Leontief's Input-Output Representation of Regression Coefficients*

Eric Iksoon Im [i]

College of Business and Economics, University of Hawai'i at Hilo

Sang-Hyop Lee [ii]

Department of Economics, University of Hawai'i at Mānoa

Tam Bang Vu [iii]

College of Business and Economics, University of Hawai'i at Hilo

Hyun Kyung Kim [iv]

Department of Economics, University of Hawai'i at Mānoa & Statistics Korea

Abstract

This paper formally establishes tighter-than-known connectivity between Least Squares (LS) estimators and Leontief's input-output (I/O) model. First, this study shows that the nexus between partial and total derivatives of the general function of a given number of independent variables is identical to the one between LS estimators of regression coefficients. Second, this study presents that the nexus between the coefficients is exactly represented by Leontief's I/O model. Third, this study shows that the Leontief matrix therefrom is positive-definite as long as regressor collinearity is less than perfect, satisfying the Hawkins-Simon condition. Finally, this study derives Cramer/Leontief alternative to Frisch/Waugh/Lovell theorem.

---

* This paper is dedicated to the memory of Eric Im, an excellent econometrician, mathematician, good teacher, and dedicated father.

i) First Author; Professor, College of Business and Economics, University of Hawaii at Hilo, Hilo, HI 96720, USA

ii) Co-Author; Professor, Department of Economics, University of Hawaii at Mānoa, Honolulu, HI 96822, USA

iii) Co-Author; Professor, College of Business and Economics, University of Hawaii at Hilo, Hilo, HI 96720, USA

iv) Corresponding Author; 1) PhD student, Department of Economics, University of Hawaii at Mānoa, Saunders Hall 542, 2424 Maile Way, Honolulu, HI 96822, USA. 2) Deputy director, Statistics Korea. Email: hkkim@hawaii.edu

First received, November 25, 2021; Revision received, February 14, 2022; Accepted, February 24, 2022.

Keywords: Cramer/Leontief theorem; Frisch/Waugh/Lovell theorem; Hawkins-Simon condition; Leontief's input-output (I/O) model; Shadow coefficients

JEL Classification: C67, C01, C19

1 Introduction

This study formally establishes tighter-than-known connectivity between Least Squares (LS) estimators and Leontief's input-output (I/O) model. The I/O methodology proposed by Leontief (1936) is used to show the net effects of final demand on the output, which means the I/O matrix shows the marginal effects. (Massón-Guerra, 2008) We note this implication in terms of the connectivity with the estimator. Even though this implication has been understood, there has not been a formal discussion that establishes the theoretical connection between the two separate approaches in previous literature. We establish this connection more theoretically in this study.

First, we show that the nexus between partial and total derivatives of the general function of a given number of independent variables is identical to the one between LS estimators of regression coefficients. Second, we present that the nexus between the coefficients is exactly represented by Leontief's I/O model (Leontief, 1951). The multiple regression coefficients and their simple counterparts correspond to the output vector and primary input vector, respectively. Third, we show that the Leontief matrix therefrom is positive definite as long as regressor collinearity is less than perfect, satisfying the so-called Hawkins-Simon condition (Hawkins and Simon, 1949). Finally, we derive Cramer/Leontief theorem alternative to Frisch/Waugh/Lovell theorem.

The remainder of the paper is structured as follows. Section 2 proves the key points listed above. Section 3 presents our results summarizing the substantiated points in terms of theorems and corollaries before concluding this study in Section 4.

2 Proof Methodology

2.1 Partial Effect of Independent Variables on Dependent Variable

In the simple case of a differentiable function depending explicitly on multiple independent variables, such as $Y = f(X_1, X_2, \cdots, X_K)$, we do have total derivatives with respect to the variable X_k:

$$\frac{dY}{dX_k} = f_1 \frac{dX_1}{dX_k} + f_2 \frac{dX_2}{dX_k} + \cdots + f_k \frac{dX_k}{dX_k} + \cdots + f_{K-1} \frac{dX_{K-1}}{dX_k} + f_K \frac{dX_K}{dX_k} = f_k + \sum_{l=1}^{K} \frac{dX_l}{dX_k} f_l \tilde{\delta}_{kl} \tag{1}$$

where $f_k \equiv \partial Y / \partial X_k$ and $\tilde{\delta}_{kl} \equiv 1 - \delta_{kl}\ (k, l = 1, 2, \cdots, K)$ with δ_{kl} denoting Kronecker delta.

Rewriting (1) for f_k yields

$$\frac{dY}{dX_k} = f_k + \sum_{l=1}^{K} \frac{dX_l}{dX_k} f_l \tilde{\delta}_{kl} \Leftrightarrow f_k = \frac{dY}{dX_k} - \sum_{l=1}^{K} \frac{dX_l}{dX_k} f_l \tilde{\delta}_{kl} \tag{2}$$

Equation (2) shows that the partial effect of each independent variable on the dependent variable, f_k, is divided into two parts: (1) its total effect on the dependent variable, dY/dX_k, and (2) net of its indirect partial effects on the dependent variable through other independent variables on the dependent variable, $\sum_{l=1}^{K} (dX_l / dX_k) f_l \tilde{\delta}_{kl}$.

2.2 Rudimentary Scalar Derivation of Multiple Regression Coefficients

The rudimentary scalar derivation of LS estimators of multiple regression coefficients with no calculus and no matrix is limited almost exclusively to the bivariate case as in Ehrenberg (1983). The algebraic workout is increasingly laborious as the number of regressors increases, therefore normally circumvented by turning to calculus and matrix. However, we push it forward for its unique merit until we arrive at an equation system in exactly the same structure as (2) in three groups of LS estimators of regression coefficients: simple regression coefficients, multiple regression coefficients, and simple regression coefficients between regressors.

Writing the general form of a multiple linear regression model in terms of estimated coefficients and residuals,

$$Y_i = b_0 + X_{1i}b_1 + \cdots + X_{Ki}b_K + e_i = b_0 + \sum_{k=1}^{K} X_{ki}b_k + e_i \tag{3}$$

$$i = 1, 2, \cdots, n; \quad k = 1, 2, \cdots, K; \quad n > K$$

where Y_i is an observed regress and, X_{ki} is an observed *k-th* regressor, b_0 is an estimated intercept, b_k is an estimated *k-th* (slope) coefficient, and e_i is an estimated residual.

Also, when centered observations of Y_i and X_{ki} with their respective sample means denoted by $\bar{Y}$ and $\bar{X}_k$ are defined as $y_i = Y_i - \bar{Y}$ and $x_{ki} = X_{ki} - \bar{X}_k$, respectively, the sample mean of equation (3) can be rewritten as $\bar{Y} = b_0 + \sum_{k=1}^{K} \bar{X}_k b_k + \bar{e}$.

And then, in light of (3), we can show that the residual sum of squares (RSS) can be calculated as[1]

$$\begin{aligned} RSS &= \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left[Y_i - \left(b_0 + \sum_{k=1}^{K} X_{ki}b_k\right)\right]^2 \\ &= \sum_{i=1}^{n} \left[\left(y_i - \sum_{k=1}^{K} x_{ki}b_k\right) + \bar{e}\right]^2 \\ &= \sum_{i=1}^{n} \left(y_i - \sum_{k=1}^{K} x_{ki}b_k\right)^2 + n\bar{e}^2 \\ &= \sum_{i=1}^{n} \left[y_i - \left(\sum_{l=1}^{K} x_{li}b_l\tilde{\delta}_{kl} + x_{ki}b_k\right)\right]^2 + n\bar{e}^2 \\ &= \sum_{i=1}^{n} \left[\left(y_i - \sum_{l=1}^{K} x_{li}b_l\tilde{\delta}_{kl}\right)^2 \right. \\ &\quad \left. - 2x_{ki}\left(y_i - \sum_{l=1}^{K} x_{li}b_l\tilde{\delta}_{kl}\right)b_k + x_{ki}^2 b_k^2\right] + n\bar{e}^2 \end{aligned} \tag{4}$$

[1] The detailed formula is described in the Appendix.

$$
\begin{aligned}
&= \sum_{i=1}^{n}\left(y_i - \sum_{l=1}^{K} x_{li} b_l \tilde{\delta}_{kl}\right)^2 \\
&\quad + \left[-2\sum_{i=1}^{n} x_{ki}\left(y_i - \sum_{l=1}^{K} x_{li} b_l \tilde{\delta}_{kl}\right)\right] b_k \\
&\quad + \left(\sum_{i=1}^{n} x_{ki}^2\right) b_k^2 + n\bar{e}^2 \\
&= \pi_k + \tau_k b_k + \psi_k b_k^2 + n\bar{e}^2 \\
&= \pi_k - \frac{\tau_k^2}{4\psi_k} + \psi_k\left(b_k + \frac{\tau_k}{2\psi_k}\right)^2 + n\bar{e}^2
\end{aligned}
$$

where $\pi_k = \sum_{i=1}^{n}\left(y_i - \sum_{l=1}^{K} x_{li} b_l \tilde{\delta}_{kl}\right)^2$, $\tau_k = -2\sum_{i=1}^{n} x_{ki}\left(y_i - \sum_{l=1}^{K} x_{li} b_l \tilde{\delta}_{kl}\right)$, and $\psi_k = \sum_{i=1}^{n} x_{ki}^2$.

Now, it is clear from (4) that RSS reaches its minimum, $\pi_k - \left(\tau_k^2/4\psi_k\right)$, when

$$
b_k = -\frac{\tau_k}{2\psi_k};\ \bar{e} = 0 \tag{5}
$$

because $\psi_k = \sum_{i=1}^{n} x_{ki}^2 > 0$.

Then, substituting τ_k and ψ_k implicitly defined in (4) into b_k in (5) yields

$$
\begin{aligned}
b_k = -\frac{\tau_k}{2\psi_k} &= \frac{2\sum_{i=1}^{n} x_{ki}\left(y_i - \sum_{l=1}^{K} x_{li} b_l \tilde{\delta}_{kl}\right)}{2\sum_{i=1}^{n} x_{ki}^2} \\
&= \frac{\sum_{i=1}^{n} x_{ki} y_i}{\sum_{i=1}^{n} x_{ki}^2} - \frac{\sum_{l=1}^{K}\sum_{i=1}^{n} x_{ki} x_{li} b_l \tilde{\delta}_{kl}}{\sum_{i=1}^{n} x_{ki}^2} \\
&= \frac{\sum_{i=1}^{n} x_{ki} y_i}{\sum_{i=1}^{n} x_{ki}^2} - \sum_{l=1}^{K}\left(\frac{\sum_{i=1}^{n} x_{ki} x_{li}}{\sum_{i=1}^{n} x_{ki}^2}\right) b_l \tilde{\delta}_{kl} \\
&\equiv g_k - \sum_{l=1}^{K} c_{kl} b_l \tilde{\delta}_{kl}
\end{aligned} \tag{6}
$$

where both $g_k = \sum_{i=1}^{n} x_{ki} y_i / \sum_{i=1}^{n} x_{ki}^2$ and $c_{kl} = \sum_{i=1}^{n} x_{ki} x_{li} / \sum_{i=1}^{n} x_{ki}^2$ are LS estimators of simple regression coefficients: the former results when Y is regressed on X_k, and the latter results when X_l is regressed on X_k, with the intercept retained in both cases.

It is rather clear from the comparison of (2) and (6) that LS estimators of multiple regression coefficients b_k $(k = 1, 2, \cdots, K)$ in (6) depend on LS estimators of simple regression coefficients g_k in precisely the same way as partial derivatives f_k in (2) depend on total derivatives dY/dX_k, with c_{kl} in (6) playing the same role as dX_l/dX_k in (2). To wit, what b_k and g_k in (6) measure are exactly what f_k (partial impact) and dY/dX_k (total impact) in (2) measure. A couple of points may be noteworthy. First, the correspondence between (2) and (6) is the fact that b_k and g_k in (6) are outcomes of minimizing the sum of squared residuals, whereas f_k and dY/dX_k in (2) are not outcomes of minimization of any sort. Second, the correspondence holds for any given set of independent variables chosen and regardless of the functional form in (2).

2.3 Leontief's I/O Representation of LS Estimators

The primary purpose of Section 2.2 was to obtain LS estimators of multiple regression coefficients in the same form as (2) but in terms of LS estimators of regression coefficients, which are indeed the solutions for multiple regression coefficients in *implicit* forms. With that accomplished, now the K simultaneous equations in K unknowns easily convert to a matrix equation, as shown below, which is Leontief's I/O model.

By stacking b_k $(k = 1, 2, \cdots, K)$ in (6), we can easily convert the system of K linear equations to a matrix equation:

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{K-1} \\ b_K \end{pmatrix} = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_{K-1} \\ g_K \end{pmatrix} - \begin{bmatrix} 0 & c_{12} & \cdots & c_{1(K-1)} & c_{1K} \\ c_{21} & 0 & \cdots & c_{2(K-1)} & c_{2K} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{(K-1)1} & c_{(K-1)2} & \cdots & 0 & c_{(K-1)K} \\ c_{K1} & c_{K2} & \cdots & c_{K(K-1)} & 0 \end{bmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{K-1} \\ b_K \end{pmatrix} \quad (7)$$

which is equivalently rearranged to

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{K-1} \\ b_K \end{pmatrix} = \left(\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & c_{12} & \cdots & c_{1(K-1)} & c_{1K} \\ c_{21} & 1 & \cdots & c_{2(K-1)} & c_{2K} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{(K-1)1} & c_{(K-1)2} & \cdots & 1 & c_{(K-1)K} \\ c_{K1} & c_{K2} & \cdots & c_{K(K-1)} & 1 \end{bmatrix} \right) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{K-1} \\ b_K \end{pmatrix} + \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_{K-1} \\ g_K \end{pmatrix} \tag{8}$$

Also, both $g_k\left(=\sum_{i=1}^{n} x_{ki}y_i/\sum_{i=1}^{n} x_{ki}^2\right)$ and $c_{kl}\left(=\sum_{i=1}^{n} x_{ki}x_{li}/\sum_{i=1}^{n} x_{ki}^2\right)$ can be rewritten in vectors for the notational economy as $g_k = x_k'y/x_k'x_k$; $c_{kl} = x_k'x_l/x_k'x_k$ $(k, l = 1, 2, \cdots, K)$, and we can rewrite (8) compactly in matrices,

$$b = (I_K - C)b + g \equiv Ab + g \tag{9}$$

which is, in essence, the same matrix equation as the original Leontief's I/O model with b corresponding to output vector, A to input coefficient matrix, and g to the primary input vector.

Then, $L \equiv I_K - A$ is the so-called Leontief matrix, and we can rewrite (9) even more compactly in terms of $L = C = [x_k'x_l/x_k'x_k]$ as

$$Lb = g \Longrightarrow b = L^{-1}g \tag{10}$$

where $C = [x_k'x_l/x_k'x_k]$ is a square matrix whose *(k, l)-th* element is the LS estimator of regression coefficient that results when the regressor x_l is regressed on another regressor x_k, and so is L since $L = C$.[2]

[2] Note that L is not symmetric in general since $c_{kl} \neq c_{lk}$: $c_{kl} = x_k'x_l/x_k'x_k \neq x_l'x_k/x_l'x_l = c_{lk}$

2.4 Positive Definite Leontief Matrix and Hawkins-Simon Condition

It is worth noting that Leontief's I/O model in our context is a particular case of the original Leontief's model in the sense that it consists of LS estimators of simple regression coefficients between regressors themselves. Furthermore, the Leontief matrix is positive definite, which is not the case in general with the original context of Leontief's I/O model, satisfying the so-called Hawkins-Simon condition (1949). Thus, we can state that if g is *non-negative* when off-diagonals of the Leontief matrix are all non-positive, then b is *non-negative*.[3]

To facilitate the proof for $L >_L 0$, where "$>_L$" denotes Löwner ordering, we restructure the Leontief matrix equivalently. Since $[\![\sqrt{x_l' x_l}]\!] = [\![\sqrt{x_k' x_k}]\!] = [\![\sqrt{x_k' x_k}]\!]'$ for $k, l = 1, 2, \cdots, K$ when $[\![\cdot]\!]$ denotes diagonal matrix with the indexed argument representing the diagonal elements, we obtain[4]:

$$
\begin{aligned}
L = C = \left[\frac{x_k' x_l}{x_k' x_k}\right] &= [\![\sqrt{x_k' x_k}]\!]\left([\![\frac{1}{x_k' x_k}]\!]\left[\frac{x_k' x_l}{\sqrt{x_k' x_k}\sqrt{x_l' x_l}}\right]\right)[\![\sqrt{x_l' x_l}]\!] \\
&= [\![\sqrt{x_k' x_k}]\!]\left([\![\frac{1}{x_k' x_k}]\!]\left[\frac{x_k' x_l}{\sqrt{x_k' x_k}\sqrt{x_l' x_l}}\right]\right)[\![\sqrt{x_k' x_k}]\!] \\
&= [\![\sqrt{x_k' x_k}]\!]\left([\![\frac{1}{x_k' x_k}]\!]\begin{bmatrix} 1 & r_{12} & \cdots & r_{1(K-1)} & r_{1K} \\ r_{21} & 1 & \cdots & r_{2(K-1)} & r_{2K} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ r_{(K-1)1} & r_{(K-1)2} & \cdots & 1 & r_{(K-1)K} \\ r_{K1} & r_{K2} & \cdots & r_{K(K-1)} & 1 \end{bmatrix}\right)[\![\sqrt{x_k' x_k}]\!]' \\
&\equiv D^{1/2}(D^{-1}R)D^{1/2'}
\end{aligned}
\tag{11}
$$

In (11), $D = [\![x_k' x_k]\!]$ is a positive definite diagonal matrix. In addition, since we already know that $[\![1/\sqrt{x_l' x_l}]\!] = [\![1/\sqrt{x_k' x_k}]\!]$ for $k, l = 1, 2, \cdots, K$, we obtain:

[3] This statement also holds for *non-positive* (in place of *non-negative*) as well in light of (10): $Lb = g \Leftrightarrow L(-b) = (-g)$.

[4] If $r_{kl} = 0$ in (11) for all k, $l\ (= 1, 2, \cdots, K; k \neq l)$, then $R = L = I_K$.

$$R \equiv \left[\frac{x_k' x_l}{\sqrt{x_k' x_k}\sqrt{x_l' x_l}}\right] = \left(x\left[\!\left[\frac{1}{\sqrt{x_k' x_k}}\right]\!\right]\right)'\left(x\left[\!\left[\frac{1}{\sqrt{x_l' x_l}}\right]\!\right]\right)$$
$$= \left(x\left[\!\left[\frac{1}{\sqrt{x_k' x_k}}\right]\!\right]\right)'\left(x\left[\!\left[\frac{1}{\sqrt{x_k' x_k}}\right]\!\right]\right) \geq_L 0 \tag{12}$$

where x $(n \times K)$ denotes centered observation matrix for K regressors whereas x_k $(n \times 1)$ and x_l $(n \times 1)$ are its *k-th* and *l-th* column vectors, respectively. Furthermore, under the assumption of $r_{kl}^2 < 1$ $(k \neq l)$ as usual in practice, R becomes positive definite: $R >_L 0$.

Given D and R, both symmetric and positive definite, eigenvalues of $D^{-1}R$ are identical with those of $D^{-1/2}RD^{-1/2}$ in light of (18.79) in Puntanen et al. (2011), i.e. $\lambda_k(D^{-1}R) = \lambda_k\left(D^{-1/2}RD^{-1/2}\right)$ for $k = 1, 2, \cdots, K$ where $\lambda_k(\cdot)$ denotes the *k-th* largest eigenvalue of the argument matrix.

In addition, from $\lambda_k(D^{-1}R) = \lambda_k(D^{-1/2}RD^{-1/2})$, we readily follow:

$$D^{-1/2}RD^{-1/2} = (D^{-1/2}R^{1/2})(D^{-1/2}R^{1/2})' >_L 0$$
$$\Leftrightarrow \lambda_k(D^{-1/2}RD^{-1/2}) > 0 \Leftrightarrow \lambda_k(D^{-1}R) > 0 \Leftrightarrow D^{-1}R >_L 0 \tag{13}$$

Also, from $L = D^{1/2}(D^{-1}R)D^{1/2\prime}$ for $k = 1, 2, \cdots, K$ in (11) in light of (13),

$$L >_L 0 \tag{14}$$

which implies that all principal minors of L, $|\tilde{L}_k|$, are positive, and all eigenvalues are positive:

$$|\tilde{L}_k| > 0 \Leftrightarrow \lambda_k(L) > 0 \quad (k = 1, 2, \cdots, K) \tag{15}$$

Finally, equation (15) confirms that the Hawkins-Simon condition is satisfied. Note that if regressors are all independent of each other, $r_{kl} = 0$ for all $k, l = 1, 2, \cdots, K$ $(k \neq l)$, then Leontief matrix is reduced to identity matrix: $L = I_K$ so that $b = g$, or $b_k = g_k$.

2.5 Cramer/Leontief Theorem Alternative to Frisch/Waugh/Lovell Theorem

Each element in the multiple regression coefficient vector is represented by

$$b_k = \left(x_k' M_{z_k} x_k\right)^{-1} x_k' M_{z_k} y \tag{16}$$

where $M_{z_k} = I_{K-1} - z_k(z_k' z_k)^{-1} z_k'$ in which $z_k \ (n \times (K-1))$ denotes $x \ (n \times K)$ with $x_k \ (n \times 1)$ column eliminated, for $k = 1, 2, \cdots, K$.

This rather concise representation is credited to the papers by Frisch and Waugh (1933) and Lovell (1963), and stated in their honor as Frisch/Waugh/Lovell Theorem in Lovell (2008).

An equally concise analytical expression in terms of Leontief matrix L is readily obtained from (10) by Cramer's rule.

$$b_k = \left|L_{(k,g)}\right| / |L| \tag{17}$$

where $L_{(k,g)} \ (K \times K)$ denotes the Leontief matrix with its *k-th* column replaced by LS estimator of simple regression coefficient vector g in (9), for $k = 1, 2, \cdots, K$.[5]

Since both alternative representations of b_k look quite dissimilar to each other, a formal proof of their equivalence may have its own merit. To facilitate a short-cut proof of their equivalence, define a diagonal matrix of order K:

$$V \equiv \begin{bmatrix} x_1' x_1 & 0 & \cdots & 0 & 0 \\ 0 & x_2' x_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & x_{K-1}' x_{K-1} & 0 \\ 0 & 0 & \cdots & 0 & x_K' x_K \end{bmatrix} \tag{18}$$

Then, the LS estimator of multiple regression coefficient vector in (10) can be expressed alternatively as

[5] Note that Leontief matrix L has a virtue of having all of its components represented by a simple general expression: $L = [x_k' x_l / x_k' x_k] \ (k, l = 1, 2, \cdots, K)$.

$$
\begin{aligned}
b &= L^{-1}g = (VL)^{-1}Vg \\
&= \begin{bmatrix} x_1'x_1 & x_1'x_2 & \cdots & x_1'x_{K-1} & x_1'x_K \\ x_2'x_1 & x_2'x_2 & \cdots & x_2'x_{K-1} & x_2'x_K \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{K-1}'x_1 & x_{K-1}'x_2 & \cdots & x_{K-1}'x_{K-1} & x_{K-1}'x_K \\ x_K'x_1 & x_K'x_2 & \cdots & x_K'x_{K-1} & x_K'x_K \end{bmatrix}^{-1} \begin{pmatrix} x_1'y \\ x_2'y \\ \vdots \\ x_{K-1}'y \\ x_K'y \end{pmatrix} \\
&= [x'x]^{-1}x'y
\end{aligned} \tag{19}
$$

Then, partitioning $x\ (n \times K)$ into $x = [x_k \quad z_k]$ where $x_k\ (n \times 1)$ denotes *k-th* column vector in x whereas $z_k\ \big(n \times (K-1)\big)$ denotes x with a column vector x_k excluded, we can re-express b in (19) as

$$
b = \begin{pmatrix} b_k \\ b_z \end{pmatrix} = \begin{bmatrix} x_k'x_k & x_k'z_k \\ z_k'x_k & z_k'z_k \end{bmatrix}^{-1} \begin{pmatrix} x_k'y \\ z_k'y \end{pmatrix} \tag{20}
$$

Applying Cramer's rule to (20), for $k = 1, 2, \cdots, K$, yields

$$
\begin{aligned}
b_k = \begin{vmatrix} x_k'y & x_k'z_k \\ z_k'y & z_k'z_k \end{vmatrix} \Big/ \begin{vmatrix} x_k'x_k & x_k'z_k \\ z_k'x_k & z_k'z_k \end{vmatrix} &= \frac{x_k'y - x_k'z_k(z_k'z_k)^{-1}z_k'y}{x_k'x_k - x_k'z_k(z_k'z_k)^{-1}z_k'x_k} \\
&= \frac{x_k'[I_n - z_k(z_k'z_k)^{-1}z_k']y}{x_k'[I_n - z_k(z_k'z_k)^{-1}z_k']x_k} \\
&= \left(x_k'M_{z_k}x_k\right)^{-1}x_k'M_{z_k}y
\end{aligned} \tag{21}
$$

in which the second equality is based on the general formula for determinants of partitioned matrices and M_{z_k} is as defined earlier in (16).

3. Results and Discussion: Summary in Theorems and Corollaries

Key theorems and corollaries in this section heretofore may be summarized as follows without repeating the proofs already provided and notational definitions already made where appropriate.

Theorem 1 (Correspondence):

Given the number of regressors $K(\geq 1)$ in LS estimation, let b_k denote LS estimator of k-th multiple regression coefficient, g_k the simple regression counterpart, and c_{kl} LS estimator of simple regression coefficients between the regressors, and $\tilde{\delta}_{kl} \equiv 1 - \delta_{kl}$ with δ_{kl} denoting Kronecker delta. Then,

$$b_k = g_k - \sum\nolimits_{l=1}^{K} c_{kl} b_l \tilde{\delta}_{kl} \quad (k, l = 1, 2, \cdots, K)$$

The structure of which is exactly the same as the one relating partial derives (f_k), total derivatives (dY/dX_k), and total derivatives between independent variables (dX_l/dX_k) of the general function $f(X_1, X_2, \cdots, X_k)$:

$$f_k = \frac{dY}{dX_k} - \sum\nolimits_{l=1}^{K} \frac{dX_l}{dX_k} f_l \tilde{\delta}_{kl} \quad (k, l = 1, 2, \cdots, K)$$

Theorem 2 (Leontief's I/O Representation):

Let $b(K \times 1)$ and $g(K \times 1)$ define LS estimators of multiple and simple regression coefficient vectors, respectively. Then, the multiple regression coefficient vectors can be represented concisely as Leontief's I/O model:

$$Lb = g \quad \Leftrightarrow \quad b = L^{-1} g$$

where Leontief matrix is defined as $L = [x_k' x_l / x_k' x_k] \ (k, l = 1, 2, \cdots, K)$.

Theorem 3 (Positive Definite Leontief Matrix):

Under the assumption of $r_{kl}^2 < 1 \ (k \neq l)$, Leontief matrix in the context of LS estimators:

$$L = [x_k' x_l / x_k' x_k] \quad (k, l = 1, 2, \cdots, K)$$

is positive definite, thereby satisfying the Hawkins-Simon condition.

Corollary 1 (Cramer/Leontief Synthesis):

LS estimators of individual multiple regression coefficients can be expressed as

$$b_k = \left|L_{(k,g)}\right|/|L| \quad (k = 1, 2, \cdots, K)$$

where $L_{(k,g)}$ denotes the Leontief matrix $L = [x_k'x_l/x_k'x_k]$ with its *k-th* column replaced by LS estimator of simple regression coefficient vector g, which is equivalent to

$$b_k = \left(x_k' M_{z_k} x_k\right)^{-1} x_k' M_{z_k} y$$

as in Frisch/Waugh/Lovell Theorem.

4. Conclusion

This study used a rudimentary scalar approach with no calculus and no matrix in deriving implicit solutions for LS estimators of multiple regression coefficients that easily convert to a matrix equation. Thus, the obtained matrix equation turned out to be Leontief's I/O model in which LS estimators of multiple regression coefficients and their simple counterparts, respectively, correspond to outputs and primary inputs. This connectivity implies that the net effects of final demand on the output, the average marginal effects, provided by the I/O matrix can be identified by LS estimators of multiple regression coefficients. Subsequently, we showed that the Leontief matrix is positive definite unless regressors are perfectly linear. As a consequence, the Hawkins-Simon condition is satisfied, which implies that signs of simple regression coefficients are all non-negative (or non-positive) with all non-positive off-diagonals in the Leontief matrix, the signs of their multiple regression counterparts are also non-negative (or non-positive). In addition, we derived the Cramer/Leontief version of analytical expression of LS estimators of multiple regression coefficients as an alternative to that of the Frisch/Waugh/Lovell theorem, which may have its own analytical virtue.

In sum, this study formally establishes tight connectivity between LS estimators and Leontief's I/O model that has never been elaborated elsewhere in previous literature. We expect that this connectivity will help to understand Leontief's I/O model and further increase accessibility to this model. In addition, this study will contribute to broadening the scope of understanding of the Leontief matrix by identifying the condition that satisfies the Hawkins-Simon condition, as well as by presenting the Cramer/Leontief Theorem as an alternative to the Frisch/Waugh/Lovell Theorem.

References

Ehrenberg, A.S.C., "Deriving the least squares regression equation," *The American Statistician* 37(3), 1983, 232-232.

Frisch, R. and F.V. Waugh, "Partial time regressions as compared with individual trends," *Econometrica: Journal of the Econometric Society* 1(4), 1933, 387-401.

Hawkins, D. and H.A. Simon, "Note: some conditions of macroeconomic stability," *Econometrica, Journal of the Econometric Society* 17(3), 1949, 245-248.

Leontief, W.W., "Quantitative Input and Output relations in the Economic System of the United States," *The Review of Economic Statistics* 18(3), 1936, 105–125.

Leontief, W.W., *The structure of American economy, 1919-1939: an empirical application of equilibrium analysis* (No. HC106. 3 L3945 1951), 2nd ed. Oxford University Press, 1951.

Massón-Guerra, J.L., "The Multipliers and Key Sectors of Entrepreneurship Spillover: An input-output approach (October 9, 2008)," Available at SSRN 1281429: https://ssrn.com/abstract=1281429 or http://dx.doi.org/10.2139/ssrn.1281429, 2008.

Lovell, M.C., "Seasonal adjustment of economic time series and multiple regression analysis," *Journal of the American Statistical Association* 58(304), 1963, 993-1010.

Lovell, M. C., "A simple proof of the FWL theorem," *The Journal of Economic Education* 39(1), 2008, 88-91.

Puntanen, S., Styan, G.P., and J. Isotalo, *Matrix tricks for linear statistical models: our personal top twenty*. Springer Science & Business Media, 2011.

Appendix

Details of Eq. (4)

$$
\begin{aligned}
RSS = \sum_{i=1}^{n} e_i^2 &= \sum_{i=1}^{n}\left[Y_i-\left(b_0+\sum_{k=1}^{K} X_{ki}b_k\right)\right]^2 \\
&= \sum_{i=1}^{n}\left[(\bar{Y}+y_i)-\left(b_0+\sum_{k=1}^{K}(\bar{X}_k+x_{ki})b_k\right)\right]^2 \\
&= \sum_{i=1}^{n}\left[\left(y_i-\sum_{k=1}^{K} x_{ki}b_k\right)+\left(\bar{Y}-b_0-\sum_{k=1}^{K}\bar{X}_k b_k\right)\right]^2 \\
&= \sum_{i=1}^{n}\left[\left(y_i-\sum_{k=1}^{K} x_{ki}b_k\right)+\bar{e}\right]^2 \\
&= \sum_{i=1}^{n}\left[\left(y_i-\sum_{k=1}^{K} x_{ki}b_k\right)^2+2\left(y_i-\sum_{k=1}^{K} x_{ki}b_k\right)\bar{e}+\bar{e}^2\right] \\
&= \sum_{i=1}^{n}\left(y_i-\sum_{k=1}^{K} x_{ki}b_k\right)^2+2\sum_{i=1}^{n}\left[\bar{e}\left(y_i-\sum_{k=1}^{K} x_{ki}b_k\right)\right] \\
&\quad +\sum_{i=1}^{n}\bar{e}^2 = \sum_{i=1}^{n}\left(y_i-\sum_{k=1}^{K} x_{ki}b_k\right)^2+n\bar{e}^2 \\
&= \sum_{i=1}^{n}\left(y_i-\sum_{l=1}^{K} x_{li}b_l\right)^2+n\bar{e}^2 \\
&= \sum_{i=1}^{n}\left[y_i-\left\{\left(\sum_{l=1}^{K} x_{li}b_l-x_{ki}b_k\right)+x_{ki}b_k\right\}\right]^2+n\bar{e}^2 \\
&= \sum_{i=1}^{n}\left[y_i-\left(\sum_{l=1}^{K} x_{li}b_l\tilde{\delta}_{kl}+x_{ki}b_k\right)\right]^2+n\bar{e}^2 \\
&= \sum_{i=1}^{n}\left[\left(y_i-\sum_{l=1}^{K} x_{li}b_l\tilde{\delta}_{kl}\right)^2-2x_{ki}\left(y_i-\sum_{l=1}^{K} x_{li}b_l\tilde{\delta}_{kl}\right)b_k \right.\\
&\quad \left. +x_{ki}^2b_k^2\right]+n\bar{e}^2 \\
&= \sum_{i=1}^{n}\left(y_i-\sum_{l=1}^{K} x_{li}b_l\tilde{\delta}_{kl}\right)^2 \\
&\quad +\left[-2\sum_{i=1}^{n} x_{ki}\left(y_i-\sum_{l=1}^{K} x_{li}b_l\tilde{\delta}_{kl}\right)\right]b_k+\left(\sum_{i=1}^{n} x_{ki}^2\right)b_k^2 \\
&\quad +n\bar{e}^2 = \pi_k+\tau_k b_k+\psi_k b_k^2+n\bar{e}^2 \\
&= \pi_k-\frac{\tau_k^2}{4\psi_k}+\psi_k\left(b_k+\frac{\tau_k}{2\psi_k}\right)^2+n\bar{e}^2 \geq \pi_k-\frac{\tau_k^2}{4\psi_k}\ (\because \psi_k>0)
\end{aligned}
$$