University of Victoria
Department of Economics

Econometrics Working Paper EWP0517

ISSN 1485-6441

THE BIAS OF ELASTICITY ESTIMATORS IN LINEAR REGRESSION: SOME ANALYTIC RESULTS

Chen Qian
&
David E. Giles

Department of Economics, University of Victoria
Victoria, B.C., Canada V8W 2Y2

December, 2005

Abstract

Using small-disturbance expansions, we derive analytic expressions for the bias of the OLS estimator of an elasticity in a linear model, both at an individual sample point and at the sample mean. The magnitudes of these biases are illustrated with Australian expenditure data.

Keywords: Elasticity; bias; small-disturbance asymptotics

JEL Classifications: C13; C20; D12

Author Contact:
David E. Giles, Dept. of Economics, University of Victoria, P.O. Box 1700, STN CSC, Victoria, B.C., Canada V8W 2Y2; e-mail: dgiles@uvic.ca; Phone: (250) 721-8540; FAX: (250) 721-6214

1. Introduction

The linear regression model is frequently used to estimate elasticities between variables of interest. When the model is linear in the variables and parameters, the elasticities and their estimators are nonlinear functions of the random data. Accordingly, these estimators are biased in finite samples – a point that is invariably ignored in practice. This paper explores the magnitude of this bias. Specifically, suppose that the model is

$$y = X\beta + \sigma u \tag{1}$$

where X is $(n \times k)$, non-random and of full rank, and $u \sim N[0, I_n]$. At observation '*i*' the OLS estimator of the elasticity of y with respect to the j'th. regressor is

$$\hat{\eta}_{ij} = \hat{\beta}_j \frac{x_{ij}}{y_i} \quad ; \tag{2}$$

or if we estimate the elasticity at the sample mean,

$$\bar{\hat{\eta}}_j = \hat{\beta}_j \frac{\bar{x}_j}{\bar{y}} \quad , \tag{3}$$

where $\hat{\beta}_j$, the OLS estimator of β_j (j = 1, 2, …., k), is itself a function of the y data. Under mild conditions, these estimators are consistent. Of greater interest is their behaviour in *finite* samples. The *exact* biases of the estimators are complicated by the fact that (2) and (3) are non-linear functions of the 'y' data.

.

The plan of the rest of the paper is as follows. In the next section we discuss small-sigma approximations and present the principal results that we use in our derivations. Section 3 presents the main theoretical results, and interprets their implications. An empirical example that illustrates these implications is provided in section 4, and section 5 concludes.

2. Small-Disturbance Approximations

The biases of the elasticity estimators can be approximated in various ways. One option is to use an approximation based on an analytic expansion whose accuracy improves as 'n', the sample size, grows. Such approximations, proposed by Nagar (1959) in the econometric context, involve an expansion of the sampling error such that the successive terms are in decreasing order of 'n', in probability. When used to determine the moments of an estimator, this approach yields the moments of the Edgeworth expansion of that estimator's distribution (Ullah, 2004, p.29).

An alternative (*e.g.*, Kadane, 1971) is to approximate the finite-sample moments of the estimator via an expansion of the sampling error such that successive terms are in decreasing order of the population standard deviation, σ, in probability. These 'small-disturbance', or 'small-σ', approximations have been found to be extremely valuable for a number of problems in econometrics. They are valid for any sample size, and they do not require any additional assumptions about the behaviour of the sample moments of the data as n increases.[1] A good recent discussion of 'small-disturbance' expansions is given by Ullah (2004, 36-45).

Consider the data-generating process:

$$y_i = \mu + \sigma u_i \quad ; \qquad i = 1, 2, \ldots, n \tag{4}$$

where $\mu \neq 0$ and the u_i's are independently and identically distributed with

$$E(u_i) = 0 \quad ; \quad E(u_i^2) = 1 \quad ; \quad E(u_i^3) = \gamma_1 \quad ; \quad E(u_i^4) = \gamma_2 + 3. \tag{5}$$

So, the skewness and excess kurtosis of the population distribution are γ_1 and γ_2 respectively. Our objective is to determine the magnitude of the biases of the elasticity estimators in (2) and (3), under very mild assumptions about the population distribution, to $O(\sigma^4)$. We use the following result.

Lemma 1 (Ullah, 2004, 38.)

Let 'y' be an n-element random vector, with $y = \mu + \sigma u$, where 'u' satisfies the conditions in (5) above, and the non-zero mean, μ, is a function of a parameter vector, θ. Let $\hat{\theta} = h(y)$ be an estimator of θ, where $h(y)$ and its derivatives exist in a neighborhood of μ. Then[3]

$$E(\hat{\theta} - \theta) = \sigma^2 \Delta_2 + \sigma^3 \gamma_1 \Delta_3 + \sigma^4 (\gamma_2 \Delta_4 + 3\Delta_{22}) \ , \tag{6}$$

where, for s = 2, 3, 4:

$$\Delta_s = \frac{1}{s!} \sum_{k=1}^{n} \left[\frac{\partial^s h(y)}{\partial y_k^s} \right]_{y=\mu} \qquad ; \qquad \Delta_{22} = \frac{1}{4!} \sum_{k=1}^{n} \sum_{l=1}^{n} \left[\frac{\partial^4 h(y)}{\partial y_k^2 \partial y_l^2} \right]_{y=\mu} .$$

3. Results

Now consider model (1) under the conditions of Lemma 1. Using standard partitioning results,

$$\hat{\beta}_j = (x_j' M_{-j} x_j)^{-1} (x_j' M_{-j} y) ; \qquad j = 1, 2, ..., k ; \tag{7}$$

where $M_{-j} = I - X_{-j} (X_{-j}' X_{-j})^{-1} X_{-j}'$, x_j is the *j*'th. column of X, and X_{-j} is the regressor matrix excluding x_j .

3.1 Elasticity at an Individual Data Point

So the OLS estimator of the elasticity of y with respect to x_j for the *i*'th. observation is:

$$\hat{\eta}_{ij} = \hat{\beta}_j \frac{x_{ij}}{y_i} = x_{ij} (x_j' M_{-j} x_j)^{-1} \frac{(x_j' M_{-j} y)}{y_i} = x_{ij} (x_j' M_{-j} x_j)^{-1} h(y) \tag{8}$$

$$h(y) = \frac{(x_j' M_{-j} y)}{y_i} = x_j' M_{-j} y^* \tag{9}$$

$$y^* = \frac{y}{y_i} \quad . \tag{10}$$

Then the small-σ asymptotic approximation to the bias of this elasticity estimator is :

Theorem 1

$$\begin{aligned} \text{Bias}(\hat{\eta}_{ij}) = x_{ij} \{ & \sigma^2 [\beta_j \sum_i (x_i'\beta)^{-3} - (x_j' M_{-j} x_j)^{-1} \sum_i [(x_j' M_{-j} i_i) / (x_i'\beta)^2]] \\ & - \sigma^3 \gamma_1 [\beta_j \sum_i (x_i'\beta)^{-4} - (x_j' M_{-j} x_j)^{-1} \sum_i [(x_j' M_{-j} i_i) / (x_i'\beta)^3]] \\ & + \sigma^4 (\gamma_2 + 3) [\beta_j \sum_i (x_i'\beta)^{-5} - (x_j' M_{-j} x_j)^{-1} \sum_i [(x_j' M_{-j} i_i) / (x_i'\beta)^4]] \} , \end{aligned}$$

where i_i is an $(n \times 1)$ column vector whose *i*'th. element is unity, and all other elements are zero. Summations run from 1 to *n*.

Proof

From equation (8),

$$\text{Bias}(\hat{\eta}_{ij}) = x_{ij} (x_j' M_{-j} x_j)^{-1} \text{Bias}[h(y)] \quad , \tag{11}$$

and from (9),

$$\frac{\partial^s h(y)}{\partial y_k^s} = x_j' M_{-j} \frac{\partial^s y^*}{\partial y_k^s} \quad . \tag{12}$$

Some tedious but straightforward partial differentiation yields the following results. For s = 2, 3, 4, …..

$$\frac{\partial^s h(y)}{\partial y_k^s} = (-1)^s (s!)[x_j' M_{-j}(y - y_k i_k)] / y_k^{s+1} \qquad ; k = i$$

$$= 0 \qquad ; k \neq i$$

$$\frac{\partial^4 h(y)}{\partial y_k^2 \partial y_l^2} = (4!)[x_j' M_{-j}(y - y_k i_k)] / y_k^5 \qquad ; i = k = l$$

$$= 0 \qquad ; \text{otherwise}$$

So from Lemma 1, and recalling that $E(y_i) = x_i'\beta$ (i = 1, 2,……, n):

$$\Delta_s = (-1)^s \{(x_j' M_{-j} x_j \beta_j) \sum_k (x_k'\beta)^{-(s+1)} - \sum_k [(x_j' M_{-j} i_k)/(x_k'\beta)^s]\};$$

for s = 2, 3, 4; and

$$\Delta_{22} = \Delta_4 \ ,$$

Finally, applying Lemma 1 and (11), the result in Theorem 1 follows immediately, and the sign and magnitude of the bias are indeterminate.

Corollary 1

If the model has only a single regressor and is fitted through the origin, then the bias of the elasticity estimator for any particular observation is:

$$\text{Bias}(\hat{\eta}_i) = x_i [\frac{\sigma^2}{\beta^2}(\sum_i \frac{1}{x_i^3} - \sum_i \frac{1}{x_i} \Big/ \sum_i x_i^2) - \frac{\sigma^3 \gamma_1}{\beta^3}(\sum_i \frac{1}{x_i^4} - \sum_i \frac{1}{x_i^2} \Big/ \sum_i x_i^2)$$
$$+ \frac{\sigma^4(\gamma_2 + 3)}{\beta^4}(\sum_i \frac{1}{x_i^5} - \sum_i \frac{1}{x_i^3} \Big/ \sum_i x_i^2) .$$

Again, the sign and magnitude of the bias are indeterminate, in general. However, using Hölder’s inequality (Kendall and Stuart, 1977, p.39), for positive regressors the bias is positive to $O(\sigma^2)$, and also positive to $O(\sigma^4)$ for symmetric and leptokurtic (or mesokurtic) data.

3.2 Elasticity at the Sample Mean

Theorem 1 provides us with the bias of the elasticity estimator with respect to one particular independent variable for any particular observation. In practice, it is common to report a representative elasticity by evaluating it at the mean of the sample size. In this case, we have:

$$\bar{\hat{\eta}}_j = \hat{\beta}_j \frac{\bar{x}_j}{\bar{y}} = \bar{x}_j \, (x_j' M_{-j} x_j)^{-1} \frac{(x_j' M_{-j} y)}{\bar{y}} = \bar{x}_j \, (x_j' M_{-j} x_j)^{-1} \, g(y) \tag{13}$$

$$g(y) = \frac{(x_j' M_{-j} y)}{\bar{y}} = x_j' M_{-j} y^* \tag{14}$$

$$y^* = \frac{y}{\bar{y}} \quad . \tag{15}$$

Theorem 2

$$\text{Bias}(\bar{\hat{\eta}}_j) = \bar{x}_j \, [\beta_j - (\bar{X}\beta) \frac{(x_j' M_{-j} i)}{(x_j' M_{-j} x_j)^{-1}}][\frac{\sigma^2}{n(\bar{X}\beta)^3} - \frac{\sigma^3 \gamma_1}{n^2 (\bar{X}\beta)^4} + \frac{\sigma^4}{n^3 (\bar{X}\beta)^5} (\gamma_2 + 3n)]],$$

where i is an $(n \times 1)$ one column vector.

Proof

From (13),

$$\text{Bias} \, (\bar{\hat{\eta}}_j) = \bar{x}_j \, (x_j' M_{-j} x_j)^{-1} \text{Bias} \, [\, g(y) \,] \tag{16}$$

For s = 2, 3, 4, …..

$$\frac{\partial^s g(y)}{\partial y_k^s} = (-1)^s (s!)[x_j' M_{-j} (y - n\bar{y} i_k)] / n^s \, \bar{y}^{(s+1)}$$

$$\frac{\partial^4 g(y)}{\partial y_k^2 \partial y_l^2} = (4!)[x_j' M_{-j} (y - 0.5 n\bar{y} i_k - 0.5 n\bar{y} i_l)] / n^4 \bar{y}^5 \qquad ; \; k \neq l$$

$$= (4!)[x_j' M_{-j} (y - n\bar{y} i_l)] / n^4 \bar{y}^5 \qquad ; \; k = l$$

$$\Delta_s = (-1)^s [(x_j' M_{-j} x_j \beta_j) - (\bar{X}\beta) \sum_k (x_j' M_{-j} i_k)] / n^{s-1} (\bar{X}\beta)^{(s+1)}$$

$$\Delta_{22} = [(x_j' M_{-j} x_j \beta_j) - (\bar{X}\beta) \sum_k (x_j' M_{-j} i_k)] / n^2 (\bar{X}\beta)^5 \; ,$$

where $\bar{X}$ is a $(1 \times k)$ vector whose i'th. element is the mean value of the i'th. column of X. Finally, applying Lemma 1 and (16), the result in Theorem 2 follows immediately.

Corollary 2

If the model has only a single regressor and is fitted through the origin, then the bias of the elasticity evaluated at the sample mean is:

$$\text{Bias}(\bar{\hat{\eta}}) = \left(\frac{1}{\bar{x}^2} - \frac{n}{\sum_i x_i^2}\right)\left[\frac{\sigma^2}{n\beta^2} - \frac{\sigma^3\gamma_1}{n^2\bar{x}\beta^3} + \frac{\sigma^4(\gamma_2 + 3n)}{n^3\bar{x}^2\beta^4}\right].$$

Interpreting this result, first by Jensen's inequality, $\bar{x}^2 \leq \sum_i x_i^2 / n$. So to $O(\sigma^2)$, regardless of the underlying distribution of the errors and the sample values, the elasticity estimator is biased upwards. A sufficient, but not necessary, condition for the bias to be upwards, to $O(\sigma^4)$ is:

$$1 + \frac{\sigma^2(\gamma_2 + 3n)}{n^2\bar{x}^2\beta^2} > \frac{\sigma\gamma_1}{n\bar{x}\beta} \; .$$

If the regressor takes only positive values, and if the errors follow a symmetric and leptokurtic distribution, Bias($\bar{\hat{\eta}}$) > 0, to $O(\sigma^4)$.

4. Empirical Illustration

To illustrate these results we have estimated linear Engel curves for alcoholic beverages and marijuana, using Australian data from Clements and Zhao (2005). Expenditure on each good is explained by total expenditure on the group of goods.[2] The OLS-estimated elasticities and percentage biases appear in Table 1, from which bias-adjusted elasticities are readily deduced.[3] The estimated biases are positive, ranging from 0% to 10% for individual sample points, but are negligible for elasticities evaluated at the sample mean.[4] The bias estimates to $O(\sigma^2)$ are very close to those to $O(\sigma^4)$, and their signs partly reflect the positive data and regression coefficient estimates.

5. Conclusions

We have derived analytic expressions for the bias of the OLS estimator of the point elasticity in a linear model, both at an individual sample point and at the sample mean. These expressions are based on small-disturbance expansions. These biases can be substantial enough to warrant attention in practice, especially in the case of point elasticities. Additional empirical evaluations undertaken by the authors corroborate this finding.

Table 1: Estimated Elasticities and Biases

	Beer			Wine		
Year	**Elasticity**	**% Bias**		**Elasticity**	**% Bias**	
		$O(\sigma^2)$	$O(\sigma^4)$		$O(\sigma^2)$	$O(\sigma^4)$
1988	0.6760	0.0744	0.0743	1.4234	3.6568	3.7783
1989	0.6868	0.0774	0.0773	1.4977	3.6772	3.7994
1990	0.6852	0.0814	0.0813	1.6239	3.5513	3.6693
1991	0.6991	0.0824	0.0823	1.6449	3.6255	3.7460
1992	0.7007	0.0802	0.0801	1.5507	3.7538	3.8788
1993	0.6982	0.0804	0.0803	1.5099	3.8563	3.9848
1994	0.7056	0.0817	0.0816	1.4804	4.0455	4.1805
1995	0.6967	0.0836	0.0835	1.4371	4.2166	4.3576
1996	0.7072	0.0862	0.0861	1.4257	4.4607	4.6102
1997	0.7167	0.0874	0.0873	1.3696	4.7863	4.9473
1998	0.7333	0.0878	0.0877	1.3507	4.9964	5.1647
Mean	0.7011	0.0006	0.0006	1.4748	0.0257	0.0258

	Spirits			Marijuana		
Year	**Elasticity**	**% Bias**		**Elasticity**	**% Bias**	
		$O(\sigma^2)$	$O(\sigma^4)$		$O(\sigma^2)$	$O(\sigma^4)$
1988	2.5389	5.4676	5.7492	0.5653	3.4044	3.2512
1989	2.4314	6.0742	6.3890	0.5486	3.7219	3.5540
1990	2.4268	6.4001	6.7328	0.5298	4.0536	3.8702
1991	2.5241	6.3563	6.6866	0.5061	4.3997	4.1999
1992	2.3772	6.5941	6.9376	0.5308	4.0770	3.8925
1993	2.0773	7.6130	8.0135	0.5757	3.7447	3.5757
1994	1.9807	8.2451	8.6815	0.5885	3.7612	3.5914
1995	1.9640	8.4130	8.8591	0.6123	3.6477	3.4832
1996	2.0035	8.6546	9.1145	0.5956	3.9379	3.7598
1997	1.9886	8.9874	9.4666	0.6013	4.0108	3.8293
1998	1.9518	9.4478	9.9538	0.5942	4.1764	3.9871
Mean	2.1694	0.0425	0.0427	0.5671	0.0260	0.0260

Note: ‘Mean’ refers to evaluation at the sample mean.

References

Clements, K. W., Zhao, X., 2005. Economic aspects of marijuana. Invited address, Economics Ph.D. Conference, University of Western Australia.

Kendall, M., Stuart, A., 1977. The Advanced Theory of Statistics. Charles Griffin, London.

Nagar, A. L., 1959. The bias and moments matrix of the general k-class estimators of the parameters in structural equations. Econometrica 27, 575-595.

Kadane, J., 1971. Comparison of k-class estimators when the disturbances are small. Econometrica 39, 723-737.

Ullah, A., 2004. Finite Sample Econometrics. Oxford University Press, Oxford.

Footnotes

1. Assumptions of the latter type are required for the validity of 'large-*n*' expansions, and can be difficult to verify in practice.
2. The regressions include an intercept and the sample covers 1988 – 1998.
3. The true elasticity values are unknown. Estimated '% bias' is 100 times the ratio of the estimated bias to the 'bias-adjusted' elasticity estimate.
4. Differences of this order between the individual and mean results can be explained from Theorems 1 and 2, using Jensen's inequality; and by considering $(\partial Bias(\hat{\eta}_{ij})/\partial x_{ij})$. Some experimentation with other data yields similar results.