Scientific
Research
Publishing

Computing Confidence Intervals for the Postal Service's Cost-Elasticity Estimates

Bzhilyanskaya Y. Lyudmila, Margaret M. Cigno, Soiliou D. Namoro*

US Postal Regulatory Commission, Washington DC, USA
Email: *soiliou.namoro@prc.gov

Abstract

This paper provides methods for assessing the precision of cost elasticity estimates when the underlying regression function is assumed to be polynomial. Specifically, the paper adapts two well-known methods for computing confidential intervals for ratios: the delta-method and the Fieller method. We show that performing the estimation with mean-centered explanatory variables provides a straightforward way to estimate the elasticity and compute a confidence interval for it. A theoretical discussion of the proposed methods is provided, as well as an empirical example based on publicly available postal data. Possible areas of application include postal service providers worldwide, transportation and electricity.

Keywords

1. Introduction

Around the world, posts are confronted with declining volumes and revenues. In such an environment, it is vitally important that price signals are properly constructed using accurate cost data at the product level. Because posts are multi-product firms they exhibit complex cost behaviors. The presence of common costs and the thin line between fixed and variable costs make it challenging for posts to accurately determine cost elasticity by product. The United States Postal Service (Postal Service) uses quadratic and translog regression equations, among others, for estimating volume variabilities of cost-related variables, such as street time or vehicle capacity, with respect to diverse cost drivers, such as mail volume [1] [2] [3]. The term "volume variability" (or variability) is used by the Postal

Service as a substitute for "cost elasticity" [2]. The Postal Service uses both terms as synonyms to mean "the percentage change in cost caused by a percentage change in volume (or other relevant cost driver)" [4].

The present paper is motivated by the fact that confidence intervals are not reported with the Postal Service's variability estimates. If available, confidence intervals allow assessment of the precision of the estimates and help gauge their stability across alternative independent data samples. In addition, the computation of the relevant confidence intervals provides direct information on the potential effects of covariate multicollinearity on the variability estimates.

Considering that cost elasticities are estimated not only by the United States Postal Service, but also by posts worldwide, the methods described in the paper could be of wide interest [5]. In addition, the described analysis of confidence intervals for cost elasticity estimates could be applicable in other industries, such as transportation or electricity.

The Postal Service computes variability in two steps. In the first step, the Ordinary Least Squares or the generalized least Squares method is used to estimate a chosen econometric model linking the cost-related (dependent) variable to a set of explanatory variables. The estimated relation is treated as the cost-related function. In the second step, the variability is computed as the elasticity of the cost-related function with respect to the variable of interest. Since the results vary across observations, a fixed variability estimate is computed as the elasticity at the sample average of the explanatory vector. The obtained elasticity is a ratio, the numerator and denominator of which are both linear combinations of the estimated parameter vector, with coefficients depending on the sample averages of the explanatory variables. These coefficients converge in probability to constants, thanks to the weak law of large numbers (WLLN) and provided that the involved variables have finite expectations. Hence, the asymptotic normality of the estimator of the parameter vector of the model provides the basic information needed to test the significance of the variability estimate and construct a confidence interval for the implicit variability parameter. Of course, the ratio nature of the estimator raises issues about the coverage probability when the implicit variability parameter is undefined, for example because its denominator is statistically null. Indeed, as shown later in the paper, the denominator can be obtained as the estimate of the constant term in the regression of the dependent variable on the mean-centered versions of the same explanatory variables. There is, however, a possibility that this estimate is insignificant.

This paper presents a method to compute the confidential interval and conduct the relevant tests using the estimates of the regression model with mean-centered explanatory variables, *i.e.*, the main variables are all centered and the interactions variables are computed using the centered variables. As Echambadi and Hess [6] showed, for linear (in parameters) models containing interaction variables, there is a linear and one-to-one correspondence between the OLS (or the generalized least-squares (GLS)) estimates obtained with non-centered variables

versus mean-centered variables. The statistical inference for the two models is also shown in the cited paper to be exactly the same, provided that the linear transformation between the parameter vectors in the two models is accounted for in the inference. This is an example of equivariant estimators [7].

A practical advantage of using the centered regression compared to the one with non-centered explanatory variables is that any statistical software will directly provide the numerator (up to a factor equal to the mean of the cost driver) and the denominator of the variability estimates and compute the confidence intervals for the corresponding parameters. Hence, a byproduct of the proposed method is that it also provides a simpler alternative, and numerically equivalent way to compute the variability parameter. The analysis conducted in this paper is of potential interest to any application involving elasticity estimates and it falls in the literature on elasticity estimation, as illustrated in [8] [9] [10] and the references therein.

The purpose of the paper is to present a simpler way to estimate quadratic-form regression functions and two algorithms for computing confidence intervals for the Postal Service's cost elasticity estimates. More generally, the paper purports to present within the context of elasticity estimation, tools for computing confidence intervals for a parameter that is estimated as a ratio of a pair of asymptotically normal estimators.

In Section 2, the quadratic model is briefly discussed with the goal of providing the precise formal setting in which the analysis is conducted. In Section 3, the confidence interval for the variability parameter is derived from the delta method. Section 4 discusses the inferential problems related to the delta method in this context and shows how to derive an alternative confidence interval based on Fieller's method. The empirical computation of the two types of confidence interval is discussed in Section 5. The last section concludes the paper.

2. The Formal Setting

A class of regression equation often considered in applications has the form:

$$y = \beta_0 + x'\beta + x'Ax + u. \tag{1}$$

Regression functions that are not polynomial in the explanatory vector will require a less straightforward analysis than the one proposed in the present paper. In the relation (1), vectors are understood as column vectors and the prime sign denotes transposition.

For example, in the Postal Service's model, y is a cost-related variable, x is a row-vector of explanatory variables including mail volume, and u is the usual error term. In this paper, the explanatory variables are the components of the vector x. These components, together with their interactions are referred to as the covariates.

Assuming that u is mean-independent of x, the right-hand side of (1), excluding the error term, is the conditional expectation of y, given $X = x$, where x is an observation of X. To avoid complicating the notation, expectation operators

will be applied indifferently to both a random variable and its observation and will mean the same thing even though the latter is not random. Hence, for example, $E(X)$ and $E(x)$ will both denote the mathematical expectation of X and this use will be contextually unambiguous. For two explanatory variables, the model is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + a_{11} x_1^2 + a_{22} x_2^2 + a_{12} x_1 x_2 + u \qquad (2)$$

$$\beta = (\beta_1, \beta_2)', A = \begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix} \qquad (3)$$

A point worth noting is that the equality $x'Ax = \dfrac{x'(A+A')x}{2}$, which uses the symmetric part of A, *i.e.*, $B = A + A'$, allows us to write model (1) equivalently as

$$y = \beta_0 + x'\beta + \frac{x'Bx}{2} + u = g_\theta(x) + u. \qquad (4)$$

The index θ is a vector the component of which are the parameters of the model. Model (4) is referred to in this paper as the quadratic regression model. The version of it in which y and all components of x are in the logarithm form is referred to as the translog regression model, after the function introduced by Christensen *et al.* [11]. It is easily seen that although quadratic in x, the model is linear in the parameters.

3. The Delta-Method Confidence Interval

The developments in this section start from the observation that the quadratic function $g_\theta(x)$ is analytic and, hence, it is its own Taylor expansion centered at any point of its domain. From this observation, it follows that one can equivalently write

$$g_\theta(x) = \alpha_0 + (x-e)'\alpha + \frac{(x-e)' B (x-e)}{2}, \qquad (5)$$

for any interior point e of the domain of g_θ, where

$$\alpha = \left(\frac{\partial g_\theta(x)}{\partial x_1}, \cdots, \frac{\partial g_\theta(x)}{\partial x_d} \right)' \Bigg|_{x=e} = \nabla_{g_\theta}(e) \text{ is the gradient of } g_\theta \text{ at } e \text{ and } d \text{ is the}$$

dimension of the vector x. The motivation for representing the argument of the function g_θ as the deviation of x from e is the direct meaning of α as the gradient $\nabla_{g_\theta}(e) = \beta + Be$ (obtained from taking the derivative with respect to x in (4) and substituting e for x in the result), which, computationally has the advantage of being read directly from an estimation output, using any of the common statistical software packages. Let z denote the component of x, which represents the explanatory variable of interest, and x_{-z} be the remaining sub-vector of x, after discarding z. The following equality results from these notations: $x = (z, x'_{-z})'$, assuming that z is the first component of x. The formal expression of the elasticity to be computed at some chosen value of x, say $x_0 = (z_0, x'_{-z_0})'$, is

610

$$\gamma(x_0) = \frac{\dfrac{\partial g_\theta(x)}{\partial z}}{g_\theta(x)} z \Bigg|_{z=z_0,\, x_{-z}=x_{-z_0}} . \tag{6}$$

The value x_0 can be taken to be the expected value of x, *i.e.*, $x_0 = E(x)$, which is the case that is considered in this paper. In other words, the parameter of interest in the analysis is

$$\gamma(E(x)) = \frac{\dfrac{\partial g_\theta(x)}{\partial z}}{g_\theta(x)} z \Bigg|_{x=E(x)} := \frac{\dfrac{\partial g_\theta(E(x))}{\partial z}}{g_\theta(E(x))} E(z) \tag{7}$$

The least-square estimation of the model

$$y = \alpha_0 + (x - E(x))' \alpha + \frac{(x - E(x))' B (x - E(x))}{2} + u = g_\delta(x - E(x)) + u, \tag{8}$$

where $g_\delta(x - E(x)) \equiv g_\theta(x)$ and δ denotes the vector of the parameters in the expression on right-hand side of the first equality in (8), is performed by mean-centering each component of x in (4).

Echambadi and Hess [6] showed that there is a linear one-to-one correspondence between $\hat{\theta}$ and $\hat{\delta}$, the ordinary least squares (OLS) estimators in the regression in the regression with non-centered and centered variables, respectively. These two regressions will be henceforth referred to as the non-centered and the centered regression, respectively. So, assuming again that z is the first component of x and its sample mean is $\bar{z}$, directly computing the OLS estimate of α_z, the coefficient of $z - \bar{z}$ in (8), is numerically equivalent to first computing the OLS estimate of θ in (4) and then computing $\hat{\beta}_z + \hat{B}_z \bar{x}$, where $\hat{\beta}_z$ and $\hat{B}_z$ are, respectively, the estimate of the coefficient of z and the first row of $\hat{B}$, the estimate of B.

The least-square estimator of $\gamma(E(x))$, which is denoted by $\hat{\gamma}(E(x))$, only involves a ratio, $\dfrac{\hat{\alpha}_z}{\hat{\alpha}_0}$ where the numerator is the estimated coefficient of $z - \bar{z}$, and the denominator is the estimated constant term, equal to $g_\theta(\bar{x})$. Because $\bar{x}$ converges in probability to $E(x)$ (by the WLLN), the continuous mapping theorem (see theorem 1 in Borovkov [12]) implies that the estimator

$$\hat{\gamma}(E(x)) = \frac{\hat{\alpha}_z}{\hat{\alpha}_0} \bar{z}$$

is consistent (for estimating $\gamma(E(x))$, provided that the latter is defined, *i.e.*, that its denominator is non null [10]. Further, the couple $(\hat{\alpha}_z, \hat{\alpha}_0)$ is asymptotically normal. Assuming that its asymptotic distribution is

$$\begin{pmatrix} \hat{\alpha}_z \\ \hat{\alpha}_0 \end{pmatrix} \sim N\left[\begin{pmatrix} \alpha_z \\ \alpha_0 \end{pmatrix}, \begin{pmatrix} \hat{\sigma}_z^2 & \hat{\sigma}_{zm} \\ \hat{\sigma}_{zm} & \hat{\sigma}_0^2 \end{pmatrix} \right], \tag{9}$$

the corresponding asymptotic distribution of $\hat{\gamma}(E(x))$ is

$$\hat{\gamma}(E(x)) \sim N\left[\gamma(E(x)), \bar{z}^2\left(\frac{1}{\hat{\alpha}_0}, -\frac{\hat{\alpha}_z}{\hat{\alpha}_0^2}\right)\begin{pmatrix}\hat{\sigma}_z^2 & \hat{\sigma}_{zm} \\ \hat{\sigma}_{zm} & \hat{\sigma}_0^2\end{pmatrix}\begin{pmatrix}\frac{1}{\hat{\alpha}_0} \\ -\frac{\hat{\alpha}_z}{\hat{\alpha}_0^2}\end{pmatrix}\right] \tag{10}$$

$$N\left[\gamma(E(x)), \left(\frac{\bar{z}}{\hat{\alpha}_0}\right)^2\left[\hat{\sigma}_z^2 + \hat{\sigma}_0^2\left(\frac{\hat{\alpha}_z}{\hat{\alpha}_0}\right)^2 - 2\hat{\sigma}_{zm}\left(\frac{\hat{\alpha}_z}{\hat{\alpha}_0}\right)\right]\right]. \tag{11}$$

Using the notation $\hat{w} = \left[\hat{\sigma}_z^2 + \hat{\sigma}_0^2\left(\frac{\hat{\alpha}_z}{\hat{\alpha}_0}\right)^2 - 2\hat{\sigma}_{zm}\left(\frac{\hat{\alpha}_z}{\hat{\alpha}_0}\right)\right]$, the delta-method

based 95%-level confidence interval for $\gamma(E(x))$ is written as

$$\left[\hat{\gamma}(E(x)) - 1.96\sqrt{\hat{w}\left(\frac{\bar{z}}{\hat{\alpha}_0}\right)^2}; \hat{\gamma}(E(x)) + 1.96\sqrt{\hat{w}\left(\frac{\bar{z}}{\hat{\alpha}_0}\right)^2}\right] \tag{12}$$

A robust version of (12) can be obtained by using, for example, the White's heteroskedasticity-consistent covariance matrix estimator for the covariance matrix in (9) [13].

4. The Fieller-Method Confidence Interval

Several researchers have expressed doubt pertaining to a class of confidence-intervals for a ratio, including the one obtained by the delta-method. Doubt about the confidence interval estimated using the delta method arises from the fact that the delta method suggests that a ratio of two normally distributed random variables will be normally distributed, which is not true. The delta method also relies crucially on the assumption that the denominator of the "true" value of the ratio parameter of interest is not null.

To get more insight about a potential problem posed by the delta-method confidence interval, a few definitions are necessary. The presentation here is based on Gleser and Hwang [14]. Suppose some variable of interest X has the parametric probability distribution $A \mapsto P_\theta(X \in A)$, where the finite-dimensional parameter vector θ lies in the set Θ. A confidence interval for a function $\gamma(\theta)$ of the parameter vector has the form $[L(X), U(X)]$. It is a random interval, because its bounds depend on the random variable X. Of course, for a particular sample, x of X, $[L(x), U(x)]$ is a fixed (non-random) interval. However, the statistical properties of the confidence interval are studied based on the random interval. One important property is the coverage probability of the interval, which, for a given $\theta \in \Theta$, is defined as the probability p_θ, that the random interval contains the parameter of interest, $p_\theta = P_\theta[L(X) \le \gamma(\theta) \le U(X)]$. The lower bound, $(1-\alpha) = \inf_{\theta \in \Theta}(p_\theta)$, of the coverage probability, where α is here a positive real falling in the interval $(0,1)$, is called the confidence level of the confidence interval and represents the theoretical lowest probability that the random interval contains the parameter of interest. To assess the confidence interval, one additional piece of information is needed, namely, its expected length,

which, for a given θ, is the expectation $l(\theta)=E_{\theta}\left[U(X)-L(U)\right]$. Clearly, a larger $(1-\alpha)$ and a lower $l(\theta)$ across the set Θ, makes the interval more appealing for the researcher than it would be otherwise.

In some important cases, such as building a confidence interval for a ratio of two OLS regression parameters, these desirable properties of a confidence interval may be out of reach when the intervals are obtained by a class of methods that includes the delta method. Confidence intervals of the form (12) fall in a class that Dufour [15] (page 1366) refers to as Wald-type confidence sets.

Specifically, Gleser and Hwang [14], Koschat [16], and Dufour [15] have proved what is known in the relevant literature as *impossibility theorems*. Recall that in the context of the present paper, α_z and α_0 are the expectations of the limiting jointly normal distribution of $(\hat{\alpha}_z, \hat{\alpha}_0)$, and the variability estimator is based on the ratio between the least-squares estimators of α_z and α_0. Koschat [16] shows that, in general, "[T]here is no procedure that with probability 1 gives bounded α-level confidence intervals for the ratio". Dufour [15] also shows that any confidence interval of the form $\left[\left(\dfrac{\hat{\alpha}_z}{\hat{\alpha}_0}\right)\bar{z}\pm\tau^*\sqrt{\hat{w}\left(\dfrac{\bar{z}}{\hat{\alpha}_0}\right)^2}\right]$, (say,

$\tau^*=1.96$), derived, for example, from the estimation of a linear regression equation, such as (12), has zero coverage probability.

Fortunately, there is a way out of these impossibility theorems. Both Koschat and Dufour indicate that the Fieller confidence interval is a better alternative for estimating a ratio [17]. The general principle that underlies the procedure for obtaining a Fieller confidence interval is clearly explained in Rao [18]. This procedure is used here to provide an alternative confidence interval for the variability estimator. Setting aside the expectation $E(z)$, the ratio $\lambda=\left(\dfrac{\alpha_z}{\alpha_0}\right)$ is the parameter of interest and we assume again that the asymptotic covariance matrix of the vector $\begin{pmatrix}\hat{\alpha}_z\\\hat{\alpha}_0\end{pmatrix}$ is $\begin{pmatrix}\hat{\sigma}_z^2 & \hat{\sigma}_{zm}\\\hat{\sigma}_{zm} & \hat{\sigma}_0^2\end{pmatrix}$. The equality $\lambda=\left(\dfrac{\alpha_z}{\alpha_0}\right)$ implies

$\alpha_z-\lambda\alpha_0=0$ and, if λ is the "true" ratio, the least-squares estimator of $\alpha_z-\lambda\alpha_0$ is $\hat{\alpha}_z-\lambda\hat{\alpha}_0$, with $E(\hat{\alpha}_z-\lambda\hat{\alpha}_0)=0$, and its asymptotic variance is $\hat{\sigma}_z^2+\lambda^2\hat{\sigma}_0^2-2\lambda\hat{\sigma}_{zm}$. Letting t denote the Student's t-statistic for the $(1-\alpha)100\%$ confidence level, the procedure consists of determining the set of values of λ for which the inequality $\dfrac{(\hat{\alpha}_z-\lambda\hat{\alpha}_0)^2}{\hat{\sigma}_z^2+\lambda^2\hat{\sigma}_0^2-2\lambda\hat{\sigma}_{zm}}\leq t^2$ holds with a probability equal to $(1-\alpha)$:

$$P\left(\frac{(\hat{\alpha}_z-\lambda\hat{\alpha}_0)^2}{\hat{\sigma}_z^2+\lambda^2\hat{\sigma}_0^2-2\lambda\hat{\sigma}_{zm}}\leq t^2\right)$$
$$=P\left[(\hat{\alpha}_z-\lambda\hat{\alpha}_0)^2-t^2\left(\hat{\sigma}_z^2+\lambda^2\hat{\sigma}_0^2-2\lambda\hat{\sigma}_{zm}\right)\leq 0\right] \qquad (13)$$
$$=1-\alpha$$

In other words, the confidence set is the set of values of λ over which the null

hypothesis $H_0 : \alpha_z - \lambda\alpha_0 = 0$ is not rejected against the alternative $H_0 : \alpha_z - \lambda\alpha_0 \neq 0$ in an *F*-test. It is worth recalling here that the square of the Student's *t* is equal (in distribution) to the Fisher's *F*. The expression on the left-hand side of the last inequality (in the probability) can be written as a quadratic form in λ:

$$Q(\lambda) = \left(\hat{\alpha}_z - \lambda\hat{\alpha}_0\right)^2 - t^2\left(\hat{\sigma}_z^2 + \lambda^2\hat{\sigma}_0^2 - 2\lambda\hat{\sigma}_{zm}\right) = \mu\lambda^2 + \rho\lambda + \delta, \qquad (14)$$

with $\mu = \hat{\alpha}_0^2 - t^2\hat{\sigma}_0^2$, $\rho = 2\left(\hat{\sigma}_{zm}t^2 - \hat{\alpha}_z\hat{\alpha}_0\right)$, $\delta = \hat{\alpha}_z^2 - t^2\hat{\sigma}_z^2$.

Provided that the quadratic form $\mu\lambda^2 + \rho\lambda + \delta$ in λ has two distinct real-valued roots, these roots represent the two bounds of the ratio parameter $\dfrac{\alpha_z}{\alpha_0}$. It is worth noting here that if the non-centered regression is used instead, a similar equation in matrix form determines the bounds of the Fieller confidence interval [19]. The conditions determining the nature and the signs of the roots of $Q(\lambda)$ are discussed in Scheffé [20]. In fact, from

$$\Delta = \left(\frac{\rho}{2}\right)^2 - \mu\delta = \left(\hat{\sigma}_{zm}t^2 - \hat{\alpha}_z\hat{\alpha}_0\right)^2 - \left(\hat{\alpha}_0^2 - t^2\hat{\sigma}_0^2\right)\left(\hat{\alpha}_z^2 - t^2\hat{\sigma}_z^2\right),$$

it is clear that if $\Delta > 0$, then $Q(\lambda)$ has two distinct real-valued roots. Any value of λ falling between these two roots is such that $Q(\lambda)$ is of the opposite sign of μ. Hence, if $\Delta > 0$ and $\mu > 0$, then $Q(\lambda) < 0$ and one obtains a bounded confidence interval for $\gamma\left(E(x)\right)$. However, if $\Delta > 0$ and $\mu < 0$, the confidence interval will be the (unbounded) complement set of a bounded interval. If $\Delta < 0$, then $Q(\lambda)$ is of the sign of μ. So, if, in addition to $\Delta < 0$, one has $\mu < 0$, then the confidence interval is the entire real line. The case ($\Delta < 0$, $\mu > 0$) is mathematically impossible, as shown by Scheffé [20], who also notes that the cases $\Delta = 0$ and $\mu = 0$ are zero-probability (negligible) events.

The event $A = \left\{\mu = \hat{\alpha}_0^2 - t^2\hat{\sigma}_0^2 < 0\right\}$ determines, therefore, the boundedness of the Fieller interval. This event has the meaning that the Student test for α_0 leads to no rejection of the null $H_0 : \alpha_0 = 0$. Under this null assumption, the event *A* has the probability $P(A) = 1 - \alpha$. For example, if the null $H_0 : \alpha_0 = 0$ is not rejected at 5% significance level in a *t*-test, then the 95%-level confidence interval will be unbounded. This is in sharp contrast to the always bounded confidence interval resulting from the delta method, which assumes away $H_0 : \alpha_0 = 0$. By doing so, it imposes a sure rejection ($P(A) = 0$) of H_0 in the described Student test, whence, $\alpha = 1$. In other words, it imposes a zero coverage probability for the confidence interval. By doing so, it imposes a sure rejection ($P(A) = 0$) of H_0 in the described Student test, whence, $\alpha = 1$. In other words, it imposes a zero coverage probability for the confidence interval.

In the context of the Postal Service, if for example street time is the dependent and volume the cost-driver of interest, the possibility that the best-prediction of street time given average mail-volume—the meaning of the parameter α_0—be equal to zero is hard to imagine. However, unless $\alpha_0 \neq 0$ is imposed through some additional modeling as an estimation constraint, there is no guaranty that

the corresponding estimate will be statistically significant. It seems reasonable, therefore, to always report both confidence intervals, delta and Fieller, regardless of the researcher's prior belief about the possibility that α_0 be null.

5. Results and Discussion

5.1. Empirical Examples of Delta and Fieller Intervals

The above theory is applied to a data set that is publicly available at https://www.prc.gov/dockets/document/109489. In the example, the dependent variable is "street time" and labelled as "thours". Table 1 displays the table of the estimation results, along with all the information needed to compute the confidence intervals for the elasticity estimate. The variable with respect to which the elasticity is computed is "volume". Its mean-centered version is labelled as "CVol" (centered volume) in the top larger table containing the regression result. Its coefficient and the constant term are labelled in the bottom small table respectively as "Alpha_Vol" and "Alpha_Zero". The small table also contains the sample means of the volume variable and the covariance matrix of the pair (Alpha_Vol, Alpha_Zero). Recall here that the constant term is equal to $g_{\hat{\theta}}(\bar{x})$ and is the denominator of the elasticity. The numerator of the elasticity is the product of Alpha_Vol by the sample average of volume. The volume-elasticity is estimated as 0.62064233. Using the described information, the delta-method 95% confidence interval for the volume elasticity, denoted below by *DCI*, is

$$DCI = \left[\hat{\gamma}(E(x)) - 1.96\sqrt{\hat{w}\left(\frac{\bar{z}}{\hat{\alpha}_0}\right)^2}\;;\hat{\gamma}(E(x)) + 1.96\sqrt{\hat{w}\left(\frac{\bar{z}}{\hat{\alpha}_0}\right)^2}\right]$$
$$= \left[0.592578355; 0.648706305\right].$$

The Fieller 95%-confidence interval, denoted below by *FCI*, is

$$FCI = \left[0.592581517; 0.648733034\right].$$

Both confidence intervals are computed after estimating the centered regression estimation and collecting the estimates $\hat{\alpha}_z$ (Alpha_Vol) and $\hat{\alpha}_0$ (Alpha_Zero), along with the corresponding estimate of their variance-covariance matrix. The volume-elasticity is computed as $\hat{\gamma}(E(x)) = \dfrac{\hat{\alpha}_z}{\hat{\alpha}_0}\bar{z}$, where $\bar{z}$ is the sample mean of the volume variable. The delta-method interval, *DCI*, is obtained from (12). To compute the Filler confidence interval for the ratio $\dfrac{\alpha_z}{\alpha_0}$, the expressions μ, ρ, and δ are first computed as they appear in (14). The discriminant is computes as $\Delta = \left(\dfrac{\rho}{2}\right)^2 - \mu\delta$ and, if both Δ and μ are positive, the left bound of the interval (respectively, the right bound of the interval) is computed as $L = \dfrac{1}{\mu}\left(-\dfrac{\rho}{2} - \sqrt{\Delta}\right)$ (Respectively, $L = \dfrac{1}{\mu}\left(-\dfrac{\rho}{2} + \sqrt{\Delta}\right)$). The final interval, *FCI*, is obtained for $\gamma(E(x))$ by simply multiplying the above bounds by $\bar{z}$.

Table 1. Regression results with centered explanatory variables.

	Number of Obs: 3146					
	F(29, 3116) = 250.53					
	Prob > F = 0.0000					
	R-squared = 0.8710					
	Root MSE = 22.194					
thours = dependent	Coefficients	Robust Std. Err.	t	P > t	[95% Conf.]	
Covariates						
Cvol	0.0813301	0.0022094	36.810	0.000	0.0769981	0.0856622
vol2	−0.0000053	0.0000006	−8.530	0.000	−0.0000065	−0.0000041
volcurb	0.0270784	0.0204113	1.330	0.185	−0.0129426	0.0670994
volcbu	−0.0725671	0.0359772	−2.020	0.044	−0.1431086	−0.0020257
volcent	0.0072648	0.0197153	0.370	0.713	−0.0313915	0.0459212
voldoor	−0.0124164	0.0152493	−0.810	0.416	−0.0423161	0.0174833
volbr	−0.0003747	0.0206299	−0.020	0.986	−0.0408243	0.0400748
voliam	−0.0462756	0.0134243	−3.450	0.001	−0.0725971	−0.0199542
volfdp	−0.0550723	0.0151079	−3.650	0.000	−0.0846947	−0.0254498
volifr	0.0306551	0.0168253	1.820	0.069	−0.0023346	0.0636449
volcv	−0.0000378	0.0000171	−2.220	0.027	−0.0000713	−0.0000044
Ccv	0.0095768	0.0086620	1.110	0.269	−0.0074070	0.0265606
cv2	0.0000176	0.0000030	5.930	0.000	0.0000118	0.0000235
Ccurb	1.1207270	7.2289530	0.160	0.877	−13.0532700	15.2947200
Ccbu	−19.9691900	11.1230400	−1.800	0.073	−41.7784300	1.8400450
Ccent	10.7491700	6.5281810	1.650	0.100	−2.0508020	23.5491400
Cdoor	−2.8270550	5.0978520	−0.550	0.579	−12.8225400	7.1684330
Cbr	−2.9026710	6.1822770	−0.470	0.639	−15.0244200	9.2190780
Ciam	−28.2502300	3.7681750	−7.500	0.000	−35.6385900	−20.8618800
Cfdp	−35.3579100	4.0966100	−8.630	0.000	−43.3902400	−27.3255800
Cifr	1.6938050	5.0887680	0.330	0.739	−8.2838720	11.6714800
curb2	3.8599290	5.6633330	0.680	0.496	−7.2443130	14.9641700
cbu2	−7.7040280	6.6341400	−1.160	0.246	−20.7117500	5.3037000
cent2	−43.3282700	9.1974540	−4.710	0.000	−61.3619500	−25.2945800
door2	−8.2948350	5.2232330	−1.590	0.112	−18.5361600	1.9464920
br2	−18.9822600	8.5685330	−2.220	0.027	−35.7828000	−2.1817160
iam2	−2.8435370	5.5828790	−0.510	0.611	−13.7900300	8.1029570
fdp2	1.0518450	5.9419160	0.180	0.860	−10.5986200	12.7023100
ifr2	−16.0448000	6.9424530	−2.310	0.021	−29.6570500	−2.4325600
_cons	49.3815800	0.7297492	67.670	0.000	47.9507400	50.8124200

	Estimates	Covariances	
		Alpha_Vol	Alpha_Zero
Alpha_Vol	0.0813301	4.88E−06	0.00085156
Alpha_Zero	49.38158	0.00085156	0.53253391
Sample Mean		Volune-Elasticitiy of thours	
376.8382		0.62064233	

As discussed above, the main difference between the two confidence intervals is that *FCI* accounts for the possibility that α_0 be null, while *DCI* does not. In this example, the two intervals are seen to be very close to one another. The difference between the radius of the Fieller and the delta-method intervals is in the order of 10^{-5}, the delta-method confidence interval being very slightly narrower. In particular, the Fieller confidence interval is bounded, because $\Delta > 0$ and $\mu > 0$.

5.2. The Translog Regression Function

Because the translog functional form is quadratic in the logs of the components of *x*, the above analysis carries over to this model, but in a rather trivial way. The model considered becomes

$$\ln(T) = g_\theta(\ln x_1, \cdots, \ln x_d) + u \tag{15}$$

with g_θ defined, as above, to be quadratic. To add to the notation used so far, the elasticity of a function *f* of the vector $(z, x_2, \cdots, x_d)$, with respect to the argument *z*, computed at $(z, x_2, \cdots, x_d) = (a_1, a_2, \cdots, a_d)$, will be denoted by $\mathcal{E}_f(z)\big|_{(z, x_2, \cdots, x_d) = (a_1, a_2, \cdots, a_d)}$. The sample geometric mean of *z* will be denoted by $\bar{z}^g$ or, simply, $\mathcal{E}_f(z)\big|_{(a_1, a_2, \cdots, a_d)}$. Assuming that the model (15) is estimated with centered explanatory variables and writing $\widehat{\ln(T)} = g_{\hat{\theta}}(\ln x_1, \cdots, \ln x_d)$, one has

$$\overline{\ln(z)} = \ln(\bar{z}^g) \tag{16}$$

$$\mathcal{E}_{e^{\widehat{\ln(T)}}}(z)\bigg|_{\left(\overline{\ln(z)}, \overline{\ln(x_2)}, \cdots, \overline{\ln(x_d)}\right)} = \frac{\partial g_{\hat{\theta}}(\ln x_1, \cdots, \ln x_d)}{\partial \ln(z)}\Bigg|_{\left(\overline{\ln(z)}, \overline{\ln(x_2)}, \cdots, \overline{\ln(x_d)}\right)} \tag{17}$$

$$= \frac{\partial g_{\hat{\theta}}(\ln x_1, \cdots, \ln x_d)}{\partial \ln(z)}\Bigg|_{\left(\ln(\bar{z}^g), \ln(\overline{x_2}^g), \cdots, \ln(\overline{x_d}^g)\right)} \tag{18}$$

The right-hand side of (18) is the coefficient of $\ln(z) - \overline{\ln(z)}$ in the centered regression result:

$$\hat{\alpha}_{\ln z} = \frac{\partial g_{\hat{\theta}}(\ln x_1, \cdots, \ln x_d)}{\partial \ln(z)}\Bigg|_{\left(\ln(\bar{z}^g), \ln(\overline{x_2}^g), \cdots, \ln(\overline{x_d}^g)\right)}$$

It is the elasticity of the predicted street time, $e^{\widehat{\ln(T)}}$, with respect to *z*, computed at the vector of log geometric means. So, the inference on this elasticity is provided directly by the statistical software package.

6. Conclusion

The present paper has provided two alternative ways to compute confidence intervals for cost-elasticity (variability) when the underlying functional form in the regression equation is quadratic. The polynomial aspect of the assumed regression function has considerably simplified the computations. A similar analysis can be conducted on any functional form that is linear in the parameters, but will be a bit more complex if the function is not polynomial in its arguments. Both the delta-method and the Fieller confidence intervals have been derived and illustrated with an empirical example based on publicly available Postal Service data. For the sake of completeness and transparency, both confidence intervals should be reported in applications.

Acknowledgements

The authors would like to thank Irina Murtazashvili for her valuable comments. All errors are the authors'.

Conflicts of Interest

The views expressed in this paper are solely those of the authors and do not reflect the views of the Postal Regulatory Commission.

References

[1] United States Postal Service (2014) Report on the City Carrier Street Time Study. Postal Regulatory Commission, Washington DC, 25-26.
https://www.prc.gov/docs/90/90869/prop.13.city%20carrier.report.pdf

[2] Bradley, D.M. (2016) Research on Estimating the Variability of Purchased Highway Transportation Capacity with Respect to Volume. Postal Regulatory Commission, Washington DC, 2
https://www.prc.gov/docs/96/96940/Research.Report.Proposal.Four.pdf

[3] Bradley, D.M. (2019) A New Study of Special Purpose Route Carrier Costs. Postal Regulatory Commission, Washington DC.
https://www.prc.gov/docs/109/109484/spr.public.study.report.pdf

[4] United States Postal Service (2012) Postal Service Report Regarding Cost Studies: Response to PRC Order No. 1626 (April 18, 2012). Postal Regulatory Commission, Washington DC, 6.
https://www.prc.gov/docs/86/86858/Report_Response_Order_1626.pdf

[5] Cazals, C., Duchemin, P., Florens, J-P., Roy, B. and Vialaneix, O. (2002) An Econometric Study of Cost Elasticity in the Activities of Post Office Counters. In: Crew, M.A. and Kleindorfer, P.R., Eds., *Postal and Delivery Services*, Kluwer Academic Publishers, Boston, 161-170. https://doi.org/10.1007/978-1-4613-0253-7_9

[6] Echambadi, R. and Hess, J.D. (2007) Mean-Centering Does Not Alleviate Collinearity Problems in Moderated Multiple Regression Models. *Marketing Science*, **26**, 438-445. https://doi.org/10.1287/mksc.1060.0263

[7] Lehmann, E.L. and Casella, G. (1998) Theory of Point Estimation. 2nd Edition, Springer, New York, 150. https://doi.org/10.1007/b98854

[8] Miller, S.E., Capps Jr., O. and Wells, G.J. (1984) Confidence Intervals for Elasticities and Flexibilities from Linear Equations. *American Journal of Agricultural Econom-*

ics, **66**, 392-396. https://doi.org/10.2307/1240807

[9] Anderson, R.G. and Thursby, J.G. (1986) Confidence Interval for Elasticity Estimators in Translog Models. *The Review of Economics and Statistics,* **68**, 647-656.

[10] Hirschberg, J.G., Lye, N. and Slottje, D.J. (2008) Inferential Methods for Elasticity Estimates. *Journal of Econometrics,* **147**, 299-315.

[11] Christensen, L.R., Jorgenson, D.W. and Lau, L.J. (1973) Transcendental Logarithmic Production Frontiers. *The Review of Economics and Statistics*, **55**, 28-45. https://doi.org/10.2307/1927992

[12] Borovkov, A.A. (1998) Mathematical Statistics. First Edition, Theorem 1, Gordon and Breach Science Publishers, New York, 13.

[13] White, H. (1980) A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, **48**, 817-838. https://doi.org/10.2307/1912934

[14] Gleser, L.J. and Hwang, J.T. (1987) The Nonexistence of $100(1-a)$% Confidence Sets of Finite Expected Diameter-in-Errors in Variables and Related Models. *The Annals of Statistics*, **15**, 1351-1362. https://doi.org/10.1214/aos/1176350597

[15] Dufour, J.-M. (1997) Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models. *Econometrica*, **65**, 1365-1387. https://doi.org/10.2307/2171740

[16] Koschat, M.A. (1987) A Characterization of the Fieller Solution. *The Annals of Statistics*, **15**, 462-468. https://doi.org/10.1214/aos/1176350282

[17] Fieller, E.C. (1940-1941) The Biological Standardization of Insulin. *Supplement to the Journal of the Royal Statistical Society*, **7**, 1-64. https://doi.org/10.2307/2983630

[18] Radhakrishna, R.C. (1973) Linear Statistical Inference and Its Applications. 2nd Edition, Wiley, Hoboken, 241.

[19] Hirschberg, J.G.J., Lye, N. and Slottje, D.J. (2008) Inferential Methods for Elasticity Estimates. *Journal of Econometrics*, **147**, 299-315. https://doi.org/10.1016/j.jeconom.2008.09.037

[20] Scheffé, H. (1970) Multiple Testing versus Multiple Estimation. Improper Confidence Sets. Estimation of Directions and Ratios. *The Annals of Mathematical Statistics*, **41**, 1-29. https://doi.org/10.1214/aoms/1177697184