An Anatomy of the Distribution of Urban Income
A Tale of Two Cities in Colombia

Rakesh Mohan

WORLD BANK STAFF WORKING PAPERS
Number 650
An Anatomy of the Distribution of Urban Income
A Tale of Two Cities in Colombia

Rakesh Mohan

The World Bank
Washington, D.C., U.S.A.
This is a working document published informally by the World Bank. To present the results of research with the least possible delay, the typescript has not been prepared in accordance with the procedures appropriate to formal printed texts, and the World Bank accepts no responsibility for errors. The publication is supplied at a token charge to defray part of the cost of manufacture and distribution.

The views and interpretations in this document are those of the author(s) and should not be attributed to the World Bank, to its affiliated organizations, or to any individual acting on their behalf. Any maps used have been prepared solely for the convenience of the readers; the denominations used and the boundaries shown do not imply, on the part of the World Bank and its affiliates, any judgment on the legal status of any territory or any endorsement or acceptance of such boundaries.

The full range of World Bank publications is described in the Catalog of World Bank Publications; the continuing research program of the Bank is outlined in World Bank Research Program: Abstracts of Current Studies. Both booklets are updated annually; the most recent edition of each is available without charge from the Publications Sales Unit of the Bank in Washington or from the European Office of the Bank, 66, avenue d'Iéna, 75116 Paris, France.

Rakesh Mohan is an economist in the Development Research Department of the World Bank.

Library of Congress Cataloging in Publication Data
Mohan, Rakesh, 1948-
Anatomy of the distribution of urban income.
(World Bank staff working papers no. 650)
Bibliography: p.
1. Income distribution--Colombia--Bogotá. 2. Income distribution--Colombia--Cali. I. Title. II. Series.
HC198.B5M62 1984 339.2'2'098614 84-11989
ABSTRACT

This paper suggests an approach for analysis of distributional and other change in incomes in fast growing cities in the developing world. It notes that between 1973 and 1978 there appeared to be a tendency toward some improvement in the distribution of income in the two Colombian cities of Bogota and Cali. However, the level of inequality in the two cities remained very high, possible among the highest in the world.

Among the distinguishing features of this study is the utilisation of primary data from different household surveys, of comparable quality spread over six years. Hence there is a high degree of confidence in the results obtained. The paper uses different samples, inequality indices, income concepts and ranking procedures in its analysis. It is found that, overall, there is a higher level of inequality in Bogota than in Cali. Particular attention is paid to the existence of spatial inequality among different parts of the two cities by using an index of spatial income segregation (ISIS). ISIS is found to be very high when the cities are disaggregated spatially into radical sectors rather than into concentric rings. The main disquieting feature of the findings is that ISIS appears to be increasing over time while inequality within the sectors is declining; this suggests that the cities are becoming more spatially segregated over time.

The inequality in labor earnings is also calculated and decomposed. The paper finds that the industry of activity and the size of the firm of the worker contribute very little to overall inequality in earnings, in contrast to much of the labor segmentation literature. Instead, education, occupation and the location of residence (or background) of the worker contributed most to inequality in earnings. The results of the study suggest that increases in spatial inequality in household as well as labour earnings could have deleterious effects on the distribution of income in the future.

Finally, in addition to sixteen tables and eight appendices, the paper carries a special computer program call EQUALISE which can compute income distribution indices directly from household data and decompose income inequality according to any characteristics.
En el presente documento se propone un método para el análisis de los cambios en la distribución y otros aspectos de los ingresos en las ciudades del mundo en desarrollo en proceso de rápido crecimiento. Se observa en él que, entre 1973 y 1978, parecía haber una tendencia hacia una mejor distribución de los ingresos en las ciudades colombianas de Bogotá y Cali. Sin embargo, el nivel de desigualdad en las dos ciudades siguió siendo muy grande, y posiblemente sea uno de los más altos del mundo.

Una de las características distintivas de este estudio es la utilización de datos primarios de diversas encuestas de unidades familiares de calidad comparable esparcidas en el curso de seis años. De allí que haya un alto grado de confianza en los resultados obtenidos. El documento usa en su análisis diferentes muestras, índices de desigualdad, conceptos de ingresos y procedimientos de jerarquización. Llega a la conclusión de que, en general, hay un mayor nivel de desigualdad en Bogotá que en Cali. Dedica atención preferente a la existencia de una desigualdad espacial entre diversas partes de las dos ciudades al utilizar un índice de desagregación espacial de los ingresos. Halla que este índice es muy alto cuando las ciudades se desglosan espacialmente en sectores radiales y no en anillos concéntricos. La principal característica perturbadora de las conclusiones es que el índice de desagregación espacial de los ingresos parece aumentar con el tiempo en tanto que la desigualdad dentro de los sectores está disminuyendo; esto indica que con el correr el tiempo las ciudades se están volviendo más segregadas espacialmente.

También se calcula y desglosa en el documento la desigualdad de los ingresos de los trabajadores. Se llega a la conclusión de que el tipo de industria de la actividad y el tamaño de la empresa a que pertenece el trabajador contribuyen muy poco a la desigualdad general de los ingresos, contrariamente a lo que se sostiene en gran parte del material escrito sobre el desglose de la mano de obra. En cambio la educación, la ocupación y el lugar de residencia (o los antecedentes) del trabajador son elementos que han contribuido más a la desigualdad de los ingresos. Los resultados del estudio indican que los aumentos de la desigualdad espacial de los ingresos tanto de las unidades familiares como de los trabajadores podrían tener efectos perjudiciales sobre la distribución de los ingresos en el futuro.

Finalmente, además de 16 cuadros y 8 apéndices, el documento presenta un programa especial de computadora llamado EQUALISE con el que pueden calcularse índices de distribución de los ingresos directamente a partir de los datos de las unidades familiares y descomponer la desigualdad de los ingresos de acuerdo con cualquier características.

La présente étude se caractérise, entre autres, par l'utilisation de données primaires tirées de diverses enquêtes sur les ménages de qualité comparable, effectuées sur une période de six ans. Les résultats obtenus sont donc jugés très fiables. L'auteur utilise différents échantillons, indices d'inégalité, concepts de revenu et procédures de classement dans le cadre de son analyse. Selon lui, les inégalités sont, dans l'ensemble, plus prononcées à Bogota qu'à Cali. Il accorde une attention particulière aux inégalités qui existent entre différents quartiers des deux villes et, à cette fin, utilise un indice de ségrégation spatiale des revenus. Cet indice prend une valeur très élevée lorsque les villes sont découpées en secteurs radiaux plutôt qu'en cercles concentriques. Le résultat le plus inquiétant est que cet indice paraît augmenter dans le temps alors même que les inégalités au sein de chacun des secteurs diminuent; la ségrégation géographique semblerait donc s'intensifier avec les années.
L'auteur a également calculé et décomposé les inégalités enregistrées au niveau des rémunérations de la main-d'œuvre. Il en conclut que le type d'activité industrielle et la taille des entreprises ne contribuent que peu à l'inégalité globale des salaires contrairement à ce qu'avancent la plupart des études consacrées à la segmentation de la main-d'œuvre. Ce sont le niveau d'éducation, le type d'emploi et le lieu de résidence (ou l'environnement) de l'employé qui contribuent le plus à l'inégalité des revenus salariaux. L'étude suggère que tout accroissement des inégalités géographiques enregistrées au niveau des revenus des ménages ainsi que des salaires pourrait avoir un effet défavorable sur la répartition des revenus dans les années à venir.

Enfin, outre seize tableaux et huit annexes, le lecteur trouvera joint à la présente étude un programme informatique intitulé EQUALISE, qui permet de calculer directement les indices de répartition de revenu à partir des données sur les ménages et de décomposer les inégalités des revenus en fonction de n'importe quel facteur caractéristique.
This paper is part of the City Study research program (RPO 671-47) which has been carried out in Bogota and Cali, Colombia by the World Bank in collaboration with the Corporacion Centro Regional de Poblacion of Bogota. The goal of the City Study is to increase the understanding of the workings of five major urban sectors—housing, transport, employment location, labor markets, and public finance—in order that the impact of policies and projects can be assessed more accurately.

A rather large amount of data handling, documentation, and preparation has gone into the work leading to this paper. I would like to thank Alan Carroll for the original cleaning and documentation of the 1975 and 1977 Household Surveys; Nelson Valverde for making the 1973 Population Census Sample ready for use; Nancy Hartline for wrestling with the 1977 Household Survey, preparing its final documentation, and putting the income imputing method into effect; Sungyong Kang and M.Wilhelm Wagner for making the 1978 sample ready for use.

The computations appearing in this paper have been prepared by Sungyong Kang and Robert Marshall. A word of special thanks goes to them.

My intellectual debt to Sudhir Anand is obvious from the text of the paper. Original programming of parts of the program EQUALISE appearing as appendix III was done by Sudhir Anand and Sherman Robinson. I am grateful to them for making available their programs. I have also benefited from discussions with and the pioneering works on income distribution in Colombia of Miguel Urrutia and R. Albert Berry.

A special word of appreciation goes to the staff of DANE (the Colombian Statistical Office) for having painstakingly conducted all the surveys in Colombia and for making them available in almost raw form so readily. The high quality of the 1978 City Study DANE Household Survey owes much to the diligence of Roberto Pinilla and Maria Cristina Jimenez of DANE and Gary Losee Alfredo Aliaga, Alvaro Pachon and Jairo Arias for CCRP-World Bank.

Other papers in the City Study Labor Market and Income Distribution series are:


# Table of Contents

## I. INTRODUCTION .......................................................... 1
   1.1 Objectives.............................................................. 1
   1.2 Income Concepts used............................................ 3
   1.3 Measures of Income Inequality................................. 6
   1.4 Decomposition of Inequality.................................... 6

## II. OVERALL TRENDS IN THE DISTRIBUTION OF INCOME............... 23
   2.1 Trends in Colombia as a Whole................................. 23
   2.2 The Distribution of Income in Bogota and Cali............ 28

## III. THE SPATIAL DISTRIBUTION OF INCOME............................. 40

## IV. THE DECOMPOSITION OF LABOUR EARNINGS.......................... 54
   4.1 The Spatial Distribution of Personal Earnings............. 55
   4.2 The Decomposition of Labor Earnings According to Various Personal Characteristics.................. 59
   4.3 Summary............................................................. 74

## V. SUMMARY AND CONCLUSIONS............................................. 77

APPENDIX I Statistical Appendix....................................... 82
APPENDIX II The Data..................................................... 89
APPENDIX III EQUALISE: A Program for Income Distribution Studies.................. 99

BIBLIOGRAPHY............................................................. 131
<table>
<thead>
<tr>
<th>Table No.</th>
<th>Title</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The Distribution of Income in Colombia 1964-1974.</td>
<td>25</td>
</tr>
<tr>
<td>2.</td>
<td>The Distribution of Income In Bogota and Cali: Income Shares 1973-1978</td>
<td>31</td>
</tr>
<tr>
<td>3a.</td>
<td>Inequality Indices for Bogota Households 1973-1978.</td>
<td>33</td>
</tr>
<tr>
<td>3b.</td>
<td>Inequality Indices for Bogota's Population Ranked by Household Income Per Capita 1973-1978</td>
<td>34</td>
</tr>
<tr>
<td>4.</td>
<td>Inequality Indices for Cali 1973-1978.</td>
<td>36</td>
</tr>
<tr>
<td>5.</td>
<td>Inequality Indices for Bogota 1978. Results form Using Different Numbers of Observations in the Lorenz Curve</td>
<td>38</td>
</tr>
<tr>
<td>6a.</td>
<td>The Spatial Distribution of Income in Bogota 1973-1978.</td>
<td>45</td>
</tr>
<tr>
<td>6b.</td>
<td>The Spatial Distribution of Income in Cali 1973-1978.</td>
<td>46</td>
</tr>
<tr>
<td>7.</td>
<td>The Spatial Distribution of Income Distinguished by Rings and Sectors: Bogota and Cali 1978.</td>
<td>49</td>
</tr>
<tr>
<td>8a.</td>
<td>Spatial Inequality in Bogota: Individuals Ranked by HINCAP 1973-1978.</td>
<td>51</td>
</tr>
<tr>
<td>Table No.</td>
<td>Title</td>
<td>Page No.</td>
</tr>
<tr>
<td>----------</td>
<td>----------------------------------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>8b.</td>
<td>Spatial Inequality in Cali Individuals</td>
<td>52</td>
</tr>
<tr>
<td>9a.</td>
<td>The Spatial Distribution of Labour Earnings</td>
<td>56</td>
</tr>
<tr>
<td>9b.</td>
<td>The Spatial Distribution of Labour Earnings</td>
<td>57</td>
</tr>
<tr>
<td>10.</td>
<td>The Decomposition of Labour Earnings 1978.</td>
<td>61</td>
</tr>
<tr>
<td>11.</td>
<td>The Distribution of Earnings by Age Groups 1978.</td>
<td>62</td>
</tr>
<tr>
<td>12.</td>
<td>The Distribution of Earnings by Education Groups 1978.</td>
<td>64</td>
</tr>
<tr>
<td>13.</td>
<td>The Distribution of Earnings by Employment Status 1978.</td>
<td>66</td>
</tr>
<tr>
<td>14.</td>
<td>The Distribution of Earnings by Occupational Categories 1978.</td>
<td>68</td>
</tr>
<tr>
<td>15.</td>
<td>The Distribution of Earnings by Industry of Activity 1978.</td>
<td>70</td>
</tr>
<tr>
<td>16.</td>
<td>The Distribution of Earnings by Firm Size 1978.</td>
<td>72</td>
</tr>
</tbody>
</table>
# List of Maps

<table>
<thead>
<tr>
<th>Map No.</th>
<th>Title</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Bogota: Ring and Sector System</td>
<td>42</td>
</tr>
<tr>
<td>2.</td>
<td>Cali: Ring and Sector System</td>
<td>43</td>
</tr>
</tbody>
</table>
I. INTRODUCTION

1.1 Objectives

This paper is part of the program of research on urban labour markets and income distribution in the cities of Bogota and Cali that has been conducted under the rubric of the City Study. Earlier papers\(^1\) have developed a profile of the distribution of income as well as the labour market somewhat descriptively. One of the striking results was the appearance of significant differences between different areas of the city in terms of income and occupation. The existence of these spatial differences was also utilized as an additional variable analysing the determinants of income in Mohan (1981). This paper attempts to measure the distribution of income somewhat more systematically and provides a spatial decomposition of the inequality indices estimated. In so doing a measure of the spatial inequality of income within cities suggests itself.

The availability of micro data sets for 4 different years from 1973 to 1978 permits the estimation of inequality indices for this whole period. It is unusual to have access to such micro data and to be able to make such estimations across data sets consistently. The comparisons between different years give a sense of the robustness (or otherwise) of the estimates derived. They also suggest caution in the overinterpretation of small changes in such indices between different surveys. Nonetheless the estimates for these years are compared with those available for earlier years.

Mohan (1981) examined in detail the determinants of personal earnings. Those findings are supplemented here by the decomposition of the earnings distribution according to various characteristics of the

\(^1\) See Mohan (1980) and Mohan and Hartline (1980)
labour force. The results are largely similar but the decompositions do
give further insight into the distribution of labour earnings.

Most income distribution studies use grouped data to calculate
the inequality indices. Sometimes, only published data are available
which give data for the distribution of the population by specific income
groups: there is then no choice but to use such grouped data. Even when
micro data are available, computation is usually done after the data are
grouped because of the expense involved in making direct calculations from
the micro data. A similar procedure is used here but Appendix III lists
EQUALISE, the computer program used, which takes micro data directly as
input and permits any level of disaggregation desired for the computation
of inequality indices. An extension of the same program is used to do the
decomposition of inequality.1/

The objective of this paper is therefore to document the
distribution of income in Bogota and Cali and to give some indications of
the pattern of change that may have occurred in the recent past. A novel
feature of this paper is the computation of spatial inequality within a
city. This has been done particularly because spatial inequality was
found to be quite marked in the cities under investigation.

The computation of inequality indices has been done for both
household income as well as household income per capita (henceforth
referred to as HINCAP) and later for earned incomes. The concepts and

1/ Parts of the program were originally developed by
Sherman Robinson and Sudhir Anand in earlier studies of
income distribution for the World Bank.
methodology used are laid out in the next parts of this introduction. The sections following present the results that have been obtained.

1.2  **Income Concepts Used**

Appendix II gives a brief description of the sources of data used in this study. The incomes used are all monthly incomes—derived in the different surveys as detailed in the appendix. All the incomes measured are current incomes and therefore some caution needs to be exercised in interpreting them as measures of welfare. Household consumption expenditures are, in general, regarded as a better proxy of permanent income and hence of welfare levels of households and individuals. There is little choice, here, since there have been no recent surveys of consumption in Colombia. The computation of inequality by using current household income or household income per capita (HINCAP) also suffers from yet another problem. It fails to take into account the differential tax incidence on different incomes according to the prevailing tax structure in the country. Similarly, no adjustment is made of the differential accrual to different income groups of the benefits due to public expenditure on public services such as health, education, sanitation, water, etc. 1/

Although, as is detailed in Appendix II, somewhat different income questions were asked in the four sources used—the 1973 Census, 1/

---

1/ See Selowsky (1979) for explicit consideration of these in the computation of inequality in Colombia.
the 1975 and 1977 DANE household surveys and in the 1978 DANE-World Bank City Study Household Survey, the basic income concepts used here have been the Monthly Household Income (HHY) and the Monthly Household Income per Capita (HINCAP). HINCAP is merely the HHY divided by the household size (HHSIZE). In principle, HHY includes the labour earnings of all members of the household, including earnings in kind, as well as non-labour income. Earnings in kind are not taken to include imputed earnings for household work. The coverage of labour earnings in all the surveys is much better than the coverage of non-labor income. The 1978 survey included detailed questions on non-labour income which were addressed specifically according to types of non-labour income. However, it is suspected that the coverage was still less than desirable. HHY does not include imputed income from housing for owner occupying households.

Section IV of the paper analyses the distribution of labour earnings of individuals. Here, naturally, only individual labour earnings are used in the computation. Non-labour earnings are completely ignored.

One important point worthy of note is that all households with zero incomes have been excluded in the computation of income inequality. Similarly, other households with miscoded location were also excluded. For the workers, the definition included those individuals who gave their primary activity as working and others who worked at least 15 hours a week. The households with zero incomes were essentially those whose members happened to be unemployed. As is mentioned in the Appendix II on

---
data, income was systematically imputed to those workers who were employed but who did not give information on their incomes.

Studies on inequality have to choose the population unit as well as the income concept by which they should be ranked in order to derive income distributions. The most common distribution used has been that of households ranked by household income. This distribution has come under attack increasingly in recent works (e.g. Anand, 1982; Datta and Meerman, 1980) as not providing a good indication of differences in the level of living in the population. As argued by Anand, our general concern is with the welfare of individuals rather than households. Secondly, households vary by size as well as by composition in terms of age and sex. Hence, for true cross household comparisons, adjustments should be made to account for these variations. In some of the literature concerning the measurement of poverty based on nutrition levels, adjustments have been attempted by deriving "adult equivalent" measures of consumption. While some norms may be found for deriving adult equivalents based on age and sex for food consumption purposes, it is difficult to do so for consumption as a whole. For example, it may well be that, while a child may require less nutrition than an adult, expenditures on his/her education and health might well exceed those for adults. It is therefore difficult to estimate adult equivalents for total consumption purposes. The next best strategy is followed therefore and HINCAP (household income adjusted for household size) is used to rank households and individuals for welfare comparisons. This is certainly better than the use of HHY even though economies of scale in consumption are also ignored in the
computation of HINCAP. It is assumed implicitly that distribution within the household is even—though some might argue that this is usually not the case. In any case, there is no information which could be used to make any other assumption.

Although it is being argued that it is individual welfare that we are interested in, the basic income concept from which HINCAP is derived is HHY. There is really no other choice since the household is the basic consumption unit where consumption allocations are typically made. Furthermore, non-labour income usually accrues to the household rather than the individual.

The distributions that are analyzed then in sections II and III are:

i) The distribution of households ranked by household income.

ii) The distribution of households ranked by household income per capita.

iii) The distribution of individuals ranked by household income per capita.

Section IV gives the analysis of the distribution of workers ranked by labour income, which includes income in kind.

1.3. Measures of Income Inequality

Various inequality measures have been utilized in this study and then two used for decomposition purposes. The measures displayed in the tables are:

i) Gini coefficient

ii) Theil inequality index
iii) The standard deviation of logarithms
iv) Atkinson index
v) Coefficient of variation

The Gini coefficient, the standard deviation of logs and the coefficient of variation need no explanation since they are the most commonly used and well known measures. Although Theil\(^1\) and Atkinson\(^2\) are now well known, their use in the computation of inequality is still not common. Anand (1982) gives a comprehensive theoretical introduction to all these measures so only a brief explanation of Theil and Atkinson is offered here.\(^3\)

Two important properties of inequality measures that are regarded as desirable are:

1) They be mean-independent (i.e. their value remain unchanged if everyone's income is increased by the same proportion).

(ii) They satisfy the "Pigou-Dalton" condition which requires that any transfer from a richer person to a poorer person (which does not change their relative ranks) be reflected in a reduction of the inequality index. Simply put, a transfer from a richer to a poorer person without changing their ranks should imply higher equality as measured by the index.

The Gini coefficient satisfies these two conditions, as does the coefficient of variation (the square root of variance divided by the mean).

---

1/ First proposed in Theil (1967) and applied in Theil (1972).
The variance of incomes, which would be an obvious measure of the inequality in a distribution, does satisfy the Pigon-Dalton condition but is obviously not mean independent. The log variance of income, (or the variance of the logarithm of incomes) does, however, satisfy both these conditions:

$$\text{Var} (y) = \frac{1}{n} \sum_{i=1}^{n} (y_i - \mu)^2$$

hence $$\text{Var} (\lambda y) = \lambda^2 \text{Var} (y)$$

but $$\text{Var} (\log \lambda y) = \text{Var} (\log y)$$

where $y_i$ is income of unit $i$,

$\mu$ is the mean

and $\lambda$ is the a proportionate change in all incomes. The log variance of incomes is therefore a convenient measure to use.\(^1\) The main problem in its use is that the measure is not defined when there are zero incomes in the sample. This is the main reason for the exclusion of all zero incomes in the sample. It is also an attractive measure to use because of its decomposition properties - as will be detailed below.

1.31 An Introduction to the Theil Index

The Theil index is based on the concept of expected information in a distribution. Essentially, when the probability of an event

\(^1\) The Pigon–Dalton condition is satisfied for log variance for all incomes except those over $\mu e$ where $\mu$ is the geometric mean income and $e$ is the base of natural logarithms.
occurring is low, the information contained in a message stating that it has occurred is high. Intuitively, then, information can be regarded as the converse of probability. As a precise measure of information contained in a message that states that an event with prior probability \( p \) has occurred, the function suggested is:

\[
h(p) = \log \frac{1}{p}
\]  

(1)

which ranges from (when \( p \) is small, the event is unlikely and hence when it does occur, information in the message stating that it has occurred is high) to zero (when \( p=1 \), the event is certain). For a distribution of events, each with probability \( p_i \) (\( i = 1, \ldots, n \)) and

\[
\sum_{i \neq j} p_i = 1
\]

The expected information of the message on the occurrence of one of these events is

\[
H = \sum_{i \neq j} p_i \log \frac{1}{p_i}
\]  

(2)

i.e. the weighted average of the information contained in each message.

The minimum of this distribution is when

\[
p_j = 1 \text{ and all } p_i = 0 \ (i \neq j)
\]

i.e. \( H = \log 1 = 0 \)

and the maximum is when
\[ p_j = p_i = \frac{1}{n} \quad \text{for all } i = j \]

Then \[ H = \sum_{i=1}^{n} \frac{1}{n} \log \frac{1}{1/n} = \log n. \]

\( H \) is also known as entropy. In its application to the measurement of income distribution, it can be interpreted as a measure of uncertainty. Higher the uncertainty in a distribution, greater the expected information in the message stating that an event has occurred from the distribution of possible events, hence the entropy of a distribution may be said to measure the "amount of uncertainty" in it. It is similar to variance which is the more common measure used to measure uncertainty.

Now consider an income distribution, \( z_1, z_2 \ldots z_n \)
i.e. \( n \) individuals with income \( z_i \).

Let \( Z = \sum_{i=1}^{n} z_i \)

Then \( \sum_{i=1}^{n} \frac{z_i}{Z} = 1. \)

The entropy of the income distribution can then be defined

Let \( q_i = \frac{z_i}{Z} \)

hence \( \sum q_i = 1. \)
and all $q_i > 0$

Then

$$H(q) = \sum_{i=1}^{n} q_i \log \frac{1}{q_i}$$

as for any distribution of probabilities

$$H(q) = \sum_{i=1}^{n} \frac{z_i}{Z} \log \frac{1}{z_i/Z}$$

but $Z = n \bar{z}$

where $\bar{z}$ is the mean of the distribution

Then

$$H(q) = \sum_{i=1}^{n} \frac{z_i}{nz} \log \frac{1}{z_i/nz}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \left( \frac{z_i}{z} \log \frac{n}{z_i/nz} \right)$$

$$= \frac{1}{n} \sum_{i=1}^{n} \left( \frac{z_i}{z} \log n + \frac{z_i}{z} \log \frac{1}{z_i/z} \right)$$

$$= \frac{1}{n} \sum_{i=1}^{n} \frac{z_i}{z} \log n + \frac{1}{n} \sum_{i=1}^{n} \left( \frac{z_i}{z} \log \frac{1}{z_i/z} \right)$$

$$= \log n + \frac{1}{n} \sum_{i=1}^{n} \left( \frac{z_i}{z} \log \frac{1}{z_i/z} \right)$$

(3)

Now if, $z_i = z_j$ for all $i = j$

$z_i = \bar{z}$

and $H(q) = \log n$.

and if all income is earned by one individual $j$, 

\[ z_j = Z = nz \]

\[
H(q) = \log n + \frac{1}{n} \left( \frac{nz}{z} \log \frac{1}{nz/z} \right)
\]

\[
= \log n + \frac{1}{n} \left( n \log \frac{1}{n} \right)
\]

\[
= \log n - \log n
\]

\[
= 0.
\]

Hence entropy is high with low inequality while it is low with high inequality. Intuitively then, if we want an index which is low for low inequality and high for high inequality we can arbitrarily define

\[ T = \log n - H(q) \]

(4)

to measure the departure of an income distribution from an even distribution.

In fact the Theil Index (T) of income inequality is defined as "The expected information of the message which transforms population shares into income shares." Consider again an event \( E \) with prior probability \( p \). If a message is now received that the probability of this event has changed to \( q \), what is the information contained in this message? The prior probability of the event occurring is \( p \) and the posterior probability is \( q \). Now when the event does occur, the information provided by that message would be \( h(q) \). Hence, if the information content of the intermediary message changing the probability
to q, is
\[ I + h(q) = h(p) \]
or \[ I = h(p) - h(q) = \log \frac{q}{p} \]

Again, this can be applied to a distribution of probabilities
\[ (p_1, p_2, \ldots, p_n) \]
being changed to
\[ (q_1, q_2, \ldots, q_n). \]
The expected information or entropy of the message changing the probabilities is then
\[ I(q:p) = \sum_{i=1}^{n} q_i \log \frac{q_i}{p_i} \]
i.e. weighted average of the information for each event - the weighting being done by the posterior probability. "I" may be regarded as a measure of the "differentness" of the second distribution \( (q_i) \) from the first \( (p_i) \).

Note that if \( p_i = \frac{1}{n} \) i.e.

the case of equiprobability,
\[ I(q:p) = \sum_{i=1}^{n} q_i \log \frac{q_i}{1/n} \]
\[ = \log n - \sum_{i=1}^{n} q_i \log \frac{1}{q_i} \]
which is the same form as \( T \), in equation (4). The interpretation is that \( T \) is a measure of the departure from complete equality. Since an individual's share in a population of \( n \) is \( \frac{1}{n} \), the distribution of
population shares \( (p_1, p_2 \ldots p_n) \) is an even distribution. When we are then provided with their income distribution, their share of income may be regarded as the posterior distribution and we arrive at the original definition of the Theil index as "The expected information of the message which transforms population shares into income shares."

**Methodology Used in EQUALISE**

It is clear then that the value of the Theil index varies from zero for perfect equality to \( \log n \) for the extreme case where one individual gets all the income and the other \((n-1)\) individuals earn nothing. This raises a problem of comparability of the Theil index between different populations since it is dependent on the magnitude of \( n \). One method of normalization is to compute \( T \) as a proportion of \( \log n \), its feasible maximum value for the population at hand. This has not been utilized in this paper since all the calculations have been done on the basis of 20 quantiles drawn from each sample. The procedure adopted in **EQUALISE** is to compute each inequality index on the basis of the income share received by each quantile of population. As mentioned earlier, the program makes it possible to use all the observations available or to aggregate them into quantiles. Experiments were done using different numbers of quantiles and 20 was chosen as a convenient number large enough to give reasonable accuracy in the computation. As evident from equations (3) and (4) the only information needed for the calculation of the Theil index is

1) income shares and
ii) population shares

for a defined number of population groups. In the case of 20 quantiles, 

\[ n = 20 \]

and, in principle, the index is calculated as if there were 20 individuals in the population. The magnitudes of the Theil indices for the different years and different sized samples are therefore directly comparable in this study.

1.32 The Atkinson Index

Atkinson's index was originally suggested by Atkinson (1970) and is different in that it is based on a social welfare function evaluating alternative income distributions. The principle is to measure the departure from the existing per capita income from that "equally distributed income" which would achieve the same level of social welfare as the given distribution. The definition is independent of the social welfare function chosen but Atkinson restricts his set of welfare functions to additively separable and symmetric functions. The "equally distributed income" \( z_E \) is defined by 

\[ n U(z_E) = \sum_{i=1}^{nE} U(z_i) \]  

(5)

where the right hand side \( U \) is the social welfare function chosen. Then the Atkinson index

\[ A = 1 - \frac{z_E}{z} \]  

(6)

---

where $\bar{z}$ is the mean income of the distribution of incomes

$$(z_1, z_2, \ldots, z_n)$$

as before.

For the function to satisfy the first condition, i.e. be mean-independent, $U(z)$ must be of the constant marginal utility form i.e.

$$U(z) = \begin{cases} 
\frac{1}{1-\varepsilon} z^{1-\varepsilon} & \varepsilon \neq 1 \\
\log z & \varepsilon = 1. \end{cases}$$

For values of $\varepsilon > 0$, $U(z)$ is concave

and $Z_E < \bar{z}$.

$\varepsilon$ measures "inequality aversion": the greater that $\varepsilon$ is more weight is attached to income transfers to the poor. As reported by Anand (1982) there seems to be general agreement at placing the value of $\varepsilon$ between 1.5 and 2.5: EQUALISE uses $\varepsilon = 2.0$. \footnote{The Atkinson index cannot be computed for values $\varepsilon \geq 1$ of $|\varepsilon|$ if there are zero incomes in the sample.}

Now $A = 1 - \frac{Z_E}{\bar{z}}$

Hence, the lower the $Z_E$, or greater the departure of $Z_E$ from $\bar{z}$, greater is the implied welfare loss from inequality.

1.4 Decomposition of Inequality

A special interest in this study is the measurement of spatial inequality. But the problem is a general one where one wants to measure the inequality between groups as opposed to interpersonal inequality. The traditional method of doing this has been the measurement of differences
between the means of incomes between the groups. This naturally ignores
the distribution of incomes within the groups and does not give a sense of
the contribution of between group inequality to total inequality. It is
quite possible, for example, for within group inequality to be so large
for each group that even with widely differing means between the groups,
the contribution of between group inequality to total inequality would be
small. This was found to be the case, for example, for inter-regional
inequality in Colombia, by Fields and Schultz (1977). Obviously the more
homogeneous each group is higher the contribution of between group
inequality. The decomposition of inequality between and within groups can
therefore also be used as a test for the degree of homogeneity within a
group. The decomposition of labour earnings, for example, as has been
done in Section IV, between occupations and other personal characteristics
is useful in analysing the homogeneity of earnings within occupations.

Essentially, the between group component of inequality is the
value of the inequality index when the within group inequality is
suppressed. In the methodology for the Theil index outlined above as used
in EQUALISE, for example, the index is calculated by dividing the
population into quantiles. The index calculated is essentially the
between group inequality between quantiles assuming that everyone in the
quantile earns the mean of the quantile. Overall inequality is therefore
somewhat understated - the larger the number of quantiles, the smaller
would be the error. Similarly, the within group component is then the
inequality assuming that there is no between group inequality - for a mean
independent index it is merely a weighted average of the within group
inequalities. The weighting in the Theil index is by the income shares of each group while the weighting in the log-variance is by population shares. Both the log-variance and Theil indices are additive decomposable i.e. the sum of the between group and within group inequalities gives the overall inequality. This is not true for the Gini coefficient.

Decomposition of the Theil Index

To understand the decomposition of the Theil index, it is useful to recall the derivation of the Theil index as an information expectation.1/ Consider n events $E_1, E_2 ... E_n$ and G sets of events $S_1, S_2, ... S_G$ such that each set $S_g$ has $k_g$ events where

$$n = \sum_{g=1}^{G} k_g$$

The probability of the set $S_g$ occurring is

$$P_g = \prod_{i \in S_g} p_i$$

and the entropy at the level of sets is

$$H_0 = \sum_{g=1}^{G} P_g \log \frac{1}{P_g}$$

This is "between group entropy n - analogous to between group inequality.

Now "total" entropy at the level of events is

$$H = \sum_{i=1}^{n} p_i \log \frac{1}{p_i}$$

1/ From Theil (1972) pp. 20-22.
\[
= \sum_{g=1}^{G} \sum_{i \in S_g} p_i \log \frac{1}{p_i}
\]
\[
= \sum_{g=1}^{G} p_g \sum_{i \in S_g} \frac{p_i}{p_g} \left( \log \frac{1}{p_g} + \log \frac{p_g}{p_i} \right)
\]
\[
= \sum_{g=1}^{G} p_g \sum_{i \in S_g} \frac{p_i}{p_g} \log \frac{1}{p_g}
\]
\[
\quad + \sum_{g=1}^{G} p_g \left( \sum_{i \in S_g} \frac{p_i}{p_g} \log \frac{p_g}{p_i} \right)
\]
\[
= \sum_{g=1}^{G} p_g \log \frac{1}{p_g} + \sum_{g=1}^{G} p_g \left( \sum_{i \in S_g} \frac{p_i}{p_g} \log \frac{1}{p_i} \right).
\]

i.e.
\[
H = H_0 + \sum_{g=1}^{G} p_g H_g \tag{10}
\]

where
\[
H_g = \sum_{i \in S_g} \frac{p_i}{p_g} \log \frac{1}{p_i/p_g} \tag{11}
\]

\(H_g\) is the entropy within set \(S_g\) and hence \(\sum p_g H_g\) is the average within group entropy. The total entropy is then the sum of between group entropy \(H_0\) and the average within group entropy — exactly analogous to the within and between group inequality decomposition.

This result can be extended to the decomposition of information expectation of a message transforming prior probabilities \(p_i\) to \(q_i\).

Let
\[
P_g = \sum_{i \in S_g} p_i, \quad Q_g = \sum_{i \in S_g} q_i
\]
with \(p_i \ (i = 1, \ldots, n)\) and \(q_i \ (i = 1, 2, \ldots, n)\) being the prior and
posterior probabilities respectively which are aggregated into G sets correspondingly. At the set level, the information expectation is then

\[ I_o(q:p) = \sum_{g=1}^{G} Q_g \log \frac{Q_g}{P_g} \]  

(12)

but

\[ I(q:p) = \sum_{i \neq 1}^{n} q_i \log \frac{q_i}{p_i} \]

\[ = \sum_{g=1}^{G} Q_g \left( \sum_{i \in g} q_i \log \frac{Q_g}{P_g} + \log \frac{q_i}{p_i} \right) \]

\[ = \sum_{g=1}^{G} Q_g \log \frac{Q_g}{P_g} + \sum_{g=1}^{G} Q_g \left( \sum_{i \in g} q_i \log \frac{q_i}{Q_g} \frac{Q_g}{p_i} \right) \]

i.e. \[ I(q:p) = I_o(q:p) + \sum_{g=1}^{G} Q_g I_g(q:p) \]  

(13)

where

\[ I_g(q:p) = \sum_{i \in g} q_i \log \frac{q_i/Q_g}{p_i/P_g} \]  

(14)

hence the expected information of a message transforming prior probabilities to posterior probabilities is decomposable in an exactly analogous manner as the entropy for groups.

We can now return to income inequality where the \( p_i \) are population shares and \( q_i \) are income shares with \( p_i \) now being aggregated into groups of individuals with population shares \( P_g \) and income shares \( Q_g \).
Now
\[ p_G = \frac{n_g}{n} \] where \( n_G \) is the number of individuals in the group and \( n \) the total number of individuals.

\[ Q_g = \sum_{i \in S_g} q_i = \text{Income share of the group.} \]

Hence, using (12), (13), and (14)

\[ T = T_o + \sum_{g=1}^{G} Q_g T_g \] (15)

where \( T_o \) is the between group Theil index, \( T_g \) is the within group Theil index and \( T \) is the overall Theil index

\[ T_o = \sum_{g=1}^{G} Q_g \log \left( \frac{Q_g}{p_g} \right) \] (16)

and

\[ T_g = \sum_{i \in S_g} \frac{q_i}{Q_g} \log \left( \frac{q_i/Q_g}{p_i/p_g} \right) \] (17)

Hence,

\( Q_g \) is the income share of group \( S_g \), and \( p_g \) is the population share of group \( S_g \).

Thus the overall Theil index is the sum of the between group Theil index (or contribution to inequality) and a weighted average of the within group Theil indices - the weights used being the income shares.
Q_g. Income shares are used because they are regarded as the posterior probabilities. When the groups used are quantiles, it is easily seen that the within group inequality is suppressed and regarded as zero because everyone is taken to receive the mean income within the quantile.

The decomposition of log variance may be done just like the decomposition of any variance into between group and within group components.\footnote{Anand (1982) gives a good exposition of the derivation of the decomposition.} It may be noted, however, that the weighting of the within group components, unlike for the Theil decomposition, is by the population shares of each group. Log variance is also additively decomposable such that total inequality is the sum of between group and within group inequality. The log variance is very sensitive to small changes at low income levels - hence the inclusion of one or a few very low incomes in a sample would tend to heighten inequality as measured by log variance. One advantage of using log variance is that the ratio of between-group and within-group variance follows an F-Distribution. Hence the ratio can be tested for statistical significance. The measures described here are essentially descriptive - they have normative content. It is difficult to find any means of evaluating the comparative magnitudes of the percentage between group contributions to total inequality as measured by the Theil index and the log variance. Hence both measures are given in this study: the closeness of results between the two then gives greater confidence in the results and the converse.
II. OVERALL TRENDS IN THE DISTRIBUTION OF INCOME

2.1 Trends in Colombia as a Whole

Good data on the distribution of income in Colombia have become available only in the last 15 years or so. The 1973 census was the first census in the country to include a question on income. The first household surveys were conducted by CEDE in the mid to the late sixties but these were largely confined to urban areas. DANE conducted a nationwide household survey in 1970 which probably gave the first real nationwide estimates of income distribution. It is therefore difficult to make comparisons of the current distribution of income with historical trends. However, in a pioneering book on the subject, Berry and Urrutia (1976) pieced together information from a multitude of sources in an attempt to reach conclusions on trends since the late nineteen thirties.

They used 1964 as the bench mark year and derived the distribution of income for the economically active population. The 1964 census provided a generally accurate distribution of occupations and activities. They used the 1967-1969 CEDE unemployment survey to derive income estimates for urban areas and applied the 1964 occupational distribution to obtain a notional distribution of earnings in urban areas. For rural areas they used the agricultural census of 1960 and various later sources for the
distribution of land and of agricultural production. Table 1 summarizes their findings for rural areas, urban areas and the country as a whole. Their results were somewhat surprising in that they revealed a distribution of income in the country which was among the most unequal in the world. The top 5 percent of income earners earned as much as 30 - 40 percent of total income and the Gini coefficient ranged from 0.55 to 0.57. Inequality was somewhat higher in rural areas than in urban areas due to a highly skewed pattern of land holdings. In an international comparison reported by Berry and Urrutia themselves, only one estimate for Brazil reported an income distribution more unequal than that in Colombia.

Selowsky (1979) conducted a nationwide survey in 1974 and his results on income distribution for the population as a whole ranked by HINCAP are also reported in Table 1. The distribution derived by him was substantially better for rural areas than for urban areas: a finding contrary to the Berry-Urrutia estimates for 1964. Few conclusions can be drawn about trends from these two studies, however, because the Berry-Urrutia data were estimates pieced together from diverse set of sources while the Selowsky estimates come from one survey. His sample size was small and questions have also been raised about the quality of responses to the sole income question in his survey.1

1/ 4,000 thousand households were sampled in which rural areas were somewhat underrepresented. (Selowsky 1979, p. 15).
<table>
<thead>
<tr>
<th></th>
<th>1964(^1) Economically Active Population(^4)</th>
<th>1974(^2) Population(^3) Malaysia(^5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Urban        Rural         Total            Urban        Rural         Total        Total</td>
<td></td>
</tr>
<tr>
<td>Bottom 20 percent</td>
<td>2.5          4.1           4.5             3.2          5.4           3.6            4.3</td>
<td></td>
</tr>
<tr>
<td>Bottom 40 percent</td>
<td>9.8          10.8          9.6             9.9          15.8          10.8           12.3</td>
<td></td>
</tr>
<tr>
<td>Top 20 percent</td>
<td>49.9         63.5          63.1            60.6         49.4          60.2           54.8</td>
<td></td>
</tr>
<tr>
<td>Top 5 percent</td>
<td>27.1         40.4          33.7            31.5         24.4          32.8           28.5</td>
<td></td>
</tr>
<tr>
<td>Gini Coefficient</td>
<td>0.55         0.57          0.57            0.54         0.42          0.50           0.50</td>
<td></td>
</tr>
</tbody>
</table>

Notes:
1. From Selowsky (1979) Table 2.2 derived from Berry and Urrutia (1976).
2. From Berry and Soligo (1980) Table 1.1 derived from Selowsky (1979).
3. Population ranked by Household Income per Capita (HINCAP).
4. Economically active population ranked by earnings.
5. From Anand (1982) Table 3.2 for individuals ranked by HINCAP.
The two estimations are agreed on one point: The degree of income inequality in Colombia is high and has been for some time with Gini coefficients ranging between 0.50 and 0.57 for the country as a whole. For purposes of comparison, Table 1 also gives estimates for Malaysia in 1970. In comparison, the distribution for Colombia is significantly worse for 1964 but only marginally so for 1974.

Berry and Urrutia also examined the available data going back up to the 1930s. They essentially compiled wage series for various groups of workers and then attempted to derive the income distribution for the population. Their main conclusions were

"(1) That income distribution in agriculture has worsened throughout the period since the mid-1930s.

(2) That the non-agricultural income distribution probably worsened from the mid 1930s until some time in the early 1950s, then improved till some time in the mid 1960s, and then tended to level off." 1/

They speculate that the distribution of income generally tends to worsen during periods of rapid growth while it improves during periods of slow growth. The late 1940s and the early 1950s were a period of rapid industrialization and creation of new industries under conditions of substantial protection. If these protected industries can be assumed to be relatively highly capital intensive such a pattern of industrialization could be expected to

1/ Berry and Urrutia (1976) p. 89.
lead to a lower labor share of total earnings. Protected industries would, moreover, tend to generate monopoly profits. With the ownership of capital being highly skewed, such industrialization can therefore be expected to lead a worsening of income distribution. Rural-urban migration was also very high during this period (Mohan, 1980) and particularly in the late forties up to the mid fifties during the period of the "violencia". Thus unskilled, blue collar wages may have been kept down during this period because of this unlimited supply of labor into the cities. The burst of industrialization also created a sudden demand for highly skilled and educated workers and consequently their real earnings rose at very high rates between 1945 and 1953. The whole period from the 1930s to the mid 1950s can then be characterized as one involving major structural changes in the Colombian economy, in particular in the shift from a predominantly agricultural rural economy to an urban based increasingly industrial and service economy. These structural changes continued apace until the mid 1960s but the later period witnessed dramatic increases in blue collar wages. Berry-Urrutia hypothesize that during the later period the degree of protected import substitution may have decreased and that small scale competitive enterprises may have become more successful. The rapid expansion of primary education in the 1950s may have also contributed to higher average worker skill levels. The later period between the mid 1950s and mid 1960s was one of slower economic growth and may be regarded as a period of consolidation. That lower skilled workers' incomes rose at high rates is surprising since rural-urban migration continued at high rates.
Since the mid 1960s Colombia has had a relatively high and sustained real rate of growth of g.n.p. (about 5.5. percent a year) but it is difficult to draw any firm conclusions about the trend of the distribution of income from the available data. With declines in rates of fertility and the progressively smaller proportion of people left in the rural areas, the rate of rural-urban migration and the rate of urbanization have slowed down. With the coffee boom and the "success" of the drug trade, it may be that incomes of rural unskilled workers have been rising -- and consequently those of urban workers as well. As has been shown elsewhere, it is clear that between 1973 and 1978 the earnings of the most unskilled categories of the urban labor force have indeed been increasing at relatively high rates. It is possible that a slow down in urbanization rates has begun to improve the overall distribution of income.

2.2 The Distribution of Income in Bogota and Cali 1973 - 1978

Unlike most other studies of income distribution which usually take a country as the unit of observation this study is mainly concerned with the urban income distribution as evidenced from data available from Bogota and Cali in Colombia. Table 2(A) gives shares of income for Bogota households ranked by household income for different years. The evidence is mixed but there is some indication that the distribution of income may have improved during this period. As has been documented

elsewhere \(^1\) the key problem in comparing estimates from different years is that the overall coverage of income is different in different surveys. The 1973 census seemed to have covered only about 50 percent of total personal income while the 1978 survey may have covered as much as 90 percent. It is, however, difficult to find out whether the undercoverage is consistent across income groups or if it is skewed towards particular groups. The evidence is conflicting. An examination\(^1\) of the labour earnings of the lowest skilled occupational categories suggests that the census may have covered the earnings of poor male workers well. It does appear, however, that many working females were not captured by the census. The participation rates of women in the census are substantially lower in 1973 than in other years. If it is the case that this under-coverage is greater for women in the lower income households then the income coverage of these households would be lower as compared with the richer groups. Since we know that there is substantial undercoverage over all, it is probably the case that the under coverage at the low end because of the neglect of some female workers does not cause atypical under-coverage of income among those households. Moreover, participation rates at the low income end are low anyway.

Table 2 shows the shares of income received by selected decile groups. It reflects the striking inequality that exists in Colombia. The income share of the bottom 40 percent of households

\(^1\) See Mohan, Wagner and Garcia (1981)
is less than half of the share received by the top 5 percent.

Even in 1978, when the income coverage is estimated to over 90 percent, it appeared that there was substantial under-coverage of income from capital. This income is likely to accrue to the richer households. Consequently, the actual distribution might be even worse than that apparent from Table 2. The poor have, however, gained from economic growth at least proportionately if not at somewhat higher rates than the rich. The picture is quite similar for Cali. There appears to be a slight worsening between 1977 and 1978 but it is difficult to distinguish changes due to sample errors from real changes. That there may have been real changes would be supported by the fact that the worsening seems to be very similar for both cities.

Table 2(B) gives similar results but here individuals are ranked by household income per capita (HINCAP). As shown by Datta and Meerman (1980) the rankings of particular individuals change substantially when ranked by HINCAP as opposed to household income (HHY), but the overall result does not. The particular rankings are clearly important when the aim is to find out the characteristics of the poor or of the rich in order to design any policies aimed at poverty alleviation. The distribution itself does not appear to be different. There seems to be overall improvement between 1973 and 1978 but, again, there is a suggestion of worsening after 1977.

---

1/ As, for example, in Mohan and Hartline (1980).
Table 2: THE DISTRIBUTION OF INCOME IN BOGOTA AND CALI: INCOME SHARES 1973-1978

A. HOUSEHOLDS RANKED BY HOUSEHOLD INCOME

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottom 20 percent</td>
<td>3.3</td>
<td>3.8</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td>Bottom 40 percent</td>
<td>9.8</td>
<td>11.0</td>
<td>11.6</td>
<td>11.9</td>
</tr>
<tr>
<td>Top 20 percent</td>
<td>62.5</td>
<td>56.9</td>
<td>56.3</td>
<td>55.5</td>
</tr>
<tr>
<td>Top 5 percent</td>
<td>30.2</td>
<td>24.9</td>
<td>23.8</td>
<td>25.1</td>
</tr>
</tbody>
</table>

B. INDIVIDUALS RANKED BY HOUSEHOLD INCOME PER CAPITA

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottom 20 percent</td>
<td>3.3</td>
<td>3.9</td>
<td>4.2</td>
<td>4.0</td>
</tr>
<tr>
<td>Bottom 40 percent</td>
<td>9.9</td>
<td>11.5</td>
<td>12.0</td>
<td>11.5</td>
</tr>
<tr>
<td>Top 20 percent</td>
<td>62.6</td>
<td>57.4</td>
<td>56.0</td>
<td>58.0</td>
</tr>
<tr>
<td>Top 5 percent</td>
<td>30.8</td>
<td>24.0</td>
<td>24.5</td>
<td>29.0</td>
</tr>
</tbody>
</table>

Sources:
1973 Population Census
1975 DANE Household Survey EH8E.
1977 DANE Household Survey EH15.
Tables 3 and 4 present various inequality indices for all the years for Bogota and Cali respectively. The results are surprisingly consistent over different income concepts, different ranking procedures and over the two cities. The Gini coefficient ranges from 0.57 to 0.50 and Theil index from 0.68 to 0.44. The distribution of households appears to be worse if ranked by household income per capita rather than household income itself. A similar result was obtained by Anand (1982) for Malaysia as a whole but he had found that the plotted Lorenz curves for the two distributions (Households by HHY and Households by HINCAP) intersected each other. He therefore concluded that the HINCAP distribution could not be said to be unambiguously worse and that the use of different indices could produce different results. Here it may be noted that all the indices used yield higher inequality of households when ranked by HINCAP. This implies that even though households with higher incomes tend to be of larger size, the large size is not compensating enough and the distribution of HINCAP becomes worse than HHY because of the reordering of households. Since HINCAP is a better measure of welfare than HHY (as argued by Anand (1980), among others) the inequality in HINCAP is the more relevant indicator of disparities in welfare. If we are concerned with the welfare of individuals, rather than households, then the ranking of individuals by HINCAP is the one of greater interest. The results in Table 3b are as might be expected. The inequality is somewhat less than for households ranked by HINCAP but it is worse than households ranked by HHY. As was shown in Mohan (1980), household income increases with household size but HINCAP decreases.
Table 3a: INEQUALITY INDICES FOR BOGOTA HOUSEHOLDS 1973-1978

<table>
<thead>
<tr>
<th>Inequality Index</th>
<th>Ranked by Household Income</th>
<th>Ranked by Household Income per Capita</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gini</td>
<td>0.565</td>
<td>0.514</td>
</tr>
<tr>
<td>Theil</td>
<td>0.610</td>
<td>0.467</td>
</tr>
<tr>
<td>Standard Deviation of Logs</td>
<td>1.023</td>
<td>0.951</td>
</tr>
<tr>
<td>Atkinson</td>
<td>0.654</td>
<td>0.586</td>
</tr>
<tr>
<td>Coefficient of Variation</td>
<td>1.403</td>
<td>1.142</td>
</tr>
</tbody>
</table>

Mean Household Income 3/ 5/

Mean Household Income per Capita

Sample Size (No. of Households)

Expanded Sample Size 4/ (Households) (thousands)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Household Income 3/ 5/</td>
<td>3,323</td>
<td>5,692</td>
<td>8,229</td>
<td>13,405</td>
</tr>
<tr>
<td>Mean Household Income per Capita</td>
<td>895</td>
<td>1,249</td>
<td>1,800</td>
<td>3,683</td>
</tr>
<tr>
<td>Sample Size (No. of Households)</td>
<td>41,282</td>
<td>3,620</td>
<td>2,934</td>
<td>2,991</td>
</tr>
<tr>
<td>Expanded Sample Size 4/ (Households) (thousands)</td>
<td>440,000</td>
<td>636,000</td>
<td>696,000</td>
<td></td>
</tr>
</tbody>
</table>

Sources: As in Table 2.

Notes:

1/ Mean taken over households.
2/ Not expanded because expansion factors were not available.
3/ Current Colombian pesos.
4/ i.e. Estimated No. of Households in Bogota.
5/ Consumer Price Index (1970 = 100) (approx.)
   1975 - 150
   1975 - 240
   1977 - 380
   1978 - 400
Table 3b: INEQUALITY INDICES FOR BOGOTA'S POPULATION RANKED BY HOUSEHOLD INCOME PER CAPITA 1973-1978

<table>
<thead>
<tr>
<th>Inequality Index</th>
<th>1973</th>
<th>1975</th>
<th>1977</th>
<th>1978</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gini</td>
<td>0.568</td>
<td>0.515</td>
<td>0.499</td>
<td>0.522</td>
</tr>
<tr>
<td>Theil</td>
<td>0.649</td>
<td>0.499</td>
<td>0.442</td>
<td>0.508</td>
</tr>
<tr>
<td>Standard Deviation of Logs</td>
<td>1.007</td>
<td>0.949</td>
<td>0.911</td>
<td>0.920</td>
</tr>
<tr>
<td>Atkinson</td>
<td>0.646</td>
<td>0.601</td>
<td>0.561</td>
<td>0.570</td>
</tr>
<tr>
<td>Coefficient of Variation</td>
<td>1.582</td>
<td>1.309</td>
<td>1.117</td>
<td>1.255</td>
</tr>
</tbody>
</table>

Mean Household Income Per Capita

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>697</td>
<td>1,012</td>
<td>1,573</td>
<td>2,844</td>
</tr>
</tbody>
</table>

Expanded Sample Size (Individuals) (millions)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.899</td>
<td>2/</td>
<td>3.328</td>
<td>3.279</td>
</tr>
</tbody>
</table>

Sources: Same as Table 2.

Notes

1/ Mean taken over individuals.
2/ Not expanded because expansion factors were not available.
3/ Current Colombian pesos.
4/ Estimated population of Bogota.
5/ The estimated population for Bogota is less in 1978 as compared with 1977 because of differences in the expansion factors used. The 1978 survey estimates are based on a new sample frame taken during that year, while the 1977 estimates relied in an old sample frame merely blown-up every year.
6/ See footnotes to Table 3a for Consumer Price Indices from 1973 to 1978.
The increase in income earners does not compensate adequately for the increase in dependents -- on average. Nonetheless, these results are reassuring from a methodological point of view. The level of inequality deduced from either of these income concepts and ranking criteria gives somewhat similar results. Moreover, the changes over different years and different income concepts are also consistent for the different indices. All show an improvement between 1973 and 1977 and slight worsening between 1977 and 1978.

Table 4 gives the comparable results for Cali but reports the Gini and Theil indices only. The conclusions are the same as for Bogota but note that there is less overall inequality in Cali. One would expect that a larger city has more higher paid specialized professionals as well as a larger number of people with high non-labour income residing in it. As mentioned earlier, the coverage of non-labour income is not good in all the sources of data considered, hence the former reason is regarded as the more likely explanation. This is corroborated by the analysis of the distribution of labour earnings in Section IV.

One technical detail about the computation of inequality indices is worth a mention here. The usual method for calculating inequality indices is to group the population in quantiles and then to use the resulting piecewise linear Lorenz curve for calculating the relevant indices. Usually data on income distribution are available in one of two formats: either income shares of different quantile groups or
Table 4: INEQUALITY INDICES FOR CALI 1973-1978

<table>
<thead>
<tr>
<th>Inequality Index</th>
<th>1973</th>
<th>1977</th>
<th>1978</th>
</tr>
</thead>
<tbody>
<tr>
<td>Households Ranked by Household Income</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gini</td>
<td>0.553</td>
<td>0.475</td>
<td>0.487</td>
</tr>
<tr>
<td>Theil</td>
<td>0.601</td>
<td>0.403</td>
<td>0.429</td>
</tr>
<tr>
<td>Households Ranked by Household Income per Capita</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gini</td>
<td>0.568</td>
<td>0.487</td>
<td>0.524</td>
</tr>
<tr>
<td>Theil</td>
<td>0.665</td>
<td>0.430</td>
<td>0.510</td>
</tr>
<tr>
<td>Individuals Ranked by Household Income per Capita</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gini</td>
<td>0.554</td>
<td>0.481</td>
<td>0.500</td>
</tr>
<tr>
<td>Theil</td>
<td>0.632</td>
<td>0.420</td>
<td>0.470</td>
</tr>
</tbody>
</table>

Mean Household Income\(^3/\) \(^6/\)

\[
\begin{array}{ccc}
\text{Mean Household Income per Capita} & 2,632 & 7,192 & 11,321 \\
\text{Sample Size} & 11,520 & 1,016 & 974 \\
\end{array}
\]

Expanded Sample Households and Individuals\(^1/\)

\[
\begin{array}{ccc}
\text{Households} & 1/ & 191,421 & 216,633 \\
\text{Individuals} & 1/ & 1,019,766 & 1,011,298 \\
\end{array}
\]

Sources: Same as in Table 2.

Notes:

1/ Not expanded because of absence of expansion factors.
2/ Mean taken over individuals.
3/ Current Colombian Pesos.
4/ Estimated No. of Households and Population in Cali.
5/ See footnote 5 to Table 5b. The same applies to Cali.
6/ See footnote 5 to Table 3a for Consumer Price Indices from 1973 to 1978.
population (or household) group falling within specified income ranges. An approximation of the Lorenz curve is obtained from either of these two kinds of data. The essential point is that data are usually available in a form such that only a small number (10 to 20) of observations can be plotted on the Lorenz curve. A novel feature of this study is that income data were available for individual households from as many as four different surveys, with one of them being a census sample. It is therefore of interest to find out the nature of error involved when as many as 3000 available observations on income are grouped into 10 to 20 quantile groups. The program \textbf{EQUALISE} (described in Appendix III) provides for flexibility in the number of observations used on the Lorenz curve.

It is expensive to utilize all the information available from a sample survey: all observations need to be ranked for an index to be computed. Typically, then, quantile averages are taken and the relevant indices calculated. Table 5 gives some idea of the error involved in using 20 quantiles rather than all observations. As would be expected, inequality is slightly understated by the use of quantiles. The Gini coefficient and the standard deviation of logarithms appear to be relatively insensitive to the number of observations used. Clearly, the coefficient of variation is affected significantly by outliers. The Theil index and the Atkinson index are also affected but not as much. It is difficult to find the optimal level of aggregation but Table 5 indicates that the errors caused by the use of 20 quantiles are probably tolerable for most purposes.
Table 5: INEQUALITY INDICES FOR BOGOTA 1978
RESULTS FROM USING DIFFERENT NUMBERS OF OBSERVATIONS IN THE LORENZ CURVE

<table>
<thead>
<tr>
<th>Inequality Index</th>
<th>Households by Household Income</th>
<th>Households by HINCAP</th>
<th>Individuals by HINCAP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20 Obs.</td>
<td>All 1/</td>
<td>20 Obs.</td>
</tr>
<tr>
<td>Gini</td>
<td>0.507</td>
<td>0.511</td>
<td>0.558</td>
</tr>
<tr>
<td>Theil</td>
<td>0.458</td>
<td>0.477</td>
<td>0.589</td>
</tr>
<tr>
<td>Standard Deviation of Logs</td>
<td>0.926</td>
<td>0.953</td>
<td>0.996</td>
</tr>
<tr>
<td>Atkinson</td>
<td>0.569</td>
<td>0.647</td>
<td>0.626</td>
</tr>
<tr>
<td>Coefficient of Variation</td>
<td>1.146</td>
<td>1.229</td>
<td>1.351</td>
</tr>
</tbody>
</table>

1/ About 3,000 observations.

2/ About 17,000 observations.
In summary, the levels of overall inequality found in Bogota and Cali are not substantially different from earlier estimates reported in the last section. There is no striking move towards greater equality though there do appear to be some tendencies toward slight improvement. The distribution of income in Cali is somewhat better than in Bogota. The use of different indices, nor of different income concepts and nor of different rankings causes any changes in these broad conclusions.
III. THE SPATIAL DISTRIBUTION OF INCOME

The last section demonstrated the great degree of inequality that exists in Bogota and Cali. The issue being addressed now is: how are these incomes distributed across space within these cities and are there any discernible trends in spatial inequality? The earlier papers\(^1\) have established descriptively that both Bogota and Cali exhibit very distinct patterns of income segregation by space. It was shown that the ratio between mean incomes in different zones of the city could be as much as 1 to 6: means being taken over relatively large zones. This kind of spatial disparity in incomes is even unusual between the different regions of a country (though it exists in Colombia itself as documented by Fields and Schultz (1977)). There is very little information of this type for other cities so it is difficult to say if such a pattern of income differentiation within a city is unusual. Merely observing differences between means of different populations can be misleading, however, and it is therefore necessary to examine the full distributions.

Discussing the extent of spatial inequality within a city is difficult because there are no natural units for analysis. Unlike the world, which can be divided into countries, or a country which can be divided into states, there are no natural divisions within a city. In the U.S. much discussion is conducted by contrasting the characteristics of central cities with suburbs - but even in this case, there is no natural definition of what constitutes a central city. It is clear then that

any division of a city would be arbitrary but we need to find a
zonification system which is convenient for analysis.

Two geometric patterns suggest themselves as possible
methods of zonification. For a circular city, we can either have rings,
or radial sectors. In the U.S., for example, suburbs are often
considered to be quite distinct from the central city. Urban economic
theory1 also suggests that under certain assumptions the poorest
would be expected to locate themselves in the centre and incomes
would increase as one proceeds outwards from the centre. In such a
case it would be useful to divide the city into rings and to study
the spatial distribution of income in terms of these rings. The
other obvious pattern, that of radial sectors was suggested by Homer Hoyt
(1939, 1966) among others. The city can be divided into pie slices
and the distribution of income studied in terms of these divisions.
He had traced the historical development of a large number of North
American cities and concluded that income groups tended to
locate themselves near like groups. Hence, if the rich historically
located themselves in one section of the city, the development of the
city took place such that the rich continued to locate themselves in
the same direction as the city continued to expand. Such a historical
process resulted in different pie slices or radial sectors emerging
with different characteristics. Peter Amato (1968) focused on the
elite of Bogota and showed how they have tended to locate predominantly
in the North of Bogota and have continued to move in that direction.

1/ e.g. Muth (1969)
Maps (1) and (2) illustrate the geography of Bogota and Cali and the spatial disaggregation system is described in Appendix (II). Both the cities can be divided conveniently into semi-circular rings and radial sectors. There is no good reason why Bogota has been divided into eight radial sectors and six rings except that the DANE geo-coding system makes these numbers the obvious choices. The rings and sectors are constituted from smaller units - comunas - of which there are 38 in Bogota and 25 in Cali. The discussion of spatial inequality is then conducted in terms of these geographical divisions which are particular to these cities but the principle can be generalized to other cities.

Tables 6a and 6b describe the distribution of mean incomes and of the population in Bogota and Cali by radial sectors from 1973 to 1978. That the proportions of population in each sector are similar in all the years gives confidence in the sample distribution of the household surveys. The sample frame of the 1978 sample was updated and the distribution indicates that sector 8 may have gained in terms of population over the period 1973 to 1978. The 1975 and 1977 surveys were samples drawn from the same basic sample frame so no conclusions can be drawn from them on any change in the spatial distribution of population. The mean household income per capita is given in terms of the overall mean. The multiples remain broadly similar over the whole time period. HINCAP of the poorest sector - sector 2 in the south of the city - is about a fifth or sixth of the richest - sector 8 in the north of the city. With the exception of sector 6, mean incomes increase as one moves clockwise from the South of Bogota toward
Table 6a: THE SPATIAL DISTRIBUTION OF INCOME
IN BOGOTA 1973-1978

<table>
<thead>
<tr>
<th>Sector</th>
<th>1973 % Pop.</th>
<th>Mean HINCAP</th>
<th>1975 % Pop.</th>
<th>Mean HINCAP</th>
<th>1977 % Pop.</th>
<th>Mean HINCAP</th>
<th>1978 % Pop.</th>
<th>Mean HINCAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sector 1 (city center)</td>
<td>2.3</td>
<td>1.07</td>
<td>1.2</td>
<td>2.05</td>
<td>1.9</td>
<td>1.38</td>
<td>2.4</td>
<td>0.80</td>
</tr>
<tr>
<td>Sector 2</td>
<td>18.2</td>
<td>0.57</td>
<td>16.5</td>
<td>0.66</td>
<td>18.6</td>
<td>0.56</td>
<td>20.6</td>
<td>0.50</td>
</tr>
<tr>
<td>Sector 3</td>
<td>26.0</td>
<td>0.68</td>
<td>28.2</td>
<td>0.71</td>
<td>25.5</td>
<td>0.76</td>
<td>25.2</td>
<td>0.73</td>
</tr>
<tr>
<td>Sector 4</td>
<td>9.3</td>
<td>0.90</td>
<td>9.0</td>
<td>0.92</td>
<td>8.2</td>
<td>1.08</td>
<td>7.2</td>
<td>0.98</td>
</tr>
<tr>
<td>Sector 5</td>
<td>7.3</td>
<td>0.93</td>
<td>6.5</td>
<td>1.45</td>
<td>6.5</td>
<td>1.11</td>
<td>6.2</td>
<td>1.03</td>
</tr>
<tr>
<td>Sector 6</td>
<td>18.2</td>
<td>0.80</td>
<td>20.0</td>
<td>0.95</td>
<td>19.0</td>
<td>0.87</td>
<td>20.0</td>
<td>0.91</td>
</tr>
<tr>
<td>Sector 7</td>
<td>12.0</td>
<td>1.71</td>
<td>10.3</td>
<td>1.38</td>
<td>13.5</td>
<td>1.39</td>
<td>9.1</td>
<td>1.44</td>
</tr>
<tr>
<td>Sector 8</td>
<td>6.6</td>
<td>2.87</td>
<td>8.2</td>
<td>1.90</td>
<td>7.0</td>
<td>2.33</td>
<td>9.1</td>
<td>2.66</td>
</tr>
<tr>
<td>Total Mean</td>
<td>100.0</td>
<td>697.0</td>
<td>100.0</td>
<td>1012.0</td>
<td>100.0</td>
<td>1573.0</td>
<td>100.0</td>
<td>2843.0</td>
</tr>
</tbody>
</table>

Notes:

1/ Percent of the city's total population living in Sector.
2/ Mean Household Income Per Capita taken across individuals in the Sector as a multiple of Overall Mean HINCAP.
3/ Mean Household Income per Capita taken across individuals for Bogota in current Colombian Pesos.

Sources: Same as Table 2.
Table 6b: THE SPATIAL DISTRIBUTION OF INCOME IN CALI 1973-1978

<table>
<thead>
<tr>
<th>Sector</th>
<th>1973</th>
<th>1977</th>
<th>1978</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% Pop in sector</td>
<td>Mean HINCAP</td>
<td>% Pop in sector</td>
</tr>
<tr>
<td>City Center (Sector 1)</td>
<td>4.3</td>
<td>1.03</td>
<td>3.2</td>
</tr>
<tr>
<td>Sector 2</td>
<td>4.2</td>
<td>3.68</td>
<td>7.3</td>
</tr>
<tr>
<td>Sector 3</td>
<td>19.5</td>
<td>0.89</td>
<td>13.4</td>
</tr>
<tr>
<td>Sector 4</td>
<td>21.2</td>
<td>0.69</td>
<td>16.8</td>
</tr>
<tr>
<td>Sector 5</td>
<td>35.1</td>
<td>0.68</td>
<td>37.1</td>
</tr>
<tr>
<td>Sector 6</td>
<td>11.1</td>
<td>1.61</td>
<td>16.5</td>
</tr>
<tr>
<td>Sector 7</td>
<td>4.5</td>
<td>1.37</td>
<td>2.7</td>
</tr>
<tr>
<td>Total Mean</td>
<td>100.0</td>
<td>594</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Notes:

1/ Percent of total population living in Sector.

2/ Mean Household Income per Capita taken across individuals in the sector as a multiple of Overall mean HINCAP

3/ Mean Household Income per Capita taken across individuals in Cali in current Colombian pesos.

4/ See footnote 5 to Table 3a for Consumer Pride Indices for 1973 to 1978.

Sources: Same as in Table 2.
the north. (Sector 1 is the city center). The poor sectors 2, 3, and 6 account for almost 65 percent of the total population. There is no noticeable trend toward a narrowing of the differences in mean income between the sectors. In Cali, the data indicate that some changes might be taking place in that city. First, the division there seems to be more between the Eastern and Western parts of the city: the Eastern sectors, 3, 4 and 5 being relatively poorer than the Western sectors 2, 6, 7. Mean HINCAP in the poorest sector is about a sixth as much as the richest in 1973 and about a quarter in 1974. The multiples are not as stable across time as in Bogota: this may be because of smaller sample sizes and hence larger sample errors. It does appear, though, that sector 2 is becoming relatively less rich. If the changes in sector 1 results are not caused by sampling errors, they indicate that the city center may be gaining higher income people.

These patterns had been established in Mohan (1980) for Bogota. The question then is how much is being hidden behind the means. Are there large variances around these means or do the means capture the essence of the differing characteristics of these sectors. Section 1 has laid out the decomposition method followed in this study to investigate the spatial distribution of income. To recapitulate, the variance in incomes is being decomposed into two parts. Households or people are being grouped by sectors or rings. Overall inequality in the city is then decomposed into within group and between group contributions. The within group contribution is essentially a weighted average of the indices of inequality within each group and the between group contribution is a measure of the differences in
the means appropriately weighted by the group population for log variance and group snares in income for the Theil index.

In both Bogota and Cali, the radial sectors are more distinct in terms of income patterns than the rings. This is confirmed in Table 7. Households are ranked by both household income as well as household income per capita. Two inequality indices are decomposed: the Theil index and log variance. For Bogota more than a quarter of overall inequality as measured by the Theil index may be attributed to space; and about 20 percent as measured by log variance. The spatial contributions are somewhat lower for Cali. The between group contribution is very small when households are grouped by rings. This finding confirms systematically earlier conjectures that the sectors are far more distinct in characteristics than rings in Bogota and Cali. Note that the between group contribution of rings increases substantially when households are ranked by HINCAP: this is due to the fact that the average household size is larger in the periphery than in the center (Mohan, 1980) and hence the household income per capita differences are greater between the rings than total household income. The between group contribution can be regarded as an index of spatial inequality in incomes: it may be termed an index of spatial income segregation (ISIS). Table 7 also shows ISIS for individuals ranked by HINCAP as well as workers ranked by labor earnings. The ISIS for workers is substantially lower than for households and for individuals. This finding indicates that some of the spatial differences are due to locational differences arising from differences in household characteristics rather than spatial differences in the labor market.
Table 7: THE SPATIAL DISTRIBUTION OF INCOME DISTINGUISHED BY RINGS AND SECTORS: BOGOTA AND CALI

<table>
<thead>
<tr>
<th>Ranking Criterion</th>
<th>Zonification System</th>
<th>Bogota Theil</th>
<th>Bogota Log Variance</th>
<th>Cali Theil</th>
<th>Cali Log Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Households</td>
<td>Rings</td>
<td>4</td>
<td>3</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Ranked by Household Income</td>
<td>Sectors</td>
<td>26</td>
<td>17</td>
<td>23</td>
<td>13</td>
</tr>
<tr>
<td>Households</td>
<td>Rings</td>
<td>9</td>
<td>6</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Ranked by HINCAP</td>
<td>Sectors</td>
<td>32</td>
<td>22</td>
<td>21</td>
<td>14</td>
</tr>
<tr>
<td>Individuals</td>
<td>Rings</td>
<td>6</td>
<td>4</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>Ranked by HINCAP</td>
<td>Sectors</td>
<td>27</td>
<td>17</td>
<td>23</td>
<td>14</td>
</tr>
<tr>
<td>Workers</td>
<td>Rings</td>
<td>6</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Ranked by Labor Earnings</td>
<td>Sectors</td>
<td>20</td>
<td>12</td>
<td>15</td>
<td>6</td>
</tr>
</tbody>
</table>
Tables 8a and 8b give the pattern of income inequality across and within radial sectors in Bogota and Cali respectively. The Theil index is shown for each sector as well as for the city as a whole. The two tables reveal quite interesting trends. First, the level of inequality within each sector is less than overall inequality. Second, the city center (sector 1) is the most heterogeneous in both the cities. Third, the poorest sectors (sector 2 in Bogota and sector 4 in Cali) show low indexes of inequality. Fourth, the richest sectors, sector 8 in Bogota and sector 2 in Cali are also tending to show low indexes of inequality. In Bogota there is a clear indication that ISIS is increasing over time from about 22 percent in 1973 to 27 percent in 1978 (with lower figures in intervening years). The last line in both the tables shows the ISIS if comunas are used as the spatial division. The between group contribution rises to 35 to 40 percent indicating that comunas are relatively homogeneous units: between comuna differences account for almost 40 percent of the overall inequality. ISIS for these also increases over time.

These results indicate strongly that the pattern of income segregation in Bogota is tending to get worse with time: there is certainly no evidence of the spatial disparities reducing. The last section showed that overall there is little evidence of increasing inequality, however. If spatial segregation of income groups has long term deleterious effects on the chances of the poor in the labor market as suggested in Mohan (1981), then it is a cause for concern that Bogota is being

1/ Similar tables are given in the appendix as Tables A-1 and A-2 using log variance as the inequality index.
Table 8a: SPATIAL INEQUALITY IN BOGOTA:  
INDIVIDUALS RANKED BY HINCAP 1973-78

(Inequality Index: Theil)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1/</td>
<td>2/</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ranking</td>
<td>Ranking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sector 1</td>
<td>3</td>
<td>0.635</td>
<td>8</td>
<td>0.678</td>
</tr>
<tr>
<td>Sector 2</td>
<td>1</td>
<td>0.429</td>
<td>3</td>
<td>0.440</td>
</tr>
<tr>
<td>Sector 3</td>
<td>2</td>
<td>0.390</td>
<td>1</td>
<td>0.309</td>
</tr>
<tr>
<td>Sector 4</td>
<td>5</td>
<td>0.407</td>
<td>2</td>
<td>0.273</td>
</tr>
<tr>
<td>Sector 5</td>
<td>6</td>
<td>0.553</td>
<td>5</td>
<td>0.504</td>
</tr>
<tr>
<td>Sector 6</td>
<td>4</td>
<td>0.514</td>
<td>4</td>
<td>0.439</td>
</tr>
<tr>
<td>Sector 7</td>
<td>7</td>
<td>0.572</td>
<td>6</td>
<td>0.462</td>
</tr>
<tr>
<td>Sector 8</td>
<td>8</td>
<td>0.602</td>
<td>7</td>
<td>0.542</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Overall</th>
<th>0.649</th>
<th>0.499</th>
<th>0.442</th>
<th>0.508</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Between Group Contribution (Sectors)</th>
<th>0.140</th>
<th>0.069</th>
<th>0.087</th>
<th>0.137</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Between Group (percent) Contribution (Sectors)</th>
<th>21.6</th>
<th>13.7</th>
<th>19.6</th>
<th>26.9</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Between Group (percent) Contribution (Comunas)</th>
<th>32.6</th>
<th>26.1</th>
<th>30.5</th>
<th>37.8</th>
</tr>
</thead>
</table>

Notes:

1/ Theil Index
2/ Ranking by ascending order of inequality.
3/ Between group contribution if grouped by comunas.
4/ Ascending order of income by Mean HINCAP. See Map 1.
Table 8b: SPATIAL INEQUALITY IN CALI:
INDIVIDUALS RANKED BY HINCAP 1973-78

(Inequality Index: Theil)

<table>
<thead>
<tr>
<th>Sector</th>
<th>1973</th>
<th>1977</th>
<th>1978</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T&lt;sup&gt;1/&lt;/sup&gt; Ranking</td>
<td>T&lt;sup&gt;2/&lt;/sup&gt; Ranking</td>
<td>T</td>
</tr>
<tr>
<td>Sector 1</td>
<td>5</td>
<td>0.573</td>
<td>5</td>
</tr>
<tr>
<td>Sector 2</td>
<td>7</td>
<td>0.368</td>
<td>2</td>
</tr>
<tr>
<td>Sector 3</td>
<td>3</td>
<td>0.429</td>
<td>3</td>
</tr>
<tr>
<td>Sector 4</td>
<td>1</td>
<td>0.328</td>
<td>1</td>
</tr>
<tr>
<td>Sector 5</td>
<td>2</td>
<td>0.435</td>
<td>4</td>
</tr>
<tr>
<td>Sector 6</td>
<td>4</td>
<td>0.731</td>
<td>6</td>
</tr>
<tr>
<td>Sector 7</td>
<td>6</td>
<td>0.781</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Group Contribution (Sectors)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Group (percent) Contribution (Sectors)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Group (percent) Contribution (Comunas)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:
1/ Theil Index
2/ Ranking by ascending order of Inequality.
3/ Between group contribution if grouped by comunas (See Map 2)
4/ Income ranking in ascending order of income by mean HINCAP
increasingly segregated by income. To the extent that policy-makers are relatively richer, increasing segregation can have the longer term effect of making them oblivious to the conditions of the poor in the city. Job opportunities are generally greater in the richer areas than in poorer areas. Increasing segregation might then also mean increasing average commutes for the poor who work in the richer areas of town. If such trends continue dualism will become a reality.

The methods employed in analyzing the spatial distribution of income can be extended to other cities in order to find out if Bogota and Cali are atypical in their degree of income segregation. If such studies are carried out in many cities it may also be possible to get at the reasons behind higher and lower levels of spatial differentiation between cities.
IV. THE DECOMPOSITION OF LABOUR EARNINGS

The previous sections have traced the inequality in incomes in Bogota and Cali using households as units of observation. It has been argued that of the income concepts used, the household income per capita (HINCAP) is the best measure of welfare for the household and for the individual. Indices of inequality have been computed for different income concepts and different criteria of ranking. These calculations are essentially of a descriptive nature: they are meant to give an idea of the extent of inequality prevailing among households and individuals but they are not useful in gaining insights into the causes of inequality. The only decompositions attempted of the inequality measures were by location of residence: the between group contribution varied between 20 to 25 percent of the total inequality. In other words, spatial inequality contributed about a quarter of the total inequality in Bogota and Cali.

This section takes the worker as the unit of observation and investigates the distribution of labor earnings. The greater part of the earnings of households comes from labor earnings. Hence the decomposition of labor earnings is useful in providing insights into the sources of inequality in incomes across households and individuals. The other sources of inequality between households is the difference in participation rates and in differences in household size. Thus, if participation rates are correlated with the different groups of people being considered for decomposition analysis, it is likely that the between group contribution to inequality in labour earnings would be less than for household income per capita.
The work presented in this section complements the analysis of earnings functions in Mohan (1981). There, the determinants of labor earnings are investigated systematically and education and years of experience are found to be the key determinants of the variance in earnings. The results in this section confirm those findings, but in addition, give further insights into the distribution of earnings within and between groups.

4.1 The Spatial Distribution of Personal Earnings

Tables 9a and 9b present the within sector and between sector indices of inequality in labor earnings in Bogota and Cali respectively. As might be expected the pattern is broadly similar to the distribution of household income and of HINCAP but with some differences. By and large, in Bogota the poorer sectors (1, 2 and 3) have lower levels of inequality than the richer sectors (7 and 8). Another way of describing the same phenomenon is to say that the richer sectors have poor people as well as many rich people while the poorer sectors have few rich people living there. There is about a 1 to 4 ratio between the mean earnings in the poorest sector (2) and in the richest sector (8). The ratio was about 1 to 5 for HINCAP. Thus the spatial disparities are less pronounced for labor earnings than for HINCAP as measured by the differences between means. The between-group contribution to the Theil index is also lower for labor earnings than for HINCAP. It is clear then that the extent of spatial disparity in incomes in Bogota is not caused solely by the maldistribution of labor earnings by location of residence. The results from the earnings functions analysis in Mohan (1981) indicated that the location of residence (acting as a proxy for other unmeasured characteristics) was a significant determinant of labor earnings but did not contribute much in magnitude to the explanation of log variance, after controlling for other factors. Appendix Tables A.3 and A.4 give the
Table 9a: THE SPATIAL DISTRIBUTION OF LABOUR EARNINGS IN BOGOTA - 1973-1978  
(Inequality Index: Theil Index)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T&lt;sup&gt;1/&lt;/sup&gt;</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td></td>
<td>Ranking&lt;sup&gt;2/&lt;/sup&gt;</td>
<td>Ranking</td>
<td>Ranking</td>
<td>Ranking</td>
</tr>
<tr>
<td>Sector 1</td>
<td>0.565</td>
<td>0.392</td>
<td>0.645</td>
<td>0.228</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>Sector 2</td>
<td>0.380</td>
<td>0.363</td>
<td>0.271</td>
<td>0.268</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Sector 3</td>
<td>0.349</td>
<td>0.281</td>
<td>0.324</td>
<td>0.259</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Sector 4</td>
<td>0.426</td>
<td>0.306</td>
<td>0.446</td>
<td>0.355</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Sector 5</td>
<td>0.523</td>
<td>0.462</td>
<td>0.465</td>
<td>0.415</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Sector 6</td>
<td>0.521</td>
<td>0.430</td>
<td>0.416</td>
<td>0.423</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Sector 7</td>
<td>0.636</td>
<td>0.487</td>
<td>0.484</td>
<td>0.486</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Sector 8</td>
<td>0.675</td>
<td>0.549</td>
<td>0.460</td>
<td>0.449</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>8</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Overall</td>
<td>0.634</td>
<td>0.452</td>
<td>0.460</td>
<td>0.481</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7592</td>
</tr>
<tr>
<td>Between Group Contribution (Sectors)</td>
<td>0.113</td>
<td>0.043</td>
<td>0.046</td>
<td>0.096</td>
</tr>
<tr>
<td>Between Group Contribution (percent)</td>
<td>17.9</td>
<td>9.5</td>
<td>9.9</td>
<td>20.0</td>
</tr>
<tr>
<td>Between Group Contribution (Comunas)</td>
<td>26.4</td>
<td>18.0</td>
<td>18.4</td>
<td>25.9</td>
</tr>
</tbody>
</table>

Notes:
1/ Theil Index
2/ Ranking by ascending order of inequality. (See Map 1).
3/ Between group contribution if grouped by comunas.
4/ Mean Earnings in Sector as proportion of overall mean.
5/ Ranking by mean earnings in sector.
6/ Overall Mean Earnings in 1978 Colombian pesos.

Sources: As in Table 2.
### Table 9B: THE SPATIAL DISTRIBUTION OF LABOUR EARNINGS IN CALI 1973-1978
(Inequality Index: Theil Index)

<table>
<thead>
<tr>
<th></th>
<th>1973</th>
<th></th>
<th>1977</th>
<th></th>
<th>1978</th>
<th>Mean</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T</td>
<td>Ranking</td>
<td>T</td>
<td>Ranking</td>
<td>T</td>
<td>Ranking</td>
<td></td>
</tr>
<tr>
<td>Sector 1</td>
<td>0.524</td>
<td>4</td>
<td>0.312</td>
<td>3</td>
<td>0.590</td>
<td>7</td>
<td>1.36</td>
</tr>
<tr>
<td>Sector 2</td>
<td>0.628</td>
<td>5</td>
<td>0.499</td>
<td>6</td>
<td>0.480</td>
<td>4</td>
<td>2.00</td>
</tr>
<tr>
<td>Sector 3</td>
<td>0.400</td>
<td>2</td>
<td>0.284</td>
<td>1</td>
<td>0.272</td>
<td>2</td>
<td>0.85</td>
</tr>
<tr>
<td>Sector 4</td>
<td>0.337</td>
<td>1</td>
<td>0.340</td>
<td>4</td>
<td>0.255</td>
<td>1</td>
<td>0.74</td>
</tr>
<tr>
<td>Sector 5</td>
<td>0.439</td>
<td>3</td>
<td>0.308</td>
<td>2</td>
<td>0.291</td>
<td>3</td>
<td>0.77</td>
</tr>
<tr>
<td>Sector 6</td>
<td>0.760</td>
<td>7</td>
<td>0.559</td>
<td>7</td>
<td>0.500</td>
<td>5</td>
<td>1.40</td>
</tr>
<tr>
<td>Sector 7</td>
<td>0.640</td>
<td>6</td>
<td>0.372</td>
<td>5</td>
<td>0.511</td>
<td>6</td>
<td>1.58</td>
</tr>
<tr>
<td>Total</td>
<td>0.592</td>
<td></td>
<td>0.454</td>
<td></td>
<td>0.440</td>
<td></td>
<td>6322</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>Between Group Contribution (Sectors)</th>
<th>13.0</th>
<th>11.3</th>
<th>14.6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Between Group (percent) Contribution (Sectors)</td>
<td>21.4</td>
<td>21.3</td>
<td>23.1</td>
</tr>
</tbody>
</table>

**Notes:**

1/ Theil Index
2/ Ranking by ascending order of inequality
3/ Between group contribution if grouped by comunas (see Map 2)
4/ Mean Earnings in Sector as proportion of overall mean.
5/ Ranking by Mean Earnings in Sector
6/ Overall mean earnings in 1978 Colombian pesos

**Sources:** As in Table 2.
analysis information as in Table 9 but using log variance as the index: the spatial contribution is smaller than for the Theil index since the log variance is more sensitive to low incomes. These results are naturally consistent with the earnings function estimates.

That the between sector contribution to inequality is lower for labor earnings indicates that dependency ratios are higher among the households in the poorer sectors. The spatial maldistribution of earnings is then compounded by either larger families or lower participation rates or higher unemployment rates. Any of these would produce higher dependency ratios. Cause and effect here is not easy to disentangle: do the presence of children inhibit poorer women from working because of child care responsibilities? Are they out of the labor force or are they unemployed because they can only take part time jobs with less rigid work patterns?

The spatial concern is important in this respect. If it is the case, as is being demonstrated in Bogota, that large areas of the city -- the South -- have predominantly low income people, then it is likely that those areas have comparatively fewer jobs per resident population than the richer areas. The purchasing power in these areas is simply much lower. It is certainly a fact (as demonstrated in Mohan (1980)) that there is a great amount of out commuting from the poorer sectors to the richer sectors but very little in reverse. It is also true, however, that unemployment rates as well as participation rates do not differ too much between sectors -- though there are differences whose significance is difficult to assess statistically. That large differences do not exist does not mean that the poorer people would not have higher participation rates and therefore be better off
were these large spatial disparities not there. These questions cannot be answered by the evidence at hand: only suggestive implications can be drawn.

The Theil index decreased between 1973 and 1978 from 0.38 to 0.27 in Sector 2 and 0.68 to 0.45 in Sector 8. Thus both the poorest and richest sectors are getting more homogeneous and therefore more distinct from each other over time. Indeed, declines in inequality are observed in all the sectors although the overall index of inequality increased between 1975 and 1978. In earlier years the Theil index was greater for some sectors than the overall index: in 1978 within sector inequality in all sectors is lower than overall inequality in the city. Similar trends are observed in Cali though the spatial disparity is somewhat less pronounced there. The ratio of mean earnings in the poorest sector (4) to those in the richest sector (2) is less than 1 to 3 in Cali. There is clearer indication that overall inequality has been declining in that city.

One must therefore conclude from these data that the spatial disparities are getting more pronounced in both these cities in terms of both labour earnings distributions as well as household income and household income per capita. If the spatial separation of income groups can be seen as harmful for the opportunities in the future for the poor as well as for social cohesion in the present, then these findings are truly a cause for concern.

4.2 The Decomposition of Labour Earnings According to Various Personal Characteristics

The distribution of labour earnings is dissected further according to conventional classifications. Between group and within
group contributions are computed by age, education, employment status, occupation, industry of activity and size of firm. These calculations merely complement the earnings functions work but also give a better sense of the distributions within these various categories.

Table 10 summarizes the decomposition of the distribution of labour earnings according to the various classifications utilized. The succeeding pages give the details of the classifications and of the individual distributions. It is clear at a glance that the between group contributions are lower for Cali than for Bogota for all the classifications. A smaller city has a somewhat less variegated or heterogeneous labour market. Occupations are, perhaps, less specialized at the top end. The highest financial and governmental functions are typically in the political capital of a country or in the key business center: Bogota serving as both in the case of Colombia. It has been observed earlier that overall inequality is also lower in Cali.

Age and Education

It is somewhat surprising that age group differences contribute only about 9 to 11 percent to total inequality in labour earnings. The implication is that for the majority of workers, their age-earnings profiles are pretty flat. As was shown in Mohan (1980) the age earnings profiles are very steep for higher educated people but much less for the less educated. Table 12 shows that almost half of the labour force in Bogota still has only primary education or less. Moreover, Table 11 shows that after age 25 the overall age-earnings profile is rather flat. At the same time, inequality within
Table 10: THE DECOMPOSITION OF LABOUR EARNINGS - 1978

<table>
<thead>
<tr>
<th>Variable</th>
<th>City</th>
<th>Number of Groups</th>
<th>Between Groups Contribution</th>
<th>Contribution</th>
<th>Log Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Theil (Percent)</td>
<td>Theil (Percent)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>Bogota</td>
<td>7</td>
<td>11.4</td>
<td>8.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cali</td>
<td></td>
<td>11.0</td>
<td>9.2</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>Bogota</td>
<td>4</td>
<td>38.0</td>
<td>31.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cali</td>
<td></td>
<td>26.6</td>
<td>22.4</td>
<td></td>
</tr>
<tr>
<td>Employee Status</td>
<td>Bogota</td>
<td>3</td>
<td>8.0</td>
<td>7.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cali</td>
<td></td>
<td>7.5</td>
<td>7.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bogota</td>
<td>6</td>
<td>19.1</td>
<td>17.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cali</td>
<td></td>
<td>13.4</td>
<td>12.6</td>
<td></td>
</tr>
<tr>
<td>Occupation</td>
<td>Bogota</td>
<td>7</td>
<td>40.3</td>
<td>32.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cali</td>
<td></td>
<td>24.8</td>
<td>17.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bogota</td>
<td>13</td>
<td>41.2</td>
<td>33.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cali</td>
<td></td>
<td>31.1</td>
<td>24.5</td>
<td></td>
</tr>
<tr>
<td>Industry of Activity</td>
<td>Bogota</td>
<td>8</td>
<td>4.6</td>
<td>4.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cali</td>
<td></td>
<td>3.9</td>
<td>3.3</td>
<td></td>
</tr>
<tr>
<td>Firm Size</td>
<td>Bogota</td>
<td>6</td>
<td>7.5</td>
<td>11.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cali</td>
<td></td>
<td>7.4</td>
<td>11.7</td>
<td></td>
</tr>
<tr>
<td>Sector</td>
<td>Bogota</td>
<td>8</td>
<td>20.0</td>
<td>11.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cali</td>
<td></td>
<td>14.6</td>
<td>6.5</td>
<td></td>
</tr>
<tr>
<td>Comunas</td>
<td>Bogota</td>
<td>38</td>
<td>25.9</td>
<td>16.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cali</td>
<td>25</td>
<td>23.1</td>
<td>12.7</td>
<td></td>
</tr>
</tbody>
</table>
Table 11: THE DISTRIBUTION OF EARNINGS BY AGE CATEGORIES - 1978

<table>
<thead>
<tr>
<th>Age Category</th>
<th>Bogota</th>
<th></th>
<th>Cali</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Earnings</td>
<td>Theil</td>
<td>Percent</td>
</tr>
<tr>
<td>12 - 14</td>
<td>0.29</td>
<td>0.145</td>
<td>1.0</td>
</tr>
<tr>
<td>15 - 24</td>
<td>0.55</td>
<td>0.207</td>
<td>30.2</td>
</tr>
<tr>
<td>25 - 34</td>
<td>1.15</td>
<td>0.438</td>
<td>30.0</td>
</tr>
<tr>
<td>35 - 44</td>
<td>1.27</td>
<td>0.459</td>
<td>20.4</td>
</tr>
<tr>
<td>45 - 54</td>
<td>1.28</td>
<td>0.532</td>
<td>12.1</td>
</tr>
<tr>
<td>55 - 64</td>
<td>1.23</td>
<td>0.548</td>
<td>4.7</td>
</tr>
<tr>
<td>65 +</td>
<td>0.95</td>
<td>0.514</td>
<td>1.5</td>
</tr>
<tr>
<td>Total</td>
<td>7600</td>
<td>0.481</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Total No. of Workers.

<table>
<thead>
<tr>
<th></th>
<th>1,180,000</th>
<th>350,000</th>
</tr>
</thead>
</table>

Between Group Theil Contribution

<table>
<thead>
<tr>
<th></th>
<th>11.4</th>
<th>11.0</th>
</tr>
</thead>
</table>

Notes:

1/ Mean earnings of Age Group expressed as a multiple of the overall mean earnings.

2/ Overall mean in 1978 Colombian Pesos.
age groups increases consistently along with age except for the over 65s. This is true for both Bogota and Cali. Note that earnings are very egalitarian in the young age group 15-24: The disparities increase with age as those with more education increase their earnings - along with others who are successful as well - while the rest have to live with almost the same earnings for most of their lives. This is of great importance for the study of life-cycle effects on poverty as has been shown in Mohan and Hartline (1980) and Mohan, Garcia and Wagner (1981). Those who have relatively flat age-earnings profiles experience decreases in their real household income per capita as the process of family formation takes place and there are more mouths to feed. Many studies show a relatively higher incidence of malnutrition among children: flat age earnings profiles would be one of the key causes for such a phenomenon.

The differences in educational achievements account for 32 to 38 percent of the overall inequality in Bogota and about 22 to 27 percent in Cali. Only 4 groups are utilized: those with no education, any primary, any secondary and any higher education. The information in Table 12 is of great interest. First, clearly Bogota has a much larger proportion of the labour force with higher education. The capital city presumably attracts - has jobs for - the more highly educated. This partly accounts for the relatively higher level of inequality in Bogota. Second, the highest level of inequality is found among the secondary level educated group. Clearly, those who
### Table 12: THE DISTRIBUTION OF EARNINGS BY EDUCATION GROUP - 1978

<table>
<thead>
<tr>
<th>Education Category</th>
<th>Bogota</th>
<th></th>
<th>Cali</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (^1)/ Earnings</td>
<td>Theil Index</td>
<td>Percent in Category</td>
<td>Mean (^1)/ Earnings</td>
<td>Theil Index</td>
</tr>
<tr>
<td>None</td>
<td>0.38</td>
<td>0.160</td>
<td>3.6</td>
<td>0.44</td>
<td>0.176</td>
</tr>
<tr>
<td>Primary</td>
<td>0.56</td>
<td>0.239</td>
<td>42.1</td>
<td>0.68</td>
<td>0.283</td>
</tr>
<tr>
<td>Secondary</td>
<td>0.92</td>
<td>0.356</td>
<td>38.4</td>
<td>1.16</td>
<td>0.388</td>
</tr>
<tr>
<td>Higher</td>
<td>2.49</td>
<td>0.286</td>
<td>15.9</td>
<td>2.67</td>
<td>0.267</td>
</tr>
<tr>
<td>Total</td>
<td>7600(^2)/</td>
<td>0.481</td>
<td>100.0</td>
<td>6350(^2)/</td>
<td>0.440</td>
</tr>
</tbody>
</table>

Total No. of Workers: 1,180,000 (Bogota) 350,000 (Cali)

Between Group Theil: 38.0 (Bogota) 26.6 (Cali)

Contribution Log Variance (percent): 31.7 (Bogota) 22.4 (Cali)

**Notes:**

1/ Mean earnings of Age Group expressed as a multiple of the overall mean.
2/ Overall mean in 1978 Colombian Pesos.
receive higher education have uniformly high incomes and those with only primary, uniformly low incomes. Much of the unexplained variance in earnings must be among the secondary level educated workers. They have the potential for higher earnings: some make it but many do not. Third, because of the presence of more higher educated workers the mean earnings in Bogota are rather higher than those in Cali and even the mean of the secondary educated is less than the overall mean. The between group contribution enumerated above is rather large and attests to the importance of education in determining earnings and earnings potential.

Employment Status

Table 13 gives the decomposition of labour earnings according to groups of workers by employment status. Two classifications were attempted. The first with only 3 groups merely distinguished between all employees, owners or bosses and the self-employed. The second classification broke down the employees category into blue collar workers, white collar workers, domestic workers and "non-remunerated". The greatest extent of inequality is found among the self-employed. Skills in entrepreneurship, the exercise of initiative and other unmeasured characteristics are likely to be important within this group: some are quite poor while others can be very well off. The Theil index for inequality within this group is higher than the index for all the workers taken together. It is also of interest that the lower skilled group - blue collar workers and domestic workers - have the lowest within group index.\footnote{Ignore "unremunerated workers" as a misclassified and small category.} The labour market
TABLE 13: THE DISTRIBUTION OF EARNINGS BY EMPLOYMENT STATUS - 1978

<table>
<thead>
<tr>
<th>Employment category</th>
<th>Mean earnings(^5)</th>
<th>Theil index</th>
<th>Percent in category</th>
<th>Mean earnings</th>
<th>Theil index</th>
<th>Percent in category</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bogota</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3 categories)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employees 3/</td>
<td>0.97</td>
<td>0.445</td>
<td>75.3</td>
<td>0.95</td>
<td>0.369</td>
<td>72.2</td>
</tr>
<tr>
<td>Owners or Boss</td>
<td>2.25</td>
<td>0.338</td>
<td>5.5</td>
<td>2.31</td>
<td>0.354</td>
<td>4.7</td>
</tr>
<tr>
<td>Self-employed</td>
<td>0.76</td>
<td>0.516</td>
<td>19.3</td>
<td>0.85</td>
<td>0.574</td>
<td>22.8</td>
</tr>
<tr>
<td>(6 categories)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blue collar 1/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>employees</td>
<td>0.57</td>
<td>0.169</td>
<td>21.6</td>
<td>0.74</td>
<td>0.242</td>
<td>25.8</td>
</tr>
<tr>
<td>White collar 2/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>employees</td>
<td>1.25</td>
<td>0.429</td>
<td>45.7</td>
<td>1.18</td>
<td>0.387</td>
<td>39.1</td>
</tr>
<tr>
<td>Non-remunerated 4/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Owners or Boss</td>
<td>2.25</td>
<td>0.338</td>
<td>5.5</td>
<td>2.31</td>
<td>0.354</td>
<td>4.7</td>
</tr>
<tr>
<td>Self-employed</td>
<td>0.76</td>
<td>0.516</td>
<td>19.3</td>
<td>0.85</td>
<td>0.574</td>
<td>22.8</td>
</tr>
<tr>
<td>Domestic workers</td>
<td>0.42</td>
<td>0.167</td>
<td>7.8</td>
<td>0.51</td>
<td>0.098</td>
<td>7.3</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>7600 (^6)</td>
<td>0.481</td>
<td>100.0</td>
<td>6350 (^6)</td>
<td>0.440</td>
<td>100.0</td>
</tr>
<tr>
<td>Total no. of workers</td>
<td>1,180,000</td>
<td></td>
<td></td>
<td>350,000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between group 7/</td>
<td>Theil 19.1</td>
<td></td>
<td></td>
<td>13.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contribution (percent)</td>
<td>log variance 17.6</td>
<td></td>
<td></td>
<td>12.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: 1/ Obrero loosely translated as blue collar workers.  
2/ Empleado loosely translated as white collar workers.  
3/ Includes obreros, empleados, non-remunerated workers and domestic workers.  
4/ Clearly misclassified since mean earnings, though low, are not zero.  
5/ Mean earnings of employment group expressed as a multiple of the overall mean.  
7/ For 6 categories.
for these categories must be homogeneous as well as competitive. These low paid groups comprise almost 30 percent of the labour force. Owners and bosses earnings are also distributed in a relatively egalitarian fashion: The Theil index is only about 0.34 to 0.35. This is a bit surprising since one would expect that there would be many small relatively poor owners as well as rich ones. The between group contribution for the first classification is only about 10 percent while for the second it is about 20 percent: the latter would therefore suggest that an expansion of categories is worth examination.

**Occupational Status**

Occupational differences account for about 33 to 40 percent in the inequality in earnings in Bogota and 17 to 25 in Cali. The between group contribution does not increase much if one changes the classification from the 1 digit level (7 categories) to more using 13 categories. The one-digit level is therefore a fairly adequate level of disaggregation for occupations. Ignoring agricultural workers, the level of inequality within occupations is substantially lower than the overall inequality except for sales workers, as shown in Table 14. Appendix Table A.5 gives the analogous information for the more dissaggregated classification scheme. Even when sales workers are separated into sales managers and others the index of inequality does not decrease appreciably. There is clearly a high variance of earnings among sales establishments and among the people who work in these establishments – a
Table 14: THE DISTRIBUTION OF EARNINGS BY OCCUPATIONAL CATEGORIES - 1978

<table>
<thead>
<tr>
<th>Occupational category</th>
<th>ILO code</th>
<th>Mean earnings (^1)</th>
<th>Theil index</th>
<th>Percent in category</th>
<th>Mean earnings</th>
<th>Theil index (^1)</th>
<th>Percent in category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professional and technical</td>
<td>1-19</td>
<td>2.30</td>
<td>0.300</td>
<td>12.0</td>
<td>2.43</td>
<td>0.304</td>
<td>6.2</td>
</tr>
<tr>
<td>Administrative and manager</td>
<td>20-29</td>
<td>3.53</td>
<td>0.137</td>
<td>2.9</td>
<td>3.11</td>
<td>0.288</td>
<td>2.2</td>
</tr>
<tr>
<td>Clerk and typist</td>
<td>30-39</td>
<td>0.92</td>
<td>0.219</td>
<td>14.3</td>
<td>1.05</td>
<td>0.250</td>
<td>12.5</td>
</tr>
<tr>
<td>Sales</td>
<td>40-49</td>
<td>1.06</td>
<td>0.523</td>
<td>16.0</td>
<td>1.07</td>
<td>0.529</td>
<td>16.7</td>
</tr>
<tr>
<td>Service workers</td>
<td>50-59</td>
<td>0.49</td>
<td>0.255</td>
<td>20.0</td>
<td>0.60</td>
<td>0.249</td>
<td>20.3</td>
</tr>
<tr>
<td>Agriculture workers</td>
<td>60-69</td>
<td>1.47</td>
<td>0.579</td>
<td>0.9</td>
<td>1.97</td>
<td>0.586</td>
<td>1.0</td>
</tr>
<tr>
<td>Production workers</td>
<td>70-98</td>
<td>0.61</td>
<td>0.197</td>
<td>33.8</td>
<td>0.80</td>
<td>0.291</td>
<td>41.4</td>
</tr>
<tr>
<td>Total</td>
<td>1-98</td>
<td>7600(^2)</td>
<td>0.482</td>
<td>100.0</td>
<td>6350(^2)</td>
<td>0.440</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Total no. of workers 1,180,000 350,000

Between group Theil 40.5 24.8
Contribution (percent) Log variance 32.7 17.5

Notes 1 Mean earnings in occupation expressed as a multiple of the overall mean.
2 1978 Colombian Pesos.
finding consistent with the high level of inequality found among the self-employed since many of the latter are likely to be engaged in trade. It is quite surprising to find that the large category comprising production workers (35 to 40 percent of the labour force) is quite homogeneous with a Theil index of only about 0.2 in Bogota, and somewhat higher in Cali. Within white collar occupations also - clerks and typists and the administrators and managers - the earnings distribution seems to be in a relatively narrow band. Almost all the within occupation contributions to inequality are higher for Cali. The evidence from these data is clearly that except for sales, choice of occupation is very important as a predictor of potential earnings: indeed, the between group contribution of occupational categories to income inequality is of the same order of magnitude as that of education.

Industry of Activity

The industry of activity is a far less interesting classification of workers from the viewpoint of income inequality. The one digit level of classification yields only a 4 to 5 percent between group contribution to inequality: essentially negligible. When the manufacturing industries are disaggregated and 20 industrial categories are used the between group contribution increases to 11 to 15 percent. Nonetheless, as shown in Table 15 every industrial category obviously has a rather heterogeneous composition of workers. It is the only relatively small category of utilities which has a low Theil index of between
<table>
<thead>
<tr>
<th>Industrial category</th>
<th>SIC code</th>
<th>Mean earnings $^1$</th>
<th>Theil index</th>
<th>Percent in category</th>
<th>Mean earnings</th>
<th>Theil index $^1$</th>
<th>Percent in category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture and mining</td>
<td>1-29</td>
<td>1.77</td>
<td>0.474</td>
<td>1.4</td>
<td>1.62</td>
<td>0.669</td>
<td>1.4</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>30-39</td>
<td>0.89</td>
<td>0.469</td>
<td>23.7</td>
<td>1.02</td>
<td>0.436</td>
<td>31.5</td>
</tr>
<tr>
<td>Utilities</td>
<td>40-49</td>
<td>1.14</td>
<td>0.214</td>
<td>0.5</td>
<td>1.45</td>
<td>0.249</td>
<td>1.0</td>
</tr>
<tr>
<td>Construction</td>
<td>50-59</td>
<td>0.72</td>
<td>0.394</td>
<td>7.3</td>
<td>0.78</td>
<td>0.281</td>
<td>6.2</td>
</tr>
<tr>
<td>Trade and commerce</td>
<td>60-69</td>
<td>0.92</td>
<td>0.514</td>
<td>23.4</td>
<td>0.89</td>
<td>0.476</td>
<td>24.2</td>
</tr>
<tr>
<td>Transport and communication</td>
<td>70-79</td>
<td>1.12</td>
<td>0.333</td>
<td>5.9</td>
<td>1.07</td>
<td>0.271</td>
<td>6.4</td>
</tr>
<tr>
<td>Financial establishments</td>
<td>80-89</td>
<td>1.61</td>
<td>0.424</td>
<td>8.2</td>
<td>1.81</td>
<td>0.436</td>
<td>3.8</td>
</tr>
<tr>
<td>Public administration and social services</td>
<td>91-96</td>
<td>0.98</td>
<td>0.475</td>
<td>29.4</td>
<td>0.93</td>
<td>0.417</td>
<td>25.6</td>
</tr>
<tr>
<td>Total</td>
<td>1-96</td>
<td>7600$^2$</td>
<td>-0482</td>
<td></td>
<td>6350$^2$</td>
<td>0.441</td>
<td>100</td>
</tr>
</tbody>
</table>

Total no. of workers 1,180,000 350,000

Between group
Contribution
Log variance 4.4 3.3

Notes
1 Mean earnings in occupation expressed as a multiple of the overall mean.
2 1978 Colombian Pesos.
0.2 and 0.25. Construction and transport and communication are the other two industrial categories which are somewhat homogeneous. Appendix Table A.6 gives the same information for the expanded number of industrial categories - (20). Essentially, the manufacturing categories have been dissgregated.

As might be expected, the within group contributions to inequality decrease and the between group contribution rises. These results suggest that, overall, industrial categories cannot be regarded as a very useful classification in the study of inequality in labour earnings. Occupational categories are more distinct.

Segmentation is often sought in the labour market between industries: it would appear that in Bogota and Cali all industries are quite heterogeneous in that within industry inequality is almost as great the overall inequality for most industries.

**Size of Firm**

The last decomposition attempted was by categories of firm size since it is often suggested that the labour market is segmented according to firm size. It is argued that the protected sector consists of the larger firms and, because they pay considerably higher wages, much of the inequality observed in urban areas is due to the existence of this protected sector. Six categories of firm size are distinguished in Table 16: 1, 2-5, 6-10, 11-25, 25 to 100 and 100+.

The between group contribution to inequality is only about 7 to 12 percent -- a proportion lower than that gound for spatial disparity in earnings. As many as 45 to 50 percent of the workers
Table 16: THE DISTRIBUTION OF EARNINGS BY FIRM SIZE - 1978

<table>
<thead>
<tr>
<th>Category of Firm size</th>
<th>Bogota</th>
<th></th>
<th></th>
<th>Cali</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean earnings&lt;sup&gt;2&lt;/sup&gt;</td>
<td>Theil index</td>
<td>Percent in category</td>
<td>Mean earnings&lt;sup&gt;2&lt;/sup&gt;</td>
<td>Theil index</td>
<td>Percent in category</td>
</tr>
<tr>
<td>1</td>
<td>0.58</td>
<td>0.449</td>
<td>23.2</td>
<td>0.64</td>
<td>0.429</td>
<td>25.7</td>
</tr>
<tr>
<td>2-5</td>
<td>0.90</td>
<td>0.497</td>
<td>21.9</td>
<td>0.98</td>
<td>0.455</td>
<td>26.2</td>
</tr>
<tr>
<td>6-10</td>
<td>1.07</td>
<td>0.506</td>
<td>11.8</td>
<td>0.93</td>
<td>0.439</td>
<td>9.5</td>
</tr>
<tr>
<td>11-25</td>
<td>1.19</td>
<td>0.460</td>
<td>12.1</td>
<td>1.08</td>
<td>0.450</td>
<td>10.5</td>
</tr>
<tr>
<td>25-100</td>
<td>1.25</td>
<td>0.438</td>
<td>14.3</td>
<td>1.37</td>
<td>0.458</td>
<td>12.4</td>
</tr>
<tr>
<td>100+</td>
<td>1.25</td>
<td>0.374</td>
<td>16.6</td>
<td>1.31</td>
<td>0.335</td>
<td>15.7</td>
</tr>
<tr>
<td>Total</td>
<td>7505&lt;sup&gt;3&lt;/sup&gt;</td>
<td>0.486</td>
<td>100.0</td>
<td>6358&lt;sup&gt;3&lt;/sup&gt;</td>
<td>0.459</td>
<td>100.0</td>
</tr>
<tr>
<td>Total no.&lt;sup&gt;1&lt;/sup&gt;</td>
<td>of workers</td>
<td>1,140,000</td>
<td></td>
<td>324,000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between group</td>
<td>Theil</td>
<td>7.5</td>
<td></td>
<td>7.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contribution (percent)</td>
<td>Log variance</td>
<td>11.6</td>
<td></td>
<td>11.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes
1. The discrepancy between the total number of workers in this table as compared with other tables is due to non-responses for the firm size question. The overall mean income is also, therefore, slightly different.
2. Mean earnings of workers in firm size category expressed as a multiply of the overall mean.
3. 1978 Colombian Pesos.
in Bogota and Cali work in firms of less than 5 people. The mean earnings of workers in single employee establishments -- presumably mostly self-employed service and trade activities -- are considerably lower than the other categories. There is not much difference between the means beyond the size of 10. Moreover, within category levels of inequality are not very different from the largest category: firms employing over 100 people. The Theil index for them is about 0.34 to 0.38. The mean level of earnings is somewhat higher and they are distributed within a relatively narrower band. But the overall evidence, consistent with the results from the earnings functions, is that firm size is not an important feature of the labour market in Bogota and Cali. The mean levels are not very different between categories, the between group contribution is low, and the within group contributions are not very different from the overall level of inequality.
4.3. Summary

This section has probed the distribution of labour earnings in a detailed manner. An analysis of the determinants of labour earnings has been reported elsewhere (Mohan, 1981) and the decomposition of earnings is merely another method of looking at the labour market. Not surprisingly, the results are consistent with those found in the earlier paper.

The overall spatial inequality in labour earnings was found to be less than that for HINCAP. This is probably because of higher dependency rates for the households in the poorer parts of the city. The spatial maldistribution of earnings is then compounded by the existence of either larger families, lower participation rates or higher unemployment rates - to produce a worse distribution of HINCAP. It was also found that the inequality in earnings within sectors has been decreasing consistently over time while that between sectors has been increasing. Thus the different sectors are tending to become relatively more homogeneous. Large areas of Bogota - mainly the South - are inhabited predominantly by poor people. Because of the resulting low purchasing power of the area as a whole, there is therefore also lower demand for labour. There are comparatively fewer jobs per resident labour as compared with the richer sectors.

The other types of categories which were found to exhibit the most between group inequality were education groups and occupation groups. Within group inequality was low for both higher educated people as well as lower educated people. The secondary educated group exhibited the greatest - within group inequality which, moreover, was greater than the overall inequality. The fact that between group inequality was low confirms the relatively flat age-earnings profile that exists in Bogota as a whole because
of the majority of people still being educated up to the primary level or less. Naturally, since the higher educated people have relatively steep age-earnings profile the degree of inequality increases within age groups along with age. Interestingly, there is very low inequality within the young age group of 15 to 24 years.

Among occupations, production workers were the most homogeneous while sales workers exhibited the most within group inequality. Overall, the between group contribution was as high as about 40 percent for occupational categories. It is somewhat surprising that production workers exhibit low inequality as a group since one would have expected a wide variety of skill levels within the group. Significantly, the industry of activity, as well as the size of firm show very low between group contributions to inequality. The labor market segmentation literature has most often focused on these two types of categorization for identifying segmentation in the labour market. These results would indicate that the labour market in Bogota cannot be said to be segmented by either size of firm or by industry of activity.

The level of education and choice of occupation are clearly the most important factors in the determination of labour earnings in Bogota and Cali on the basis of these results. To the extent that poor education as well as low aspirations (e.g. in the choice of education) result at least partially from living in a poor part of the city, the increasing spatial segregation measured by increases in the ISIS is truly a cause for concern.

The last result of interest is that the between group contribution to inequality was consistently lower for Cali as compared with Bogota, as was
the overall level of inequality. Clearly, a smaller city is expected to have a much less heterogeneous and variegated labour market, and particularly at the highest skill levels. This is evidenced by the fact that Bogota has a higher proportion of people with higher education.
V. SUMMARY AND CONCLUSIONS

This paper has attempted to document the distribution of income and the changes therein in two fast growing cities in the developing world as part of a larger study that seeks to explain the economic structure of cities in poor countries. Earlier papers have given some of the same information in a more descriptive fashion. Conjectures offered in these papers concerning the distribution of income and the pattern of labour earnings have been confirmed in this paper.

Overall, there appears to be a tendency toward some improvement in the distribution of income between 1973 and 1978 in Bogota and Cali. The quality of the data is not robust enough to be able to state this conclusively but such a tendency would be consistent with overall changes in the national economy over this period. With over 60 percent of the population now already in urban areas there is a natural slow down in the rate of urbanization. This has been reflected in the growth of Bogota and Cali as well. To the extent that such a slow down in the growth of the cities also implies a tightening of the labour market, pushing up unskilled wages, an improvement in the distribution of income may also be expected. Further, although over half of the population in these cities is still educated up to only the primary level or less there has been a continuing deepening of the educational level of the labour force. As the inequality in educational levels declines so should it in the distribution of earned income. The one major caveat in all these
observations is that if the coverage of non-labour income in all the surveys is bad, as is suspected, some of the improvement in the distribution of income could be illusory.

Despite any improvement in the distribution that may have occurred, the level of inequality in the two cities remains very high and which may possibly be among the more unequal in the world. Measured in any way—the distribution of households ranked according to household income or by household income per capita, or individuals by household income per capita, earners by labour earnings, the Gini coefficient comes to be between about 0.55 and 0.50 with a tendency towards improvement over the 1973 to 1978 time period for which data were analyzed. There is a high degree of confidence in these results since they are robust over the different samples, different indices, different income concepts and different ranking procedures. Overall, there is a higher level of inequality in Bogota, as might be expected, than the smaller city, Cali.

A special feature of this study is the calculation of ISIS—the Index of the Spatial Income Segregation defined as the percentage between-group contribution of spatial zones to overall inequality in a city. As had been suggested in earlier papers, Bogota and Cali both have more distinct radial sectors in terms of income patterns rather than rings as is the case with North American cities. The between-group contribution is very low when the grouping is done by rings. The ISIS in both the cities is substantially lower for labour earnings as compared with household income per capita, implying that some of the inequality is due to the locational differences in
the composition of households. The main disquieting feature of these findings is that ISIS appears to be increasing over time while inequality within sectors is declining. Thus each sector is becoming more homogeneous – the rich sectors as well as the poor ones. The city is therefore becoming more spatially segregated by income over time.

If, as may be suspected, greater segregation by income has deleterious effects on social welfare mainly through the operation of labour markets the increasing spatial disparities in Bogota are truly a case for concern. Mohan (1981) had established the location of residence as a significant factor in the determinants of labour earnings within the city. The between group contribution of sectors to the overall inequality in labour earnings supports that finding. The interpretation of the location of residence variable was as a proxy for family background, educational quality and other similar determinants of earnings. If the poor environmental factors faced by a poor child are reinforced by higher segregation, then the increase in ISIS gives major cause for concern for the distribution of earnings in the future. It is typically the case in most cities that richer areas are much better served by public utilities than poorer areas. If the richer and poorer areas are mixed – as they are in many cities – then the poorer areas also benefit from the pressures brought by the rich for good sanitation, water supply, roads, etc. It is then likely that if very large areas, such as the South of Bogota and East of Cali, have very few better off people then the likelihood of neglect by public authorities increases. The quality of schooling can be very seriously affected by such segregation since it may be expected that both the quality of teachers as well as average quality of students would be low. Self-reinforcing and long term effects would then result in terms of spatial disadvantage being suffered by the children living
in these areas which would manifest itself later on in the labour market. Although the problem of consistent zonification remains for comparisons to be made across cities, ISIS can be used as a measure of income segregation and could easily be applied to different cities. If the same number of quantiles as well as the same number of zones were used the results would be quite comparable across cities.

The decomposition of labour earnings has lent support to the earlier work on the determinants of labour earnings. The industry of activity and the size of firm - the variables that are often used as segmenting variables in labour market segmentation literature - were found to contribute very little to the overall inequality in earnings. The main variables contributing to inequality were found to be education and occupation - which could be taken as a proxy of skill.

Overall, Bogota was found to exhibit higher inequality than Cali as well as higher ISIS. These are very suggestive results on the characteristics and role of large cities as compared with smaller cities. It would seem that a larger city can be expected to have a much more variegated and heterogeneous labour market leading to higher inequality. It is only in the largest cities that the highest professional and other skills can be utilized and therefore the highest professional incomes would be found only in the larger cities. There would then also be greater pressures, particularly in a high overall inequality situation, for higher spatial segregation. The relatively rich would want to live near others of like qualities and in a larger city there would be a critical mass for them to be able to do so.
Appendix I. Tables A.1 to A.6

<table>
<thead>
<tr>
<th>No.</th>
<th>Title</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.1</td>
<td>Spatial Inequality in Bogota. Individuals Ranked by HINCAP 1973-1978.</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>(Using Log Variance)</td>
<td></td>
</tr>
<tr>
<td>A.2</td>
<td>Spatial Inequality in Cali: Individuals Ranked by HINCAP 1973-1978</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>(Using Log Variance)</td>
<td></td>
</tr>
<tr>
<td>A.3</td>
<td>The Spatial Distribution of Labour Earnings in Bogota 1973-1978</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>(Using Log Variance)</td>
<td></td>
</tr>
<tr>
<td>A.4</td>
<td>The Spatial Distribution of Labour Earnings in Cali 1973-1978</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>(Using Log Variance)</td>
<td></td>
</tr>
<tr>
<td>A.5</td>
<td>The Distribution of Earnings by Occupational Categories 1978</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>(13 Categories)</td>
<td></td>
</tr>
<tr>
<td>A.6</td>
<td>The Distribution of Earnings by Industry of Activity 1978</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>(20 Categories)</td>
<td></td>
</tr>
</tbody>
</table>
Table A.1: SPATIAL INEQUALITY IN BOGOTA
INDIVIDUALS RANKED BY HINCAP 1973-1978
(Inequality Index: Log Variance)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sector</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.101</td>
<td>1.514</td>
<td>1.301</td>
<td>0.645</td>
</tr>
<tr>
<td>2</td>
<td>0.718</td>
<td>0.767</td>
<td>0.551</td>
<td>0.474</td>
</tr>
<tr>
<td>3</td>
<td>0.686</td>
<td>0.629</td>
<td>0.599</td>
<td>0.546</td>
</tr>
<tr>
<td>4</td>
<td>0.769</td>
<td>0.598</td>
<td>0.819</td>
<td>0.694</td>
</tr>
<tr>
<td>5</td>
<td>0.917</td>
<td>1.062</td>
<td>0.886</td>
<td>0.750</td>
</tr>
<tr>
<td>6</td>
<td>0.822</td>
<td>0.814</td>
<td>0.639</td>
<td>0.715</td>
</tr>
<tr>
<td>7</td>
<td>1.247</td>
<td>1.054</td>
<td>0.905</td>
<td>1.001</td>
</tr>
<tr>
<td>8</td>
<td>1.990</td>
<td>1.439</td>
<td>0.970</td>
<td>1.282</td>
</tr>
<tr>
<td>Total</td>
<td>1.013</td>
<td>0.901</td>
<td>0.830</td>
<td>0.846</td>
</tr>
<tr>
<td>Between Group Contribution (Sectors)</td>
<td>0.109</td>
<td>0.066</td>
<td>0.115</td>
<td>0.147</td>
</tr>
<tr>
<td>Between Group Contribution (Percent) (Sectors)</td>
<td>10.7</td>
<td>7.3</td>
<td>13.9</td>
<td>17.3</td>
</tr>
<tr>
<td>Between Group Contribution (Percent) (Comunas)</td>
<td>20.3</td>
<td>18.3</td>
<td>23.9</td>
<td>28.5</td>
</tr>
</tbody>
</table>

Notes:
1. Log Variance
2. Ranking by ascending order of inequality.
3. Between Group contribution if grouped by comunas. See Map 1.
Table A.2: SPATIAL INEQUALITY IN CALI: INDIVIDUALS RANKED BY HINCAP - 1973-1978

(Inequality Index: Log Variance)

<table>
<thead>
<tr>
<th></th>
<th>1973</th>
<th></th>
<th>1977</th>
<th></th>
<th>1978</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L Ranking</td>
<td></td>
<td>L Ranking</td>
<td></td>
<td>L Ranking</td>
</tr>
<tr>
<td>Sector 1</td>
<td>0.925 4</td>
<td></td>
<td>1.117 7</td>
<td></td>
<td>1.212 7</td>
</tr>
<tr>
<td>Sector 2</td>
<td>0.970 5</td>
<td></td>
<td>0.446 1</td>
<td></td>
<td>0.604 3</td>
</tr>
<tr>
<td>Sector 3</td>
<td>0.792 3</td>
<td></td>
<td>0.502 2</td>
<td></td>
<td>0.534 1</td>
</tr>
<tr>
<td>Sector 4</td>
<td>0.616 1</td>
<td></td>
<td>0.536 3</td>
<td></td>
<td>0.573 2</td>
</tr>
<tr>
<td>Sector 5</td>
<td>0.774 2</td>
<td></td>
<td>0.578 4</td>
<td></td>
<td>0.626 4</td>
</tr>
<tr>
<td>Sector 6</td>
<td>1.472 7</td>
<td></td>
<td>1.042 6</td>
<td></td>
<td>1.047 5</td>
</tr>
<tr>
<td>Sector 7</td>
<td>1.128 6</td>
<td></td>
<td>0.615 5</td>
<td></td>
<td>1.144 6</td>
</tr>
<tr>
<td>Overall</td>
<td>0.973 0.767 0.809</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Group Contribution (Sectors)</td>
<td>0.121 0.119 0.116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Group Contribution (Percent) Contribution (Sectors)</td>
<td>12.4 15.5 14.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Group Contribution (Percent) Contribution (Comunas)</td>
<td>26.5 32.1 30.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:
1. Log Variance.
2. Ranking by ascending order of Inequality.
3. Between group contribution if grouped by comunas (See Map 2).
Table A.3: THE SPATIAL DISTRIBUTION OF LABOUR EARNINGS IN BOGOTA 1973-1978
(Inequality Index: Log Variance)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L Ranking</td>
<td>L Ranking</td>
<td>L Ranking</td>
<td>L Ranking</td>
</tr>
<tr>
<td>Sector 1</td>
<td>0.910 6</td>
<td>1.773 8</td>
<td>1.026 8</td>
<td>0.403 2</td>
</tr>
<tr>
<td>Sector 2</td>
<td>0.577 1</td>
<td>0.770 3</td>
<td>0.454 1</td>
<td>0.402 1</td>
</tr>
<tr>
<td>Sector 3</td>
<td>0.593 2</td>
<td>0.674 2</td>
<td>0.553 2</td>
<td>0.462 3</td>
</tr>
<tr>
<td>Sector 4</td>
<td>0.707 3</td>
<td>0.591 1</td>
<td>0.661 4</td>
<td>0.644 4</td>
</tr>
<tr>
<td>Sector 5</td>
<td>0.785 5</td>
<td>1.187 6</td>
<td>0.756 5</td>
<td>0.609 6</td>
</tr>
<tr>
<td>Sector 6</td>
<td>0.725 4</td>
<td>0.881 4</td>
<td>0.595 3</td>
<td>0.646 5</td>
</tr>
<tr>
<td>Sector 7</td>
<td>1.204 7</td>
<td>1.057 5</td>
<td>0.847 6</td>
<td>0.866 7</td>
</tr>
<tr>
<td>Sector 8</td>
<td>1.911 8</td>
<td>1.024 7</td>
<td>0.938 7</td>
<td>1.226 8</td>
</tr>
<tr>
<td>Total</td>
<td>0.885</td>
<td>0.990</td>
<td>0.703</td>
<td>0.742</td>
</tr>
<tr>
<td>Between Group Contribution (Sectors)</td>
<td>0.089</td>
<td>0.047</td>
<td>0.089</td>
<td></td>
</tr>
<tr>
<td>Between Group Contribution (Percent) (Sectors)</td>
<td>7.4</td>
<td>8.9</td>
<td>6.7</td>
<td>11.9</td>
</tr>
<tr>
<td>Between Group Contribution (Percent) Comunas</td>
<td>13.8</td>
<td>18.7</td>
<td>13.9</td>
<td>16.7</td>
</tr>
</tbody>
</table>

Notes:
1. Log Variance
2. Ranking by ascending order of inequality.
3. Between group contribution if grouped by Comunas.

Sources: Same as Table 2.
Table A.4: THE SPATIAL DISTRIBUTION OF LABOUR EARNINGS IN CALI 1973-1978

(Inequality Index: Log Variance)

<table>
<thead>
<tr>
<th>Sector</th>
<th>1973</th>
<th>1977</th>
<th>1978</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L- Ranking</td>
<td>L Ranking</td>
<td>L Ranking</td>
</tr>
<tr>
<td>Sector 1</td>
<td>0.786</td>
<td>4</td>
<td>0.584</td>
</tr>
<tr>
<td>Sector 2</td>
<td>2.080</td>
<td>7</td>
<td>1.035</td>
</tr>
<tr>
<td>Sector 3</td>
<td>0.690</td>
<td>2</td>
<td>0.522</td>
</tr>
<tr>
<td>Sector 4</td>
<td>0.611</td>
<td>1</td>
<td>0.577</td>
</tr>
<tr>
<td>Sector 5</td>
<td>0.785</td>
<td>3</td>
<td>0.554</td>
</tr>
<tr>
<td>Sector 6</td>
<td>1.480</td>
<td>6</td>
<td>0.911</td>
</tr>
<tr>
<td>Sector 7</td>
<td>0.998</td>
<td>5</td>
<td>0.633</td>
</tr>
</tbody>
</table>

| Total    | 0.929  | 0.706  | 0.703  |
| Between Group Contribution (Sectors) | 0.026  | 0.037  | 0.046  |
| Between Group Contribution (Sectors) | 2.8    | 5.2    | 6.5    |
| Between Group Contribution (Percent) (Comunas) | 7.2    | 12.9   | 12.7   |

Notes:
1. Log Variance.
2. Ranking by ascending order of inequality.
3. Between group contribution if grouped by comunas.

Sources: As in Table 2.
Table A.5: THE DISTRIBUTION OF EARNINGS BY OCCUPATIONAL CATEGORIES 1978
(13 Categories)

<table>
<thead>
<tr>
<th>Occupational Category</th>
<th>BOGOTA</th>
<th></th>
<th></th>
<th>CALI</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ILO Earnings (1978 Pesos)</td>
<td>Theil Index in Category</td>
<td>Percent</td>
<td>Mean Earnings (1978 Pesos)</td>
<td>Theil Index in Category</td>
<td>Percent</td>
</tr>
<tr>
<td>Professional and Technical</td>
<td>1-19</td>
<td>17512</td>
<td>0.300</td>
<td>12.0</td>
<td>15454</td>
<td>0.304</td>
</tr>
<tr>
<td>Administrative and Manager</td>
<td>20-29</td>
<td>26834</td>
<td>0.137</td>
<td>2.9</td>
<td>19749</td>
<td>0.288</td>
</tr>
<tr>
<td>Clerk and Typist</td>
<td>30-39</td>
<td>7005</td>
<td>0.219</td>
<td>14.3</td>
<td>6636</td>
<td>0.250</td>
</tr>
<tr>
<td>Sales Managers</td>
<td>40-41</td>
<td>9370</td>
<td>0.540</td>
<td>7.2</td>
<td>15440</td>
<td>0.276</td>
</tr>
<tr>
<td>Other Sales</td>
<td>42-49</td>
<td>7004</td>
<td>0.482</td>
<td>8.8</td>
<td>5305</td>
<td>0.492</td>
</tr>
<tr>
<td>Service Workers (excl. Maids)</td>
<td>50-53</td>
<td>4023</td>
<td>0.292</td>
<td>12.3</td>
<td>4302</td>
<td>0.303</td>
</tr>
<tr>
<td>Maids</td>
<td>54</td>
<td>3237</td>
<td>0.165</td>
<td>7.7</td>
<td>3091</td>
<td>0.103</td>
</tr>
<tr>
<td>Agriculture Workers</td>
<td>60-69</td>
<td>11151</td>
<td>0.579</td>
<td>0.9</td>
<td>12509</td>
<td>0.586</td>
</tr>
<tr>
<td>Production Supervisors</td>
<td>70</td>
<td>7414</td>
<td>0.197</td>
<td>0.8</td>
<td>8285</td>
<td>0.199</td>
</tr>
<tr>
<td>Production Workers</td>
<td>71-74</td>
<td>4584</td>
<td>0.207</td>
<td>20.9</td>
<td>4776</td>
<td>0.297</td>
</tr>
<tr>
<td>Construction Workers</td>
<td>95</td>
<td>3901</td>
<td>0.151</td>
<td>5.5</td>
<td>4217</td>
<td>0.307</td>
</tr>
<tr>
<td>Transport Workers</td>
<td>96-97</td>
<td>3592</td>
<td>0.102</td>
<td>2.0</td>
<td>3811</td>
<td>0.181</td>
</tr>
<tr>
<td>Total</td>
<td>1-98</td>
<td>7600</td>
<td>0.482</td>
<td>100.0</td>
<td>6350</td>
<td>0.440</td>
</tr>
<tr>
<td>Total No. of Workers</td>
<td></td>
<td>1,180,000</td>
<td></td>
<td></td>
<td>350,000</td>
<td></td>
</tr>
</tbody>
</table>
Table A.6: THE DISTRIBUTION OF EARNINGS BY INDUSTRY OF ACTIVITY 1978  
(20 Categories)

<table>
<thead>
<tr>
<th>Industrial Category</th>
<th>SIC Code</th>
<th>Mean Earnings (1978 Pesos)</th>
<th>Theil Percent Index</th>
<th>Percent in Category</th>
<th>Mean Earnings (1978 Pesos)</th>
<th>Theil Percent Index</th>
<th>Percent in Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>10-19</td>
<td>11330</td>
<td>0.518</td>
<td>1.0</td>
<td>14271</td>
<td>0.524</td>
<td>0.9</td>
</tr>
<tr>
<td>Mining</td>
<td>20-29</td>
<td>18927</td>
<td>0.328</td>
<td>0.4</td>
<td>2098</td>
<td>0.170</td>
<td>0.4</td>
</tr>
<tr>
<td>Food Products</td>
<td>31</td>
<td>6250</td>
<td>0.344</td>
<td>3.1</td>
<td>6621</td>
<td>0.332</td>
<td>4.4</td>
</tr>
<tr>
<td>Beverages, Tobacco</td>
<td>32</td>
<td>4745</td>
<td>0.388</td>
<td>6.4</td>
<td>4831</td>
<td>0.460</td>
<td>12.0</td>
</tr>
<tr>
<td>Textiles and Footwear</td>
<td>32</td>
<td>4745</td>
<td>0.388</td>
<td>6.4</td>
<td>4831</td>
<td>0.460</td>
<td>12.0</td>
</tr>
<tr>
<td>Lumber and wood</td>
<td>33</td>
<td>4877</td>
<td>0.296</td>
<td>1.9</td>
<td>6621</td>
<td>0.332</td>
<td>4.4</td>
</tr>
<tr>
<td>Paper, Printing</td>
<td>34</td>
<td>8014</td>
<td>0.407</td>
<td>1.9</td>
<td>10137</td>
<td>0.504</td>
<td>2.6</td>
</tr>
<tr>
<td>and Publishing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mineral Products</td>
<td>36</td>
<td>6557</td>
<td>0.423</td>
<td>1.2</td>
<td>6548</td>
<td>0.500</td>
<td>1.0</td>
</tr>
<tr>
<td>Industrial Chemicals</td>
<td>35</td>
<td>10163</td>
<td>0.500</td>
<td>2.8</td>
<td>9303</td>
<td>0.366</td>
<td>3.1</td>
</tr>
<tr>
<td>Metal Industry</td>
<td>37-38</td>
<td>8559</td>
<td>0.476</td>
<td>5.5</td>
<td>6831</td>
<td>0.321</td>
<td>5.4</td>
</tr>
<tr>
<td>Other Industry</td>
<td>30-39</td>
<td>3244</td>
<td>0.163</td>
<td>0.9</td>
<td>3437</td>
<td>0.148</td>
<td>0.6</td>
</tr>
<tr>
<td>Utilities</td>
<td>40-49</td>
<td>8668</td>
<td>0.214</td>
<td>0.5</td>
<td>9208</td>
<td>0.249</td>
<td>1.0</td>
</tr>
<tr>
<td>Construction</td>
<td>50-59</td>
<td>5460</td>
<td>0.394</td>
<td>7.3</td>
<td>4955</td>
<td>0.281</td>
<td>6.2</td>
</tr>
<tr>
<td>Wholesale Trade</td>
<td>61</td>
<td>18894</td>
<td>0.367</td>
<td>1.6</td>
<td>11019</td>
<td>0.431</td>
<td>1.6</td>
</tr>
<tr>
<td>Retail Trade</td>
<td>62</td>
<td>6561</td>
<td>0.463</td>
<td>18.3</td>
<td>5516</td>
<td>0.478</td>
<td>19.1</td>
</tr>
<tr>
<td>Other Commerce</td>
<td>60-63</td>
<td>3972</td>
<td>0.272</td>
<td>3.4</td>
<td>4020</td>
<td>0.231</td>
<td>3.6</td>
</tr>
<tr>
<td>Transport and</td>
<td>70-79</td>
<td>8522</td>
<td>0.333</td>
<td>5.9</td>
<td>6778</td>
<td>0.271</td>
<td>6.4</td>
</tr>
<tr>
<td>Communication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial</td>
<td>80-89</td>
<td>12232</td>
<td>0.424</td>
<td>8.2</td>
<td>11465</td>
<td>0.436</td>
<td>3.8</td>
</tr>
<tr>
<td>Establishments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public Instruction</td>
<td>93</td>
<td>10356</td>
<td>0.351</td>
<td>7.9</td>
<td>8996</td>
<td>0.348</td>
<td>7.4</td>
</tr>
<tr>
<td>Administration and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Services</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public Instruction</td>
<td>95</td>
<td>3760</td>
<td>0.287</td>
<td>14.1</td>
<td>3598</td>
<td>0.284</td>
<td>14.6</td>
</tr>
<tr>
<td>Domestic and Personal Services</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>7594</td>
<td>0.482</td>
<td>100.0</td>
<td>6322</td>
<td>0.441</td>
<td>100.0</td>
</tr>
</tbody>
</table>
APPENDIX II

THE DATA

The City Study has assembled a data bank of existing sources of data in Bogota and Cali, Colombia in the form of copies of the original computer tapes prepared by the respective originators of the data. All have been documented in detail by Nelson Valverde (1978) and Y.J. Lee (1978). This study utilizes 4 of these data sets; all originally collected by DANE. The 1978 survey was conducted jointly by the City Study and DANE.

1973 Population Census

The last census undertaken was in 1973 and the one immediately preceding was in 1964. Unfortunately the 1964 census does not report incomes nor does it have intra-city spatial detail. There has been wide skepticism concerning the coverage of the 1973 census but Potter and Ordonez (1976) concluded after a careful demographic analysis that the information they analyzed from the advance sample appeared to be of good quality, at least in relation to previous censuses. They estimated that the overall under-enumeration for Colombia as a whole was probably about 7 percent.


2/ One example of such skepticism is Lubell and McCallum (1978) (p. 126) who regard the Bogota 1973 results as "simply not useable" and therefore place no reliance on the census. For their calculations and projections they prefer the 1972 Urban Development Study Household Survey covering 4675 households.
The public use tape provided by DANE is a 4% sample of households. However, for Bogota the tape contains all households living in the buildings which happen to have the households in the 4% sample. We have used the whole sample for tabulations in this paper but the sample has been expanded to reflect the size of the city accounting for estimated underenumeration. Since this study is particularly concerned with the spatial distributions within the city, the sample was expanded appropriately to be representative of the city. Details of this methodology have been given in Appendix (3) of Mohan (1980).

The census contains information on dwelling characteristics, household characteristics, demographic information on all individuals, labor force information for workers and fertility information on females. While I cannot comment on its coverage the overall quality of the information in the sample appears to be good and I am therefore in agreement with Potter and Ordoñez (1976). Non-responses appear to be distributed randomly; the only obvious bias is that single member households predominate in the no information categories in income and labor force information. One of the distinguishing features of this data set which makes it very useful for us is that the location of respondents is coded down to the block (manzana) of residence in the city. The following section will describe the consistent geo-coding system used by DANE in all of its work. The income was obtained from only one question "what was your income from all sources last month?" Only about 12 percent of the sample did not report income information--a proportion which compares well with non-responses in the U.S. census.
Appendix (2) of Mohan (1980) reported the results of an estimation of the income coverage of the 1973 census as well as the 1977 Household Survey (see below). A summary of the method is given in Table A.a. It appears that an aggregation of all incomes reported in the census amounted to no more than 50 percent of the estimated total personal income for Bogota. Various factors responsible for this under-reporting can be enumerated:

- 12% of the people gave no income information.
- When only one question is asked, much of the non-labor earnings are probably not reported.
- Income in kind, e.g. as received by domestic servants, is probably not reported.
- Many earners receive end-of-year bonuses; these are characteristically not covered in one-shot cross-section surveys such as the census unless the question is asked specifically.

Keeping these factors in mind it is then not surprising that the income coverage of the census was only about 50%.1/

**Household Surveys**

DANE has conducted a regular program of household surveys from 1970 with the main objective of collecting information on the labor force. Since 1975 these surveys have been quarterly and are conducted alternatingly in the four largest cities and in the seven largest cities 2/ along with

1/ This may be compared with a recent estimate of under-coverage in the Brazil 1960 and 1970 censuses where Pfefferman and Webb (1978) estimate that the censuses cover about 57 - 58 percent of the income. (p. 16).

2/ Bogota, Cali, Medellin, Bucaramanga, Barranquilla, Manizales and Pasto. The first four are the four largest.
an occasional national survey. We have obtained the computer tapes for 1972 (Encuesta de Hogares - Fuerza de Trabajo: EH 6-FT), 1975 (EH8E), and 1977 (EH15). The 1972 survey was a national one and covered 6371 households of whom 1348 were in Bogota. It contains information on housing as well as demographic and labor force characteristics. This survey does not provide the intra-city location of the respondents. The 1975 survey was a special one for the city of Bogota and sampled 3953 households and contains information on demographic and labor force characteristics only. The 1977 survey was conducted in the four largest cities and sampled 6082 households of whom 3161 households, were in Bogota. Starting with this survey DANE has begun to use "rotational sampling" such that 67% of dwelling units sampled remain in the next survey and 33% are new. Both the 1975 and 1977 surveys contain the location of residence of the respondents. The 1972 and 1975 surveys also have firm size information on the number of employees working in each respondents' work place. This question was not included in the 1977 sample. In carrying out these samples DANE classifies neighborhoods into 6 socio-economic strata: 1 low-low, 2 low, 3 medium low, 4 medium, 5 medium high and 6 high. At the conclusion of a survey weights are assigned to each of these strata which are then applied to the members of these strata for all expansions of the sample. These "expansion factors" are supposed to account for over and under sampling that might occur over the course of the survey. The expanded sample should then be correctly representative of the city as a whole.
The 1977 survey is unusual in that as many as 20 percent of the working respondents did not report their incomes: a proportion which compares very unfavorably with the census. Because of this high proportion a method was devised to impute incomes to the non-respondents. Appendix (1) of Mohan (1980) described the calculation of incomes for all respondents and imputation method for non-respondents.

The coverage of income in this survey is not much better than that in the census despite the more detailed questions asked. Labor income and non-labor income data are taken separately and income in kind is estimated as well. Even when the imputed incomes are included the survey covers only about 61 percent of estimated total personal income of Bogota. It also appears that the highest incomes are either under-reported or undersampled. If the incomes reported in the 1977 sample are converted to 1973 pesos it is found that, on average, there is little real growth in incomes, while those in the highest categories actually decline.

The 1978 City Study – DANE Survey

The World Bank City Study and DANE jointly conducted a survey of about 3000 households in Bogota and 1000 households in Cali. This was an expanded as well as more carefully conducted survey than previous ones. The survey had 5 main parts: i) Household and dwellings characteristics. ii) Demographic characteristics of all individuals. iii) Worker characteristics including information on their place of work. iv) Information on the unemployed. v) Information on vehicle ownership as well as journey to work characteristics for the workers.
A few points are worth noting with regard to this survey. A partial recount of dwelling units was done to account for the expansion of Bogota since the 1973 census (earlier surveys were all based on a 1972 sample frame). Information on income was elicited more carefully. In earlier surveys, earnings of all workers in a household were usually obtained from any adult respondent available in the household. In this survey, all worker information was obtained from each worker directly even if it required re-visits to the household. Furthermore, income questions were asked of all members of the household even if they did not work. Two questions were asked to obtain labour income: the wage and periodicity of wage payments and also total earnings in the previous month. Income in kind was imputed. Various non-labour sources of income were specifically mentioned to obtain non-labour income. As a result, the income coverage of this survey is about 90 percent (as shown in Table A.a below) which is a great improvement over previous surveys.

The percentage of no information on income was only and these have been imputed by the same method as for the 1977 Household Survey referred to earlier. All regressions have been conducted after weighting each observation with expansion factors similar to the procedure for 1973 described above. These expansion factors account for ex-post over and under sampling as compared with the sample frame.

Household Survey Samples and the Spatial Disaggregation of Bogota

Map 1 (in Section III) shows a representation of the map of Bogota. The basic socio-economic spatial unit in Bogota is a "barrio"
or neighborhood of which there were about 500 in 1973 and about 700 now as a result of rapid growth of the city. DANE geo-codes this unit in a 4 digit number of which the first 2 digits identify a comuna—a collection of barrios. The last 2 digits then identify barrios within a comuna. These were then further aggregated into "rings" and "sectors."
The boundaries of the comunas shown in the map are principal streets in Bogota. The city is bounded in the East by mountains and therefore has an approximately semi-circular shape, although it is longer going North to South as is evident from the map. As an aid to understanding the numbering system, note that the first digit goes from 1 to 9 and roughly rotates (increasing) from South to North by sectors (or pie slices). The second digit ranges from 1 to 6 and corresponds roughly to rings centered in comunas 31 and 81 and increasing from South to North.

DANE along with the Ministry of Health compiled an inventory of blocks ("manzanas") and of dwelling units within the city before the census in 1973 and that inventory has continued to form the sample frame of all subsequent surveys in Bogota. Thus none of these surveys had sampled the new neighborhoods that have developed in the past 5 years. The sampling is designed to make it equi-probable that any dwelling unit in the city according to the 1972 inventory may be selected. The basic unit of sampling (unidad primaria de muestra) is a block within which all households in all dwelling units are interviewed. Provision is made for different sizes of blocks. Since all the sampling was based on the 1972 sample frame it was difficult to trace time trends in the changes within the city. Moreover, any conclusions that are
drawn about the changing character of neighborhoods must be cautious. If different regions of the city differ from one another systematically and if one region changes character over time the later samples would no longer be representative. Sampling is based on the classification of neighborhoods into the 6 socio-economic strata. If neighborhoods change character i.e., filter up or down in the socio-economic scale, the resulting sample would then no longer be representative. Hence, drawing conclusions about fine changes in income distribution from two household surveys at two points of time is a hazardous business without detailed knowledge of the sampling procedures used. If, however, rates of change are not high, such difficulties are minimal: but even one would have less interest in tracing time trends anyway! These remarks may be extended to the coverage of national surveys where the heterogeneity of regions is perhaps typically more pronounced than within a city.

These details have been offered here since they are seldom given by users of household survey data. They became particularly important when comparisons are made between surveys of different years and information on each data source is necessary in order to be aware of biases that may arise from differences in survey design and coverage.
Table A. a.  THE COVERAGE OF INCOME IN BOGOTA IN HOUSEHOLD SURVEYS

There is considerable skepticism concerning the coverage of income in household surveys. Detailed estimates of the coverage of Bogota have been made by comparing the total personal income from regional accounts, national income accounts and incomes as revealed in surveys. 1/
The results are as follows:

<table>
<thead>
<tr>
<th>Survey</th>
<th>Income Covered</th>
</tr>
</thead>
<tbody>
<tr>
<td>1973 Census</td>
<td>49.3 percent</td>
</tr>
<tr>
<td>1977 Household Survey</td>
<td>61.3 percent</td>
</tr>
<tr>
<td>1978 Household Survey</td>
<td>92.0 percent</td>
</tr>
</tbody>
</table>

Outline of Method  

<table>
<thead>
<tr>
<th>Outline of Method</th>
<th>Millions of 1978 Colombian Pesos</th>
</tr>
</thead>
<tbody>
<tr>
<td>1978 Colombia GDP</td>
<td>(1) 870,000</td>
</tr>
<tr>
<td>Bogota GRP</td>
<td>(2) 200,100 (..22 of 1)</td>
</tr>
<tr>
<td>Personal Income</td>
<td>(3) 130,065 (.65 of 2)</td>
</tr>
<tr>
<td>From Survey</td>
<td>(4) 119,634</td>
</tr>
<tr>
<td>Survey Coverage</td>
<td>92.0% (4/3)</td>
</tr>
</tbody>
</table>

1/ Details in Rakesh Mohan (1980), Appendix 2.
APPENDIX III
EQUALISE:

A Program for Income Distribution Studies

By

Sungyong Kang+
Sudhir Anand*
Sherman Robinson+
Rakesh Mohan+

* St. Catherines College, Oxford
+ World Bank
1. Introduction

This document is intended as a user's guide to the program EQUALISE. The program is intended as an aid for income distribution studies that are increasingly being done using large cross sectional micro data sets. Its special feature is that it can use as input the micro data directly rather than grouped data as has been necessary for other existing programs computing various inequality indices.

The program performs three tasks. It has a subroutine CROST which essentially performs cross-tabulations. For example, tables can be produced for the distribution of income by any desired number of quantilies for different subgroups of the population like different regions. Cumulative frequencies are also available. Second, subroutine MEADIS computes seven selected measures of income distribution from any desired number of observations on the Lorenz curve. Thus the micro data can either be used directly to compute these indices or first grouped into quantiles by CROST and fed subsequently into MEADIS. This is accomplished internally in the program obviating the need for tedious punching of data for the computation of income distribution statistics. Lastly, subroutines ANANDT and ANANDL decompose the Theil inequality index and the log variance of incomes for different subgroups of the population. These indices are computed for each sub-group desired and measures provided for the inter-group and intra-group contribution to total inequality. EQUALISE is flexible so that any or all of the above operations can be performed in one single run. MEADIS was originally written by Sherman Robinson and ANANDT and ANANDL by Sudhir Anand. 

1/ For detailed theoretical analysis and description of the statistics provided by this program, see Robinson (1976) and Anand (1982).
The rest of the document is organized as follows; section 2 describes the general structural outline of the program; section 3 describes the input requirements; section 4 points out the changes needed in the main routine for different runs; section 5 lists the data and the results of a sample run; and section 6 lists the program for the sample run.

2. Structure of EQUALISE

The program consists of a MAIN program and four subroutines. Only the main program contains parts that need to be changed by the user. This point will be elaborated further in section 4. Also, all inputs are read in the main program and subroutines are called only in the main program. Following is the list of subroutines called by the main program.

Group 1: CROST with Entry Point MEAN

Group 2: MEADIS
ANANDT
ANANDL

The cross-sectional micro data (containing socio-economic variables including income) are read in the MAIN program and then fed into the subroutine CROST where they are cross-tabulated. When all the records are accounted for, the subroutine CROST is entered at the Entry MEAN where the mean of each category of column variable (income) is calculated. Also, the frequency table of the column variable is printed along with the mean of each category.

Subroutines in Group 2 are the routines that calculate the income distribution measures and the decomposition statistics. These
routines use the results from Group 1 as their inputs. Subroutine MEADIS produces 7 selected income distribution measures: relative mean deviation, variance of income, coefficient of variation, standard deviation of logs, the Theil index, the Atkinson measure and the Gine coefficient.

Subroutines ANANDT and ANANDL produce decomposition statistics for the Theil and Log-variance indices respectively.

The user has an option of producing either the income distribution measures from MEADIS only or the decomposition statistics from ANANDT and ANANDL or both simultaneously.

3. **Input Requirements**

There are two types of input that are needed for this program. The first is a card deck or file containing control parameters, input file name, variable names and the definition of categories for row and column variables. The second input is a data file containing socio-economic variables including an income variable. The second input requires the correct format statement and data transformations, if necessary, to be specified in the MAIN program. The rest of this section will elaborate on the first input that has to be prepared by the user.

The input data for each job to be prepared by the user are summarized in the following table, followed by a detailed discussion.
### Input Data for Each Job

<table>
<thead>
<tr>
<th>Card Type</th>
<th>No. of Cards</th>
<th>Variable</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Input file name</td>
<td>10A6</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>NOUT1, NOUT2, N</td>
<td>3I2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>ICOR1, NI, NAME1, ICOR2, NJ, NAME2</td>
<td>2(2I3, A6)</td>
</tr>
<tr>
<td>4</td>
<td>NI/5</td>
<td>C1</td>
<td>5F8.0</td>
</tr>
<tr>
<td>5</td>
<td>NJ/5</td>
<td>C2</td>
<td>5F8.0</td>
</tr>
</tbody>
</table>

Multiple jobs may be done in the same run by simply repeating card types 3, 4 and 5 for each additional job. The variable N in card type 2 should reflect the number of jobs (or matrices) to be tabulated.

Explanation of the above variables are as follows:

1) **Input File Name:** the name of the input file containing socio-economic variables. This will be printed as part of the output heading.

2) **NOUT1:** Combination of output desired.
   - 1 = MEADIS income distribution measures only.
   - 2 = Decomposition statistics only.
   - 3 = 1 and 2 both.

3) **NOUT2:** output desired.
   - 0 = frequency table of income classes not printed.
   - 1 = frequency table printed.

4) **N:** number of jobs or matrices.
5) ICOR1: position of the row variable in the array named "DATA". (i.e. if the row variable is the second variable in the array called "DATA", then ICOR1 = 2). Variables from the socio-economic data file are read into the DATA array in the MAIN program. (See below).

6) NI: number of categories in the row variable.

7) NAME1: name of the row variable. Must be six characters or less.

8) ICOR2: position of the column variable in the DATA array. The column variable must always be the income variable.

9) NJ: number of categories in the column variable.

10) NAME2: name of the column variable.

11) C1: maximum value of each category for the row variable. The number of values in these cards should equal the number of categories NI. The value zero is included in the first category unless specifically excluded in the MAIN program. All observations exceeding the maximum value of the last category are excluded and not tabulated.

12) C2: maximum value of each category for the column variable. The number of values in these cards should equal the number of categories NJ.

As noted before, the above input data can be in the form of either a card deck or a file. If it is a file, the file number 5 should be assigned to it.
4. **Changes Needed in the MAIN Program**

The program is designed in such a way that the users must change some parts of the MAIN program. This is rather unfortunate but couldn't be avoided since the input file containing socio-economic data is read in the MAIN program. In other words, since these input files will be different depending on the user, each user must supply the necessary format statement and any other input file specific information.

Basically, there are only two parts of the program that need to be changed. The first part is the cluster of dimension statements at the beginning of the main program. These should be dimensioned in the following way:

1. **Arrays needed for initializing data**
   
   FILE(10)
   
   ICOR1 (N), ICOR2 (N), NI(N), NJ(N)
   NAME1 (N), NAME2 (N), C1 (N, M1), C2 (N, M2)

2. **Arrays needed for subroutines CROST and MEADIS**
   
   MAT (N, M1, M2), MCT (N, M2), SUM (N, M2)
   MATS (M1, M2), MCTS (M2), SUMS (M2), AVGS (M2)

3. **Arrays needed for subroutines ANANDT and ANANDL**
   
   ANL (M2), ANT (M1), S (M1), T (M1)
   ANY (M1), YBAR (M1), P (M1)
   PP (M1, M2), SS (M1, M2)

4. **Arrays needed for reading input file**
   
   DATA (NVAR)
Where:

\[ N = \text{number of jobs (matrices)} \]

\[ M_1 = \text{maximum number of categories among all the row variables} \]

\[ M_2 = \text{maximum number of categories among all the column (income) variables}. \]

\[ NVAR = \text{number of variables to be read from the input file containing socio-economic data}. \]

The second set of changes start at the statement labelled 21, and end with the definition of the expansion factor, EXP. The statement 21 is a READ statement for the socio-economic data file. The user must supply correct READ and FORMAT statement appropriate for his file. These statements should be followed (if necessary) by transformations of the data. The last statement of this part defines the expansion factor. If the input data is sample data as most survey data are, then the user must define the expansion factor for each record. If the data represent the universe, the expansion factor (EXP) must be set to unity.

These changes in the MAIN program mean that the user must recompile the program each time a different file is used as input. This might be cumbersome, but the cost involved is rather negligible since the processing time is less than 3 seconds in most cases.
5. **Sample Run: Input Data and Output**

A Sample job was run using the 1978 Bogota Household Survey Data. Each record represents one household. The income variable used was HINCAP - the household income per capita. Since the calculation of inequality is done over individuals (rather than households) the expansion factor associated with the household record is multiplied by the household size.

Two row variables are used in this job. First, the sample is divided into comunas (38) and then into radial sectors (8) (see Appendix 1 for definitions). Thus the program produces the 7 indices of inequality for the whole sample from MEADIS; a cross-tabulation is provided by CROST by the 20 income groups specified and overall inequality is decomposed by ANANDT and ANANDL first by comunas and then by radial sectors. All possible outputs are produced except that the repeated printing of the frequency table is suppressed.

The following page reproduces the input data fed - except for the household survey file itself, followed by a listing of a sample of output that can be produced by EQUALISE. A program listing follows the sample output.
### SAMPLE INPUT DATA

1978 BOGOTA HOUSEHOLD SURVEY < 3062 RECORDS >

<table>
<thead>
<tr>
<th>Comuna</th>
<th>Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sector</th>
<th>Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>
SAMPLE OUTPUT FROM "EQUALISE"

<< MATRIX # 1 >>

ROW VARIABLE= COMUNA
COL VARIABLE= INCOME

THE ROW VARIABLE HAS 38 CATEGORIES.
MAXIMUM VALUE OF EACH CATEGORY ARE AS FOLLOWS:

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>23.</td>
<td>24.</td>
<td>25.</td>
<td>31.</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>41.</td>
<td>42.</td>
<td>43.</td>
<td>44.</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>51.</td>
<td>52.</td>
<td>53.</td>
<td>54.</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>56.</td>
<td>61.</td>
<td>62.</td>
<td>63.</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>65.</td>
<td>71.</td>
<td>72.</td>
<td>73.</td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>81.</td>
<td>82.</td>
<td>83.</td>
<td>84.</td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>91.</td>
<td>92.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

THE COLUMN VARIABLE HAS 20 CATEGORIES.
MAXIMUM VALUE OF EACH CATEGORY ARE AS FOLLOWS:

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>296</td>
<td>592.</td>
<td>704.</td>
<td>816.</td>
<td>936.</td>
<td></td>
</tr>
<tr>
<td>1057</td>
<td>1197.</td>
<td>1338.</td>
<td>1481.</td>
<td>1624.</td>
<td></td>
</tr>
<tr>
<td>1812</td>
<td>2000.</td>
<td>2312.</td>
<td>2625.</td>
<td>3165.</td>
<td></td>
</tr>
<tr>
<td>3704</td>
<td>5022.</td>
<td>6340.</td>
<td>95670.</td>
<td>185000.</td>
<td></td>
</tr>
</tbody>
</table>

INC CLASS NO OF PEOPLE % CUM % TOTAL INCOME % CUM % MEAN INCOME

<p>| | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>46152.</td>
<td>1.4</td>
<td>1.4</td>
<td>8174537.</td>
<td>0.1</td>
<td>0.1</td>
<td>177.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>283393.</td>
<td>8.6</td>
<td>10.0</td>
<td>135478470.</td>
<td>1.5</td>
<td>1.5</td>
<td>478.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>144565.</td>
<td>4.4</td>
<td>14.5</td>
<td>94956182.</td>
<td>1.0</td>
<td>2.6</td>
<td>656.84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>189889.</td>
<td>5.8</td>
<td>20.2</td>
<td>144534880.</td>
<td>1.5</td>
<td>4.1</td>
<td>761.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>200357.</td>
<td>6.1</td>
<td>26.4</td>
<td>173691362.</td>
<td>1.9</td>
<td>6.0</td>
<td>886.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>136460.</td>
<td>4.2</td>
<td>30.5</td>
<td>136076406.</td>
<td>1.5</td>
<td>7.4</td>
<td>997.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>154283.</td>
<td>4.7</td>
<td>35.2</td>
<td>173314339.</td>
<td>1.9</td>
<td>9.3</td>
<td>1123.35</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>176878.</td>
<td>5.4</td>
<td>40.6</td>
<td>224137374.</td>
<td>2.4</td>
<td>11.7</td>
<td>1267.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>161403.</td>
<td>4.9</td>
<td>45.5</td>
<td>225783142.</td>
<td>2.4</td>
<td>14.1</td>
<td>1398.88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>149592.</td>
<td>4.6</td>
<td>50.1</td>
<td>231081827.</td>
<td>2.5</td>
<td>16.6</td>
<td>1544.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>177150.</td>
<td>5.4</td>
<td>55.5</td>
<td>303127155.</td>
<td>3.3</td>
<td>19.8</td>
<td>1711.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>174512.</td>
<td>5.3</td>
<td>60.8</td>
<td>336176043.</td>
<td>3.6</td>
<td>23.4</td>
<td>1926.38</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>163646.</td>
<td>5.0</td>
<td>65.8</td>
<td>351655082.</td>
<td>3.8</td>
<td>27.2</td>
<td>2148.88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>147066.</td>
<td>4.5</td>
<td>70.3</td>
<td>362980699.</td>
<td>3.9</td>
<td>31.1</td>
<td>2468.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>206153.</td>
<td>6.3</td>
<td>76.6</td>
<td>600367812.</td>
<td>6.4</td>
<td>37.5</td>
<td>2912.24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>120661.</td>
<td>3.7</td>
<td>80.3</td>
<td>413272638.</td>
<td>4.4</td>
<td>42.0</td>
<td>3425.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>226468.</td>
<td>6.9</td>
<td>87.2</td>
<td>986612519.</td>
<td>10.6</td>
<td>52.6</td>
<td>4354.53</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>118216.</td>
<td>3.6</td>
<td>90.8</td>
<td>674198578.</td>
<td>7.2</td>
<td>59.8</td>
<td>5703.11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>301890.</td>
<td>9.2</td>
<td>100.0</td>
<td>3618304571.</td>
<td>38.8</td>
<td>98.6</td>
<td>11985.51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>711.</td>
<td>0.0</td>
<td>100.0</td>
<td>121703350.</td>
<td>1.4</td>
<td>100.0</td>
<td>185295.78</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TOTAL 3279445. 100.0 100.0 9325176966. 100.0 100.0 2843.52
TOTAL NUMBER OF RECORDS READ = 3062
TOTAL NUMBER OF RECORDS PROCESSED= 2991

MEASURES OF OVERALL INCOME DISTRIBUTION
BASED ON THE 20 OBSERVATIONS FOR THE LORENZ CURVE

<table>
<thead>
<tr>
<th>Measure</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative Mean Deviation</td>
<td>2228.5092</td>
</tr>
<tr>
<td>Variance of Incomes</td>
<td>17394593.1360</td>
</tr>
<tr>
<td>Coefficient of Variation</td>
<td>1.4667</td>
</tr>
<tr>
<td>Standard Deviation of Logs</td>
<td>0.9258</td>
</tr>
<tr>
<td>Theil Measure</td>
<td>0.5057</td>
</tr>
<tr>
<td>Atkinson Measure</td>
<td>0.5750</td>
</tr>
<tr>
<td>Gini Coefficient</td>
<td>0.5167</td>
</tr>
</tbody>
</table>
## Decomposition of the Theil Index

- **N(I)** = No. of observations in the Ith group
- **YBAR(I)** = Mean income of the Ith group
- **T(I)** = Theil index within Ith group

<table>
<thead>
<tr>
<th>NO.</th>
<th>N(I)</th>
<th>YBAR(I)</th>
<th>T(I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>21987.0</td>
<td>1483.14</td>
<td>0.2366</td>
</tr>
<tr>
<td>2</td>
<td>63946.0</td>
<td>2910.05</td>
<td>0.3040</td>
</tr>
<tr>
<td>3</td>
<td>145750.0</td>
<td>1110.43</td>
<td>0.1619</td>
</tr>
<tr>
<td>4</td>
<td>107534.0</td>
<td>1530.73</td>
<td>0.1927</td>
</tr>
<tr>
<td>5</td>
<td>57394.0</td>
<td>3562.92</td>
<td>0.3547</td>
</tr>
<tr>
<td>6</td>
<td>44751.0</td>
<td>1745.60</td>
<td>0.3371</td>
</tr>
<tr>
<td>7</td>
<td>63105.0</td>
<td>2472.10</td>
<td>0.3098</td>
</tr>
<tr>
<td>8</td>
<td>118115.0</td>
<td>1835.69</td>
<td>0.2384</td>
</tr>
<tr>
<td>9</td>
<td>243862.0</td>
<td>1128.51</td>
<td>0.2025</td>
</tr>
<tr>
<td>10</td>
<td>79750.0</td>
<td>2152.51</td>
<td>0.3779</td>
</tr>
<tr>
<td>11</td>
<td>94946.0</td>
<td>1637.18</td>
<td>0.3066</td>
</tr>
<tr>
<td>12</td>
<td>65428.0</td>
<td>3832.95</td>
<td>0.3713</td>
</tr>
<tr>
<td>13</td>
<td>47274.0</td>
<td>2525.34</td>
<td>0.4249</td>
</tr>
<tr>
<td>14</td>
<td>64095.0</td>
<td>2504.32</td>
<td>0.3612</td>
</tr>
<tr>
<td>15</td>
<td>176287.0</td>
<td>2367.98</td>
<td>0.2999</td>
</tr>
<tr>
<td>16</td>
<td>367526.0</td>
<td>1951.26</td>
<td>0.3557</td>
</tr>
<tr>
<td>17</td>
<td>56722.0</td>
<td>5683.19</td>
<td>0.2852</td>
</tr>
<tr>
<td>18</td>
<td>18748.0</td>
<td>1451.38</td>
<td>0.3003</td>
</tr>
<tr>
<td>19</td>
<td>17202.0</td>
<td>1915.83</td>
<td>0.2731</td>
</tr>
<tr>
<td>20</td>
<td>69433.0</td>
<td>4173.99</td>
<td>0.4034</td>
</tr>
<tr>
<td>21</td>
<td>107688.0</td>
<td>1751.39</td>
<td>0.2781</td>
</tr>
<tr>
<td>22</td>
<td>398531.0</td>
<td>2127.91</td>
<td>0.3155</td>
</tr>
<tr>
<td>23</td>
<td>43763.0</td>
<td>4396.56</td>
<td>0.2689</td>
</tr>
<tr>
<td>24</td>
<td>16770.0</td>
<td>6014.09</td>
<td>0.3057</td>
</tr>
<tr>
<td>25</td>
<td>33166.0</td>
<td>3796.77</td>
<td>0.5180</td>
</tr>
<tr>
<td>26</td>
<td>110046.0</td>
<td>1800.71</td>
<td>0.2354</td>
</tr>
<tr>
<td>27</td>
<td>59634.0</td>
<td>2522.08</td>
<td>0.2651</td>
</tr>
<tr>
<td>28</td>
<td>26305.0</td>
<td>7958.98</td>
<td>0.1679</td>
</tr>
<tr>
<td>29</td>
<td>42957.0</td>
<td>5271.74</td>
<td>0.2763</td>
</tr>
<tr>
<td>30</td>
<td>41405.0</td>
<td>4345.34</td>
<td>0.2741</td>
</tr>
<tr>
<td>31</td>
<td>39076.0</td>
<td>4900.20</td>
<td>0.3146</td>
</tr>
<tr>
<td>32</td>
<td>49482.0</td>
<td>6631.32</td>
<td>0.2797</td>
</tr>
<tr>
<td>33</td>
<td>82045.0</td>
<td>4712.54</td>
<td>0.4325</td>
</tr>
<tr>
<td>34</td>
<td>76595.0</td>
<td>10665.46</td>
<td>0.4248</td>
</tr>
<tr>
<td>35</td>
<td>47646.0</td>
<td>7908.65</td>
<td>0.1901</td>
</tr>
<tr>
<td>36</td>
<td>43078.0</td>
<td>5558.06</td>
<td>0.4251</td>
</tr>
<tr>
<td>37</td>
<td>80446.0</td>
<td>4098.76</td>
<td>0.4376</td>
</tr>
<tr>
<td>38</td>
<td>52957.0</td>
<td>1368.88</td>
<td>0.3948</td>
</tr>
</tbody>
</table>

**Total sample size** = 3279445.0

- **Between group component** = 0.1820
- **Within group component** = 0.3236
- **Theil index of total sample** = 0.5057

- **Between group contribution** = 36.00%
- **Within group contribution** = 64.00%
DECOMPOSITION OF LOG VARIANCE OF INCOME

\[ N(I) = \text{NUMBER OF OBSERVATIONS IN THE ITH GROUP} \]
\[ \text{MEANLOG}(I) = \text{MEAN LOG INCOME OF THE ITH GROUP} \]
\[ V(I) = \text{VARIANCE LOG INCOME OF THE ITH GROUP} \]

<table>
<thead>
<tr>
<th>GROUP NO</th>
<th>N(I)</th>
<th>MEANLOG(I)</th>
<th>V(I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>21987.0</td>
<td>7.0856</td>
<td>0.3851</td>
</tr>
<tr>
<td>2</td>
<td>63946.0</td>
<td>7.6822</td>
<td>0.5767</td>
</tr>
<tr>
<td>3</td>
<td>145750.0</td>
<td>6.8391</td>
<td>0.3743</td>
</tr>
<tr>
<td>4</td>
<td>107534.0</td>
<td>7.1646</td>
<td>0.3019</td>
</tr>
<tr>
<td>5</td>
<td>57394.0</td>
<td>7.8055</td>
<td>0.7579</td>
</tr>
<tr>
<td>6</td>
<td>44751.0</td>
<td>7.1564</td>
<td>0.5610</td>
</tr>
<tr>
<td>7</td>
<td>63105.0</td>
<td>7.5369</td>
<td>0.4878</td>
</tr>
<tr>
<td>8</td>
<td>118115.0</td>
<td>7.2898</td>
<td>0.4244</td>
</tr>
<tr>
<td>9</td>
<td>243862.0</td>
<td>6.8303</td>
<td>0.3932</td>
</tr>
<tr>
<td>10</td>
<td>79750.0</td>
<td>7.3252</td>
<td>0.6265</td>
</tr>
<tr>
<td>11</td>
<td>94946.0</td>
<td>7.1077</td>
<td>0.5722</td>
</tr>
<tr>
<td>12</td>
<td>65428.0</td>
<td>7.8309</td>
<td>0.9220</td>
</tr>
<tr>
<td>13</td>
<td>47274.0</td>
<td>7.4179</td>
<td>0.7713</td>
</tr>
<tr>
<td>14</td>
<td>64095.0</td>
<td>7.4953</td>
<td>0.5889</td>
</tr>
<tr>
<td>15</td>
<td>176287.0</td>
<td>7.4974</td>
<td>0.5075</td>
</tr>
<tr>
<td>16</td>
<td>367526.0</td>
<td>7.2620</td>
<td>0.5006</td>
</tr>
<tr>
<td>17</td>
<td>60722.0</td>
<td>8.3171</td>
<td>0.7103</td>
</tr>
<tr>
<td>18</td>
<td>18748.0</td>
<td>7.0014</td>
<td>0.5189</td>
</tr>
<tr>
<td>19</td>
<td>17202.0</td>
<td>7.2371</td>
<td>0.7048</td>
</tr>
<tr>
<td>20</td>
<td>69433.0</td>
<td>7.8408</td>
<td>1.581</td>
</tr>
<tr>
<td>21</td>
<td>107688.0</td>
<td>7.2024</td>
<td>0.5047</td>
</tr>
<tr>
<td>22</td>
<td>398531.0</td>
<td>7.3654</td>
<td>0.5560</td>
</tr>
<tr>
<td>23</td>
<td>43763.0</td>
<td>8.0894</td>
<td>0.6545</td>
</tr>
<tr>
<td>24</td>
<td>16770.0</td>
<td>8.3230</td>
<td>0.8656</td>
</tr>
<tr>
<td>25</td>
<td>33166.0</td>
<td>7.6609</td>
<td>1.1744</td>
</tr>
<tr>
<td>26</td>
<td>10046.0</td>
<td>7.2876</td>
<td>0.4484</td>
</tr>
<tr>
<td>27</td>
<td>59634.0</td>
<td>7.5522</td>
<td>0.5832</td>
</tr>
<tr>
<td>28</td>
<td>26305.0</td>
<td>8.7729</td>
<td>0.4918</td>
</tr>
<tr>
<td>29</td>
<td>42957.0</td>
<td>8.2419</td>
<td>0.7791</td>
</tr>
<tr>
<td>30</td>
<td>4405.0</td>
<td>8.0915</td>
<td>0.5767</td>
</tr>
<tr>
<td>31</td>
<td>39076.0</td>
<td>8.1357</td>
<td>0.7948</td>
</tr>
<tr>
<td>32</td>
<td>49482.0</td>
<td>8.3993</td>
<td>1.0635</td>
</tr>
<tr>
<td>33</td>
<td>82045.0</td>
<td>7.9210</td>
<td>1.2084</td>
</tr>
<tr>
<td>34</td>
<td>76595.0</td>
<td>8.9572</td>
<td>0.6003</td>
</tr>
<tr>
<td>35</td>
<td>47646.0</td>
<td>8.7230</td>
<td>0.6305</td>
</tr>
<tr>
<td>36</td>
<td>43078.0</td>
<td>7.9195</td>
<td>2.0678</td>
</tr>
<tr>
<td>37</td>
<td>80446.0</td>
<td>7.8265</td>
<td>1.0109</td>
</tr>
<tr>
<td>38</td>
<td>52957.0</td>
<td>6.8638</td>
<td>0.6924</td>
</tr>
</tbody>
</table>

\[
\text{MEAN LOG INCOME OF THE WHOLE SAMPLE} = 7.4803 \\
\text{TOTAL SAMPLE SIZE} = 3279445.0 \\
\text{BETWEEN GROUP VARIANCE} = 0.2374 \\
\text{WITHIN GROUP VARIANCE} = 0.6197 \\
\text{TOTAL VARIANCE} = 0.8570 \\
\text{BETWEEN GROUP CONTRIBUTION} = 27.70\% \\
\text{WITHIN GROUP CONTRIBUTION} = 72.30\% \\
\]
### Statistics for Income Distribution

Input file is 1978 Bogota household survey < 3062 records >

<< Matrix # 2 >>

Row variable = Sector
Col variable = Income

The row variable has 8 categories.
Maximum value of each category are as follows:

1. 2. 3. 4. 5.
6. 7. 8.

The column variable has 20 categories.
Maximum value of each category are as follows:

296. 592. 704. 816. 936.
1057. 1197. 1338. 1481. 1624.
1812. 2000. 2312. 2625. 3165.
3704. 5022. 6340. 95670. 185000.

Total number of records read = 3062
Total number of records processed = 2991

Measures of overall income distribution

Based on the 20 observations for the Lorenz curve

Relative mean deviation: 2228.5092
Variance of incomes: 17394593.1360
Coefficient of variation: 1.4667
Standard deviation of logs: 0.9258
Theil measure: 0.5057
Atkinson measure: 0.5750
Gini coefficient: 0.5167
DECOMPOSITION OF THE THEIL INDEX

\[ N(I) = \text{NO. OF OBSERVATIONS IN THE ITH GROUP} \]

\[ Y\text{BAR}(I) = \text{MEAN INCOME OF THE ITH GROUP} \]

\[ T(I) = \text{THEIL INDEX WITHIN ITH GROUP} \]

<table>
<thead>
<tr>
<th>NO.</th>
<th>N(I)</th>
<th>YBAR(I)</th>
<th>T(I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>79750.0</td>
<td>2152.51</td>
<td>0.3779</td>
</tr>
<tr>
<td>2</td>
<td>678025.0</td>
<td>1439.17</td>
<td>0.2858</td>
</tr>
<tr>
<td>3</td>
<td>827178.0</td>
<td>2164.00</td>
<td>0.3424</td>
</tr>
<tr>
<td>4</td>
<td>236431.0</td>
<td>2880.68</td>
<td>0.3746</td>
</tr>
<tr>
<td>5</td>
<td>203745.0</td>
<td>3030.00</td>
<td>0.4220</td>
</tr>
<tr>
<td>6</td>
<td>655122.0</td>
<td>2593.33</td>
<td>0.4089</td>
</tr>
<tr>
<td>7</td>
<td>300348.0</td>
<td>4136.52</td>
<td>0.4136</td>
</tr>
<tr>
<td>8</td>
<td>298846.0</td>
<td>7187.44</td>
<td>0.4128</td>
</tr>
</tbody>
</table>

TOTAL SAMPLE SIZE = 3279445.0

BETWEEN GROUP COMPONENT = 0.1231
WITHIN GROUP COMPONENT = 0.3826
THEIL INDEX OF TOTAL SAMPLE = 0.5057

BETWEEN GROUP CONTRIBUTION = 24.34 %
WITHIN GROUP CONTRIBUTION = 75.66 %
DECOMPOSITION OF LOG VARIANCE OF INCOME

N(I) = NUMBER OF OBSERVATIONS IN THE ITH GROUP
MEANLOG(I) = MEAN LOG INCOME OF THE ITH GROUP
V(I) = VARIANCE LOG INCOME OF THE ITH GROUP

<table>
<thead>
<tr>
<th>GROUP NO</th>
<th>N(I)</th>
<th>MEANLOG(I)</th>
<th>V(I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>79750.0</td>
<td>7.3252</td>
<td>0.6265</td>
</tr>
<tr>
<td>2</td>
<td>678025.0</td>
<td>7.0127</td>
<td>0.4826</td>
</tr>
<tr>
<td>3</td>
<td>827178.0</td>
<td>7.3691</td>
<td>0.5609</td>
</tr>
<tr>
<td>4</td>
<td>236431.0</td>
<td>7.5871</td>
<td>0.7409</td>
</tr>
<tr>
<td>5</td>
<td>203745.0</td>
<td>7.5950</td>
<td>0.7999</td>
</tr>
<tr>
<td>6</td>
<td>655122.0</td>
<td>7.4668</td>
<td>0.7304</td>
</tr>
<tr>
<td>7</td>
<td>300348.0</td>
<td>7.8421</td>
<td>1.0789</td>
</tr>
<tr>
<td>8</td>
<td>298846.0</td>
<td>8.3934</td>
<td>1.2527</td>
</tr>
</tbody>
</table>

MEAN LOG INCOME OF THE WHOLE SAMPLE = 7.4803
TOTAL SAMPLE SIZE = 3279445.0

BETWEEN GROUP VARIANCE = 0.1386
WITHIN GROUP VARIANCE = 0.7185
TOTAL VARIANCE = 0.8570

BETWEEN GROUP CONTRIBUTION = 16.17%
WITHIN GROUP CONTRIBUTION = 83.83%
6. **Sample Run - Program Listing**

   Following is the listing of the program used for the run shown in the sample output. Note that the corresponding sector numbers for each comuna are entered via the DATA statement.
PROGRAM "EQUALISE"

1000 C
1100 C **********************************************************************
1200 C **********MAIN PROGRAM EXECUTES FOLLOWING:**************************
1300 C
1400 C 1) INITIALIZE BY READING DATA FILE OR CARD DECK:
1500 C
1600 C --- INPUT FILE NAME
1700 C --- OUTPUT DESIRED < ROBINSON AND/OR ANAND >
1800 C --- OUTPUT DESIRED < PERCENT INCOME OF EACH Y CLASS >
1900 C --- NUMBER OF MATRICES
2000 C --- POSITIONS OF THE ROW AND THE COLUMN VARIABLES
2100 C --- IN THE 'DATA' ARRAY.
2200 C --- NUMBER OF CATEGORIES FOR ROW AND COLUMN VARIABLES
2300 C --- NAMES OF ROW AND COLUMN VARIABLES
2400 C --- MAX VALUE OF EACH CATEGORY FOR ROW & COL VARIABLES
2500 C
2600 C 2) READ INPUT DATA FILE ONE RECORD AT A TIME
2700 C
2800 C 3) PERFORM DATA TRANSFORMATIONS IF NECESSARY
2900 C
3000 C 4) EXECUTE CROST TABULATIONS BY CALLING SUBROUTINE CROST
3100 C
3200 C 5) AFTER ALL THE RECORDS ARE READ, CALCULATE MEAN INCOME
3300 C --- OF EACH INCOME CLASS BY CALLING ENTRY STATEMENT MEAN
3400 C
3500 C 6) CALL MEADIS AND/OR ANANDT--ANANDL TO PRODUCE SELECTED
3600 C --- STATISTICS FOR INCOME DISTRIBUTION AND DECOMPOSITION.
3700 C
3800 C SUBROUTINE CALLED ARE:
3900 C
4000 C 1) CROST -- PERFORMS CROSS-TABULATION FOR EACH DATA RECORD
4100 C
4200 C 2) MEADIS -- CALCULATES & PRINT OUT SELECTED STATISTICS OF
4300 C 3) ANANDT---CALCULATE DECOMPOSITION STATISTICS FOR THEIL INDEX
4400 C
4500 C THIS PROGRAM WAS WRITTEN BY SUNGYONG KANG OF WORLD BANK
4600 C IN JUNE 1980, EXCEPT WHERE NOTED OTHERWISE
4700 C
4800 C **********************************************************************
4900 C
5000 C
5100 C FILE 10(TITLE='A/B')
5200 C FILE 5(TITLE='C/D') % IF CARD DECK, DELETE THIS LINE
5300 C
5400 C ARRAYS NEEDED FOR INITIALIZING DATA
5500 C
5600 C DIMENSION FILE(10).
5700 C * ICOR1(2),ICOR2(2),NI(2),NJ(2),NAME1(2),NAME2(2).
5800 C * C1(2.38),C2(2.20)
5900 C
6000 C ARRAYS NEEDED FOR SUBROUTINES CROST AND MEADIS
6100 C
6200 C DIMENSION MAT(2.38,20),MCT(2.20),SUM(2.20).
6300 C * MATS(38,20),MCTS(20),SUMS(20),AVGS(20)
6400 C
6500 C
C ARRAYS NEEDED FOR SUBROUTINES ANANDL AND ANANDT

DIMENSION ANL(20), ANT(38), S(38), T(38), ANY(38),
YBAR(38), P(38), PP(38, 20), SS(38, 20)

C ARRAYS NEEDED FOR READING INPUT FILE

C DIMENSION DATA(8), NSEC(92)

C CORRESPONDENCE BETWEEN BOGOTA COMUNA AND SECTOR NUMBER

DATA NSEC / 10*0,
1 2, 2, 2, 2, 6*0,
2 3, 3, 3, 3, 2, 5*0,
3 1, 2, 5*0,
4 4, 4, 4, 3, 3, 5*0,
5 6, 6, 7, 6, 6, 6, 4*0,
6 5, 5, 5, 4, 5*0,
7 7, 7, 7, 7, 6*0,
8 8, 8, 8, 8, 8, 5*0,
9 7, 7/

C ********** ********** ********** ********** ********** **********

C READ INITIALIZING DATA

1) FILE = INPUT FILE NAME -- 60 CHARACTERS LONG
2) NOUT1 = OUTPUT DESIRED:
   1= S. ROBINSON'S SELECTED INCOME DISTRIBUTION STATS.
   2= S. ANAND'S DECOMPOSITION STATISTICS FOR BOTH
   THEIL AND LOG-VARIANCE INCOME DISTRIBUTION INDICES.
   3= BOTH 1 AND 2 ABOVE.
3) ICOR1= CORRESPONDING DATA NUMBER TO THE 1ST VARIABLE
4) ICOR2 = CORRESPONDING DATA NUMBER TO THE 2ND VARIABLE
5) NI = NO OF CATEGORIES IN THE 1ST VARIABLES OF N MATRICES
6) NJ = NO OF CATEGORIES IN THE 2ND VARIABLES OF N MATRICES
7) NAME1= NAMES OF THE N FIRST VARIABLES
8) NAME2= NAMES OF THE N SECOND VARIABLES
9) N = NUMBER OF MATRICES TO BE CREATED
10) NOUT2 = OUTPUT DESIRED:
    0= PERCENT INCOME OF EACH INCOME CLASS NOT PRINTED
    1= PERCENT INCOME OF EACH INCOME CLASS PRINTED OUT

READ(5,1001) FILE, NOUT1, NOUT2, N

C ********** ********** ********** ********** ********** **********

M1=1
M2=1
DO 11 I=1, N
READ(5,1002) ICOR1(I), NI(I), NAME1(I), ICOR2(I), NJ(I), NAME2(I)
1002 FORMAT (2(213,A6))
C
K1=NI(I)  % NUMBER OF CATEGORIES IN THE 1ST VARIABLE
K2=NJ(I)  % NUMBER OF CATEGORIES IN THE 2ND VARIABLE
M1=AMAXO(M1,K1)  % MAX NUMBER OF CATEGORIES FOR 1ST VARIABLE
M2=AMAXO(M2,K2)  % MAX NUMBER OF CATEGORIES FOR 2ND VARIABLE
C
READ(5,1003) (C1(I,J),J=1,K1)
READ(5,1003) (C2(I,J),J=1,K2)
1003 FORMAT (5F8.0)
11 CONTINUE
C
C ** **
******** ** ********
C READ DATA FILE AND PERFORM ANY TRANSFORMATION OF DATA
C IF NECESSARY
C
21 READ (10,1004,END=400) ICOM,(DATA(J),J=1,6)
1004 FORMAT (12X.12.5X,F2.0,261X.4F6.0,6X.F6.0)
C
ICOM = COMUNA NUMBER
DATA(1)= FAMILY SIZE INCLUDING MAID
DATA(2)= FAMILY LABOR INCOME IN PESOS
DATA(3)= FAMILY NON-LABOR INCOME IN PESOS
DATA(4)= FAMILY IMPUTED HOUSING INCOME--IMPUTED RENT
DATA(5)= DOMESTIC SERVANT'S INCOME
DATA(6)= EXPANSION FACTOR IN 1/100
C
C *************** ********** ********** ********** **********
C EXECUTE CROSTAB TABULATION, ACCUMULATE COLUMN TOTAL
KNT1=KNT1+1  % NUMBER OF RECORDS READ--MUST BE PRESENT
C
DATA(2)=DATA(2)+DATA(3)  % FAMILY LABOR+NON LABOR INCOME
IF (DATA(2).EQ.0) GO TO 21
C
IF (ICOM.EQ.0.OR.ICOM.GT.92) GO TO 21
C
FS=DATA(1)  % FAMILY SIZE
IF (DATA(5).GT.0) FS=FS-1
DATA(2)=DATA(2)/FS  % PER CAPITA INCOME
C
DATA(7)=ICOM  % COMUNA NUMBER
C
NS=NSEC(ICOM)  % CORRESPONDING SECTOR NUMBER
IF (NS.EQ.0) GO TO 21
C
C *************** ********** ********** ********** **********
**FORTRAN程序**

```
18100 C OF BOTH TABULATION AND INCOME VALUE FOR N MATRICES.
18200 C ********** ********** ********** ********** ********** **********
18300 C
18400 C DD 31 K=1,N
18500 K1=NI(K)  % # OF CATEGORY FOR 1ST VARIABLE
18600 K2=NJ(K)  % # OF CATEGORY FOR 2ND VARIABLE
18700 C NV1=ICOR1(K)
18800 NV2=ICOR2(K)
19000 C
19100 C D1=DATA(NV1)  % DATA FOR 1ST VARIABLE
19200 D2=DATA(NV2)  % DATA FOR 2ND VARIABLE
19300 C CALL SUBROUTINE CROST TO PERFORM CROSS-TABULATION
19400 C OF THE PEOPLE IN THE CELL(I,J) OF MAT(N,I,J), DETERMINED
19500 C BY CATEGORY BOUNDARIES OF CI(N,K1) AND C2(N,K2).
19600 C CALL CROST(D1,D2,EXP.N,M1,M2,K,K1,K2,C1,C2,MAT,MCT,SUM,AVGS)
19800 C
19900 C 31 CONTINUE
20000 GO TO 21
20100 C
20200 C ********** ********** ********** ********** ********** **********
20300 C AFTER ALL THE RECORDS IN THE DATA FILE HAVE BEEN ACCOUNTED FOR,
20400 C CALCULATE MEAN INCOME IN EACH INCOME CATEGORY AND PRINT
20500 C OUT THE HEADINGS FOR EACH MATRIX.
20600 C ********** ********** ********** ********** ********** **********
20700 C 400 DO 44 K=1,N
20800 K1=NI(K)  % NO OF CATEGORIES IN THE 1ST VARIABLE
20900 K2=NJ(K)  % NO OF INCOME CLASSES
21100 C TRANSFORM MCT AND SUM INTO ONE DIMENSIONAL ARRAYS TO BE
21200 C USED IN THE SUBROUTINE MEADIS
21300 C ALSO TRANSFORM MAT INTO TWO DIMENSIONAL ARRAY
21700 C DO 43 J=1,K2
21800 MCTS(J) = MCT(K,J)
21900 SUMS(J) = SUM(K,J)
22100 C DO 43 I=1,K1
22200 MATS(I,J)=MAT(K,I,J)
22400 C 43 CONTINUE
22500 C WRITE HEADING OF THE OUTPUT AND MATRIX RELATED INFORMATION
22600 C WRITE (6,2001) FILE
22700 2001 FORMAT ('1',30X,'*** STATISTICS FOR INCOME DISTRIBUTION ***罗马尼亚',//
22800 21X,'INPUT FILE IS ',10A6//)
22900 C WRITE (6,2002) K,NAME1(K),NAME2(K)
23000 2002 FORMAT ('0',3X,'<< MATRIX #',12.' >罗马尼亚',//
23100 9X,'ROW VARIABLE= ',A6/ 9X,'COL VARIABLE= ',A6)
```

**注释**

这个程序似乎用于对人口数据进行交叉表计算，并计算每个收入类别的平均收入。它首先读取数据文件，然后计算每个收入类别的平均收入，并输出矩阵相关的信息。程序使用了FORTRAN语言编写，包含了数据读取、交叉表计算和输出部分。
23900 2003 FORMAT ('O'.,5X,'THE ROW VARIABLE HAS',I3,' CATEGORIES.'//
24000    *     6X,'MAXIMUM VALUE OF EACH CATEGORY ARE AS FOLLOWS:'//
24100    *     (9X,5F8.0)) 00024000
24200 C 00024200
24300 WRITE (6,2004) K2,(C2(K,U).J=1,K2) 00024300
24400 2004 FORMAT ('O'.,5X,'THE COLUMN VARIABLE HAS',I3,' CATEGORIES.'//
24500    *     6X,'MAXIMUM VALUE OF EACH CATEGORY ARE AS FOLLOWS:'//
24600    *     (9X,5F8.0)) 00024500
24700 C 00024700
24800 IF (K.GT.1) NOUT2=0 00024800
24900 CALL MEAN(NOUT2,N,M2,K,K2,MCT,SUM,AVGS,KNTi) 00024900
25000 C 00025000
25100 C 00025100
25200 C ********** ********** *t********
30200 ********** ********** 00025200
25300 C CALL SUBROUTINES MEADIS AND/OR ANAND FOR SELECTED STATISTICS
25400 C 2******* ********** ********** ********** **********
30400 00025400
25500 C 00025500
25600 GO TO (401,403,402), NOUTI 00025600
25700 C 00025700
25800 401 CALL MEADIS(MCTS,SUMS,AVGS,K2) 00025800
25900 GO TO 44 00025900
26000 C 00026000
26100 402 CALL MEADIS(MCTS,SUMS.AVGS,K2) 00026100
26200 403 CALL ANANDT(M1,M2,K1,K2,MATS.AVGS.ANT,S.T,ANY,YBAR,P,PP,SS) 00026200
26300 CALL ANANDL(Mi,M2,K1,K2.MAINS.AVGS.ANY,P) 00026300
26400 C 00026400
26500 44 CONTINUE 00026500
26600 C 00026600
26700 999 STOP 00026700
26800 END 00026800
26900 SUBROUTINE CROST(VI.V2.EXP,N,M1,M2,K,K1,K2,CI,C2.MAT,MCT,SUM,AVGS) 00026900
27000 C 00027000
27100 C ******************************************************t*********
30100 ****************************************************** 00027100
27200 C 00027200
27300 C SUBROUTINE CROST PERFORMS CROSS-TABULATION OF VI BY V2 (INCOME).
27400 C ALSO, SUM AND THE MEAN VALUE OF INCOME IN EACH INCOME CLASS
27500 C IS CALCULATED ALONG WITH THE NUMBER OF PEOPLE IN EACH INCOME CLASS
27600 C AFTER ALL THE RECORDS ARE READ, THIS SUBROUTINE IS ENTERED
27700 C AT THE ENTRY POINT -- ENTRY MEAN -- TO CALCULATE
27800 C THE MEAN VALUES OF INCOME IN EACH INCOME CLASS.
27900 C VARIABLES ARE AS FOLLOWS:
28000 C V1 = DATA FOR THE FIRST VARIABLE
28100 C V2 = DATA FOR THE SECOND VARIABLE (INCOME DATA)
28200 C EXP = EXPANSION FACTOR FOR EACH RECORD
28300 C N = TOTAL NUMBER OF MATRIX NUMBER
28400 C M1= MAX NUMBER OF CATEGORIES IN THE FIRST VARIABLE
28500 C M2= MAX NUMBER OF CATEGORIES IN THE SECOND VARIABLE
28600 C K = MATRIX NUMBER FOR THIS CROSTAB
28700 C K1= NUMBER OF CATEGORIES IN THE FIRST VARIABLE
28800 C K2= NUMBER OF CATEGORIES IN THE SECOND VARIABLE
28900 C C1= MAX VALUE OF EACH CATEGORY IN THE 1ST VARIABLE
29000 C C2= MAX VALUE OF EACH CATEGORY IN THE 2ND VARIABLE
29100 C
DIMENSION C1(N,M1),C2(N,M2),MAT(N,M1,M2),MCT(N,M2).

DIMENSION KNT(20)

$ SET OWN

DIMENSION KNT(20)

$ RESET OWN

C DETERMINE ROW I FOR THE FIRST VARIABLE

C IF THERE IS ONLY ONE ROW (K1=1), THEN SKIP TO 13

C

11 DO 12 II=1,K1

12 CONTINUE

RETURN

DATA IS OUT OF BOUND

C DETERMINE COLUMN J FOR THE SECOND VARIABLE (INCOME)

C IF THERE IS ONLY ONE COLUMN (K2=1), THEN SKIP TO 16

C

13 IF (K2.GT.1) GO TO 14

14 DO 15 JJ=1,K2

15 CONTINUE

RETURN

DATA IS OUT OF BOUND

C ONCE THE ROW I AND COLUMN J HAVE BEEN DETERMINED, INCREMENT THE COUNT IN THE CELL (I,J) BY EXP FACTOR DETERMINED IN THE MAIN PROGRAM.

C ALSO, ACCUMULATE THE VALUE OF THE SECOND VARIABLE -- INCOME -- IN EACH CATEGORY (J) -- INCOME CLASS --, AND THE NUMBER OF PEOPLE IN EACH CATEGORY (J).

C

MAT(K,I,J)=MAT(K,I,J)+EXP

MCT(K,J)=MCT(K,J)+EXP

C

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)

RESULTING N X K1 X K2 MATRIX OF CROSSTAB TABULATION

COLUMN TOTAL OF MAT(K,I,J) MATRIX

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)

SUM = ACCUMULATED SUM OF 2ND VARIABLE (INCOME) IN EACH INCOME CLASS

AVGS = MEAN VALUE OF EACH INCOME CLASS (K2)
35500 C 00035500
35600 SUM(K,J)=SUM(K,J)+V2*EXP % TOTAL INCOME IN EACH INCOME CLASS 00035600
35700 KNT(K)=KNT(K)+1 % NUMBER OF RECORDS PROCESSED 00035700
35800 C 00035800
35900 RETURN 00035900
36000 C 00036000
36100 ENTRY MEAN(NO,N2,K2,MCT,SUM,AVGS,KNT1) 00036100
36200 C 00036200
36300 C ************************************ 00036300
36400 C 00036400
36500 C AFTER ALL THE RECORDS IN THE DATA FILE HAVE BEEN ACCOUNTED FOR, 00036500
36600 C CALCULATE THE MEAN VALUES OF INCOME IN EACH INCOME CLASS 00036600
36700 C 00036700
36800 C ************************************ 00036800
36900 C 00036900
37000 C 00037000
37100 TY=0 00037100
37200 TT=0 00037200
37300 C 00037300
37400 DD 31 J=1,K2 00037400
37500 IF (MCT(K,J).EQ.0) GO TO 30 00037500
37600 C 00037600
37700 AVGS(J)=SUM(K,J)/MCT(K,J) % MEAN INCOME IN EACH INCOME CLASS 00037700
37800 C 00037800
37900 TY=TY+SUM(K,J) % TOTAL INCOME OF ALL PEOPLE 00037900
38000 TT=TT+MCT(K,J) % TOTAL NUMBER OF PEOPLE 00038000
38100 GO TO 31 00038100
38200 C 00038200
38300 30 AVGS(J)=0 00038300
38400 SUM(K,J)=0 00038400
38500 C 00038500
38600 31 CONTINUE 00038600
38700 C 00038700
38800 IF (NO.EQ.0) GO TO 41 00038800
38900 C 00038900
39000 WRITE (6,311) 00039000
39100 311 FORMAT ('0'/IX,'INC CLASS',3X,'NO OF PEOPLE',3X,'%','-','CUM %',3X,'TOTAL INCOME',3X,'%','-','CUM %',2X,'MEAN INCOME'/IX.78('-')) 00039100
39200 CPY=0 00039200
39300 CPT=0 00039300
39400 C 00039400
39500 WRITE (6,321) J,MCT(K,J),PT,CPT,SUM(K,J),CPY,CPT,TMEAN 00039500
39600 321 FORMAT (7X.12,2(F15.0,F6.1,IX,F6.1),F12.2) 00039600
39700 C 00039700
39800 32 CONTINUE 00039800
39900 TMEAN=TY/TT % MEAN INCOME FOR ALL CLASSES 00039900
40000 WRITE (6,322) TT,CPT,CPT,CPY,CPY,TMEAN 00040000
40100 322 FORMAT ('0','8(I5.0,F6.1,IX,F6.1)',F12.2) 00040100
40200 C 00040200
40300 CPY=CPY+PY % CUMULATIVE PERCENTAGE OF INCOME DISTRIBUTION 00040300
40400 CPT=CPT+PT % CUMULATIVE PERCENTAGE OF POP DISTRIBUTION 00040400
40500 WRITE (6,321) J,MCT(K,J),PT,CPT,SUM(K,J),PY,CPY,AVGS(J) 00040500
40600 321 FORMAT (7X,I2.2(F15.0,F6.1,IX,F6.1),F12.2) 00040600
40700 C 00040700
40800 32 CONTINUE 00040800
40900 C 00040900
41000 TMEAN=TY/TT % MEAN INCOME FOR ALL CLASSES 00041000
41100 WRITE (6,322) TT,CPT,CPT,CPY,CPY,TMEAN 00041100
41200 322 FORMAT ('0','8(I5.0,F6.1,IX,F6.1)',F12.2) 00041200
C

41300  F12.2)
41400  C
41500  41  WRITE (6,411) KNT1,KNT(K)
41600  411  FORMAT ('0',1X,'TOTAL NUMBER OF RECORDS READ =','I8/ 
41700  *  2X,'TOTAL NUMBER OF RECORDS PROCESSED=','I8)
41800  C
41900  RETURN
42000  END
42100  SUBROUTINE MEADIS(EMALL,YALL,YPALL,NNUM)
42200  C
42300  C THIS PROGRAM IS WRITTEN BY SHERMAN ROBINSON IN 
42500  C MEADIS COMPUTES SEVERAL INCOME DISTRIBUTION MEASURES.
42600  C FOR ATKINSON'S MEASURE, A FRISCH PARAMETER OF .5 IS ASSUMED
42700  C EMALL = NUMBER OF PEOPLE IN EACH INCOME CLASS.
42800  C YALL = TOTAL INCOME OF EACH INCOME CLASS
42900  C YPALL = PER CAPITA INCOME OF EACH INCOME CLASS
43000  C NNUM = NUMBER OF INCOME CLASSES
43100  C
43200  DIMENSION EMALL(NNUM),YPALL(NNUM),YALL(NNUM)
43300  C ARITHMETIC MEAN INCOME AND TOTAL HOUSEHOLDS
43400  C DUM = 0.0
43500  POPUL = 0.0
43600  DO 100 N=1,NNUM
43700  POPUL = POPUL + EMALL(N)
43800  100 DUM = DUM + YALL(N)
43900  AMU = DUM/POPUL
44000  C RELATIVE MEAN DEVIATION
44100  C DUM = 0.0
44200  DO 101 N=i,NNUM
44300  DUMY = AMU - YPALL(N)
44400  101 DUM = DUM + EMALL(N)*ABS(DUMY)
44500  RMDEV = DUM/POPUL
44600  C VARIANCE
44700  C DUM = 0.0
44800  DO 102 N=1,NNUM
44900  DUMY = (AMU-YPALL(N))*(AMU-YPALL(N))
45000  102 DUM = DUM + EMALL(N)*DUMY
45100  VARIAN = DUM/POPUL
45200  C COEFFICIENT OF VARIATION
45300  C COEVAR = SORT(VARIAN)/AMU
45400  C STANDARD DEVIATION OF LOGARITHMS ABOUT ARITHMETIC MEAN
45500  C
47100 DUM = 0.0
47200 DO 103 N=1,NNUM
47300 DUMY = (ALOG(AMU) - ALOG(YPALL(N)))*(ALOG(AMU) - ALOG(YPALL(N)))
47400 DUM = DUM + EMALL(N) + DUMY
47500 VLOGA = DUM/POPUL
47600 SDLOGA = SQRT(VLOGA)
47700 C GINI COEFFICIENT
47800 C
47900 C
48000 DUM = 0.0
48100 DO 104 N=1,NNUM
48200 DO 104 K=1,NNUM
48300 DUMY = YPALL(N) - YPALL(K)
48400 DUM = DUM + ABS(DUMY) * EMALL(N) * EMALL(K)
48500 GINI = DUM/(2.0*POPUL*POPUL*AMU)
48600 C THEIL'S ENTROPY MEASURE
48700 C
48800 C
48900 C
49000 C
49100 DUM = 0.0
49200 DO 105 N=1,NNUM
49300 DUMY = DUMY + EMALL(N) * YPALL(N)
49400 DUMX = ((EMALL(N) * YPALL(N)) / DUMY)
49500 DUML = (POPUL*DUMX)/EMALL(N)
49600 DUM = DUM + DUMX*ALOG(DUML)
49700 THEIL = DUM
49800 C STANDARD DEVIATION OF LOGARITHMS ABOUT GEOMETRIC MEAN
49900 C
50000 C
50100 GMU = 0.0
50200 DO 120 N=1,NNUM
50300 GMU = GMU + (EMALL(N)*ALOG(YPALL(N)))/POPUL
50400 DUM = 0.0
50500 DO 121 N=1,NNUM
50600 DUMY = (GMU-ALOG(YPALL(N)))*(GMU-ALOG(YPALL(N)))
50700 DUM = DUM + EMALL(N) + DUMY
50800 VLOGG = DUM/POPUL
50900 SDLOGG = SQRT(VLOGG)
51000 C ATKINSON'S EQUALLY DISTRIBUTED EQUIVALENT INCOME MEASURE.
51100 C CONSTANT ELASTICITY UTILITY = A + (B/(1-C))*(Y**C-1). C=2
51200 C
51300 C
51400 DUM = 0.0
51500 DO 150 N=1,NNUM
51600 DUM = DUM + EMALL(N)/YPALL(N)
51700 YEQU = POPUL/DUM
51800 ATKIN = 1.0 - YEQU/AMU
51900 C WRITE OUT RESULTS.
52000 C
52100 C
52200 WRITE(6.902)
52300 WRITE(6.903) NNUM
52400 WRITE(6.1010) RMDEV,VARIAN,COEVAR,SDLLOGG,THEIL
52500 1010 FORMAT(1H,5X,'RELATIVE MEAN DEVIATION',T40,F14.4,/, 00052500 138X,'VARIANCE OF INCOMES'.T40,F14.4./, 00052600 138X,'COEFFICIENT OF VARIATION',T40,F14.4./, 00052700 138X,'STANDARD DEVIATION OF LOGS',T40,F14.4./, 00052800
* 6X,'THEIL MEASURE',F4.14,4 ) 00052900
53000 WRITE(6,1020) ATKIN,GINI 00053000
53100 1020 FORMAT(6X,'ATKINSON MEASURE',F4.14,4,/) 00053100
53200 * 6X,'GINI COEFFICIENT',F4.14,4 ) 00053200
53300 C- 00053300
53400 902 FORMAT(1HO,4.0H MEASURES OF OVERALL INCOME DISTRIBUTION) 00053400
53500 903 FORMAT('O','BASED ON THE','15, 00053500
53600 * ' OBSERVATIONS FOR THE LORENZ CURVE'//) 00053600
53700 RETURN 00053700
53800 END 00053800
53900 SUBROUTINE ANANDT(M1,M2,K1,K2,ANN,Y,AN,S,T,ANY,YBAR,P,SS) 00053900
54000 C 00054000
54100 C * 00054100
54200 C 00054200
54300 C THIS SUBROUTINE CALCULATES DECOMPOSITION OF THEIL INDEX 00054300
54400 C 00054400
54500 C --- WRITTEN BY SUDHIR ANAND 00054500
54600 C AND MODIFIED BY SUNGYDNG KANG IN JULY 1980 00054600
54700 C 00054700
54800 C INPUT NEEDED ARE; 00054800
54900 C M1 = MAX NUMBER OF ROW GROUPS 00054900
55000 C M2 = MAX NUMBER OF COL GROUPS 00055000
55100 C 00055100
55200 C K1 = NUMBER OF ROW GROUPS 00055200
55300 C K2 = NUMBER OF COL.--- INCOME CLASS 00055300
55400 C 00055400
55500 C ANN= CROSTAB TABULATION OF INPUT DATA IN K1 BY K2 MATRIX 00055500
55600 C Y = MEAN INCOME IN EACH INCOME CLASS 00055600
55700 C 00055700
55800 C 00055800
55900 C 00055900
56000 C DIMENSION ANN(M1,M2),Y(K2), 00056000
56100 C PP(K1,K2),SS(K1,K2), 00056100
56200 C AN(K1),S(K1),T(K1),ANY(K1),YBAR(K1),P(K1) 00056200
56300 C 00056300
56400 C CAPN=0. 00056400
56500 DO 11 I=1,K1 00056500
56600 ANY(I)=0. 00056600
56700 AN(I)=0. 00056700
56800 C 00056800
56900 DO 11 J=1,K2 00056900
57000 AN(I)=AN(I)+ANN(I,J) % TOTAL POP IN ROW GROUP I 00057000
57100 CAPN=CAPN+ANN(I,J) % TOTAL POPULATION 00057100
57200 ANY(I)=ANY(I)+(ANN(I,J)*Y(J)) % TOTAL INCOME OF ROW GROUP I 00057200
57300 11 CONTINUE 00057300
57400 C 00057400
57500 CAPY=0. 00057500
57600 DO 122 I=1,K1 00057600
57700 ANY(I)=0. 00057700
57800 AN(I)=0. 00057800
57900 C 00057900
58000 AN(I)=AN(I)+ANN(I,J) % TOTAL POP IN ROW GROUP I 00058000
58100 CAPN=CAPN+ANN(I,J) % TOTAL POPULATION 00058100
58200 ANY(I)=ANY(I)+(ANN(I,J)*Y(J)) % TOTAL INCOME OF ROW GROUP I 00058200
58300 122 CONTINUE 00058300
58400 C 00058400
58500 P(I)=AN(I)/CAPN % SHARE OF POP BY ROW GROUP 00058500
58600 120 P(I)=AN(I)/CAPN % SHARE OF POP BY ROW GROUP 00058600
58700 C YBAR(I)=ANY(I)/AN(I) % MEAN INCOME OF ROW GROUP I
58800 C
58900 C
59000 PP(I,J)=ANN(I,J)/AN(I) % ROW DIST OF POPULATION
59100 C
59200 C
59300 IF (PP(I,J).EQ.0.) GO TO 121
59400 C
59500 C C CALCULATE THEIL INDEX FOR EACH ROW GROUP I
59600 C
59700 C
59800 T(I)=SS(I,J)+ALOG(SS(I,J)/PP(I,J))+T(I)
59900 C
60000 121 CONTINUE

60100 C
60200 IF (SS(I,.J).EQ.0.) GO TO 121

60300 C
60400 C
60500 C C CALCULATE THE BETWEEN GROUP COMPONENT OF THE
60600 C
60700 C THEIL INDEX
60800 C
60900 C
61000 TB=O.
61100 C
61200 DO 14 I=1,K1
61300 C
61400 C
61500 CAPT=O.
61600 C
61700 CAPT=(S(I)*T(I))+CAPT
61800 C
61900 C
62000 C CAPT=TB+CAPT % THEIL INDEX OF TOTAL SAMPLE
62100 C
62200 TBET=(TB/CAPT)*100 % BETWEEN GROUP THEIL INDEX CONTRIBUTION
62300 C
62400 C
62500 WRITE (6,100)
62600 100 FORMAT ('1', ' DECOMPOSITION OF THE THEIL INDEX ')
62700 C
62800 WRITE (6,101)
62900 101 FORMAT ('1', '3X.'N(I) NO. OF OBSERVATIONS IN THE ITH GROUP'/
63000 * '4X.'YBAR(I)= MEAN INCOME OF THE ITH GROUP'/
63100 * '4X.'T(I) = THEIL INDEX WITHIN ITH GROUP'/
63200 * '4X.'NO.',7X,.'N(I)',9X,.'YBAR(I)',12X,.'T(I)'/
63300 * '1X,6S('-')'/
63400 C
63500 C WRITE (6,102) (1,AN(I),YBAR(I),T(I),I=1,K1)
63600 C
63700 102 FORMAT (4X.12,3X.FIO.l,5X.FIo.2.7X.FIO.4)
63800 C
63900 C WRITE (6,103) CAPN,TB,TW,CAPT,TBET,TWIT
64000 103 FORMAT ('1', 'TOTAL SAMPLE SIZE =',F10.1//
64100 * '1X,'BETWEEN GROUP COMPONENT =',F10.4//
64200 * '1X,'WITHIN GROUP COMPONENT =',F10.4//
64300 * '1X,'THEIL INDEX OF TOTAL SAMPLE =',F10.4//
64400 * '1X,'BETWEEN GROUP CONTRIBUTION =',F10.2,' %'/
64500 * '1X,'WITHIN GROUP CONTRIBUTION =',F10.4,' %'//
SUBROUTINE ANANDL(M1,M2,K1,K2,ANN,Y,AN,S,ANY,P)  

C
**********..*************************************..*****.

DIMENSION ANN(M1,M2),Y(K2),AN(K2),S(K1),ANY(K1),P(K1)

DO 11 U=1,K2  
AN(U)=ALOG(Y(J))  
% MEAN INCOME IN LOG FOR INCOME CLASS 
II CONTINUE

CAPN=O.  
DO 122 I=1,K1  
ANY(I)=O.  
S(I)=O.  

DO 121 J=1,K2  
ANY(I)=(ANN(I,J)*AN(J))+ANY(I)  
% TOTAL INCOME OF I IN LOG 
S(I)=ANN(I,J)+S(I)  
% NUMBER OF PEOPLE IN ROW GROUP 
121 CONTINUE

IF (S(I).EQ.O) GO TO 132  

CAPN=CAPN+S(I)  
ANY(I)=ANY(I)/S(I)  
% MEAN INCOME IN LOG FOR ROW GROUP I
122 CONTINUE

GBAR=O.  
DO 132 I=1,K1  
GBAR=(S(I)*ANY(I))+GBAR  
% LOG OF TOTAL INCOME 
P(I)=O.  
IF (S(I).EQ.O) GO TO 132 

C  

IF (S(I).EQ.O) GO TO 122  

CAPN=CAPN+S(I)  
ANY(I)=ANY(I)/S(I)  
% MEAN INCOME IN LOG FOR ROW GROUP I
122 CONTINUE

GBAR=O.  
DO 132 I=1,K1  
GBAR=(S(I)*ANY(I))+GBAR  
% LOG OF TOTAL INCOME 
P(I)=O.  
IF (S(I).EQ.O) GO TO 132 

C  

CALCULATE THE VARIANCE LOG INCOME OF THE ITH GROUP
DO 131 J=1,K2
P(I)=(ANN(I,J)+((AN(J)-ANY(1)**2))+P(I)
CONTINUE
C
P(I)=P(I)/S(I)
CONTINUE
C
GBAR=GBAR/CAPN  \% LOG OF MEAN INCOME
VB=O.
CAPV=O.
DO 14 I=1,K1
VB= (S(I)+((ANY(I)-GBAR)**2)+VB  \% BETWEEN GROUP VARIANCE
CAPV=(S(I)+P(I))+CAPV  \% TOTAL VARIANCE
CONTINUE
C
VB=VB/CAPN
CAPV=VB+(CAPV/CAPN)  \% BETWEEN GROUP CONTRIBUTION IN PERCENT
VBET=(VB/CAPV)*100  \% WITHIN GROUP CONTRIBUTION IN PERCENT
VWIT=100-VBET  \% WITHIN GROUP CONTRIBUTION
WRITE (6,201)
201 FORMAT ('1', 'DECOMPOSITION OF LOG VARIANCE OF INCOME')
WRITE (6,202)
202 FORMAT ('0', 'NUMBER OF OBSERVATIONS IN THE ITH GROUP'/
* 3X,'N(I)'  \% MEAN LOG INCOME OF THE ITH GROUP'/
* 4X,'MEANLOG(1)'  \% VARIANCE LOG INCOME OF THE ITH GROUP'/
* 4X,'V(I)'  \% TOTAL VARIANCE'
WRITE (6,203) (I,S(I),ANY(I),P(I),I=1,K1)
WRITE (6,204) GBAR,CAPN,VB,CAPV,CAPV-VBET,CAPV-VWIT
203 FORMAT (4X.12.2X,Fi0.1.5X,F10.4,7X.Fi0.4)
REFERENCES


BIBLIOGRAPHY


The Determinants of Labor Earnings in Developing Metropolises: Estimates from Bogota and Cali, Colombia
Rakesh Mohan
Stock No. WP 0498. $5.

Differences in Income, Nutrition, and Poverty within Brazil
Vinod Thomas
Staff Working Paper No. 505. 1982. 91 pages (including references, map).
Stock No. WP 0505. $3.

The Distribution of Income in Brazil
Guy P. Pfeffermann and Richard C. Webb
Stock No. WP 0356. $5.

Economic Growth and Employment in China
Thomas G. Rawski
Examines how China's agricultural sector has made significant strides toward full employment for a labor force of nearly half a billion persons.
Oxford University Press, 1979. 208 pages (including maps, bibliography, index).
Stock Nos. OX 520151, $18.95 hardcover; OX 520152, $7.95 paperback.

Employment Patterns and Income Growth
Joseph J. Stern and Jeffrey D. Lewis
Staff Working Paper No. 419. 1980. 70 pages (including bibliography, 2 appendixes).
Stock No. WP 0419, $3.

Employment Policy in Developing Countries: A Survey of Issues and Evidence
Lyn Squire
Low rates of growth in industrial employment, high rates of unemployment among new entrants to the urban labor market, and low levels of labor productivity and remuneration are the three issues addressed in this study. The author identifies the important determinants of labor demand and supply and the extent to which the growth of labor demand has been constrained—and labor supply advanced—by inappropriate policies. On the demand side, industrial trade policy, agricultural growth, and the operation of capital markets are discussed; on the supply side, attention is focused on population and education policy.
Oxford University Press, 1981. 242 pages (including bibliography, index).
Stock Nos. OX 520266, $16.95 hardcover; OX 520267, $7.95 paperback.

Growth and Equity in Semi-Industrialized Countries
Joel Bergsman
Stock No. WP 0351. $5.

Growth with Equity: The Taiwan Case
John C. H. Fei, Gustav Ranis, and Shirley W.Y. Kuo
Introduces a method for tracing the inequality of family income to the inequality of various kinds of factor income and the shares of factor income in family income.
Oxford University Press, 1980. 444 pages (including index).
Stock Nos. OX 520115, $27.50 hardcover OX 520116, $12.95 paperback.

Household Income or Household Income per Capita in Welfare Comparisons
Gautam Datta and Jacob Meerman
Stock No. WP 0378. $3.

How Segmented is the Bogota Labor Market?
Gary S. Fields
Stock No. WP 0434. $3.

Incidence of Poverty and the Characteristics of the Poor in Peninsular Malaysia, 1973
Pravin Visaria
Stock No. WP 0460. $10.

Income Distribution and Poverty in Mexico
Joel Bergsman
Stock No. WP 0395. $3.

Income Distribution Policy in the Developing Countries: A Case Study of Korea
Irmad Adelman and Sherman Robinson
Seeks to answer the question of how much can actually be done to improve income distribution by means of a dynamic general equilibrium model for investigating the potential impact of standard policy instruments and programs intended to improve the relative and absolute incomes of the poor.

Income Inequality and Poverty: Methods of Estimation and Policy Applications
Nanak C. Kakwani
Deals with income distribution meth-
Inequality and Poverty in Malaysia: Measurement and Decomposition
Sudhir Anand
An account of income inequalities and poverty in Malaysia. The research is policy oriented and the findings, to which the author’s statistical technique is applied, are thoroughly discussed.
A range of issues is covered, from data problems to conceptual questions arising with respect to measurement. Employment rates, also looks at disadvantage factors such as age, sex, caste, and health. Suggests that access to physical and human assets may need to be redistributed to improve labor markets.

Labor Productivity: Un Tour d’Horizon
Susan Horton and Timothy King

Manpower and International Labor Migration in the Middle East and North Africa
Ismail Serageldin, James A. Socknat, Stace Birks, Bob Li, and Clive A. Sinclair
This study of labor market trends in the region from 1975 to 1985 analyzes the demand for and supply of manpower by country, sector, occupation, sex, educational level, and ethnic composition. It indicates the sources and destinations of workers and examines the social and economic effects of migration for both importers and exporters, particularly the implications for government policy and planning.
Oxford University Press. 1983. 152 pages (including appendixes, index).

A Model for Income Distribution, Employment, and Growth: A Case Study of Indonesia
Syamprasad Gupta
Presents a quantitative framework to explore the implications of alternative policies for investment, employment, income distribution, and fiscal activity for medium-term and long-term growth.

Patterns in Household Demands and Savings
Constantin Lluch, Alan A. Powell, and Ross A. Williams
Stock Nos. OX 920097, $22.50 hardcover; OX 920100, $12.95 paperback.

Poverty and Growth in Kenya
Paul Collier and Deepak Lal
Stock Nos. OX 520096, $27.50 hardcover; OX 520097, $12.50 paperback.

Poverty and Unemployment in India: An Analysis of Recent Evidence
Pravin Visaria

Public Expenditure in Malaysia: Who Benefits and Why
Jacob Meerman
A methodologically novel study of the household distribution of public services in relation to incomes. The study examines public spending for education, medical care, agriculture, public utilities, and welfare transfers and the use of these services by households.
Oxford University Press, 1979. 404 pages (including map, bibliography, index).
Stock Nos. OX 520406, $27.50 hardcover; OX 520407, $12.50 paperback.
Rural Enterprise and Nonfarm Employment
Dennis Anderson and Mark Leiserson

Examines the importance of rural nonfarm activities as a source of employment and income for the poorest groups of the world’s rural population.
Stock Nos. BK 9087 (English), BK 9057 (French), BK 9059 (Spanish). $5.

---

Saudis in Transition: The Challenges of a Changing Labor Market
Naiem A. Sherbiny, Ismail A. Sirageldin, and M. Ismail Serageldin

Presents the first systematic and comprehensive examination of a rapidly changing labor market in a developing country that does not face the typical capital constraints. Using hitherto unavailable data on many labor activities and processes in Saudi Arabia, the authors examine the push of modernity and the pull of tradition and measure their effects on skill formation, wage levels, and geographic and occupational mobility. They review the basic conceptual and methodological issues involved in the analysis of the Saudi labor market and present a detailed account of the dynamics of labor supply and demand and the characteristics of both the workers and their employers.

Size of Land Holding, Living Standards and Employment in Rural Western India, 1972-73
Pravin Visaria

Stock No. WP 0459. $5.

Some Aspects of Relative Poverty in Sri Lanka, 1969-70
Pravin Visaria

Stock No. WP 0461. $10.

Unskilled Labor for Development: Its Economic Cost
Orville John McDiarmid

Estimates the economic cost of unskilled labor by examining market conditions that cause a significant spread between economic and market wages.
The Johns Hopkins University press. 1977. 218 pages (including 5 appendices, bibliography).
Stock Nos. JH 1938, $14 hardcover; JH 1949, $5.50 paperback.

The Urban Labor Market and Income Distribution: A Study of Malaysia
Dipak Mazumdar

An integrated analysis of differences in income among urban households, the determinants of employment rates of household members, and the major factors affecting personal earnings.
Oxford University Press. 1981. 392 pages (including index).
Stock Nos. OX 520213, $26; OX 520214, $9.95.

Who Benefits from Government Expenditure? A Case Study of Colombia
Marcelo Selowsky

Clearly identifies which income group benefit from public expenditures and attempts to separate the factors behind the supply of and demand for a public service as a means of understanding the present pattern of consumption.
Stock Nos. OX 520098, $20; OX 520099, $9.95.

Women in the Urban Labor Markets of Africa: The Case of Tanzania
Nwanganga Shields

Stock No. WP 0380. $5.

Staff Working Papers

Staff Working Papers are documents based on research conducted within the World Bank or by its consultants on issues affecting the economic and social development of developing countries. The texts are released in an informal manner to make the results of this research available quickly. On average, sixty such working papers are issued annually. The price of a year’s output of Staff Working Papers is $180.

Some 200 Staff Working Papers not available from the Bank are available through the National Technical Information Service (NTIS).
The World Bank Publications Order Form

SEND TO: YOUR LOCAL DISTRIBUTOR OR TO WORLD BANK PUBLICATIONS
P.O. BOX 37525
WASHINGTON, D.C. 20013 U.S.A.

Date __________________________

Name __________________________ Ship to: (Enter if different from purchaser)

Title __________________________ Name __________________________

Firm __________________________ Title __________________________

Address __________________________ Firm __________________________

City State Postal Code __________________ Address __________________________

Country Telephone (_____) __________ City State Postal Code __________

Purchaser Reference No. ________________ Country Telephone (_____) ______

Check your method of payment.
Enclosed is my □ Check □ International Money Order □ Unesco Coupons □ International Postal Coupon.
Make payable to World Bank Publications for U.S. dollars unless you are ordering from your local distributor.

Charge my □ VISA □ MasterCard □ American Express □ Choice. (Credit cards accepted only for orders addressed
to World Bank Publications.)

Credit Card Account Number __________________________ Expiration Date __________________________ Signature __________________________

□ Invoice me and please reference my Purchase Order No. __________________________

Please ship me the items listed below.

<table>
<thead>
<tr>
<th>Stock Number</th>
<th>Author/ Title</th>
<th>Customer Internal Routing Code</th>
<th>Quantity</th>
<th>Unit Price</th>
<th>Total Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All prices subject to change. Prices may vary by country. Allow 6-8 weeks for delivery.

Subtotal Cost $____

Total copies _____ Air mail surcharge if desired ($2.00 each) $____

Postage and handling for more than two complimentary items ($2.00 each) $____

Total $____

Thank you for your order.
Distributors of World Bank Publications

ARGENTINA
Carlos Hirsch, SRL
Attn: Ms. Monica Bustos
Florida 165 4° piso
Galeria Guemes
Buenos Aires 1307

AUSTRALIA, PAPUA NEW GUINEA, FIJI, SOLOMON ISLANDS, WESTERN SAMOA, AND VANUATU
The Australian Financial Review Information Service (AFRIS)
Attn: Mr. David Jamieson
235-243 Jones Street
Broadway
Sydney, NSW 20001

BELGIUM
Publications des Nations Unies
Attn: Mr. Jean de Lannoy
av. du Roi 202
1060 Brussels

CANADA
Le Diffuseur
Attn: Mrs. Suzanne Vermette
C.P. 85, Boucherville J4B 5E6
Quebec

COSTA RICA
Libreria Trejos
Attn: Mr. Hugo Chamberlain
Calle 11-13, Av. Fernandez Guell
San Jose

DENMARK
Sanfundslitteratur
Attn: Mr. Wilfried Roloff
Rosenderns Alle 11
DK-1970 Copenhagen V.

EGYPT, Arab Republic of
Al Ahram
Attn: Mr. Sayed El-Gabri
Al Galaa Street
Cairo

FINLAND
Akateeminen Kirjakauppa
Attn: Mr. Karl Litmanen
Keskuskatu 1, SF-00100 Helsinki 10

FRANCE
World Bank Publications
66, avenue d’Iena
75116 Paris

GERMANY, Federal Republic of
UNO-Verlag
Attn: Mr. Joachim Krause
Simrockstrasse 23
D-5300 Bonn 1

HONG KONG, MACAU
Asia 2000 Ltd.
Attn: Ms. Gretchen Wearing Smith
6 Fl., 146 Prince Edward Road
Kowloon

INDIA
UBS Publishers’ Distributors Ltd.
Attn: Mr. D.P. Veer
5 Ansari Road, Post Box 7015
New Delhi 110002
(Branch offices in Bombay, Bangalore, Kanpur, Calcutta, and Madras)

INDONESIA
Pt. Indira Limited
Attn: Mr. Bambang Wahyudi
Jl. Dr. Sam Ratulangi No. 37
Jakarta Pusat

IRELAND
TDC Publishers
Attn: Mr. James Booth
12 North Frederick Street
Dublin 1

JAPAN
Eastern Book Service
Attn: Mr. Terumasa Hirano
37-3, Hongo 3-Chome, Bunkyu-ku
Tokyo

KENYA
Africa Book Services (E.A.) Ltd.
Attn: Mr. M.B. Dar
P.O. Box 45245
Nairobi

KOREA, REPUBLIC OF
Pan Korea Book Corporation
Attn: Mr. Yoon-Sun Kim
P.O. Box 101, Kwanghwamun
Seoul

MALAYSIA
University of Malaya Cooperative Bookshop Ltd.
Attn: Mr. Mohammed Fahim Htj Yacob
P.O. Box 1127, Jalan Pantai Baru
Kuala Lumpur

MEXICO
INFOTEC
Attn: Mr. Jorge Cepeda
San Lorenzo 153-11, Col. del Valle,
Deleg. Benito Juarez
03100 Mexico, D.F.

NETHERLANDS
MEB BV
Attn: Mr. Gerhard van Bussell
Noorderwol 38,
7241 BL Lochem

NORWAY
Johan Grundt Tanum A.S.
Attn: Ms. Randi Mikkelborg
P.O. Box 1177 Sentrum
Oslo 1

PANAMA
Ediciones Libreria Cultural Panamena
Attn: Mr. Luis Fernandez Fraguela R.
Av. 7, Esquina 16
Panama 301

PHILIPPINES
National Book Store
Attn: Mrs. Socorro C. Ramos
701 Rizal Avenue
Manila

SAUDI ARABIA
Jarir Book Store
Attn: Mr. Akram Al-Agil
P.O. Box 3196
Riyadh

SINGAPORE, TAIWAN, BURMA
Information Publications Private, Ltd.
Attn: Ms. Janet David
02-06 1st Floor, Pei-Fu Industrial
Building
24 New Industrial Road
Singapore

SPAIN
Mundi-Prensa Libros, S.A.
Attn: Mr. J.M. Hernandez
Castillo 37
Madrid

SRI LANKA AND THE MALDIVES
Lake House Bookshop
Attn: Mr. Victor Walatara
41 Wad Ramanayake Mawatha
Colombo 2

SWEDEN
ABCE Fritzson Kungl, Hovbokhandel
Attn: Mr. Eide Segerback
Regeringsgatan 12, Box 16356
S-103 27 Stockholm

SWITZERLAND
Librairie Payot
Attn: Mr. Henri de Perrot
6, rue Grenus
1211 Geneva

TANZANIA
Central Department Store, Head Office
Attn: Mrs. Ratana
308 Silom Road
Bangkok

THAILAND
Thailand Management Association
Attn: Mrs. Sunan
308 Silom Road
Bangkok

UNITED KINGDOM AND NORTHERN IRELAND
Microinfo Ltd.
Attn: Mr. Roy Selwyn
Newman Lane, P.O. Box 3
Alton, Hampshire GU34 2PG

UNITED STATES
The World Bank Book Store
600 19th Street, N.W.
Washington, D.C. 20433
(Postal address: P.O. Box 37525
Washington, D.C. 20013, U.S.A.)

Baker and Taylor Company
501 South Gladiola Avenue
Memphis, Tennessee, 38103

380 Edison Way
Reno, Nevada, 89564
50 Kirby Avenue
Somerville, New Jersey, 08876

Commerce, Georgia 30599
Bernan Associates
9730-E George Palmer Highway
Lanham, Maryland, 20706

Blackwell North America, Inc.
1001 Fries Mill Road
Blackwood, New Jersey 08012

Sidney Kramer Books
1722 Fl Street, N.W.
Washington, D.C. 20006

United Nations Bookshop
United Nations Plaza
New York, N.Y. 10017

VENEZUELA
Libreria del Este
Attn: Mr. Juan Pericas
Avda Francisco de Miranda, no. 52
Edificio Galipan, Aptido. 60.337
Caracas 1060-A
AN ANATOMY OF THE DISTRIBUTION OF URBAN...