A Consistent Least Squares Estimator Of The Economic Base Multiplier With Measurement Errors In The Export Coefficients*

SE-HARK PARK**

I. Introduction

Empirical estimations of the regional multiplier have been abundant in the literature of regional studies for many years. There have been traditionally two general approaches to the study of regional multipliers: one using the input-output model and the other the economic base approach.[1] However, a large number of structural simultaneous equation models, which have been used primarily at the level of national economy, have been recently constructed and estimated to analyze many regional economic problems, including regional multiplier effects.[2] Although input-output and regional macroeconomic models introduce a high degree of sophistication and comprehensiveness, the construction of such regional models usually entails high costs and complex data problems, especially in the case of a complete interindustry model.[3] These difficulties seem to have led many investigators to use less costly and simpler economic base models.

The economic base approach typically dichotomizes regional economic activities into two categories: those economic activities serving demand and agencies external to the region—export or basic—and those activities serving local residents and enterprises—service or local. Then it postulates that the level of locally-oriented service activities is determined by variations in export activities.[4]

In recent years, a considerable number of economic base multipliers have been estimated by least squares.[5] Yet, despite its wide application, there have been few attempts to examine the validity of the least squares method for the application in question. Recently, Park [14] set out the algebra of regression analysis of economic base multiplier models commonly found in the literature, characterized and interpreted the basic features of these models, and evaluated their advantages and limitations in practical applications. One of the many findings of the study was that the accuracy of the multiplier estimated by least squares depends critically upon the accuracy of regional export coefficients used.

*I wish to thank the referee for his helpful comments. Any findings, views, and conclusions presented herein are those of the author and do not necessarily reflect the views of the Department of Energy or Marquette University.

**Economist, Regional Socioeconomic, and Environmental Analysis Division, Department of Energy, and Professor of Economics, Marquette University.

There are numerous techniques available for estimating the levels of these export activities at the regional level, ranging from a simple arbitrary assumption as to the percentage of exports in a given industry's output to comprehensive surveys of firms and consumers. The conceptual and technical difficulties of estimating regional export activities have been discussed extensively in the literature and need not be further elaborated here.[6]

In view of the critical importance of export activities and the difficulties of estimating them in the economic base multiplier analysis, I will attempt in this paper to analyze the effects of measurement errors in the export coefficients upon the multiplier value yielded by ordinary least squares and derive a consistent least squares estimator of the economic base mulitplier.

II. The Regression Model for Economic Base Multipliers

Suppose there are k industries in a region and let Z_{tj} be an observation on total income or earnings (or employment) for the j^{th} industries in time period t.[7] Then Z_{tj} can be divided into export and local components

$$(1) \qquad Z_{tj} = a_j Z_{tj} + (1-a_j)\, Z_{tj} \quad 0 \leq a_j \leq 1, \quad j = 1, \ldots, k, \quad t = 1, \ldots, n$$

where a_j is an export coefficient of the j^{th} industry. To be more specific, an export coefficient a_j represents the proportion of income of the j^{th} industry originating from sales to buyers outside a given region. Also, it must be noted that the local component $(1-a_j)Z_{tj}$ includes not only income earned in locally-oriented service activities supported by regional exports, but also all other nonexport income affected by factors other than changes in export income.[8] In most of the base multiplier studies, the export coefficients are assumed to remain constant for the entire time period in question.[9] Summing over j both sides of equation (1), we obtain an identity equation for total area income, i.e.,

$$(2) \qquad Z_t = X_t + Y_t$$

where $Z_t = \sum_{j=1}^{k} Z_{tj}$ is total area income,

$X_t = \sum_{j=1}^{k} a_j Z_{tj}$ is income earned in the export sector, and

$Y_t = \sum_{j=1}^{k} (1 - a_j) Z_{tj}$ is income earned in the local-service sector.

Given measures of local and export activities by industry, a regression of aggregate income generated in local sectors on those in export sectors yields a quantitative measure of the export-service relationship over time. The aggregate multiplier model we are concerned with is then

$$(3) \qquad Y_t = b_o + b \sum_{j=1}^{k} a_j Z_{tj} = b_o + bX_t$$

Substituting (3) into (2) will give us

$$(4) \qquad Z_t = b_o + (1 + b) X_t$$

which shows that the regression coefficient b of equation (3) plus one, (1+b), represents the regional income multiplier which we attempt to estimate, since the multiplier is defined as the ratio of an incremental change in total income to that in the basic or export sectors.

Let us assume that total income Z_t is correctly measured and measurement errors of export and service income result only from the incorrect estimation of the export coefficients a_j, i.e.,

$$(5) \qquad a_j = a^*_j + \hat{a}_j \qquad 0 \leq a_j \leq 1 \qquad j = 1, \ldots, k$$

where a_j indicates estimated values, a^*_j true values and $\hat{a}_j$ errors. Suppose that an exact linear relationship is assumed to exist between the true variables y^*_t and x^*_t in mean deviation form

$$(6) \qquad y^*_t = bx^*_t$$

but that these true variables are unobserved because of estimation errors $\hat{a}_j$ in equation (5). Our sample consists of observations on the measured variables y_t and x_t that are related to the true variables by

$$(7) \qquad y_t = y^*_t + u_t, \; x_t = x^*_t + v_t$$

where u_t and v_t are the errors of observation.

However,

$$(8) \qquad x_t = \sum_j a_j z_{tj} = \sum_j (a^*_j + \hat{a}_j) z_{tj} = \sum_j a^*_j z_{tj} + \sum_j \hat{a}_j z_{tj} = x^*_t + v_t$$

and

$$y_t = \sum_j (1 - a_j) z_{tj} = \sum_j (1 - a^*_j) z_{tj} - \sum_j \hat{a}_j z_{tj} = y^*_t + u_t$$

where

$$v_t = \sum_j \hat{a}_j z_{tj}, \quad u_t = - \sum_j \hat{a}_j z_{tj}$$

and

$$v_t = - u_t$$

substituting (7) into (6) and rearranging yields

(9)

$$y_t = b(x_t - v_t) + u_t = bx_t + \varepsilon_t$$

where

$$\varepsilon_t = (u_t - bv_t) = - (1+b) v_t$$

The case where measurement errors ε_t include a stochastic component of behavior will be treated later. Let us specify that the errors u and v have zero means and constant variance, and are also independent of the true values x^* and y^* so that

(10)

$$E(u_t) = E(v_t) = 0, E(u_t^2) = E(v_t^2) = \sigma_v^2,$$
$$E(u_t y^*_t) = E(v_t x^*_t) = E(u_t x^*_t) = E(v_t y^*_t) = 0$$

Then the covariance of x_t and ε_t is

(11)

$$E(x_t \varepsilon_t) = E\left\{ (x^*_t + v_t)(-(1+b) v_t) \right\}$$
$$= - (1+b) \sigma_v^2$$

Thus, it is clear that the regressor is contemporaneously correlated with the error terms, that is, a dependence exists between all the measured values and the error of measurement. Obviously, the straightforward application of least squares to (9) would yield not only a biased estimate of the multiplier parameter b, but also the least squares estimate is inconsistent. The asymtotic bias in this case will be

(12)
$$\operatorname{plim} (\hat{b} - b) = \frac{\operatorname{plim} (\Sigma x \varepsilon / n)}{\operatorname{plim} (\Sigma x^2/n)}$$
$$= - \frac{(1+b) \sigma_v^2}{\sigma x^{*2} + \sigma_v^2}$$

since $E(x^2) = E(x^* + v)^2 = E(x^{*2}) + E(v^2) + 2E(x^* v)$

$$= \sigma_{x*}^2 + \sigma_v^2$$

where σ_{x*}^2 is the variance of the true value x^*.

Rewriting (12), we obtain

(13)

$$\operatorname{plim} \hat{b} = b - \frac{(1+b)\,\sigma_v^2}{\sigma_v^2 + \sigma_{x*}^2} = \frac{b\sigma_{x*}^2 - \sigma_v^2}{\sigma_{x*}^2 + \sigma_v^2}$$

Equation (13) shows that since by the multiplier hypothesis, $b > 0$, the ordinary least squares estimate of the regional multiplier underestimates the true multiplier to the extent that the variance of the true export income is small relative to the variance of measurement errors associated with export income.

Introducing the notation for the sample covariance

$$m_{xy} = \frac{1}{n} \Sigma\, (X_i - \bar{X})\, (Y_i - \bar{Y}) = \frac{1}{n} \Sigma\, xy$$

$$m_{xx} = \frac{1}{n} \Sigma\, (X_i - \bar{X})\, (X_i - \bar{X}) = \frac{1}{n} \Sigma\, x^2$$

first let us take covariances of X with all variables in (9) to derive a consistent estimate of b under the assumptions given in (10).

(14)

$$m_{xy} = b\; m_{xx} + m_{x\varepsilon}$$

Since the sample covariance $m_{x\varepsilon}$ is approximately equal to or exactly equal in the probability limit to population covariance, i.e., from (11)

$$\operatorname{plim}\, (\Sigma x\varepsilon / n) = -\, (1+b)\, \sigma_v^2$$

we can rewrite (14) in approximation as

(15)

$$m_{xy} \triangleq b\, m_{xx} - (1+b)\, \sigma_v^2$$

To correctly estimate b, we need a second estimating equation because there are two unknowns, b and σ_v^2. We, therefore, take covariance of Y with all variables in (9)

(16)

$$m_{yy} = b\, m_{xy} + m_{y\varepsilon}$$
$$= b\, m_{xy} + (1+b)\, \sigma_v^2$$

$$\text{since } \operatorname{plim}\, (\Sigma y\varepsilon/n) = \operatorname{plim}\, (\Sigma(y^* + u)\, (u - bv)/n)$$
$$= \operatorname{plim}\, (-(1+b)\, \Sigma\, y^* v/n) + \operatorname{plim}\, ((1+b)\, \Sigma\, v^2/n)$$
$$= (1+b)\, \sigma_v^2 \text{ using } (10)$$

Adding (15) and (16) eliminates term $(1 + b)\, \sigma^2{}_v$ and gives a solution for $\hat{b}$

(17)

$$\hat{b} = \frac{m_{xy} + m_{yy}}{m_{xx} + m_{xy}}$$

Thus, unlike many errors-in-variables approaches such as Johnston's work [12], b can be estimated without knowing the variance of the error term σ_v^2,[10] or the ratio of error variances. Needless to say, if $\sigma_v^2 = 0$, $\hat{b} = m_{xy}/m_{xx}$ from (15), which is the ordinary least squares estimate of b.

Now let us include the assumption of a stochastic relation between the true values. The model would then be in mean deviation form

$$
\begin{aligned}
y_t &= y^*_t + u_t \\
x_t &= x^*_t + v_t \qquad (18)\\
y^*_t &= bx^*_t + e_t
\end{aligned}
$$

The first two relations are the same as before, with u and v indicating the errors of observation and $-u = v$. The e_t term in (18) is a stochastic disturbance term. These relations might be expressed simply as

(19)

$$
\begin{aligned}
y_t &= bx_t - (1+b)\, v_t + e_t \\
&= bx_t + \varepsilon_t
\end{aligned}
$$

where this time $\varepsilon_t = -(1+b)\, v_t + e_t$

If v and e are all assumed to be independently distributed with zero expectations and variances σ_v^2 and σ_e^2, then there is an exact formal correspondence between this and the previous model, provided we treat the composite term u + e in this one as the equivalent of the error of observation in y in the previous one.

Following the same procedure used with the previous model, we can obtain covariances of X with all variables in (19)

(20)

$$
\begin{aligned}
m_{xy} &\stackrel{\wedge}{=} b\, m_{xx} - (1+b)\, \sigma_v^2 + m_{xe} \\
&\stackrel{\wedge}{=} b\, m_{xx} - (1+b)\, \sigma_v^2
\end{aligned}
$$

since $E(xe) = E(x^* + v)e = E(x^*e) + E(ve) = 0$

Likewise, we take covariances of Y with all variables in (19)

$$
\begin{aligned}
m_{yy} &\stackrel{\wedge}{=} b\, m_{xy} + (1+b)\, \sigma_v^2 + m_{ye} \qquad (21)\\
&\stackrel{\wedge}{=} b\, m_{xy} + (1+b)\, \sigma_v^2 + \sigma^2
\end{aligned}
$$

since $E(ye) = E\left\{ (bx_t - (1+b)\, v_t + e_t)\, e_t \right\}$

$$= \sigma_e^2$$

By adding (20) to (21), we can eliminate $(1+b)\, \sigma_v^2$ and hence, obtain

(22)
$$\hat{b}^{\prime} = \frac{m_{xy} + m_{yy} - \sigma_e^2}{m_{xx} + m_{xy}}$$

Rewriting (22) as

(23)
$$\hat{b}^{\prime} = \frac{m_{xy} + m_{yy}}{m_{xx} + m_{xy}} - \frac{\sigma_e^2}{m_{xx} + m_{xy}}$$
$$= \hat{b} - \frac{\sigma_e^2}{m_{xx} + m_{xy}}$$

we can show easily $\hat{b} > \hat{b}^{\prime}$, i.e., the parameter estimate $\hat{b}$ obtained under the assumption of an exact relation between the true values is larger than the one estimated under the assumption of a stochastic relation. This is because

(24)
$$m_{xx} \hat{=} m_{x^*x^*} + \sigma^2$$
$$m_{xy} \hat{=} b\ m_{xx} - (1 + b)\ \sigma_v^2$$
$$\hat{=} b\ (m_{x^*x^*} + \sigma_v^2) - (1 + b)\ \sigma_v^2$$
$$= b\ m_{x^*x^*} - \sigma_v^2$$

Therefore

$$m_{xx} + m_{xy} \hat{=} (1 + b)\ m_{x^*x^*} > 0 \text{ for } b > 0$$

and hence,

$$\sigma_v^2 / (m_{xx} + m_{xy}) > 0$$

As a result, the underestimation of the multiplier would be greater in the model which incorporates the assumption of a stochastic relation between the true values than in the one involving an exact relation. However, unlike in the previous case, an *a priori* or independent estimate of the variance of a stochastic term σ_e^2 is required to obtain a consistent estimate of the multiplier.

III. Summary and Conclusions

In this paper, we have treated the problem encountered in the use of linear regression models to estimate economic base multipliers where measurement errors are present in the export coefficients.

In all cases, the ordinary least squares estimates were shown to be biased, inconsistent, and to underestimate the true multipliers. Under the assumption of an exact relation between the true values, we developed a modified consistent estimation procedure which does not depend on a knowledge of the variance-covariance matrix of measurement errors. However, an *a priori* knowledge of the variance of the disturbance term was required for the similar modified consistent estimator when a stochastic relation was postulated between the true values. But this restriction does not appear to be too critical in practice, since the multipliers estimated under the assumption of a stochastic relation were shown to be smaller than those obtained under the assumption of an exact relation, and hence, the upper limit of the multipliers values can be readily estimated without knowing the variances of the stochastic terms or measurement errors.

Finally, the method of instrumental variables may be practical for our problem, but that possibility is not pursued in any depth in this paper.

APPENDIX

Consistency Proof of the Estimator

The modified least squares equation for the aggregate multiplier is

$$\hat{b} = (\Sigma xy + \Sigma y^2) / (\Sigma x^2 + \Sigma xy) \tag{A.1}$$

Therefore,

$$\text{plim } \hat{b} = \frac{\text{plim } (\Sigma xy/n) + \text{plim } (\Sigma y^2/n)}{\text{plim } (\Sigma x^2/n) + \text{plim } (\Sigma xy/n)} \tag{A.2}$$

Substituting $y = bx + \varepsilon$ into (A.2) and noting $\varepsilon = -(1 + b)v$ and $E(x\varepsilon) = -(1 + b)\,\sigma_v^2$ will give

$$\text{plim } (\Sigma xy/n) = b\,\sigma_x^2 - (1 + b)\,\sigma_v^2$$

and

$$\text{plim } (\Sigma y^2/n) = b^2\,\sigma_x^2 - 2b\,(1 + b)\,\sigma_v^2 + (1 + b)^2\,\sigma_v^2$$

$$\text{since plim } (\Sigma \varepsilon^2/n) = (1 + b)^2\,\sigma_v^2$$

As a result

$$\text{plim } \hat{b} = \frac{b\,(1 + b)\,(\sigma_x^2 - \sigma_v^2)}{(1 + b)\,(\sigma_x^2 - \sigma^2)} = b$$

and consistency is proved.

FOOTNOTES

[1]Some notable works employed the input-output framework are Hirsch [10], Garnick [4, 5], Bourque [2], Morrison [13], Czamanski and Malizia [3], and Isard and Czamanski [11].

[2]For example, those described by Steinnes and Fisher [19], Hill [9], and Glickman [6, 7]. Glickman's paper [7] contains a thorough bibliography of published works on regional econometric modeling.

[3]Isard and Czamanski [11] have initially shown that the aggregate multipliers derived from economic base models and input-output models are of the same order of magnitude for each number of regions. Subsequently, Billings [1] has shown the mathematical identity of the multipliers derived from the economic base model and the input-output model and Romanoff [17] has extended his results to show that the economic base model is a very special case of input-output analysis. No doubt, the major advantage of the input-output approach is that it delineates the interdependence of all sectors of the economy. In contrast, the most severe limitation of the economic base study is its inability to bring out these interindustry relations. However, this limitation is least severe when dealing with small regions whose interindustry effects are generally small.

[4]In a strict sense, the traditional economic base multiplier is an extended application of the foreign trade multiplier analysis. The scope

of base multiplier analysis is limited to an examination of the total income (or employment) effects of changes in export activities at the regional rather than national level. Recently, economic base models have been, however, defined more broadly than is customary to include, in addition to exports, a local investment sector and other variables affecting local income and employment. In this way, the simple Keynesian multiplier is incorporated into the regional multiplier model. A detailed discussion of the problem appears in Park [15].

[5]Some representative works are Hildebrand and Mace [8], Thompson [20], Sasaki [18], and Weiss and Gooding [21].

[6]See Phouts [16] for a compendium of articles related to these topics.

[7]For lack of more detailed regional income accounts than exist at present, most of regional multipliers were estimated using time series data on employment. Obviously, if available, income data would be preferable to employment data, since income data would reflect interindustry differences in wages and productivity, and could also explicitly account for interregional money and income flows. Furthermore, the employment multiplier could be derived from income-employment relationships.

[8]For example, autonomous investment in the region, import substitution, and changes in such factors as spending habits, population, and tax or transfer policies of external fiscal units.

[9]It remains highly uncertain whether these export coefficients are reasonably stable over time. This question of stability may be used as another justification for the use of errors-in-variables model in which estimation errors of export coefficients resulting from their shifts over time are also treated as measurement errors.

[10]With the additional assumption that the error v is normal, the maximum likelihood estimation solution used by Johnston [12, pp. 152-154] produces exactly the same result as [17].

REFERENCES

[1] Billings, R. B., "The Mathematical Indentify of the Multipliers Derived from the Economic Base Model and the Input-Output Model," *Journal of Regional Science,* 9 (1969), pp. 471-473.

[2] Bourque, P. J., *Income Multipliers for the Washington Economy,* Center for Urban and Regional Research, Reprint No. 14, University of Washington, Seattle, 1969.

[3] Czamanski, S. and E. E. Malizia, "Applicability and Limitations in the use of National Input-Output Tables for Regional Studies," *Papers, Regional Science Association,* 23 (1969), pp. 65-77.

[4] Garnick, D. H., "Disaggregated Basic-Service Models and Regional Input-Output Models in Multiregional Projections," *Journal of Regional Science,* 9 (1969), pp. 87-100.

[5] ______, "Differential Regional Multiplier Models," *Journal of Regional Science,* 10 (1970), pp. 35-47.

[6] Glickman, N., "An Econometric Forecasting Model for the Philadelphia Region," *Journal of Regional Science,* 11 (1971), pp. 15-32.

[7] ______, "Son of the Specification of Regional Econometric Models," *Papers of the Regional Science Association,* 32 (1974), pp. 155-178.

[8] Hildebrand, C., and A. Mace, Jr., "The Employment Multiplier in an Expanding Industrial Market: Los Angeles County, 1940-47," *Review of Economics and Statics,* 41 (1959), pp. 360-369.

[9] Hill, D., "A Growth Allocation Model for the Boston Region," *Journal of the American Institute of Planners,* 31 (1965), pp. 111-120.

[10] Hirsch, W., "Interindustry Relations of A Metropolitan Area," *Review of Economics and Statistics,* 41 (1959), pp. 360-369.

[11] Isard, W. and S. Czamanski, "Techniques for Estimating Local and Regional Multiplier Effects of Changes in the Level of Major Government Programs," *Papers, Peace Research Society* (International), 3 (1965), pp. 19-46.

[12] Johnston, J., *Econometric Methods,* New York: McGraw-Hill, 1963.

[13] Morrison, W. I., "The Development of An Urban Interindustry Model (3): Input-Output *Multipliers," Environment and Planning,* 5 (1973), pp. 545-554.

[14] Park, S. H., "Least Squares Estimates of the Regional Employment Multipliers: An Appraisal," *Journal of Regional Science,* 10 (1970), pp. 365-374.

[15] ______, "The Economic Base Identification: An Appraisal," *Land Economics,* 61 (1965), pp. 382-386.

[16] Phouts, R. W. (ed.), *The Techniques of Urban Economic Analysis,* West Trenton, New Jersey: Chandler-Davis Publishing Co., 1960.

[17] Romanoff, E., "The Economic Base Model: A Very Special Case of Input-Output Analysis," *Journal of Regional Science,* 14 (1974), pp. 121-129.

[18] Sasaki, K., "Expenditures and the Employment Multiplier in Hawaii," *Review of Economics and Statistics,* 44 (1963), pp. 298-304.

[19] Steinnes, D. N., and W. D. Fisher, "An Econometric Model of Intraurban Location," *Journal of Regional Science,* 14 (1974), pp. 65-80.

[20] Thompson, G. E., "An Investigation of the Local Employment Multiplier," *Review of Economics and Statistics,* 41 (1959), pp. 61-67.

[21] Weiss, S. J. and E. C. Gooding, "Estimation of Differential Employment Multipliers in A Small Regional Economy," *Land Economics,* 44 (1968), pp. 235-244.