORTHEAST dummy variable, then, is intended as a very loose proxy variable for demographic effects and for geographical dispersion, and takes on a zero value for the divisions labelled "Coastal West", "Lower Mainland", and "Island" (the most densely-populated areas). It takes a value of unity for the "Northern", "Interior" and "Coastal East" divisions.

Finally, it should be noted that the coded data included several categories of calls which we judged sufficiently atypical to warrant their exclusion from the model a priori. We imposed automatic deletion of any PABX/PBX (private automated branch exchange) calls from the residential Canada/U.S. sample, and coin calls, calls from hotels, centrex calls, and miscellaneous or unidentified calls from both samples. We retained only those residential calls classified as individual, two-party or multiparty service. We eliminated "collect to coin" and "coin paid" calls (less than one percent), few of which are likely to survive previous deletion criteria. We deleted "radio and two-number calls" (less than four percent). The absence of data on time of initiation was also deemed sufficient to exclude calls (less than one percent). A number of these atypical calls recorded an associated distance value of zero. In general, then, since distance is presumed to be an important explanatory variable, we opted to establish that calls spanning less than five miles could not be considered "long-distance." This criterion excluded about seven percent of the original sample.