Since $m = \Sigma \phi_j m_j$, therefore

$$L = \sum_j \frac{\phi_j m_j}{m} L_j + \sum_{i>j} \frac{\phi_i \phi_j (x_G - x_1)}{m} L_{ij} ,$$

$$(13) \quad L = \sum_j \Theta_j L_j + \sum_{i>j} \frac{\phi_i \phi_j (x_G - x_1)}{m} L_{ij} ,$$

where $\Theta_j = \phi_j m_j / m$ is the proportion of national family income enjoyed by families in the j^{th} region. This is a decomposition of the national Gini ratio as the sum of a weighted average (weights adding to one) of the regional Gini ratios and a weighted sum of all possible Gini difference ratios. Thus the first expression measures the contribution of "within-region inequality" whereas the second measures the contribution of "between-region inequality".

An informative way of presenting the decomposition might be in the form of a table, as below. The diagonal elements would be $\phi_j m_j L_j / mL$ and the lower triangular elements would be $\phi_i \phi_j (x_G - x_1) L_{ij} / mL$. The sum of all the elements would be one, and relatively large numbers would indicate the most pressing sources of income inequality.

Relative Sizes of Regional Sources of National Income Inequality

	Region 1	Region 2	. . .	Region R
Region 1	Inequality Within 1			
Region 2	Inequality Between 2 and 1	Inequality Within 2		
⋮	⋮	⋮		
Region R	Inequality Between R and 1	Inequality Between R and 2	. . .	Inequality Within R

Graphical Depiction of Between-Region Inequality

From its definition, we have

$$(x_G - x_1)L_{ij} = f_i'Pf_i + f_j'Pf_j - f_i'Pf_j - f_j'Pf_i$$

The relationship of the first two r.h.s. terms to the Lorenz diagram is clear. The first term, for instance, is the area underneath the Lorenz curve of region i, times two and times m_i. Denote this as $2m_iA_i$, where A_i is the area under the Lorenz curve marked "A_i" in Figure 1. The second r.h.s. term is therefore $2m_jA_j$, where A_j is the area under the Lorenz curve marked "A_j".

For the last two r.h.s. terms,[3/] note that

$$f_i'Pf_j = f_i'HXf_j = m_jf_i'H(1/m_j)Xf_j = m_j(f_i'Hy_j)$$

The expression in the last parentheses can be interpreted as one minus the result of a Gini ratio computed from the distribution across income classes of <u>families in the i^{th} region</u> with the distribution across income classes of

---

[3/] The sum of the last two terms is $f_i'Pf_j + f_j'Pf_i = f_i'Pf_j + f_i'P'f_j = f_i'(P + P')f_j$, where

$$(P + P') = 2\begin{bmatrix} x_1 & x_1 & \cdots & x_1 \\ x_1 & x_2 & \cdots & x_2 \\ \vdots & \vdots & & \vdots \\ x_1 & x_2 & \cdots & x_G \end{bmatrix}.$$

This expression may be useful for computational purposes.

<u>incomes received by families in the j^{th} region</u>. In terms of Figure 1, this would be twice the area underneath the curve "A_{ij}" constructed as follows: point <u>c</u> has the same abscissa as point <u>a</u> (the same proportion of families), but has the same ordinate as point <u>d</u>; and point <u>g</u> has the same abscissa as point <u>e</u>, but has the same ordinate as point <u>h</u>. The area underneath this curve is also denoted A_{ij}. The curve "A_{ji}" is correspondingly constructed, using the (cumulative) family distribution of region j against the (cumulative) income distribution of region i. The area underneath it is denoted A_{ji}.

Since

$$f_i'Pf_j = 2\, m_j A_{ij} \quad \text{and}$$

$$f_j'Pf_i = 2\, m_i A_{ji} \quad ,$$

therefore

$$(x_G - x_1)L_{ij} = 2(m_i A_i + m_j A_j - m_j A_{ij} - m_i A_{ji})$$

$$\frac{x_G - x_1}{2} L_{ij} = m_i(A_i - A_{ji}) + m_j(A_j - A_{ji})$$

Thus L_{ij} is a weighted sum of the shaded positive area $(A_i - A_{ji})$ and the shaded negative area $(A_j - A_{ji})$. Mathematically, this sum has been shown to be positive, i.e., the mean income in region i, which weights $(A_i - A_{ji})$, will be sufficiently large relative to the mean income in region j, which weights $(A_j - A_{ij})$, so that L_{ij} is positive. Actually, the Lorenz diagram is not too helpful in portraying the size of L_{ij}. The intention of this section is merely to describe the relationship between L_{ij} and the conventional diagram.

The example depicted in Figure 1 has some families in every income class, for either region, as is the usual case. If Lorenz curves of regions i and j should intersect, the discussion carries forward, the only notable difference being that each shaded area now has both positive and negative components. Thus, although intersecting Lorenz curves may imply similar intraregional inequalities, they do imply a non-zero interregional inequality.

In case one or more income classes are empty, the graphical depiction breaks down. In particular, if all families are in one income class only, then the Lorenz curve coincides with the diagonal, regardless of which

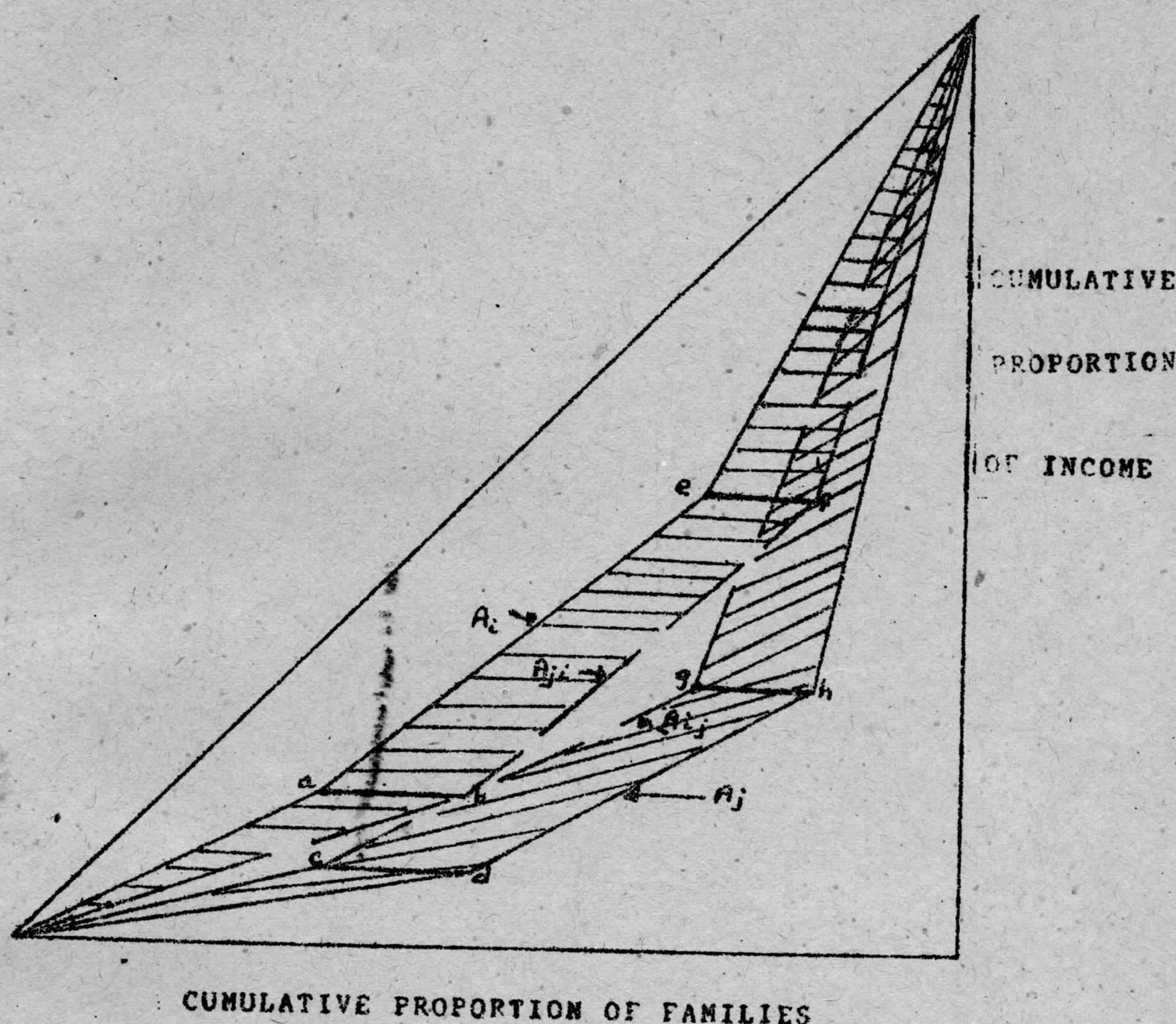
CUMULATIVE
PROPORTION
OF INCOME
e
A_i
A_ji
g
h
A_ij
a
b
c
d
A_j
CUMULATIVE PROPORTION OF FAMILIES

FIGURE I

income class it is which contains all the region's families. In this case, Lorenz curves of two regions would coincide, and yet the Gini difference-ratio would be positive.

Example 1. Hypothetical Data[3/]

Table 1 contains hypothetical data for four regions with various degrees of internal inequality. The second column contains the diagonal elements of the matrix of mean incomes X, given ten income classes. Region One is internally the most equal of the regions, and the degree of inequality grows progressively and is worst for Region Four. The hypothetical data were chosen to exaggerate somewhat the differences between regions one might expect from actual data. (The relative frequencies of families per income classes are plotted against $\ln x_k$ in Figure 2.)

Table 1. Hypothetical Data

k	X	f_1	f_2	f_3	f_4
1	1	.025	.05	.2	.3
2	2	.025	.05	.2	.3
3	3	.05	.1	.2	.125
4	5	.15	.3	.1	.075
5	9	.3	.1	.1	.05
6	15	.2	.1	.1	.05
7	25	.15	.1	.025	.025
8	40	.05	.1	.025	.025
9	80	.025	.05	.025	.025
10	150	.025	.05	.025	.025

3/ Thanks for computational help and for the chart are due to Miss Georgina Ochoa.

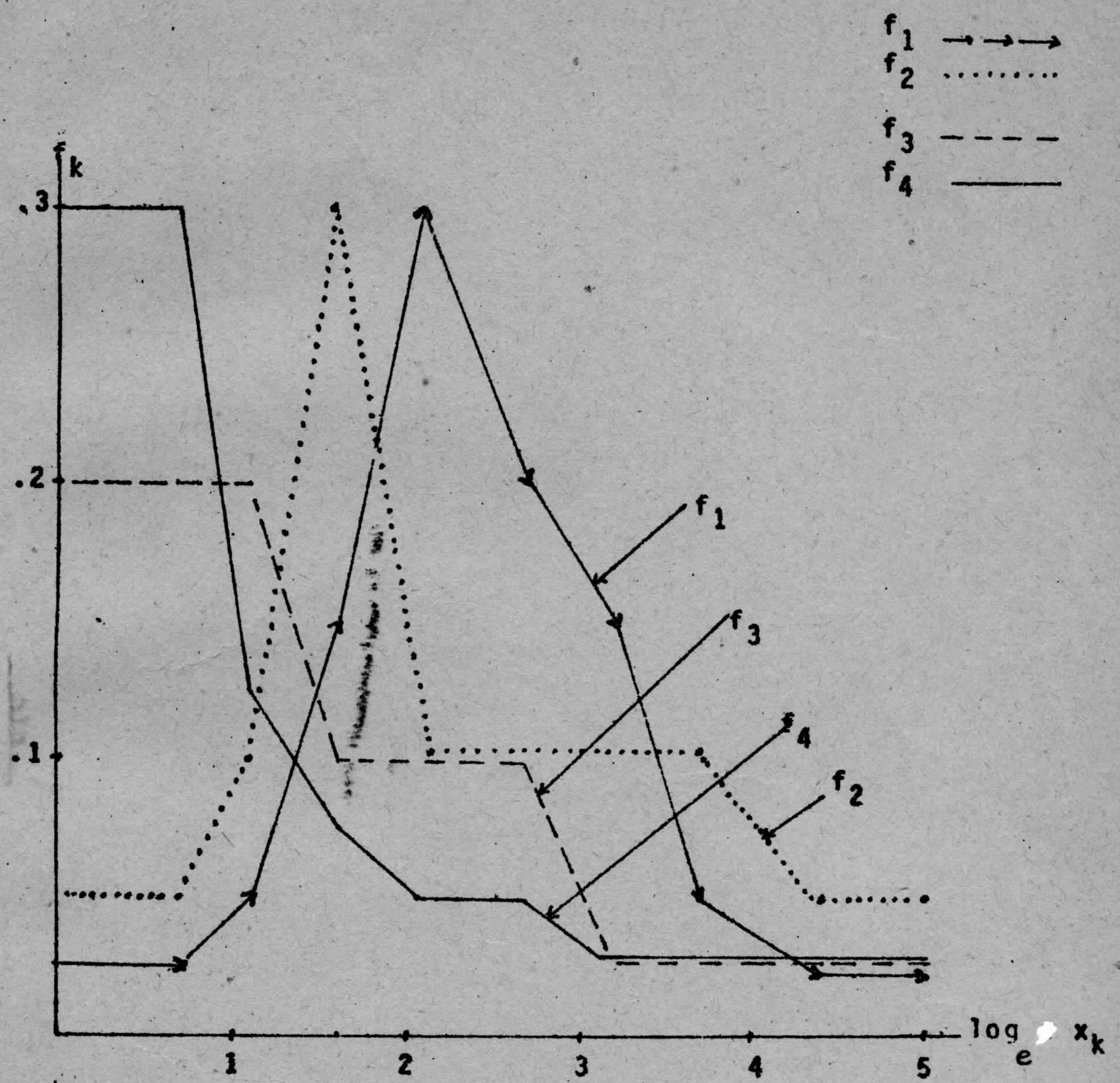

FIGURE 2

From these data the computed L_{ij} are:

i \ j	1	2	3
2	.0041		
3	.0139	.0108	
4	.0211	.0160	.0009

As desired, L_{14} is the largest Gini difference-ratio, and the ratio falls as i and j approach each other. The numerical values give a good indication of what to expect from L_{ij}. Even though L_{ij} has range (0,1), very rarely will an empirical case be found where it goes over 0.10. For instance, (a) a fifth case was considered in which 8/11 of the population was in the 7^{th} class and the rest in the 9^{th} class, and (b) a sixth case in which all families were in the 8^{th} class. (This example gives two regions the same mean family income.) Here L_{56} is .0732, which is high by the standards found for L_{ij}. Thus the Gini difference-ratio is subject to the same type of criticism as the Gini ratio (only more so): lack of numerical "room to move around" leads to the suspicion that the coefficient may be sensitive to errors in the basic data.

Example 2. Philippine Regional Data, 1971[4/]

Tables 2-4 show the results of decomposing the Gini ratio for Philippine regions in 1971, using data from the Family Income and Expenditures Survey of the Bureau of the Census and Statistics.

Table 2. Within-Region Inequality, Philippines, 1971

Region j		Income Share θ_j	Gini Ratio L_j	$\theta_j L_j$
Metropolitan Manila	1	0.17237	0.44810	0.07724
Ilocos & Mt. Prov.	2	0.04813	0.53786	0.02588
Cagayan Valley	3	0.02623	0.44270	0.01161
Central Luzon	4	0.14880	0.44357	0.06600
Southern Luzon & Is.	5	0.15874	0.47618	0.07559
Bicol	6	0.05820	0.45251	0.02634
Western Visayas	7	0.09062	0.42270	0.03826
Eastern Visayas	8	0.10523	0.51174	0.05385
Northern Mindanao	9	0.06737	0.45265	0.03050
Southern Mindanao	10	0.12447	0.44361	0.05522
	TOTAL	1.00015	TOTAL	0.46049

[4/] Thanks for programming assistance are due to Mr. Eduardo Gamboa.

Table 3. Between-Region Inequality Measured by the Gini difference-ratio (L_{ij}), Philippines, 1971

1	2	3	4	5	6	7	8	9
0.03542								
0.05182	0.00176							
0.01739	0.00446	0.01049						
0.01662	0.00400	0.01013	0.00019					
0.04018	0.00049	0.00089	0.00573	0.00545				
0.03277	0.00120	0.00293	0.00281	0.00285	0.00102			
0.04643	0.00091	0.00056	0.00841	0.00806	0.00065	0.00257		
0.03652	0.00066	0.00185	0.00385	0.00390	0.00056	0.00027	0.00144	
0.02579	0.00171	0.00525	0.00097	0.00105	0.00219	0.00050	0.00410	0.001

Table 4. Weighted Between-Region Inequality, Philippines, 1971

1	2	3	4	5	6	7	8	9
0.00134								
0.00148	0.00003							
0.00163	0.00028	0.00049						
0.00158	0.00025	0.00048	0.00003					
0.00218	0.00002	0.00002	0.00051	0.00049				
0.00241	0.00006	0.00011	0.00034	0.00035	0.00007			
0.00498	0.00006	0.00003	0.00147	0.00143	0.00007	0.00035		
0.00209	0.00002	0.00005	0.00036	0.00037	0.00003	0.00002	0.00015	
0.00233	0.00010	0.00024	0.00014	0.00016	0.00019	0.00006	0.00069	0.000
							TOTAL	0.029

The tables describe a national inequality coefficient of .49010 = .46049 (Table 1) + .02961 (Table 3). (The Gini coefficient directly computed from the national aggregates is .485, and the difference of .005 is due to the simplification of using the national mean incomes per income class as the x_k for every region.) The L_{ij} in Table 3, except for column one (Manila) are very small, as would have been expected. Thus within-region inequalities account for 93% of total national inequality. Between-region inequalities account for only 2.96 percentage points out of the total of 49.01. Inequalities between Region I (Manila) and all the other regions account for 2.00 points, or two-thirds of the total between-regions.

REFERENCES

[1] H.T. Oshima, "Income Inequality and Economic Growth: the Postwar Experience of Asian Countries," Malayan Economic Review, vol. XV, no. 2 (October 1970), pp. 7-41.

[2] M. Mangahas, "A Note on 'Income Inequality and Economic Growth: the Postwar Experience of Asian Countries,'" forthcoming in the Malayan Economic Review.