Linear Approximations of Nonlinear Relationships by the Taylor's Series Expansion Revisited

By Abner W. Womack and Jim L. Matthews

This paper examines the magnitude of error associated with linear approximations of nonlinear variables based on Taylor's Series. Little attention has been given to the error term in previous empirical studies. This paper presents the mathematical technique for the single-variable and two-variable cases. Examples are given for each situation using agricultural time-series data. Characteristics of time-series data are sometimes crucial in the selection of an evaluation point for minimum error. The importance of selecting evaluation points is illustrated for three categories of time-series data: (1) smooth trends, (2) trends with substantial variation, and (3) oscillatory series.

Key words: Mathematical analysis; nonlinear; methodology; time-series analysis; statistics; research methodology.

Nonlinear functions are commonly used in econometric analyses because of either theoretical or statistical considerations in choosing the form of the equation. As a rule, this poses no serious difficulty to the analyst. However, nonlinear specification of relationships is difficult to manage, particularly where those relations appear in a sub-block of simultaneous equations that contain linear market-clearing identities. What is frequently desired in any simultaneous equilibrium model specification is to reexpress the structural equations in terms of their reduced form equivalents. This is often difficult to do for a nonlinear system.

To achieve a solution without undue mathematical complexity, one approach suggested by Klein is to obtain linear approximations of all nonlinear endogenous variables in the system by a Taylor's Series expansion (*8*).[1] This technique was used by Gerra for a poultry model and more recently by Houck and Subotnik in a simultaneous model for the U.S. soybean economy (*5*, *7*). Several studies have employed Taylor's Series for purposes other than a linearization tool. Edwards demonstrated that linear estimation schemes could be applied to nonlinear equations iteratively in deriving B.L.U.E. estimators as well as in solving nonlinear programming problems (*3*). Burt, in 1968, applied the Taylor's Series expansion to a nonlinear identity equation ($y=x \cdot z$) to illustrate the component variances of a variable y associated with two separate random variables that appear as a product, namely, $x \cdot z$ (*2*).

[1] Italic numbers in parentheses refer to Bibliography, p. 101.

Though the Taylor's Series is well documented in most calculus texts or texts on mathematics for economists, relatively little attention has been given to its accuracy as a linearization technique for several commonly used nonlinear variables in econometric analyses (*1*, *9*).

This paper examines the accuracy and use of Taylor's Series expansions for several types of nonlinear variables commonly used in econometric analyses. Refinements in the technique are considered for several combinations of characteristics of data series for the variables linearized.

The first section gives Taylor's Theorem with a discussion of the remainder term for the linear case. The second section of the paper shows a linear approximation relation for the single-variable case, namely $\log_{10} x$. Refinements in the use of the relation are shown for two sets of sample observations taken from actual agricultural time-series data. The third section of the paper shows the linear approximations for products and ratios of variables. As in the preceding section, agricultural time-series data with markedly different sample data properties with respect to variance and trend are used to demonstrate the linear accuracy of the technique about different evaluation points, depending on the characteristic of the data series for the variable.

Taylor's Theorem

The polynomial approximation form of Taylor's Series for the single variable case is stated as:

$$(1)\quad f(x) = f(a) + f'(a)(x-a) + f''(a)(x-a)^2/2! + f'''(a)(x-a)^3/3! + \ldots + f^n(a)(x-a)^n/n! + R_n$$

where

$$R_n = f^{n+1}[a + \phi(x-a)]\,[x-a]^{n+1}/(n+1)!$$

for $0 < \phi < 1$, $x \neq a$, n a positive integer, and f a function whose nth derivative $f^{(n)}(x)$ exists for each number between x and a.

From (1) for $n = 1$ we have a linear relationship:

$$(2)\quad R_1 = f(x) - f(a) - f'(a)(x-a).$$

As indicated in (2), $R_1 \rightarrow 0$ as $x \rightarrow a$. Hence the evaluation point (a) for a particular linear approximation should be chosen such that R_1 is small. If R_1 is small, the nonlinear function $f(x)$ is approximated by the linear function $f(a) + f'(a)(x-a)$. As is often the case, the mean of a series is chosen as the point for evaluation. The next section demonstrates that the selection of the evaluation point (a) depends on the nature of the series in question and may call for a point different from the mean for a minimum R_1.

The Single-Variable Case

Though several types of nonlinear expressions for a single variable x are commonly used, most can be readily transformed to $\log_{10}x$. For this reason this section is limited to a discussion of linear approximations of $\log_{10}x$.

For the function $f(x) = \log_{10}x$, formula (1) can be used to derive a linear estimate of $\log_{10}x$ for any of the n observations for x evaluated about some selected point (a).

The linear approximation equation is expressed in general as follows:

$$(3)\quad \text{Log}_{10}x = (\log_{10}a - 0.4343) + 0.4343x/a.$$

Examination of (3) shows that when $x = a$, $\log_{10}x$ is exactly equal to $\log_{10}a$. Linear approximations of x evaluated at or near (a) will result in a good linear approximation of $\log_{10}x$.

In applying formula (3) to a given data series, selection of an evaluation point (a) is quite important if a high degree of accuracy is to be achieved. For convenience, evaluation of x about $a = \bar{x}$ may be considered. This choice of an evaluation point may not be a bad one if the data series is relatively smooth with no significant trends. However, irregular data series with or without trends may require evaluation about some point other than $a = \bar{x}$. An alternative is to use a series of choices such as the previous value of x or some moving average of recent observations on x. These series are not as convenient to use as a single point because of the required iteration routines but may be warranted if a high degree of linear accuracy in estimating x is desired.

To demonstrate some of the options open to the analyst, figures 1 to 3 show linear approximations of $\log_{10}x$ for three alternative evaluation points where x is soybeans under loan (million bushels) for the period 1954 to 1968. The series is irregular with only a slight upward trend. As shown in figure 1, evaluation about the mean ($a = \bar{x}$) is reasonably good. Largest inaccuracy occurs at the end points of the series or points furthest from the mean. An attempt to correct for this inaccuracy is shown in figures 2 and 3. In figure 3, use of the previous period value reduced the inaccuracy at the end points but magnified the errors when sharp year-to-year variations occurred. In figure 2, use of a 2-year moving average improved the fit when compared to figure 3 but was less desirable than figure 1. Longer moving averages would improve the fit in figure 2.

Based on this example, two inferences which can be drawn are: (1) for irregular series with no trend, evaluation about $a = \bar{x}$ would be the best choice, and (2) for irregular series with trend, a moving average of previous values of x should be used.

Another situation quite common in economic analysis is shown in figure 4. The series for $\log_{10}x$ is smooth with a definite trend. Evaluation about a fixed point (a = value of x in 1954) shows that substantial bias in the linear approximation of per capita disposable income in 1968 would have occurred. Use of the previous period observation on x in this situation results in a very close linear approximation. Thus, it is recommended that $a = x_{-1}$ be used for data series that are smooth with marked trends for obtaining linear approximation of x for the function $\log_{10}x$.

Products and Ratios of Variables

Linear approximations based on Taylor's Theorem for a single variable can be extended to two or more variable cases without any serious conceptual problems. Two or more variable cases commonly encountered by analysts are ratios and products of variables. This section is restricted to the two-variable case.

$$(1)\quad f(x) = f(a) + f'(a)(x-a) + f''(a)(x-a)^2/2! + f'''(a)(x-a)^3/3! + \ldots + f^n(a)(x-a)^n/n! + R_n$$

where

$$R_n = f^{n+1}[a + \phi(x-a)][x-a]^{n+1}/(n+1)!$$

for $0 < \phi < 1$, $x \neq a$, n a positive integer, and f a function whose nth derivative $f^{(n)}(x)$ exists for each number between x and a.

From (1) for $n = 1$ we have a linear relationship:

$$(2)\quad R_1 = f(x) - f(a) - f'(a)(x-a).$$

As indicated in (2), $R_1 \to 0$ as $x \to a$. Hence the evaluation point (a) for a particular linear approximation should be chosen such that R_1 is small. If R_1 is small, the nonlinear function $f(x)$ is approximated by the linear function $f(a) + f'(a)(x-a)$. As is often the case, the mean of a series is chosen as the point for evaluation. The next section demonstrates that the selection of the evaluation point (a) depends on the nature of the series in question and may call for a point different from the mean for a minimum R_1.

The Single-Variable Case

Though several types of nonlinear expressions for a single variable x are commonly used, most can be readily transformed to $\log_{10}x$. For this reason this section is limited to a discussion of linear approximations of $\log_{10}x$.

For the function $f(x) = \log_{10}x$, formula (1) can be used to derive a linear estimate of $\log_{10}x$ for any of the n observations for x evaluated about some selected point (a).

The linear approximation equation is expressed in general as follows:

$$(3)\quad \mathrm{Log}_{10}x = (\log_{10}a - 0.4343) + 0.4343x/a.$$

Examination of (3) shows that when $x = a$, $\log_{10}x$ is exactly equal to $\log_{10}a$. Linear approximations of x evaluated at or near (a) will result in a good linear approximation of $\log_{10}x$.

In applying formula (3) to a given data series, selection of an evaluation point (a) is quite important if a high degree of accuracy is to be achieved. For convenience, evaluation of x about $a = \bar{x}$ may be considered. This choice of an evaluation point may not be a bad one if the data series is relatively smooth with no significant trends. However, irregular data series with or without trends may require evaluation about some point other than $a = \bar{x}$. An alternative is to use a series of choices such as the previous value of x or some moving average of recent observations on x. These series are not as convenient to use as a single point because of the required iteration routines but may be warranted if a high degree of linear accuracy in estimating x is desired.

To demonstrate some of the options open to the analyst, figures 1 to 3 show linear approximations of $\log_{10}x$ for three alternative evaluation points where x is soybeans under loan (million bushels) for the period 1954 to 1968. The series is irregular with only a slight upward trend. As shown in figure 1, evaluation about the mean ($a = \bar{x}$) is reasonably good. Largest inaccuracy occurs at the end points of the series or points furthest from the mean. An attempt to correct for this inaccuracy is shown in figures 2 and 3. In figure 3, use of the previous period value reduced the inaccuracy at the end points but magnified the errors when sharp year-to-year variations occurred. In figure 2, use of a 2-year moving average improved the fit when compared to figure 3 but was less desirable than figure 1. Longer moving averages would improve the fit in figure 2.

Based on this example, two inferences which can be drawn are: (1) for irregular series with no trend, evaluation about $a = \bar{x}$ would be the best choice, and (2) for irregular series with trend, a moving average of previous values of x should be used.

Another situation quite common in economic analysis is shown in figure 4. The series for $\log_{10}x$ is smooth with a definite trend. Evaluation about a fixed point (a = value of x in 1954) shows that substantial bias in the linear approximation of per capita disposable income in 1968 would have occurred. Use of the previous period observation on x in this situation results in a very close linear approximation. Thus, it is recommended that $a = x_{-1}$ be used for data series that are smooth with marked trends for obtaining linear approximation of x for the function $\log_{10}x$.

Products and Ratios of Variables

Linear approximations based on Taylor's Theorem for a single variable can be extended to two or more variable cases without any serious conceptual problems. Two or more variable cases commonly encountered by analysts are ratios and products of variables. This section is restricted to the two-variable case.

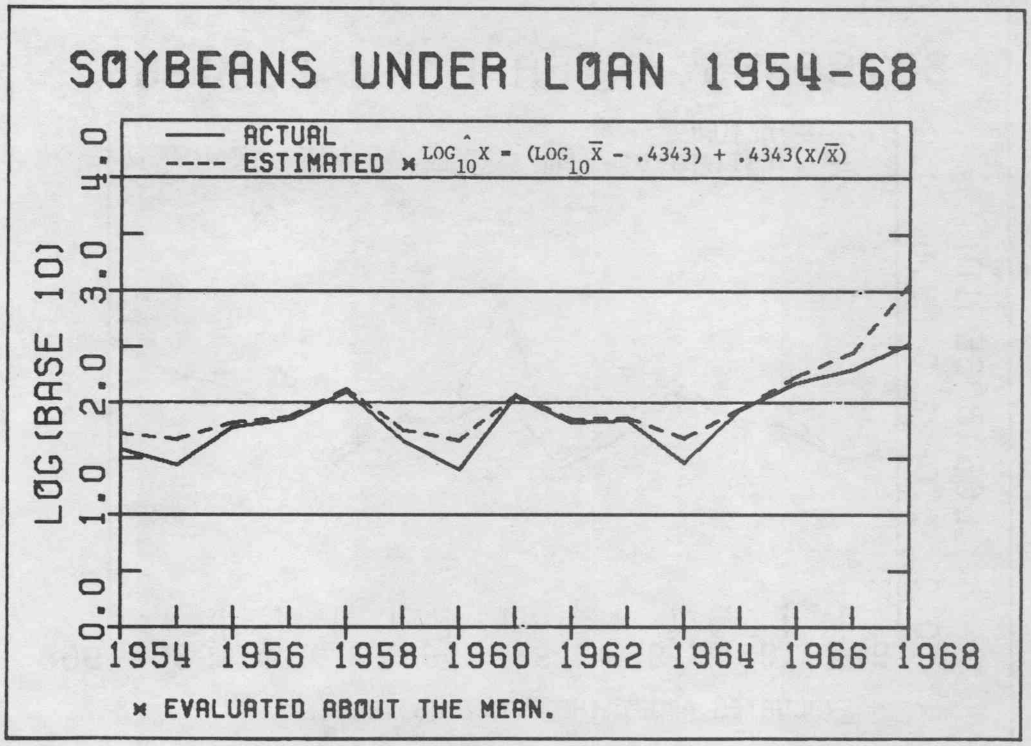

Figure 1

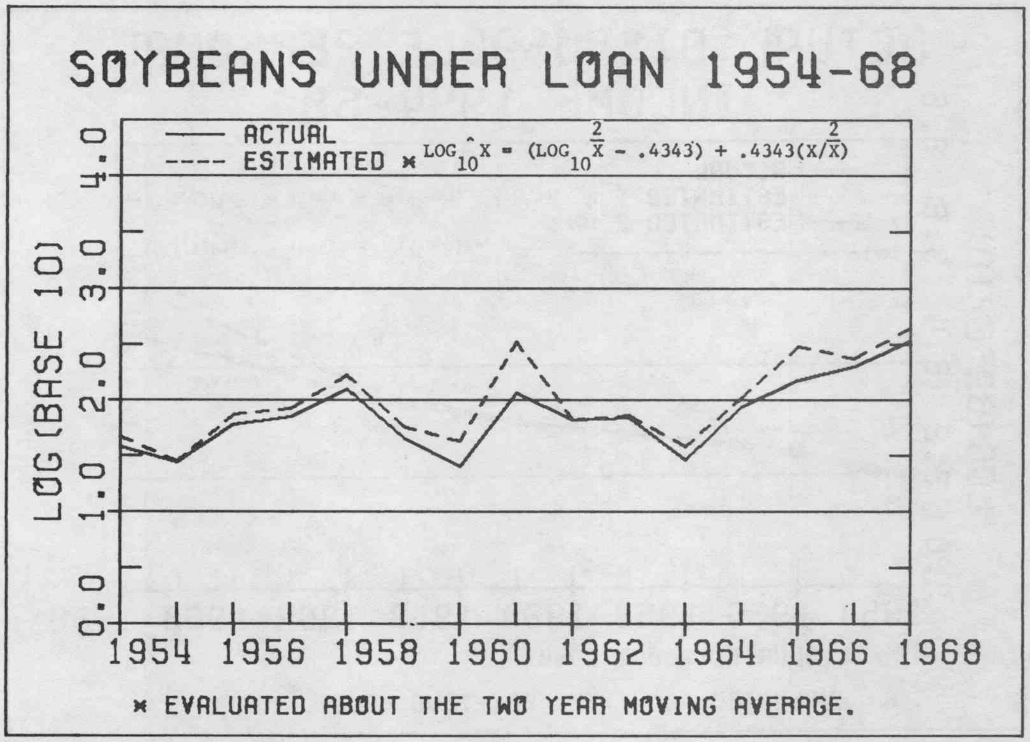

Figure 2

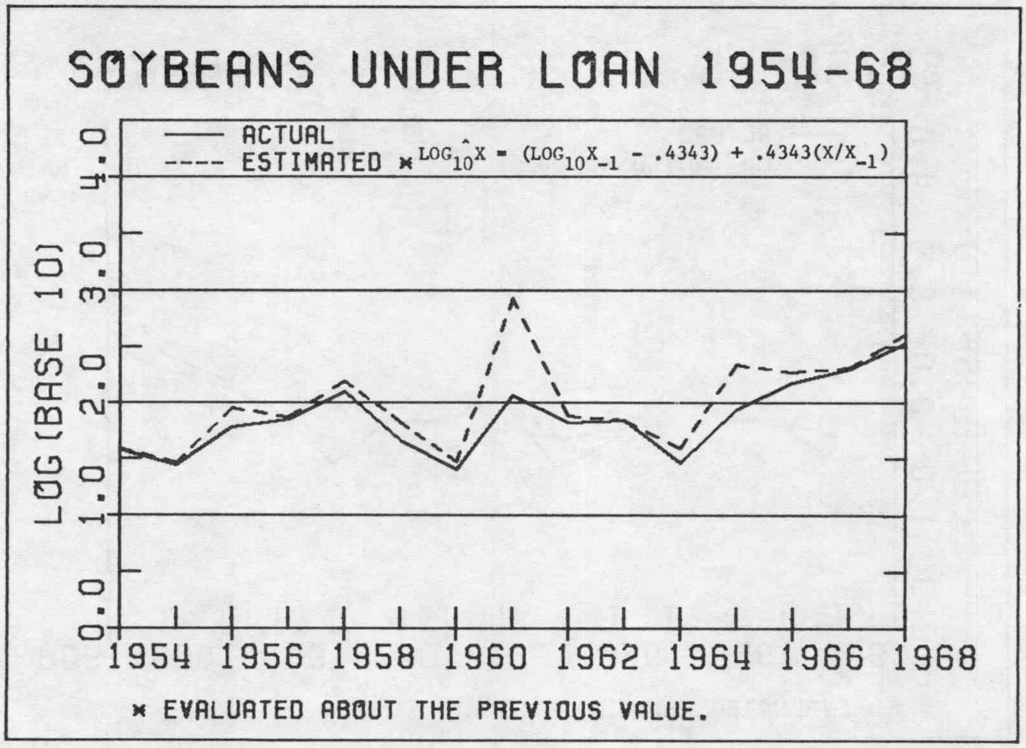

Figure 3

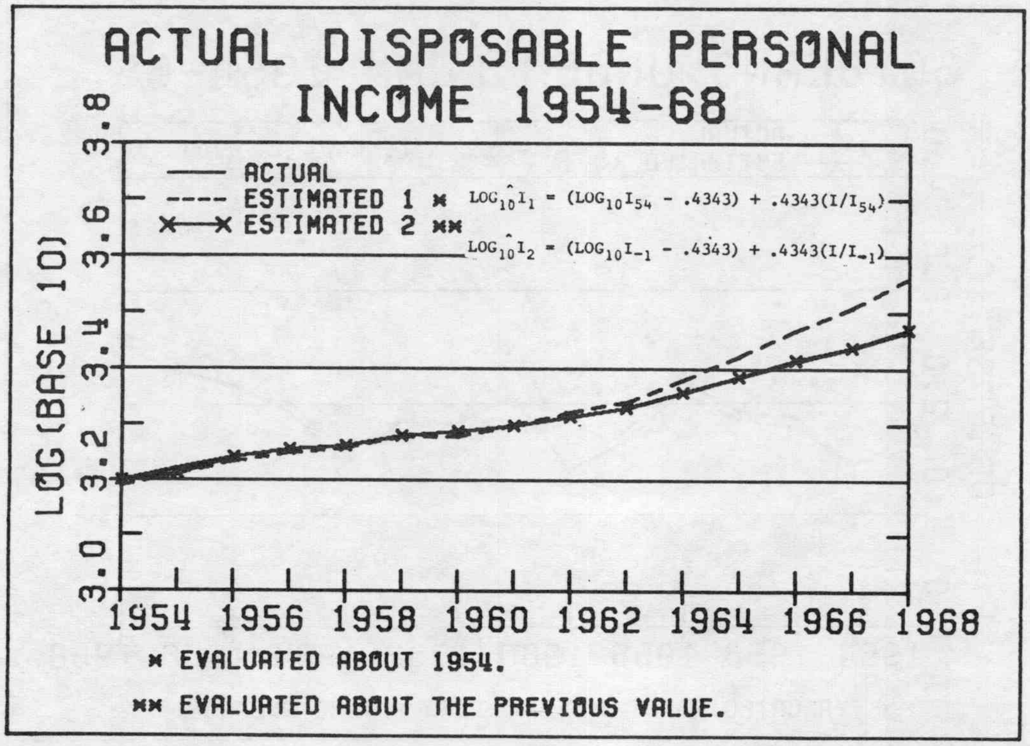

Figure 4

For the Taylor's Series expansion of a function F in two variables having continuous partial derivatives of the [illegible] order, a relation comparable to (1) can be expressed as follows:

$$(4)\quad F(a+dx, b+dy) = F(a, b) + dF(a, b) + d^2F(a, b)/2! + \ldots + d^nF(a, b)/n! + R_n$$

where

$$R_n = d^{n+1}F(C, D)/(n+1)!$$

and C is between a and $a + dx$ and D is between b and $b + dy$. dx and dy are any designated number used for differentials of the first and second variables. In this case, $dx = (x-a)$ and $dy = (y-b)$. From differential calculus, it can be shown that

$$(5)\quad dF(x,y) = F'(x,y) = \frac{\partial F(x,y)}{\partial x}\, dx + \frac{\partial F(x,y)}{\partial x}\, dy$$

For convenience of notation, the above differentials in (4) may be written as:

$$(6)\quad dF(x,y) = F'_x dx + F'_y dy$$

$$(7)\quad d^2F(x,y) = F''(x,y) = F'[F'(x,y)]$$

$$= d\left[\frac{\partial F}{\partial x}\, dx + \frac{\partial F}{\partial y}\, dy\right]$$

where $F = F(x,y)$

$$= \partial/\partial x\left[\frac{\partial F}{\partial x}\, dx + \frac{\partial F}{\partial y}\, dy\right]$$

$$+ \partial/\partial y\left[\frac{\partial F}{\partial x}\, dx + \frac{\partial F}{\partial y}\, dy\right] dy$$

$$= \frac{\partial^2 F}{\partial x^2}\,(dx)^2 + \frac{\partial^2 F}{\partial x \partial y}\,(dx)(dy)$$

$$+ \frac{\partial^2 F}{\partial x \partial y}\,(dy)(dx) + \frac{\partial^2 F}{\partial y^2}\,(dy)^2$$

$$= \frac{\partial^2 F}{\partial x^2}\,(dx)^2 + \frac{2\partial^2 F}{\partial x \partial y}\,(dx)(dy)$$

$$+ \frac{\partial^2 F}{\partial y^2}\,(dy)^2 = F''_{x^2}\,(dx)^2$$

$$+ 2F''_{xy}\,(dx)(dy) + F''_{y^2}\,(dy)^2$$

This pattern suggests that higher order derivatives may be obtained by using the corresponding expansion of the binomial distribution, $(a+b)^n$, when n represents the nth derivative. For the purpose of linearization, higher order derivatives are not required and therefore are not given.

Based on relations (4) through (7), linear approximations can be readily derived for ratios and products of variables as follows:

$$(8)\quad F(x,y) = x/y = F(a+dx, b+dy)$$

where

$$dx = (x-a) \text{ and } dy = (y-b)$$

$$F'(x,y) = (1/y)dx - (x/y^2)dy$$

$$F''(x,y) = (0)(dx)^2 + 2(-1/y^2)(dx)(dy) + (2x/y^3)(dy)^2$$

.

.

.

$$F^n(x,y) = (-1)^{n-1}(n!/y^n)(dx)(dy)^{n-1} + (-1)^n(n!\cdot x/y^{n+1})(dy)^n$$

Evaluation of the previous terms about a and b results in the following expansion:

$$(9)\quad F(x,y) = a/b + (1/b)(x-b) - (a/b^2)(y-b) + [2(-1/b^2)(x-a)(y-b) + (2a/b^3)(y-b)^2]/2! + \ldots + R_n$$

where

$$R_n = \frac{(dy)^n}{D^{n+1}}\,[(-1)^n(dx) + \frac{(-1)^{n+1}C}{D}\,(dy)]$$

for C between a and $a + dx$ and D between b and $b + dy$.

As in the single-variable case in the previous section, $R_n \to 0$ as $n \to \infty$. Thus the expansion evaluated about a and b can be taken to any desired level of accuracy. Since linearization of $F(x,y)$ is the objective, then only the first two terms of (4) can be used. Thus the linear approximation of x/y is simply the first two terms of (9) simplified as follows:

$$(10)\quad x/y = a/b + (1/b)x - (a/b^2)y + R_n$$

For the Taylor's Series expansion of a function F in two variables having continuous partial derivatives of the nth order, a relation comparable to (1) can be expressed as follows:

$$(4)\quad F(a+dx,\, b+dy) = F(a,b) + dF(a,b) + d^2F(a,b)/2! + \ldots + d^nF(a,b)/n! + R_n$$

where

$$R_n = d^{n+1}F(C,D)/(n+1)!$$

and C is between a and $a + dx$ and D is between b and $b + dy$. dx and dy are any designated number used for differentials of the first and second variables. In this case, $dx = (x-a)$ and $dy = (y-b)$. From differential calculus, it can be shown that

$$(5)\quad dF(x,y) = F'(x,y) = \frac{\partial F(x,y)}{\partial x}\,dx + \frac{\partial F(x,y)}{\partial x}\,dy$$

For convenience of notation, the above differentials in (4) may be written as:

$$(6)\quad dF(x,y) = F'_x dx + F'_y dy$$

$$(7)\quad d^2F(x,y) = F''(x,y) = F'[F'(x,y)]$$

$$= d\left[\frac{\partial F}{\partial x}\,dx + \frac{\partial F}{\partial y}\,dy\right]$$

where $F = F(x,y)$

$$= \partial/\partial x\left[\frac{\partial F}{\partial x}\,dx + \frac{\partial F}{\partial y}\,dy\right]$$

$$+ \partial/\partial y\left[\frac{\partial F}{\partial x}\,dx + \frac{\partial F}{\partial y}\,dy\right]dy$$

$$= \frac{\partial^2 F}{\partial x^2}(dx)^2 + \frac{\partial^2 F}{\partial x \partial y}(dx)(dy)$$

$$+ \frac{\partial^2 F}{\partial x \partial y}(dy)(dx) + \frac{\partial^2 F}{\partial y^2}(dy)^2$$

$$= \frac{\partial^2 F}{\partial x^2}(dx)^2 + \frac{2\partial^2 F}{\partial x \partial y}(dx)(dy)$$

$$+ \frac{\partial^2 F}{\partial y^2}(dy)^2 = F''_{x^2}(dx)^2$$

$$+ 2F''_{xy}(dx)(dy) + F''_{y^2}(dy)^2$$

This pattern suggests that higher order derivatives may be obtained by using the corresponding expansion of the binomial distribution, $(a+b)^n$, when n represents the nth derivative. For the purpose of linearization, higher order derivatives are not required and therefore are not given.

Based on relations (4) through (7), linear approximations can be readily derived for ratios and products of variables as follows:

$$(8)\quad F(x,y) = x/y = F(a+dx,\, b+dy)$$

where

$$dx = (x-a) \text{ and } dy = (y-b)$$

$$F'(x,y) = (1/y)dx - (x/y^2)dy$$

$$F''(x,y) = (0)(dx)^2 + 2(-1/y^2)(dx)(dy) + (2x/y^3)(dy)^2$$

.

.

.

$$F^n(x,y) = (-1)^{n-1}(n!/y^n)(dx)(dy)^{n-1} + (-1)^n(n!\cdot x/y^{n+1})(dy)^n$$

Evaluation of the previous terms about a and b results in the following expansion:

$$(9)\quad F(x,y) = a/b + (1/b)(x-b) - (a/b^2)(y-b) + [2(-1/b^2)(x-a)(y-b) + (2a/b^3)(y-b)^2]/2! + \ldots + R_n$$

where

$$R_n = \frac{(dy)^n}{D^{n+1}}\left[(-1)^n(dx) + \frac{(-1)^{n+1}C}{D}(dy)\right]$$

for C between a and $a + dx$ and D between b and $b + dy$.

As in the single-variable case in the previous section, $R_n \to 0$ as $n \to \infty$. Thus the expansion evaluated about a and b can be taken to any desired level of accuracy. Since linearization of $F(x,y)$ is the objective, then only the first two terms of (4) can be used. Thus the linear approximation of x/y is simply the first two terms of (9) simplified as follows:

$$(10)\quad x/y = a/b + (1/b)x - (a/b^2)y + R_n$$

or

$$x\hat{/}y = a/b + (1/b)x - (a/b^2)y$$

Linear approximations of the product of two variables is quite similar to the ratio of two variables so only the first two terms of the expansion will be given. By definition, $F(x,y) = x \cdot y$. Thus, the first two terms of the expansion are:

$$F(x,y) = x \cdot y$$

$$F'(x,y) = y \cdot dx + x \cdot dy$$

Evaluation of the above derivatives about (a,b) gives:

$$(11) \quad x \cdot y = F(x,y) = F(a+dx, b+dy)$$

where

$$dx = x-a \text{ and } dy = y-b$$

$$x \cdot y = F(a,b) + F'(a,b) + R_n$$

$$x \cdot y = a \cdot b + b \cdot x - a \cdot b + a \cdot y - a \cdot b + R_n$$

$$x \hat{\cdot} y = -a \cdot b + b \cdot x + a \cdot y$$

Use of the linear approximations given in (10) and (11) has been quite good for most applications encountered by the authors. However, these approximations are subject to the same type of biases as encountered for $\log_{10} x$. Note that when $a = x$ and $y = b$ in (10) and (11), both became exact identities. Thus, linear approximations for x and y contain only small error when the evaluation points a and b are near x and y.

For convenience, the means of the series for x and y are often used as fixed evaluation points. For example, Gerra, Houck, and Klein evaluated about the mean in deriving linear approximations for ratios of ratios of two variables. For many situations, this procedure would result in only a small amount of bias. However, as shown in figure 5, linear approximations about the means for per capita consumption of eggs based on an update of Gerra's 1958 egg study led to significant error in recent years. This bias or residual difference occurred because of the large downtrend in per capita egg consumption in the 1960's. Evaluation at some point other than the mean is indicated if a high degree of accuracy is desired. As indicated in figure 6, use of the previous period values ($a = x_{-1}$, $b = y_{-1}$) leads to a substantial improvement for a linear approximation of the ratio of egg consumption to civilian population. In contrast to the smooth trending data series on eggs, linear approximations of farm-to-retail price ratios for oranges were quite good over the sample period 1954 to 1968 when evaluated about the means of the price series. As is clear from figure 7, the price ratio series has irregular movements and no trend. Evaluation about last period values for retail and farm prices of oranges magnified the residual differences when sharp year-to-year changes occurred, as shown in figure 8. Experiments with products of two variables have resulted in similar observations on the appropriate choice of an evaluation point.

Based on the examples discussed for both the single-variable and two-variable cases, evaluation about the mean will give the best linear approximations for series which fluctuate but have little or no trend. For smooth data series which have significant trends, evaluation about the previous period value should be used to minimize residual error in the linear approximations. Finally, fluctuating series with a distinct trend will require a moving average of recent past observations to obtain the best linear approximations. The choice of the number of observations for the moving average will depend on the extent of fluctuation about the trend in the series.

The above statements should be qualified with respect to computational requirements if the mean of a series is not appropriate. In simultaneous equation systems containing some nonlinear variables, use of an evaluation point other than a fixed point substantially increases the computational requirements for solving the system. Use of the previous period value, for example, requires the matrix of endogenous coefficients to be inverted on each successive solution iteration since $b_i = f(x_{i-1}, y_{i-1})$.

To briefly indicate the nature of the computational requirements, the following derived demand equation for California fresh oranges has been used in a simultaneous equation model:

$$(12) \quad QCFO = 5.787 - 46.353\,(PCFO/PRFO) - 0.827\,WND$$

where

$QCFO$ = quantity of California oranges for fresh use

$PCFO$ = on-tree price of California fresh oranges

$PRFO$ = U.S. retail price of fresh oranges

WND = nondurable wage rate in United States or

$$x\hat{/}y = a/b + (1/b)x - (a/b^2)y$$

Linear approximations of the product of two variables is quite similar to the ratio of two variables so only the first two terms of the expansion will be given. By definition, $F(x,y) = x \cdot y$. Thus, the first two terms of the expansion are:

$$F(x,y) = x \cdot y$$

$$F'(x,y) = y \cdot dx + x \cdot dy$$

Evaluation of the above derivatives about (a,b) gives:

$$(11) \quad x \cdot y = F(x,y) = F(a+dx, b+dy)$$

where

$$dx = x-a \text{ and } dy = y-b$$

$$x \cdot y = F(a,b) + F'(a,b) + R_n$$

$$x \cdot y = a \cdot b + b \cdot x - a \cdot b + a \cdot y - a \cdot b + R_n$$

$$x \hat{\cdot} y = -a \cdot b + b \cdot x + a \cdot y$$

Use of the linear approximations given in (10) and (11) has been quite good for most applications encountered by the authors. However, these approximations are subject to the same type of biases as encountered for $\log_{10}x$. Note that when $a = x$ and $y = b$ in (10) and (11), both became exact identities. Thus, linear approximations for x and y contain only small error when the evaluation points a and b are near x and y.

For convenience, the means of the series for x and y are often used as fixed evaluation points. For example, Gerra, Houck, and Klein evaluated about the mean in deriving linear approximations for ratios of ratios of two variables. For many situations, this procedure would result in only a small amount of bias. However, as shown in figure 5, linear approximations about the means for per capita consumption of eggs based on an update of Gerra's 1958 egg study led to significant error in recent years. This bias or residual difference occurred because of the large downtrend in per capita egg consumption in the 1960's. Evaluation at some point other than the mean is indicated if a high degree of accuracy is desired. As indicated in figure 6, use of the previous period values ($a = x_{-1}$, $b = y_{-1}$) leads to a substantial improvement for a linear approximation of the ratio of egg consumption to civilian population. In contrast to the smooth trending data series on eggs, linear approximations of farm-to-retail price ratios for oranges were quite good over the sample period 1954 to 1968 when evaluated about the means of the price series. As is clear from figure 7, the price ratio series has irregular movements and no trend. Evaluation about last period values for retail and farm prices of oranges magnified the residual differences when sharp year-to-year changes occurred, as shown in figure 8. Experiments with products of two variables have resulted in similar observations on the appropriate choice of an evaluation point.

Based on the examples discussed for both the single-variable and two-variable cases, evaluation about the mean will give the best linear approximations for series which fluctuate but have little or no trend. For smooth data series which have significant trends, evaluation about the previous period value should be used to minimize residual error in the linear approximations. Finally, fluctuating series with a distinct trend will require a moving average of recent past observations to obtain the best linear approximations. The choice of the number of observations for the moving average will depend on the extent of fluctuation about the trend in the series.

The above statements should be qualified with respect to computational requirements if the mean of a series is not appropriate. In simultaneous equation systems containing some nonlinear variables, use of an evaluation point other than a fixed point substantially increases the computational requirements for solving the system. Use of the previous period value, for example, requires the matrix of endogenous coefficients to be inverted on each successive solution iteration since $b_i = f(x_{i-1}, y_{i-1})$.

To briefly indicate the nature of the computational requirements, the following derived demand equation for California fresh oranges has been used in a simultaneous equation model:

$$(12) \quad QCFO = 5.787 - 46.353(PCFO/PRFO) - 0.827\, WND$$

where

$QCFO$ = quantity of California oranges for fresh use

$PCFO$ = on-tree price of California fresh oranges

$PRFO$ = U.S. retail price of fresh oranges

WND = nondurable wage rate in United States

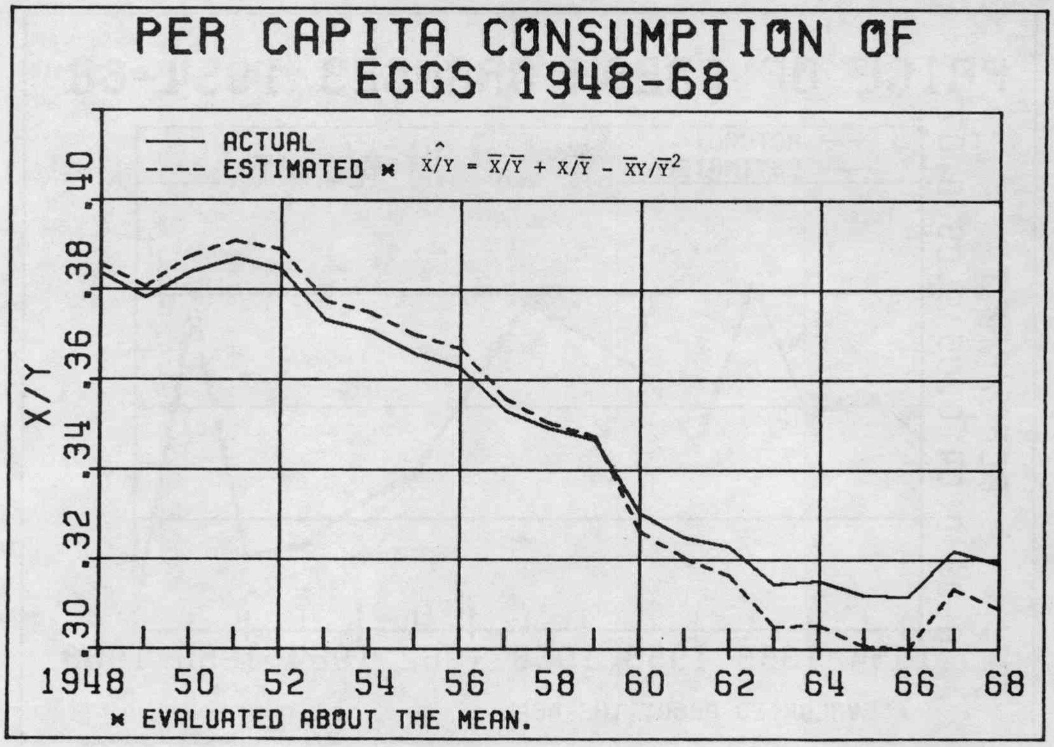

Figure 5

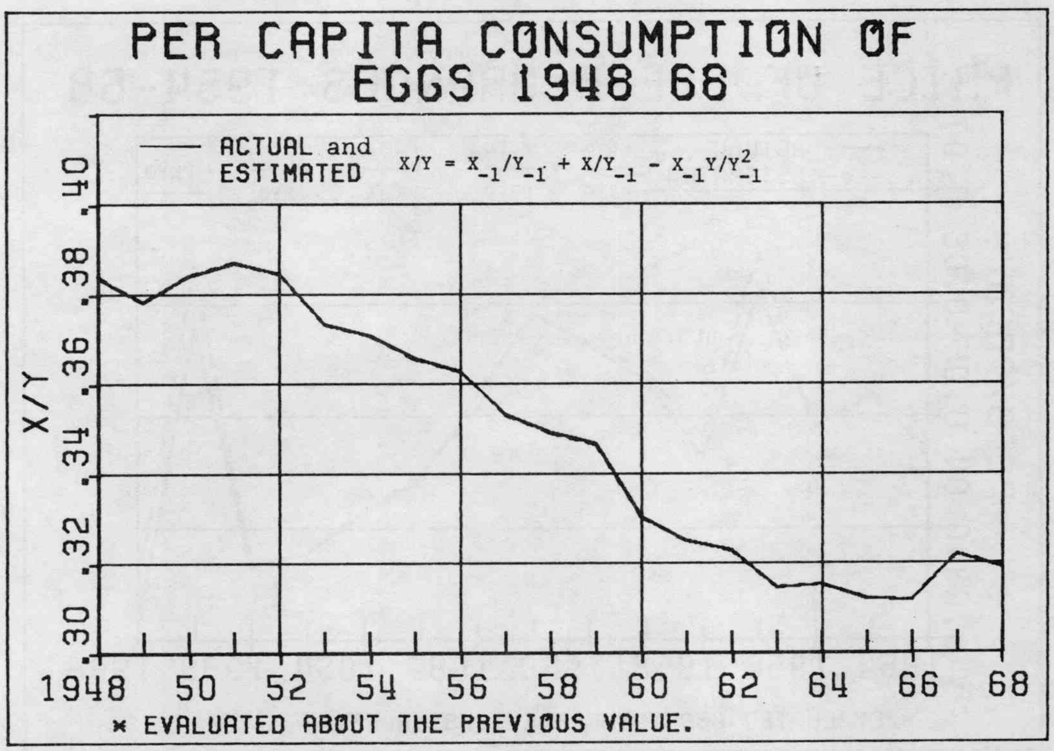

Figure 6

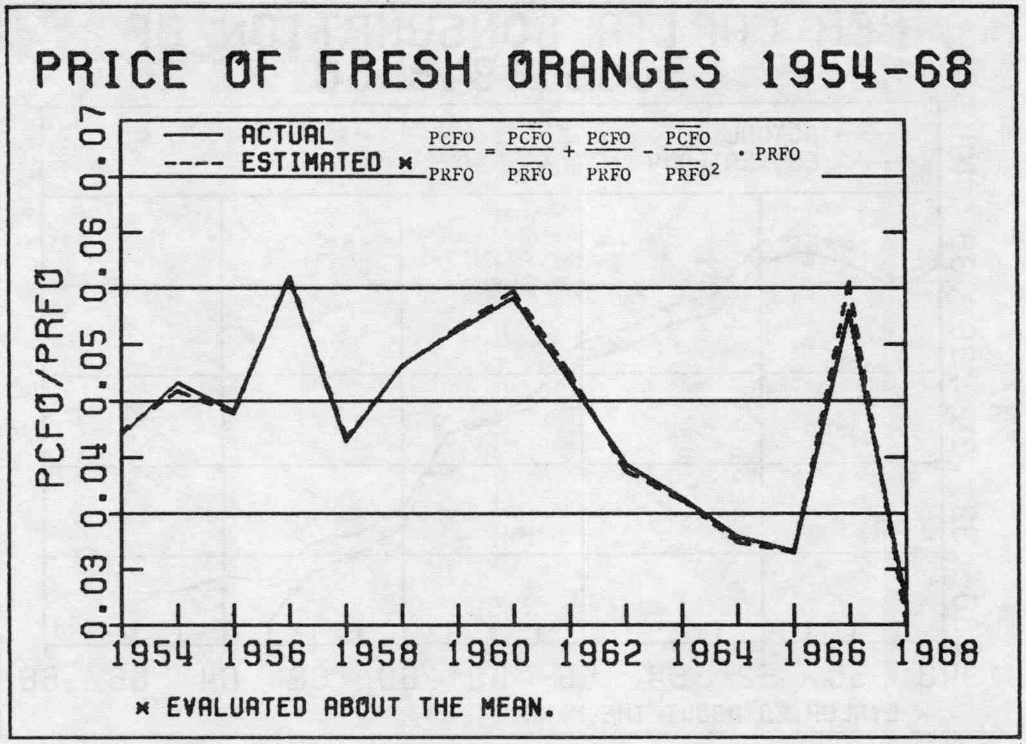

Figure 7

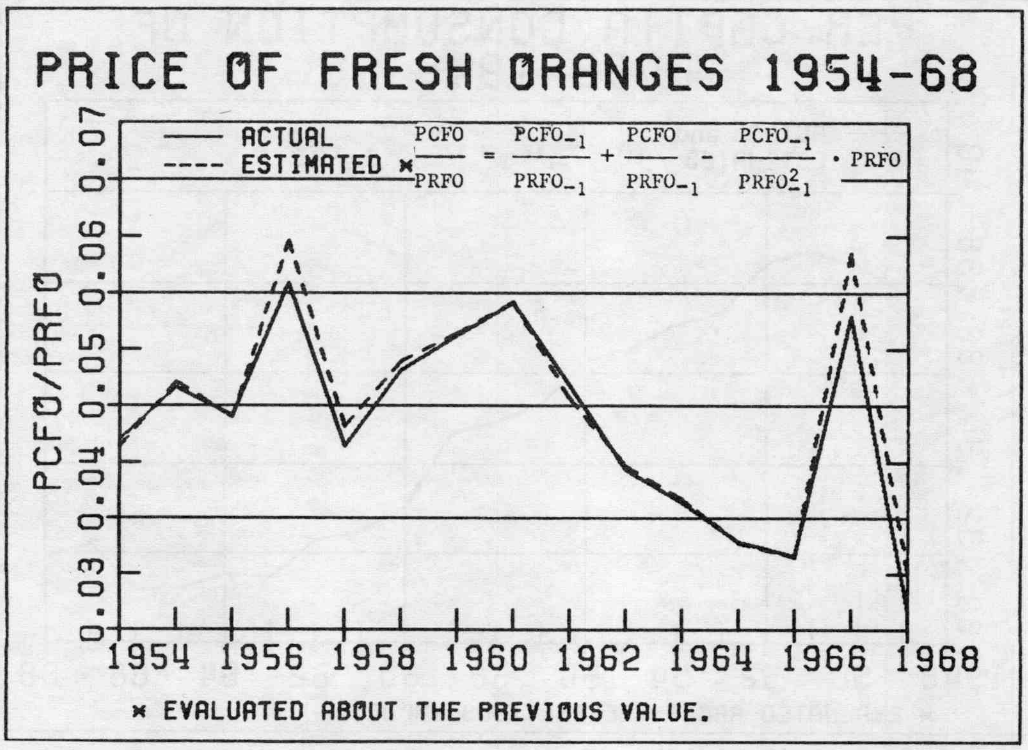

Figure 8

Since *QCFO*, *PCFO*, and *PRFO* are endogenous to the model, relation (12) can be reexpressed in linear form if reduced form solution is desired. Since a ratio of two variables is used in the equation, application of relation (10) evaluated about the means of *PCFO* and *PRFO* results in the following linear expression:

$$(13)\quad QCFO = 3.545 - 0.6118\,(PCFO) + 0.02959\,(PRFO) - 0.827\,(WND)$$

where

$$\overline{PCFO} = 3.66$$

$$\overline{PRFO} = 75.8$$

This expression in (13) is readily used in a reduced form solution and is acceptable if the bias or residual errors of the linear approximation of *PCFO/PRFO* are small as is indicated in figure 7.

Should residual error be a problem as suggested in figure 5, the use of lagged values for evaluation points might be used but the required computational procedures for a reduced form solution of the system may not be the best solution technique for the analyst to use. Other solution techniques such as the Gauss-Siedel technique should be considered (*4*, *6*). For example, evaluation of (12) about the previous period value leads to the following more complicated relation when the linear approximation is substituted for *PCFO/PRFO*:

$$(14)\quad QCFO = 5.787 - 46.353\,[PCFO_{-1}/PRFO_{-1} + (1/PRFO_{-1})PCFO - (PCFO_{-1}/PRFO^2_{-1})PRFO] - 0.827\,WND$$

or

$$QCFO = 5.787 - 46.353(PCFO_{-1}/PRFO_{-1}) - 46.353(PRFO_{-1})PCFO + 46.353(PCFO_{-1}/PRFO^2_{-1})PRFO - 0.827\,WND$$

Since the *b*'s in (14) are dependent on lagged values of *PCFO* and *PRFO*, matrix inversion of the endogenous coefficients is required each period. With the present computer capacity, more efficient solution techniques are available.

Bibliography

(1) Allen, R. G. D. Mathematical Analysis for Economists. St. Martin's Press, New York, 1938.

(2) Burt, Oscar R. and Robert M. Finley. Statistical Analysis of Identities in Random Variables. Amer. Jour. Agr. Econ. Vol. 50, No. 3, August 1968.

(3) Edwards, Clark. Non-linear Programming and Non-Linear Regression Procedures. Journ. Farm Econ. Vol. 44, No. 1, Feb. 1962.

(4) Evans, Michael K. An Agricultural Submodel for the U.S. Economy. Essays in Industrial Econometrics. Vol. 2. Wharton School of Finance and Commerce, Univ. Pa., Studies in Quantitative Economics No. 4, 1969.

(5) Gerra, Martin J. The Demand, Supply, and Price Structure for Eggs. U.S. Dept. Agr. Tech. Bul. 1204, Nov. 1959.

(6) Green, George R. Gauss-Seidel Method. Simemod, Staff Paper in Economics and Statistics, No. 17, July 1970, U.S. Dept. Commerce.

(7) Houck, James P. and Abraham S. Subotnik. A Dynamic Supply and Demand Model of the Markets for U.S. Soybeans and Their Products. Dept. Agr. Econ., Univ. Minn., Oct. 1969.

(8) Klein, Lawrence R. A Textbook of Econometrics. Univ. Mich., Dept. Econ. and Survey Res. Center, Row, Peterson and Company, 1947.

(9) Randolph, John R. Calculus and Analytic Geometry. Wadsworth Publishing Company, San Francisco, 1961.