SFC ディスカッションペーパー
SFC-DP 2013-002

On the Interpolation of Asian Input Output Data for Intermediate Years by the Use of Model: A Non-Survey Method

小坂弘行

総合政策学部名誉教授

hkosaka@sfc.keio.ac.jp

2013 年 10 月

On the Interpolation of Asian Input Output Data for Intermediate Years by the Use of Model: A Non-Survey Method

Hiroyuki Kosaka
Faculty of Policy Management, Keio University
Endo 5322,Fujisawa,Kanagawa,252-8520,Japan
E-mail: hkosaka@sfc.keio.ac.jp

1. Introduction

Construction of input output tables are usually implemented for every five years; so that, the latest table may be updated more than after five years of the present time. Asian input output table for Asian economies of nine countries plus US by Institute of Developing Economies is also for every five year encompassing 1985 to 2000. On account of the data availability, those of intermediate years have to be interpolated by some ways; one sometimes use non-survey methods for interpolation, in which the methods frequently have arbitrariness. Economic analysis using the latest data of the past more than five years may be inopportune.

The most preferable approach is believed to be done in the construction of domestic IO table every year in Korea; IT technology overcomes difficulties encountered in constructing IO table. Korean IO tables, published by the Bank of Korea (the Korean central bank), is at the time of November 16, 2012; IO table of 2010 has been published at April 31(2012), and that of 2009 at April 29(2011) ; recently, they are published at least two years later. Historically Korean IO tables, therefore, have been published; tables of 1970, 1975, 1980, 1985, 1990, 1995, 2000, 2003, 2005, 2006, 2007, 2008, 2009, 2010 are available; since 2005, they are published every year. Future preferable direction for the task of constructing IO tables should follow Korean.

Another way of coping with this difficulty is to use CGE model of utilizing the latest spot time data; IO analysis of Leontief type also is of the same kind. But, in contrast, econometric analysis uses multi-period data for determining many parameters of behavioral equations; this short note is devoted for interpolating data by the use of model information.

2. Model Estimation and Interpolation of Intermediate Year by an Bayesian Approach

Assume that we have production-related data of ten countries of Asian economies for 1985, 1990, 1995 and 2000; then, we must interpolate for 1986, 1987, 1988, 1989, 1991, 1992, 1993, 1994, 1996, 1997, 1998, 1999.

2.1 Labor Demand Function and Interpolating Production

Fortunately employment data is available ever year; yet, those of production and price are every five years. On the base of stability of estimated labor demand function depicting labor demand (employment) in close association with production, we interpolate production data. Likelihood function of labor demand of j-th industry of k-th country, in Multi-Country/Multi-Sector System(hereafter MCMS in abbreviation), is:

$$L^k_{xx,j} = \prod_{t=1985}^{2000} f^k_j(L^k_j(t), \mu^k_j(t), \sigma^k_j) \quad (2.01)$$

L^k_j: Labor Demand $\mu^k_j(t)$: Theoretical Expression of Labor Demand Equation

σ^k_j: Stochastic Term of Labor Demand of j-th Industry

of k-th Country(unknown parameter)

Idea of determining parameters by maximizing likelihood function first arisen from R.A.Fischer(see e.g. W.H.Greene(2007)). In estimating parameters of labor demand equation we take logarithm of likelihood (2.01); at this time we add another term for interpolating production of intermediate years.

$$\tilde{L}^k_{xx,j} \propto -\frac{1}{(\sigma^k_j)^2} - \sum_{t\in T_4 \cup T_{12}} \frac{(\log L^k_j - \log \mu^k_j)^2}{2(\sigma^k_j)^2} - \frac{1}{2W(\sigma^k_j)^2} \sum_{t\in T_{12}} (\log X^k_j - \log X^{k*}_j)^2 \quad (2.02)$$

W weight T_4 = (1985,1990,1995,2000)

T_{12} = (1986,1987,1988,1989,1991,1992,1993,1994,1996,1997,1998,1999)

where X^k_j is interpolation at current iteration, and X^{k*}_j is interpolation at the latest iteration; the initial value of interpolation is made by straight line fitting (afterwards it is altered). The first term of (2.02) comes from log likelihood of labor demand equation; and, the second term is expressing quadratic loss of X^k_j deviating from X^{k*}_j.

Now we maximize log likelihood of (2.02) to obtain ML estimates of parameters of labor demand equation, and to interpolate production data of intermediate years simultaneously.

a) The First Step

First, to determine unknown parameters of labor demand equation by partial derivative with respect to parameters; this step of obtaining ML estimates is the same with OLS with exception of σ^k_j.

b) The Second Step

Next, to interpolate production data of intermediate years by partial derivative with respect to production (e.g., for the year 1991).

$$\frac{\partial \tilde{L}^k_{xx,j}}{\partial X^k_j} = \frac{1}{(\sigma^k_j)^2}\left(\log L^k_j - \log \mu^k_j\right)\frac{\partial \log \mu^k_j}{\partial X^k_j} - \frac{1}{W(\sigma^k_j)^2}(\log X^k_j - \log X^{k*}_j) = 0 \quad (2.03)$$

Then, we have from (2.03):

$$\left(\log L_j^k - \log \mu_j^k\right)\frac{\partial \log \mu_j^k}{\partial X_j^k} - \frac{1}{W}(\log X_j^k - \log X_j^{k*}) = 0. \qquad (2.04)$$

Re-arranging terms give us the following:

$$\log X_j^k = \log X_j^{k*} + W\left(\log L_j^k - \log \mu_j^k\right)\frac{\partial \log \mu_j^k}{\partial X_j^k}. \quad \text{for the year 1991} \qquad (2.05)$$

Now we assume labor demand equation μ_j^k is below (see A.5 of Appendix):

$$\mu_j^k = bl_j^k \left(X_j^k\right)^{blx_j^k+1} \exp\left(blt_j^k t\right) \qquad (2.06)$$

Then, the term $\partial\mu_j^k / \partial X_j^k$ is evaluated in labor demand equation:

$$\frac{\partial \mu_j^k}{\partial X_j^k} = \left(blx_j^k + 1\right)\frac{\mu_j^k}{X_j^k}, \qquad (2.07)$$

Finally we have;

$$\frac{\partial \log \mu_j^k}{\partial X_j^k} = \frac{\left(blx_j^k + 1\right)}{X_j^k}. \qquad (2.08)$$

As a result, inserting (2.08) into (2.05), we have the following interpolation formula for T_{12}.

$$\log X_j^k = \log X_j^{k*} + W\left(\log L_j^k - \log \mu_j^k\right)\frac{\left(blx_j^k + 1\right)}{X_j^k} \quad \text{for } T_{12} \qquad (2.09)$$

In the next iteration, by the use of new production data, labor demand equation is to be re-estimated in the first step; then, the new estimated labor demand makes new interpolation of production in (2.09).

2.2 Production and Interpolating Factor Demands

The argument of the section 2.1 gives us interpolation of production X_j^k for the whole period; so that, in the next we come to interpolate factor demands. We assume disturbance of factor demand equation obeys normal distribution with mean $\log \mu_{ij}^{hk}$ and variance $(\sigma_{ij}^{hk})^2$ as in $\log x_{ij}^{hk} = \log \mu_{ij}^{hk} + u_{ij}^{hk}$; then, the density of disturbance

$u_{ij}^{hk} = \log x_{ij}^{hk} - \log \mu_{ij}^{hk} \sim N(0,(\sigma_{ij}^{hk})^2)$ is assumed: [1]

$$f_{ij}^{hkx}(x_{ij}^{hk}, \mu_{ij}^{hk}, \sigma_{ij}^{hk}) = \frac{1}{\sqrt{2\pi}\sigma_{ij}^{hk}} \exp\left\{ -\frac{(\log x_{ij}^{hk} - \log \mu_{ij}^{hk})^2}{2(\sigma_{ij}^{hk})^2} \right\} \tag{2.10}$$

The likelihood of (2.10) is below:

$$L_{x,j}^{k} = \prod_{t=1985}^{2000} \prod_{h=1}^{10} \prod_{i=1}^{6} f_{ij}^{hk}(x_{ij}^{hk}(t), \mu_{ij}^{hk}(t), \sigma_{ij}^{hk}). \tag{2.11}$$

Log likelihood with quadratic loss of factor demands is to be maximized for the interpolation of factor demands:

$$\tilde{L}_{x,j}^{k} \propto -\frac{1}{(\sigma_{ij}^{hk})^2} - \sum_{t \in T_4 \cup T_{12}} \sum_{h=1}^{10} \sum_{i=1}^{6} \frac{(\log x_{ij}^{hk} - \log \mu_{ij}^{hk})^2}{2(\sigma_{ij}^{hk})^2} - \frac{1}{2W(\sigma_{ij}^{hk})^2} \sum_{t \in T_{12}} \sum_{h=1}^{10} \sum_{i=1}^{6} (\log x_{ij}^{hk} - \log x_{ij}^{hk*})^2 \tag{2.12}$$

a) The First Step

The first step determines unknown parameters of factor demand function μ_{ij}^{hk} ; as the parameters enter only in the first term of (2.12), it is easily estimated by the use of sixteen samples of interpolated production data obtained in the section 2.1.

$$\frac{\partial \tilde{L}_{x,j}^{k}}{\partial b\tau} = \sum_{t} \frac{(\log x_{ij}^{hk} - \log \mu_{ij}^{hk})}{(\sigma_{ij}^{hk})^2} \frac{\partial \log \mu_{ij}^{hk}}{\partial b\tau} = 0 \quad \tau = null, x, t\,; i = 1,..,6; h = 1,..,10 \tag{2.13}$$

Then we have the following:

$$\sum_{t} (\log x_{ij}^{hk} - \log \mu_{ij}^{hk}) \frac{\partial \log \mu_{ij}^{hk}}{\partial b\tau} = 0 \qquad \tau = null, x, t \tag{2.14}$$

b) The Second Step

The second step interpolates factor demands for intermediate years(x_{ij}^{hk} $t \in T_{12}$). Factor demand gets involved in the both terms of (2.12):

$$\frac{\partial \tilde{L}_{x,j}^{k}}{\partial x_{ij}^{hk}} = -\frac{(\log x_{ij}^{hk} - \log \mu_{ij}^{hk})}{(\sigma_{ij}^{hk})^2} - \frac{(\log x_{ij}^{hk} - \log x_{ij}^{hk*})}{W(\sigma_{ij}^{hk})^2} = 0. \tag{2.15}$$

Then we have interpolation formula.

$$\log x_{ij}^{hk} = \frac{W \log \mu_{ij}^{hk} + \log x_{ij}^{hk*}}{W+1} \tag{2.16}$$

---

[1] As in the same way of (2.06), theoretical expression of material input of i-th product in the j-th production in the k-th county is (see A.4 of Appendix):

$$\mu_{ij}^{rk} = b_{ij}^{rk} \left(X_j^k\right)^{bx_{ij}^{rk}+1} \exp\left(bt_{ij}^{rk} t\right) + \sum_{i2 \neq i}^{N} b_{i,i2,j}^{r,r,k} \sqrt{p_{i2}^{r,k} / p_i^{rk}} \left(X_j^k\right)^{bx_j^k+1} \exp\left(bt_j^k t\right)$$

As numerical value μ_{ij}^{hk} is obtained in the first step, the formula gives us interpolation of factor demand; using the new interpolation, factor demand equation is re-estimated.

2.3 Cost Function and Interpolating Price

Dynamic formulation of price could be formulated: [2]

$$p_j^k = \alpha + \beta p_{j,-1}^k + \gamma MC_j^k + \nu_j^k = \hat{p}_j^k + \nu_j^k \tag{2.17}$$

Likelihood of price equation is below:

$$L_{p,j}^k = \prod_{t=1985}^{2000} f_j^k(\sigma_{vj}^k). \tag{2.18}$$

Log likelihood with quadratic loss of price is the following:

$$\tilde{L}_{p,j}^k \propto -\frac{1}{(\sigma_{vj}^k)^2} - \sum_{t\in T_4\cup T_{12}} \frac{(p_j^k - (\alpha + \beta p_{j,-1}^k + \gamma MC_j^k))^2}{2(\sigma_{vj}^k)^2} - \frac{W}{2(\sigma_{vj}^k)^2}\sum_{t\in T_{12}}(p_j^k - p_j^{k*})^2 \tag{2.19}$$

The first term of (2.19) comes from log of (2.18); and, the second term is quadratic loss of price deviating from the previous iterated price.

a) The First Step

The first step maximizes (2.19) with respect to the three parameters of price equation (α, β, γ), but in no connection with the second term.

b) The Second Step

The second step maximizes $\tilde{L}_{p,j}^k$ with respect to p_j^k for T_{12}.

$$\frac{\partial \tilde{L}_j^k}{\partial p_j^k} = \frac{1}{(\sigma_{vj}^k)^2}\left(p_j^k - \hat{p}_j^k\right) - \frac{W}{(\sigma_{vj}^k)^2}(p_j^k - p_j^{k*}) = 0 \tag{2.20}$$

The first order condition (2.20) yields the following.

$$\left(p_j^k - \hat{p}_j^k\right) - W(p_j^k - p_j^{k*}) = 0 \tag{2.21}$$

Finally we have interpolation formula.

$$p_j^k = \frac{\hat{p}_j^k + Wp_j^{k*}}{(1+W)} = \frac{\left(\alpha + \beta p_{j,-1}^k + \gamma MC_j^k\right) + Wp_j^{k*}}{(1+W)} \quad \text{for } T_{12} \tag{2.22}$$

We have an interpolation formula (2.22) for intermediate year; interpolation of price data for total period is transferred to the first step; then, new iteration begins.

3. Concluding Remarks

In the last, consider now determining proper W in (2.02) which is given a prior in the

[2] For detail, see section 4.3 of H.Kosaka(2013).

preceding sections. Remember that the first term is the logarithm of likelihood of labor demand equation; the second term is relating to logarithm of production. As the two terms of different units are mixed, the second term is divided by W to balance the terms. While production by price makes nominal production, labor by wage rate makes nominal wage; therefore two terms $wL \approx pX$ is balanced numerically. Hence we have $L \approx pX/w$.

$$(\log L)^2 \approx [\log(pX/w)]^2 \approx \frac{[\log(pX/w)]^2}{(\log(X))^2} \times (\log(X))^2$$

So that, W is indicated to have the following:

$$W \approx \frac{(\log(X))^2}{(\log pX/w)^2}$$

References

M.A.Fuss(1977),"The Structure of Technology Over Time," Econometrica,Vol.45(Nov.1977),1797-1822.

W.H.Greene(2007),Econometric Analysis,Prentice Hall.

S.Nakamura(1990)," A Nonhomothetic Generalized Leontief Cost Function Based on Pooled Data," The Review of Economics and Statistics,Vol.72,No.4(Nov.,1990),649-656.

H.Kosaka(2013),"Econometric Modeling of Multi-Country/Multi-Sector System for the Asian Economies," The Journal of Econometric Study of Northeast Asia, Vol.9,No.1,pp.1-18.

Appendix: **Generalized Leontief Cost Function of MCMS System**

Let the notation of prices be:

p_i^r :price of i-th industry of r-th country in local currency

$p_i^{rk} = p_i^r \left(e^k / e^r\right)$: price of i-th product in k-th country in k-th currency

It is possible to introduce export subsidy of r-th country, and import tax of k-th country into p_i^{rk} . The generalized Leontief cost function, originally proposed by M.A.Fuss(1977) for domestic IO system, can be generalized for MCMS System, expressed in k-th currency:

$$C_j^k = \sum_{r1}^{R} \sum_{r2}^{R} \sum_{i1}^{N+2} \sum_{i2}^{N+2} h_{i1,i2,j}^{r1,r2,k}(X,t) \sqrt{p_{i1}^{r1,k}} \sqrt{p_{i2}^{r2,k}} \qquad \text{(A1)}$$

where 1~N stands for material inputs, N+1 for labor input and N+2 for capital input; yet, there are no off-diagonal elements for N+1 and N+2. S.Nakamura(1990) specified $h_{i1,i2,j}^{r1,r2,k}$ in (A1) as:

$$C_j^k = \sum_{r=1}^{R}\sum_{i=1}^{N+2} b_{ij}^{rk} p_i^{rk} \left(X_j^k\right)^{bx_{ij}^{rk}} \exp\left(bt_{ij}^{rk} t\right) \tag{A2}$$

$$+\sum_{r1=1}^{R}\sum_{r2=1}^{R}\sum_{i1\neq i2}^{N+2}\sum_{i2\neq i1}^{N+2} b_{i1,i2,j}^{r1,r2,k} \sqrt{p_{i1}^{r1,k}}\sqrt{p_{i2}^{r2,k}} \left(X_j^k\right)^{bx_j^k} \exp\left(bt_j^k t\right)$$

where the upper part stands for diagonal elements and the lower part for off-diagonal elements. Then the factor demand functions in share form are:

$$x_{ij}^{rk} / X_j^k = b_{ij}^{rk}\left(X_j^k\right)^{bx_{ij}^{rk}} \exp\left(bt_{ij}^{rk} t\right) + \sum_{r2\neq r}^{R}\sum_{i2\neq i}^{N} b_{i,i2,j}^{r,r2,k} \sqrt{p_{i2}^{r2,k} / p_i^{rk}} \left(X_j^k\right)^{bx_j^k} \exp\left(bt_j^k t\right) \tag{A3}$$

The factor demand equations (A3) can be simplified to, ignoring the effect of relative prices other than the current:

$$x_{ij}^{rk} / X_j^k = b_{ij}^{rk}\left(X_j^k\right)^{bx_{ij}^{rk}} \exp\left(bt_{ij}^{rk} t\right) + \sum_{i2\neq i}^{N} b_{i,i2,j}^{r,r,k} \sqrt{p_{i2}^{r,k} / p_i^{rk}} \left(X_j^k\right)^{bx_j^k} \exp\left(bt_j^k t\right) \tag{A4}$$

Labor demand equation, a special case of (A4), in the absence of second term is:

$$L_j^k / X_j^k = bl_j^k \left(X_j^k\right)^{blx_j^k} \exp\left(blt_j^k t\right) \tag{A5}$$