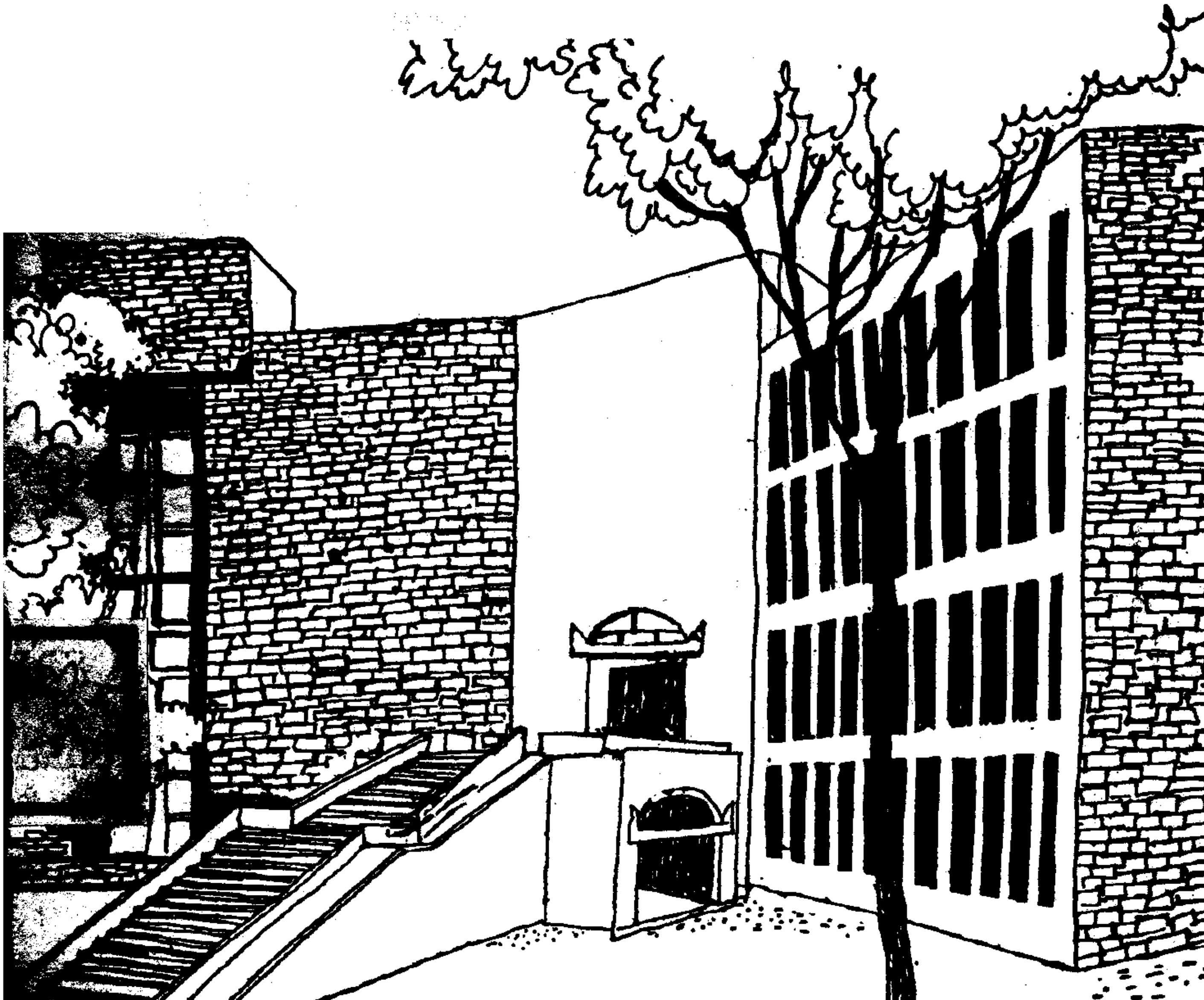
विद्याविनियोगाद्विकासः
IIM
AHMEDABAD

W.P. 599

Working Paper



ON REGIONALIZING A NATIONAL INTER-INDUSTRY TABLE

By

Nanda K. Choudhry

&

Ravindra H. Dholakia

W P No. 599

February 1986

INDIAN INSTITUTE OF MANAGEMENT
AHMEDABAD-380015
INDIA

ON REGIONALISING A NATIONAL INTER-INDUSTRY TABLE

Nanda K. Choudhry and Ravindra H. Dholakia
University of Toronto, Canada; and IIM, Ahmedabad, India.

ABSTRACT

In this paper we have proposed and evaluated a new non-survey method whereby a national input-output coefficient matrix may be decomposed into a set of regional input-output coefficients matrices. The method is based on the technique of generalised inverse. It obeys certain restrictions that have intuitive plausibility. To evaluate the method proposed here, the Canadian national and regional data for the year 1974 have been used. Dalvi and Prasad (1981, 1982) demonstrated that their method of using the Moore-Penrose generalised inverse to regionalise the national table was found more efficient than the popular RAS method. We demonstrate in this paper that the method proposed by us here is more satisfactory than the one proposed by Dalvi and Prasad both in terms of its economic implications as well as its empirical performance.

ON REGIONALISING A NATIONAL INTER-INDUSTRY TABLE

Nanda K. Choudhry and Ravindra H. Dholakia
University of Toronto, Canada; and I.I.M. Ahmedabad, India.

I. Introduction

In this paper we devise[1] and evaluate a method whereby a national Input-Output coefficients matrix may be decomposed into a set of regional input-output coefficients matrices and apply this method to Canada. The process of decomposition proposed here obeys certain restrictions that have intuitive plausibility. While the disaggregation attempted in this paper corresponds to the politically distinct units in Canada --- ten provinces plus Yukon and North West Territories, eleven in all-- the choice of the level of disaggregation is quite optional.

The need to find a method which can yield information on regional inter-industry structure without recourse to primary survey data is too well realised to require detailed comment. Comprehensive revision of even national tables can be attempted only once every few years. A parallel effort for every identifiable region in the nation is far more costly. Ready availability of regional inter-industry detail is an invaluable aid to economic planning in multi-regional societies

[1] Stimulus for the present paper comes from M.Q. Dalvi and K.N. Prasad (1981 and 1982).

where effort must be made to ensure the absence of critical bottlenecks -- and serious surpluses -- on a regional basis. The nation, after all, is only a political aggregate.

Section II of the paper outlines the method of using generalized inverse and examines the weaknesses of a solution in terms of Moore-Penrose generalized inverse suggested by Dalvi and Prasad (1981). In Section III, we outline our method, explain some of its properties and show its relationship to Dalvi and Prasad (1981). Section IV briefly describes the data used for this study and the results of our calculations. Conclusions are presented in Section V.

II. The Use of Generalized Inverse in Regionalizing a National Input-Output Table

The method outlined below allows for the fact that in the Canadian Input-Output use matrix[2] industries (denoted j) use commodities (denoted i) to produce gross value. As industries and commodities do not have one-to-one correspondence, the matrix is not a square matrix. Since a square table may be viewed as a special case of a rectangular table, a method devised for the latter would also work for the former.

Consider a nation with n commodities, m industries and k regions. Our method assumes that the national input-output coefficients matrix is known and also that, for the

[2] The Canadian Input-Output Tables are available in three parts: Input Matrix, Output Matrix and Final Demand Matrix. In the use matrix or Input Matrix, industries utilize different commodities as inputs to produce gross value. The flows in the make matrix, on the other hand, represent industries producing output in terms of commodities.

same year, data are also available on value added and gross output for each industry and region such that for each industry the national gross output (value added) is an arithmetic aggregation of regional gross-outputs (values added).[3] *Central to the 'non-survey' method being proposed* here is that the national coefficients are weighted averages of the corresponding regional coefficients, the weights being the proportion of regional gross-outputs to national gross output in a given industry (j). Thus, symbolically

$$a_{ij} = \sum_{r=1}^{k} p_j^r \, a_{ij}^r \qquad (1)$$

where a_{ij} is the national input coefficient for the i^{th} commodity in the j^{th} industry, p_j^r is the r^{th} region's share of j^{th} industry's national gross output and a_{ij}^r is the input coefficient for the i^{th} commodity in the j^{th} industry for the r^{th} region. In matrix form (1) may be represented as

$$\underset{(mxmn)}{A} = \underset{(mxmk)}{P} \cdot \underset{(mkxmn)}{A^R} \qquad (2)$$

The usual representation of the input coefficients in a *national table with n commodities and m industries is*

---

[3] In the case of several countries, official estimates of regionalvalue added by industry are available on a regular basis. See for instance, Statistics Canada (Cat. #61-202) and CSO, Govt. of India (1979).

in terms of a rectangular $n \times m$ matrix. In (2) above each column of such a matrix is written as a row block diagonal matrix. Thus

$$\underset{(m\times mn)}{A} = \begin{pmatrix} a_{11} \cdots a_{n1} & 0 \cdots\cdots 0 & 0 \cdots\cdots 0 & 0 \cdots 0 \\ 0 \cdots\cdots 0 & a_{12} \cdots a_{n2} & 0 \cdots\cdots 0 & 0 \cdots 0 \\ \vdots \quad\quad \vdots & \vdots \quad\quad \vdots & \vdots \quad\quad \vdots & \vdots \quad\quad \vdots \\ 0 \cdots\cdots 0 & 0 \cdots\cdots 0 & 0 \cdots\cdots 0 & a_{1m} \cdots a_{nm} \end{pmatrix} \tag{3}$$

P and A^R are also block-diagonal matrices defined so as to conform to the set of identities represented by (1). Thus:

$$\underset{(m\times mk)}{P} = \begin{pmatrix} P_1^1 \cdots P_1^k & 0 \cdots\cdots 0 & 0 \cdots\cdots 0 & 0 \cdots\cdots 0 \\ 0 \cdots\cdots 0 & P_2^1 \cdots P_2^k & 0 \cdots\cdots 0 & 0 \cdots\cdots 0 \\ \vdots \quad\quad \vdots & \vdots \quad\quad \vdots & \vdots \quad\quad \vdots & \vdots \quad\quad \vdots \\ 0 \cdots\cdots 0 & 0 \cdots\cdots 0 & 0 \cdots\cdots 0 & P_m^1 \cdots P_m^k \end{pmatrix} \tag{4}$$

and

$$\underset{(mk\times mn)}{A^R} = \begin{pmatrix} a_{11}^1 \cdots a_{n1}^1 & 0 \cdots\cdots\cdots\cdots 0 & 0 \cdots\cdots 0 \\ a_{11}^2 \quad\quad a_{n1}^2 & 0 \quad\quad\quad\quad 0 & 0 \quad\quad 0 \\ \vdots \quad\quad \vdots & \vdots \quad\quad\quad\quad \vdots & \vdots \quad\quad \vdots \\ a_{11}^k \cdots a_{n1}^k & 0 \cdots\cdots\cdots\cdots 0 & 0 \cdots\cdots 0 \\ 0 \quad\quad 0 & 0 \quad\quad\quad\quad 0 & a_{1m}^k \quad\quad a_{nm}^1 \\ \vdots \quad\quad \vdots & \vdots \quad\quad\quad\quad \vdots & \vdots \quad\quad \vdots \\ 0 \cdots\cdots 0 & 0 \cdots\cdots\cdots\cdots 0 & a_{1m}^k \cdots a_{nm}^k \end{pmatrix} \tag{5}$$

From (2), we can obtain,

$$\underset{(mk \times mn)}{A^R} = \underset{(mk \times m)}{P^+} \cdot \underset{(m \times mn)}{A} \qquad (6)$$

Where P^+ is a generalized inverse of P.

Dalvi and Prasad (1981, 1982) use the Moore-Penrose generalized inverse of P to get the solution as

$$a^r_{ij} = p^r_j \, a_{ij} \; / \; \sum_{r=1}^{k} (p^r_j)^2 \qquad (7)$$

The Moore-Penrose Inverse used by Dalvi and Prasad yields a unique solution because it imposes a well-known additional restriction which minimizes the norm of the solution (Ijiri, 1965; Rao and Mitra, 1971) for the solution given by (6). A possible consequence of this method is that sometimes there may arise negative values-added in certain industries in certain regions. This occurs because the solution for a^r_{ij} yielded by (7) are such that except in two extreme cases — one in which p^r_{ij} is 1 for some r and 0 for others and another in which $p^r_j = 1/k$ for all r -- the consequence of <u>relative dominance</u> for some region, r, in some industry, j, implies that for that region and industry $p^r_j \,/\, \sum_r (p^r_j)^2 > 1$. If, in such a case, $\sum_i a_{ij}$ either equals 1 or is close to 1, i.e., the ratio of value-added to gross output at national level is close to zero, then $\sum_i a^r_{ij}$ may well exceed 1, thus yielding a negative value added. In parallel fashion one may infer that regions accounting for small shares of the national gross output in a given industry may get their value added ratio to be higher than the national average. Thus, if one defines 'productive efficiency' as the share of value added to gross output then a likely consequence of using the Moore-Penrose inverse, a la Dalvi and Prasad (1981), may be a somewhat paradoxical inverse relationship between a region's dominance in an industry, as measured by its share in the national gross-output and its efficiency, as measured by its value added ratio.

III. Regionalization Based on Generalized Inverse with Value Added Constraints

In view of the aboveconsequence, we propose to use the restriction that the proportion of value added, both in each region and the nation as a whole, in each industry is known as well as non-negative. That is:

$$v_j = 1 - \sum_{i=1}^{n} a_{ij} \quad ; \qquad j = 1 \ldots m \tag{8}$$

and

$$v^r_j = 1 - \sum_{i=1}^{n} a^r_{ij} \quad ; \qquad j = 1 \ldots m, \quad r = 1 \ldots k \tag{9}$$

In Matrix form:

$$\underset{(m\times mn)}{A} \cdot \underset{(mn\times m)}{I^*} = \underset{(m\times m)}{\delta} = \begin{pmatrix} \sum_i a_{i1} & \cdots & 0 \\ 0 & \ddots & \vdots \\ \vdots & & \vdots \\ 0 & \cdots & \sum_i a_{im} \end{pmatrix} = I - V \quad (10)$$

where

$$\underset{(mn\times m)}{I^*} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & & \vdots \\ 0 & 1 & & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & & \vdots \\ \vdots & 0 & \ddots & 1 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

$$\underset{(mk\times mn)}{A^R} \cdot \underset{(mn\times m)}{I^*} = \underset{(mk\times m)}{\delta^R} = \begin{pmatrix} \sum_i a^1_{i1} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ \sum_i a^k_{i1} & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \sum_i a^1_{im} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \sum_i a^k_{im} \end{pmatrix} = I^R - V^R \quad (11)$$

I^* is a columnwise block diagonal matrix with n elements in each column as unity and the rest zeros. Each element of δ plus the corresponding element of V, the sum of the corresponding elements of δ^R and V^R, must equal unity as indicated above in (8) and (9). Equation (11) contains the unknown A^R and serves as the effective constraint on the solution of (2). In order to get a solution for P^+ consistetn with (11) we post-multiply (2) by I^* yielding:

$$\underset{(m \times mk)}{P} \quad \underset{(mk \times mn)}{A^R} \quad \underset{(mn \times m)}{I^*} = \underset{(m \times mn)}{A} \quad \underset{(mn \times m)}{I^*} \qquad (12)$$

i.e. $P\delta^R = \delta$ (using (10) and (11)) (13)

$\therefore P\delta^R\delta^{-1} = I$ since δ is always a square (mxm) diagonal matrix. (14)

Hence, we can obtain a generalised inverse of P from (14) as

$$P^{++} = \delta^R \delta^{-1} \qquad (15)$$

and substituting (15) in (6) above, we have,

$$A^R = \delta^R \delta^{-1} A \qquad (16)$$

Since δ^{-1} is a diagonal matrix and δ^R a column wise block-diagonal matrix, $\delta^R \delta^{-1}$ is also column wise block diagonal and the solution for the elements of $A^R = (a^r_{ij})$ is given by:

$$a^r_{ij} = (1-v^r_j)\, a_{ij} \,/\, (1-v_j) = \frac{\sum_i a^r_{ij}}{\sum_i a_{ij}}, \qquad (17)$$

where a^r_{ij} and a_{ij} satisfy the definitional constraint in (1).

The solution obtainedin (16) is unique since, P^{++} is the product of two unique matrixes δ^R and δ^{-1}. Moreover,

$$P\,P^{++} = P\,\delta^R\,\delta^{-1} = \delta\delta^{-1} = I$$

$$\therefore \quad PA^R I^* = P\,\delta^R = A\ I^* = \delta\ .$$

We also see that $PP^{++}\ P = P$ and $P^{++}\,P\,P^{++} = P^{++}$ and $P\,P^{++} = (P\,P^{++})'$. However, our solution does not generate $P^{++}\,P$ as a symmetric matrix which is a property of the Moore-Penrose Inverse. Thus our solution would generally differ from that used by Dalvi and Prasad (1981).

It can be seen readily that the solution yielded by (17) obviates the difficulty encountered in using (7) because it implies no relationship between a region's 'efficiency' and its 'industry-dominance.' Regional input/output coefficients vary in direct proportion to that region's relative efficiency (inefficiency) in comparison with the national average. A corollary of this implication is that input-output coefficients in a given industry would be the same for all regions if and only if every region in the nation was equally 'efficient' with respect to that industry. On the other hand, one must also note, that neither the Dalvi and Prasad (1981) method, using P^+ , nor ours, using P^{++}, allows for inter-regional variation in input-output coefficients based on substitution characteristics. Our solution, moreover, implies identical scale characteristics of the production functions across regions in the given industry. Thus depending on the observed productive efficiency differentials in an industry, we are implicitly assuming the regional differences in the production functions in the industry only in terms of the proportional factor representing the level of technology.

IV. The Data and Results

The two methods described above have been tested with Canadian data for 1974,[4] obtained from the Input-Output Division of Statistics Canada. These data were chosen because for 1974, we were able to obtain input-output tables for each of the 11 regions and for the nation (consistent with the regional tables in the sense that the regional flows sum to the corresponding national flow) at the level of disaggregation (45 x 16) based on survey data. Thus we have available to us a set of 45 x 16 x 11 'true' input-

[4]Note that Canadian input-output tables are rectangular, in that the rows represent commodities and the columns represent industries. These tables are available at various levels of disaggregation. The present study uses the most aggregated of these in which the number of commodities (n) equals 45, and number of industries (m) equals 16 and the number of regions (k) equals 11, one for each province and the last, the 11th, representing the Yukon and North West Territories together.

output coefficients against which we can assess the performance of the solutions yielded by (7) and (17).

Dalvi and Prasad (1982) compare the performance of their method based on the Moore-Penrose inverse and the RAS method and conclude that 'when survey data on regional input-output are not available, the use of generalized inverse for regionalization can be made for empirical research with a larger degree of confidence as compared to RAS method.' We have treated the conclusion of their study as the starting point for the present one. Consequently we have made no attempt to assess the relative performance of the RAS method of regionalization and the Dalvi and Prasad (1981) Generalized Inverse method with respect to the data used by us. Here we only assess the comparative performance of the two different 'Generalized Inverses,' P^+ and P^{++}.

As stated above we have a set of 7920 true input-output coefficients and the same number of coefficients generated by each of (7) and (17). We regard the first of these as a set of 7920 observations on a random variable Y, the values generated by (7) and (17) as observations on random variables X and Z respectively. Corresponding elements of X and Z, X_i and Z_i, are proxies for or estimates of Y_i.

If either or both of these proxies were to be perfect then we would have either $X_i = Y_i$ or $Z_i = Y_i$ or $X_i = Z_i = Y_i$ for all i. In actual practice, however, one may expect such correspondence to be only approximate. In order to judge whether X or Z provides a better proxy we make use of a regression model. That is,

$$Y_i = \alpha_0 + \alpha_1 X_i + \epsilon_i \quad \text{for } i=1, \dots 7920 \qquad (18)$$

and $$Y_i = \beta_0 + \beta_1 Z_i + u_i \quad \text{for } i=1, \dots 7920 \qquad (19)$$

If X_i is a good proxy for Y_i then we would expect $\hat{\alpha}_0$ and $\hat{\alpha}_1$ to be insignificantly different from zero and one, respectively, where $\hat{\alpha}_0, \hat{\alpha}_1$ are ordinary least squares estimates. Analogous expectations would be held about $\hat{\beta}_0$ and $\hat{\beta}_1$ with respect to Z_i. A final criterion of choice between (18) and (19) may also be the two coefficients of determination associated with these equations.

The estimated results are as follows:

Parameter	Estimated value	Standard Error	t(H_0: Intercept = 0, Slope = 1)
$\hat{\alpha}_0$	.0070	.0004	17.02
$\hat{\alpha}_1$	.9659	.0145	-2.35
r^2 (18)	.3606		
$\hat{\beta}_0$	.00008	.0001	.57
$\hat{\beta}_1$	.9952	.0033	-1.45
r^2 (19)	.9202		

It is clear from the above results that (19), which is derived from (17), provides a much closer set of proxies than (18), which is derived from (7). Whereas β_0 and β_1 are insignificantly different from 0 and 1, respectively, a similar conclusion is not possible for α_0 and α_1. Moreover, the explanatory power of the two regressions, (18) and (19), as given by the coefficients of determination differs substantially confirming our conclusion.

In Section II above, we also discussed a possible paradox that results from (7) in terms of the relationship between a region's 'dominance' and its 'productive efficiency' in a given industry. While the actual numbers are not so perverse as to actually produce a negative value-added share for the dominant region, they still substantiate our argument against the use of Moore-Penrose inverse for the solution. Table 1 presents the provinces with the highest and the lowest[5] shares in the national output in the six commodity producing industries along with the corresponding ratios of value added to output. The figures in Table 1 clearly confirm our reservation about the nature of the estimates generated by using the Moor-Penrose generalized inverse. By comparing the ratios of value added to output in different industries estimated on the

[5] For obvious reasons, the extremes of 1 and 0 are not considered for identifying the provinces with the highest and the lowest shares in the national output in an industry.

basis of (7) above with the actual ratios obtained from the official estimates, we find that without exception in all the six commodity producing industries, the estimated ratio is considerably understated in the 'largest' province and substantially overstated in the 'smallest' province. On the other hand, a comparison of the actual provincial ratios with the actual national ratios in these industries, does not suggest any systematic inverse relationship of this type between a region's 'relative size' and its 'productive efficiency'. Our method of regionalization, as noted earlier, does not rely on the 'size' of the province, instead, it stresses the relative 'productive efficiency' of the province in the industry. As a result, the 'implausibility' discussed above is obviated.

V. Conclusion

In this paper, we have proposed a new non-survey method based on the technique of generalized inverse with value added constraints to regionalise national inter-industry table. The method suggested by Dalvi and Prasad (1981, 1982) of using the Moore-Penrose generalized inverse to regionalize the national table is found less satisfactory than ours both in terms of its economic implications as well as its empirical performance. The Canadian data for the year 1974 have been used to demonstrate the efficiency with which the national input-output tables could be regionalized by following the method proposed here.

* * *

Table 1 : The Estimated* and Actual Value Added Ratios for the Largest and the Smallest Provinces by Commodity Producing Industries:Canada, 1974.

Industry	Province	Share of the Province in National Output	Ratio of Value Added to Output		
			Provincial Estimated*	Provincial Actual	National Actual
(1)	(2)	(3)	(4)	(5)	(6)
a) THE LARGEST PROVINCES BY INDUSTRY					
1. Agriculture	Ontario	0.2684	0.4672	0.5548	0.6076
2. Forestry	British Columbia	0.5116	0.1490	0.4054	0.4553
3. Fishing Hunting Trapping	British Columbia	0.3185	0.4760	0.6847	0.6752
4. Mines, Quarries Oil Wells	Alberta	0.4455	0.3374	0.5956	0.6180
5. Manufacturing	Ontario	0.5074	0.0143	0.3536	0.3342
6. Construction	Ontario	0.3521	0.1612	0.5020	0.4986
b) THE SMALLEST PROVINCES BY INDUSTRY					
1. Agriculture	Newfoundland	0.0018	0.9964	0.4642	0.6076
2. Forestry	P.E.I.	0.0001	0.9999	0.7143	0.4553
3. Fish, Hunt Trap	Saskatchewan	0.0163	0.9731	0.8557	0.6752
4. Mines, Quarries Oil Wells	Nova Scotia	0.0072	0.9893	0.5744	0.6180
5. Manufacturing	P.E.I.	0.0011	0.9978	0.2649	0.3342 (P.T.O.)
6. Construction	P.E.I.	0.0043	0.9897	0.5118	0.4986

Footnotes of Table 1

* Estimated on the basis of (7) above.

Note: The largest (Smallest) Province in an industry is the province accounting for the largest (smallest) share in national output in the industry.

****

References

CSO, Government of India, National Accounts Statistics, 1979, New Delhi, India

Dalvi, M.Q. and K.N. Prasad, 'Application of Generalized In erse to Regionalization of National Input-Output Tables', (Study Paper No. 17, August 1981), Government of India, Planning Commission.

Dalvi, M.Q. and K.N. Prasad, 'Application of Generalized Inverse and RAS Method to Regionalization of National Input-Output Tables', (Study Paper No. 25, January 1982), Government of India, Planning Commission (UNDP Transport Planning Project).

Ijiri, Y., Management Goals and Accounting for Control, (North Holland Publications, 1965).

Rao, C.R. and S.K. Mitra, Generalized Inverse of Matrices and Its Application. (John Wiley, 1971).

* * *