Review of Income Distribution Data:
Colombia, Mexico and Venezuela

by

Miguel Urrutia M.

Research Program in Economic Development

Woodrow Wilson School
Princeton University
Princeton, New Jersey

Discussion Paper No. 59

April 1975

This research was supported by the U.S. Agency for International Development, under the joint Princeton/Brookings Income Distribution Project.
Introduction

The purpose of this paper is to analyze the existing studies and statistical data which can be utilized for an in depth analysis of income distribution problems in Colombia, Mexico and Venezuela. In addition, at the end of each country study I have attempted to discuss briefly which distribution issues seem to be of interest in each of the three nations. In order to make the exposition more orderly, I have organized the discussion in three parts, each of them corresponding to one of the countries studied.

COLOMBIA

In Colombia interest in problems of income distribution is rather recent. Reliable quantification of income distribution dates from the second half of the decade of the sixties, and before that only foreign economic missions had discussed the problem of income distribution in the context of development planning. Furthermore, before that date very seldom were economic policy decisions taken with a specific target group within the distribution in mind, and no program or policy was evaluated from the point of view of its actual redistributive impact, among other things because there was no reliable data on the distribution of income.

Survey of Various Income Distribution Estimates

Before 1967, various foreign economic missions to Colombia tried to estimate the distribution of income, but most of these attempts limited themselves to quantifying income per capita differentials by economic
sectors. The first discussion of income distribution in Colombia was that presented in the World Bank study of the Colombian economy in 1947. Using scattered data, such as one or two rural surveys, some income tax data and wage data, the mission estimated the number of people with high, medium and low incomes. The sources for this estimate are so deficient, however, that the resulting distribution can only be considered as an important first step in the study of an important subject.

The second estimate of income distribution is that of CEPAL (U.N. Economic Commission for Latin America) in 1954-55. This study includes information on the upper 5% of the distribution from tax data, and average incomes by industrial sector, and therefore does not really present a distribution of income. A third foreign mission headed by C.J. Lebret studied distributional problems in the late fifties. Although the mission did not produce any estimate different from that of CEPAL or the National Planning Office, it did carry out surveys in various parts of the country to obtain social indicators such as standards of housing, education, health, etc.

Probably the first complete income distribution estimated for Colombia, from which income by deciles can be derived, was carried out by still another foreign mission in 1961, headed by Milton Taylor. This estimate is based on income tax data, information on salaries from national accounts, and the distribution of income of blue collar and white collar workers obtained in a 1953 family budget survey. The income of people in services and commerce was also obtained by assimilating the labor force in those sectors to the workers employed in manufacturing industry surveyed in 1953. The assumptions used to estimate rural income
are arbitrary and the basic rural wage data is of very low quality. Since it has been shown that the 1953 budget survey is not representative and tends to exclude low wage urban workers, as is shown by the relatively low average weight of food expenditures of the families surveyed, the Taylor estimate probably underestimates urban income dispersion, since it underestimates the importance of low income urban workers. In summary, this estimate does not seem reliable and cannot be usefully compared with later estimates based on better basic data. Since the Taylor estimate gives a significantly more equitable distribution than the more complete later estimates, and the methodology used would seem to bias the results in that direction, that income distribution should not be used to compare changes through time in income dispersion.

In 1968 still another mission, this time headed by Richard Musgrave, estimated an income distribution for 1964. This estimate is also largely based on income tax data, national income information, and some income statistics from survey data and rural wages. This estimate produces a more concentrated distribution than that found by Taylor, and seems better from a methodological point of view. The McLure estimate for 1954 utilizes a methodology broadly similar to that of Taylor, but with better data, including the 1964 census results of the distribution of population by sector and some more recent survey data. The McLure estimate for 1964 coincides broadly with the Urrutia-Berry tables described in detail in this chapter. Thus the McLure 1964 estimate is an interesting check on the independently estimated Urrutia-Berry distribution for the same year which is based on urban surveys, the data from the 1960 agricultural census, and a small rather local rural survey used to derive rural non-farm income.
The Urrutia-Berry 1964 income distribution is the result of integrating an urban distribution for 1967 and a rural distribution for 1960 with the information from the 1964 census of population. The data on urban incomes come from the labor force surveys carried out in various urban areas by the Centro de Estudios para el Desarrollo Economico of the Universidad de los Andes (CEDE) in 1964 and 1967-68. These are random surveys which cover cities with 48% of the urban population, and it was necessary to transform this sample to a general urban distribution by assimilating the cities not included with those surveyed on the basis of indicators such as size, productivity per capita of labor and quality of housing. In order to avoid double counting, those people whose major source of income was agriculture were excluded from the urban distribution even if they lived in towns.

The data for the rural distribution comes primarily from the agricultural census of 1960, which covered a large part of the country and data on production per hectare for different crops was obtained from a 1966 national agricultural survey. Data on agricultural wages were used to estimate the income of the rural proletariat, and a survey in a region of Boyacá was used to estimate non-agricultural rural income by relating such income to agricultural wages and applying the resulting differentials to agricultural wages in other regions of the country. It should be pointed out that the rural distribution does not include unpaid family workers.

As can be seen, the data base for the distribution involved a wide-ranging variety of sources. It should be pointed out, however, that the final results check well with national income aggregates, and that
such income data from the surveys as could be checked with other sources, as in the case of urban wages, seemed to be of rather good quality.

There is another published income distribution for the same period. In 1965-66 the Ministry of Public Health carried out a large survey of the health situation of the population which included a question on income. Unfortunately, the question was very general, the income intervals large, and the highest and open interval included a large proportion of the population. Due to the bad methodology used in obtaining the income data, total income of the population estimated by the Ministry of Health is much lower than that derived from national accounts. For example, for the city of Bogota, the Ministry data, when blown up, only accounts for 61% of the total income estimated from the CEDE unemployment sample, which in turn somewhat underestimates total income. It appears therefore that the Ministry data is very deficient and cannot be utilized for income distribution studies.

There are also two additional income distribution studies, both based on the large household sample made by the National Estatistical Department (DANE) in 1970. This sample was very ambitious and included questions on employment and labor force status, family expenditures (budget study), and incomes. The design of the urban sample seems to have been quite good and therefore the resulting urban distribution is probably of high quality. There are doubts, however, about the representativeness of the rural part of the sample. It appears that the rural areas surrounding the large cities make up a large part of the rural sample, and these areas are clearly not typical. Furthermore, while half of the nation's population is rural, only 38% of the sampled population is rural.
It thus seems that the DANE sample did not capture the poorest rural families, who in general will be found in areas located far from the principal urban centers. This would tend to bias the resulting income distribution and make it more equitable than it is in reality.

The results of the DANE survey were first published by Polibio Cordoba, and they show less concentration than that found in the Urrutia-Berry distribution; the Gini coefficient of concentration is 0.53 against 0.57. The urban distribution, however, is quite similar to that found by Urrutia-Berry. The concentration coefficient for the rural sector, on the other hand, is much lower in the DANE distribution, (0.42 against 0.57). This difference in the rural distribution is due to various factors. First, the DANE data does not include the high income farmers who live in towns, while the Urrutia-Berry distribution includes those land owners in the rural distribution. Second, the DANE survey is not representative for the rural areas, and does not include the poorest farmers in marginal areas. In summary, some reservations can be had with respect to the rural part of the DANE distribution. For that reason, no comparison can be made between the 1964 and 1970 published distributions, since the statistical basis for them are so different. However, the fact that the 1970 and 1964 urban distributions give similar results suggests that the type of concentration described in the Urrutia-Berry distribution is consistent with independent evidence from the DANE survey.

Charles McLure has also corrected his 1964 estimate by using the 1970 DANE survey results. Once again, he obtains an urban distribution similar to that of Urrutia-Berry, although the latter is
slightly more unequal. The rural distribution is less unequal than that of Urrutia-Berry, but McLure finds this logical, particularly because the DANE survey does not include in the rural sector the landowners who live in towns, and because DANE information refers to households instead of economically active persons. McLure concludes that "the Taylor group found that the bottom two-thirds of individuals received roughly 28% of income and the top 10% of individuals received slightly less than 42% of income. This is difficult to reconcile with the present study in that the Taylor group also used individuals as the basis of its estimate. Finally, Urrutia's estimates are quite consistent with those reported here and in the author's (McLure) previous study." In summary, the McLure, Urrutia-Berry and DANE studies all show similar distributions of income, and therefore reinforce the validity of the Urrutia-Berry estimates, but they cannot be used to show marginal changes in income distribution in the short period 1964-1970, since it is hard to tell what differences are due to inadequate data and what differences may reflect real underlying trends. Table A-1 illustrates some of the differences in four of the distributions discussed.

To finish this survey, we should mention the urban distribution estimated by Rafael Prieto from the ECIEL family budget study in the four largest Colombian cities. This is probably the best income survey carried out in Colombia, but it has a series of shortcomings. The interviews and questionnaires were of high quality, but the study was designed as a family budget study to obtain income elasticities for consumer goods. For that reason a stratified sample was used, with high income individuals overrepresented. In order to blow up this sample to estimate an income distribution it was necessary to weigh the three large categories of
of consuming units (high, middle and low income). These weights were derived from the Taylor distribution, and corrected by impressionistic evidence concerning the incomes of people by barrios in the cities covered. Since the Taylor distribution underestimates concentration, the weights used to obtain an income distribution from the actual sample data probably lead to an underestimation of concentration. In fact, as expected, the Gini coefficient in the Prieto study for the four cities is 0.47, against the urban coefficient of 0.55 found in the Urrutia-Berry distribution.

In addition to problems that may exist in the coefficients used to transform the stratified sample, the Prieto data refers to the income of families in three months. This also should determine a lower degree of concentration, both because family income is less concentrated than individual income, and because there should be less dispersion in quarterly income than in the weekly income which is the basis of the Urrutia estimates for the urban sector. In Colombia this would certainly be the case, since unemployment is of short duration and weighs heavily in the Urrutia results, while it would have a lower impact on the lower deciles in the Prieto data.

In summary, the Prieto data may give a good representation of the distribution of permanent income and can be very useful for in depth study of the distribution problems of the large urban centers, but due to the methodology used to transform the stratified sample into a global distribution, the published distribution may underestimate the degree of concentration.
TABLE A-1

Comparison of Results of Various Studies of Income Distribution in Colombia

% of income received by

<table>
<thead>
<tr>
<th>Author</th>
<th>Date</th>
<th>Lowest 2/3</th>
<th>Top 10%</th>
<th>Basis of estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taylor</td>
<td>1961</td>
<td>28</td>
<td>42</td>
<td>Individuals</td>
</tr>
<tr>
<td>McLure</td>
<td>1964</td>
<td>27</td>
<td>50</td>
<td>Individuals</td>
</tr>
<tr>
<td>Urrutia</td>
<td>1964</td>
<td>24</td>
<td>48</td>
<td>Individuals</td>
</tr>
<tr>
<td>McLure</td>
<td>1970</td>
<td>25</td>
<td>44</td>
<td>Households</td>
</tr>
</tbody>
</table>


Research on the determinants of income distribution

In order to analyze the determinants of the present level of income concentration it would be very useful to have data on income distribution for different points in time. Unfortunately, as has been explained, there are no distribution estimates of sufficient quality before 1964 to allow an analysis of the changes in distribution through time. Berry and Urrutia, however, have attempted to obtain a general impression of the changing profile of income distribution from the thirties to the sixties.

Given the lack of aggregate data, it was necessary to piece together diverse types of information. Major reliance was placed on the
wage series for various groups of workers; this, coupled with information on functional distribution over time and changes in the occupational and sectoral structure, permits some appraisal of a) changes in the overall personal or family income distribution and b) with somewhat greater precision, changes in the distribution of labor income. The basic conclusions of the analysis are: a) that income distribution in agriculture has worsened throughout the period since the mid-thirties, b) that non-agricultural income distribution probably worsened from the mid-thirties until some time in the early fifties, then improved till some time in the mid-sixties, and then tended to stabilize off at about the level of concentration found for the early thirties.

Although the various wage series and functional distribution data give a good idea of the trends in income distribution in the last three decades, the pre-1964 distribution data does not allow the researcher to test statistically any hypothesis concerning the possible determinants of income distribution in Colombia on the basis of time series analysis.

Berry and Urrutia have attempted, however, to identify some of the possible causes of income concentration in Colombia. For example, there is information on the extent to which the unequal distribution of rural land determines income concentration in agriculture. Some work has also been done concerning the incidence of tax and government expenditure on the distribution of income, and on the impact of the unequal distribution of education on the distribution of labor income. Some work has also been done on the possible impact of monetary and exchange policy on income distribution, as well as of the effect of government trade union policy and legislation on the incomes of workers.
Given the quality of the data, however, it has been difficult to use orthodox statistical methods to test the postulated relationships between policy variables and income distribution. Due to the lack of census material or of a representative rural-urban sample, it has also been impossible to use some interesting methodologies such as the information theory technique for decomposing total income variance into component parts attributable to sector, region, age and education (e.g., Fishlow for Brazil).

Data availability for a new estimation of income distribution

The best estimates available of Colombian income distribution are for 1964, and some research has been carried out concerning the trends in distribution up to 1964. There is therefore a high priority for estimating a new distribution and for carrying out an historical analysis of the changes that have occurred in the last decade.

It would be very important to estimate a new distribution for 1973. In this year the population census was carried out and this facilitates greatly the estimation of an income distribution. The census itself has data on the income of the population, which facilitates decomposing total income variance into component parts. Unfortunately, it is felt that the income information from the census is incomplete and probably not very reliable. But the labor force data from the census can profitably be used to generate a distribution based on income data from surveys carried out with greater care by more highly qualified interviewers.

There are in fact various very valuable sources for income distribution work for the years 1972-73. In 1972 DANE carried out its last comprehensive household survey (12,000 household units), which as
has already been pointed out, contains very complete information on income of families and members of the labor force for the urban sector. In addition, DANE carried out an extremely ambitious agricultural survey in 1973, which obtained detailed production information on 17,000 rural production units. That survey obtained detailed information on production, inputs and costs, and included data on agricultural wages and net income of the agricultural units. Although the questionnaire asked about consumption expenditures by large categories, DANE took a subsample of 5,500 rural households in order to obtain more detailed information on consumption and income.

In addition, in 1973 Banco de la Republica also carried out an independent survey of 1,800 agricultural units which also contains information on production, rural wages, consumption and net income generated by the farm. Both of these rural studies cover the major agricultural departments, and can therefore be used to ascertain regional differentials in income and production. The Banco de la Republica survey concentrates on credit availability and can therefore also be used to determine who are the beneficiaries of credit policy.

With the census information and complementary data from the above mentioned urban and rural samples, it should be possible to estimate a fairly accurate income distribution for 1973. It should be kept in mind, however, that all this information takes some time to tabulate and that it is unlikely that the Census and DANE rural survey will be published before 1976. The Banco de la Republica data will, on the other hand, be available as of December 1974.
Data availability for an estimation of recent trends in income distribution

Given the abundance of primary data for 1973, it is possible to estimate a good income distribution which would be comparable to that of 1964. In addition to seeing how income dispersion changed in the decade, which in itself would be of great interest, there is also sufficient data to make possible a study of the changing relative position of different sectors of the population.

Between 1964 and 1973 there is good yearly data on the wages and fringe benefits of industrial workers, by sectors, published by DANE. There is also good data on the income of all salaried personnel in small industry, commerce and services from the CEDE and DANE labor force surveys. The CEDE surveys are available for Bogota from 1963 to 1970 and the DANE surveys, which are comparable, run from 1970 to 1972. There is also some income data from labor force surveys for other years in various cities, but no complete series such as those for Bogota. However, there are observations for the major cities in 1967 and 1970, and for Medellín, Cali and Barranquilla for earlier years as well.

Another interesting primary source is the Social Security Institute which has income data for all of the affiliated members for most of the decade. This could cover all wage earners in modern industry, plus a large percentage of wage earners in commerce, services, banking, etc. It would not include artisans or the large group of self-employed nor many wage earners of establishments with less than 10 workers. Thus some indication concerning income distribution changes in the urban labor force can be derived from that data. At present, the data being produced by the Institute is quite detailed, and information can be
obtained by size of establishment, but there is a problem with the fact that more small establishments have been joining in the last 8 years, and it is difficult to determine whether wage changes are due to the changing composition of the members, or to average increases in wages by sector.

Statistics on rural wages exist but are unreliable, and trends in the incomes of farmers probably have to be compared by looking at the Berry rural distribution of 1964 and the production and income data derived from the Banco de la Republica and DANE agricultural surveys of 1972-73. 26

The rational accounts, complemented by the other information already mentioned, also make it possible to study the functional distribution of income in the decade, and there is some possibility of breaking the labor share into income for skilled, semi-skilled and unskilled workers by utilizing recent input-output tables.

Finally, there is some data on income tax returns for the early sixties, and detailed statistics starting in 1965. 27 This data could make possible for the first time an in depth study of the upper decile of the distribution, and of profits. In addition to the study of the characteristics of the upper income groups and of the sources of income of those groups, it would be very interesting to analyze the changing composition of the richest segments of society during the recent period of accelerated inflation.

Data of General Interest

In addition to the data already discussed, it may be useful to mention that Colombia has now some new price indexes of good quality.
There is a comprehensive wholesale price index starting in 1970, which replaced a somewhat deficient index also published by Banco de la Republica and which went back to 1952. The consumer price index is only for urban blue collar and white collar workers and goes back to 1954-55. Before that date there are less complete wholesale and consumer price indexes.

In any event, only information on urban prices exist, and this makes difficult the comparison of real incomes in the urban and rural areas. In 1974-75, however, the author plans to do research leading to a comparison between urban and rural purchasing power. (I would welcome suggestions on methodology for this research).

Of general use also are the 1967 input-output tables made by Planeacion and AID respectively. 23

Interesting subjects for Future Research

Although the first priority in Colombia at present is to study the changes in income distribution in the last decade, another very fruitful avenue to explore is that of estimating the effective impact of various policies on that distribution. In this area there are various studies which could be carried out:

1. Further research on the incidence of taxes and government expenditure on the family distribution. In particular, the incidence of government expenditures on health and education could be studied with greater detail and on the basis of surveys at the school and health center level. 29 An effort should be made also to study who benefits from transport subsidies and investments, and from expenditure on agriculture. Finally, a separate study could be made of the incidence of the Social Security
System. In principle there is some possibility that the Universidad de los Andes, in collaboration with IBRD, will do research of this type, with emphasis on the incidence of education and health expenditures. Furthermore, the McLure tax incidence study should be revised, taking into consideration the comments made by Richard C. Webb, which suggest that McLure may be underestimating the progressiveness of the tax system. 30

2. Related to the above, there is now probably sufficiently detailed income tax data to determine the incidence on income distribution of the wide range of tax exemptions in the capital market. Such a study could also be carried out simultaneously with one in which an attempt would be made to evaluate the income concentrating distortions brought about by the increasingly powerful financial groups being formed in the country.

3. Recently Colombia has started to subsidize heavily certain food items, and there is a long tradition of price controls on other food items. It would be useful to calculate the incidence of such subsidies and price controls.

4. Another vital area of study is that of the state corporations and industries. Who benefits from these state corporations? The consumer (through lower prices), those who work for them (through higher than equilibrium wages), the state in general (through profits, etc.)? It would also be useful to know who benefits from the subsidies to public utility companies.

5. There is also much that can be studied in the rural sector. Some of the subjects are:

a) Who benefits from technological change in agriculture.

b) The distributional impact of subsidized credit in agriculture.
c) The relationship between forms of rural management and labor per hectare productivity and productivity in general. The hypothesis is that absentee forms of rural management imply methods of production which are biased against labor.

6. Research can be carried out on the effectiveness and distributional impact of nutritional supplements to school and preschool children. The Instituto de Bienestar Familiar has some useful data which can be used in this research.  

In addition to the above subjects for research, at the end of this paper the reader will find a list of research topics which can best be treated by analyzing the problem for the economies of the three countries chosen.

**MEXICO**

**Survey of Existing Studies of Income Distribution**

In Mexico the interest in income distribution problems precedes by a decade similar investigations in Colombia. However, the income distribution estimates that have been made have been based on rather shaky data, and great care should be used in interpreting measured changes in distribution due to the nature of the basic data used to produce those estimates.

Interest in the distribution problem was first aroused by the statistics produced by the Comision Mixta del Gobierno Mexicano y el Banco Mundial on the functional distribution of income. The Combined Mexican Working Party produced statistics on the functional distribution of income for the period 1939 to 1959. That data showed a constant decrease
of the participation of wages and salaries in national income from 1939 to 1946, and a corresponding increase in the participation of profits. Although the statistics were quite rough, they posed the question of whether the poorest sectors of the population were benefitting from Mexico's economic growth. The interest awakened by this study led to a series of investigations on personal and family income distribution, but before turning to these, it is useful to make a short description of the statistics on functional distribution. The Combined Mexican Working Party functional distribution covers the years 1939-50, and the Banco de Mexico has published more detailed statistics for 1950-1967.33 Unfortunately, however, it appears that the Banco de Mexico discontinued its detailed national income statistics in 1968, and this makes it very difficult to update the functional distribution of income although GNP and production statistics have continued to be produced.

As in all developing countries, the great problem is to separate out from the income of family enterprises the labor incomes of the family workers. The 1939-50 data has this problem, and in addition does not include 'de brasero' income, which was probably important in the period. Even after adjusting for some of these problems the participation of wages and salaries still remains low (at about 30%). Luciano Barraza,34 however, has found independent evidence by using an input-output table for 1969 which confirms the tendency of the labor share to decrease from 1939 to 1946, only to increase to the level of 1939 by the year 1965. In summary, although the series are weak, there are indications that the functional distribution series for year previous to 1950 reflect with some accuracy general trends.
The functional distribution data for 1950 to 1967 is of much better quality, and was produced by sector. In almost all sectors, and for the aggregates, an improvement in the labor share is shown in the 1957-1967 decade.

The first attempts to estimate a family income distribution were carried out by Ifigenia M. de Navarrete, although Morris Singer mentions some earlier work based on very aggregate data. He mentions work by Flores Marquez with 1950 Census information but there is evidence that this data tend to overstate the per capita incomes received by the lower income groups. To show the problems involved in using Census data, it is illustrative to mention that according to Flores Marquez, total income in the census was 65% higher than that estimated by the Comision Mixta.

Although discrepancies of this size make any estimate of personal income distribution based on this data somewhat suspect, Ifigenia de Navarrete made a heroic effort to correct for these discrepancies and produced a personal income distribution for 1950 and 1957. Although based on shaky data, these are the least unreliable distributions yet published. The 1950 estimate is based on census statistics, adjusted so as to make the family income figures from the census comparable to those derived from the National Accounts for persons working.

In a later article, which contains a good description of her methodology, Mrs. Navarrete perfected three distribution estimates, which in her view are comparable, for 1950, 1958 and 1963. The 1958 estimate is based on a survey carried out by the Dirección General de Estadística de la Secretaría de Industria y Comercio and that for 1963 is based on a family survey carried out by the Oficina de Proyecciones
Agrícolas of the Banco de México.

The difference between personal disposable income as calculated in National Accounts and the family income derived from the census and the surveys implies an underestimation of income of 17.8% for 1950, 26.9% for 1958 and 19.6% in 1963. The adjustments made by Mrs. Navarrete consist primarily of assigning this missing income to the lowest and highest income groups, where it is supposed most of the undeclared income can be found. The method of adjustment is arbitrary, but there are good reasons for thinking that such an adjustment is justified. A 1956 survey was used to calculate income in kind for the other years, assigning income in kind to those groups whose consumption exceeded their declared income.

As can be seen by the larger underestimation found using the Dirección General de Estadística survey, the data from this source seems to be the least reliable. Various informants in Mexico confirmed this impression, and suggested that surveys of the Dirección General de Muestreo have had serious statistical problems.

Despite Mrs. Navarrete's commendable efforts, it must be admitted that her distributions are based on quite dissimilar data, and therefore any comparison of trends through time is risky. Richard Weisskoff\textsuperscript{40} comments, however, that despite his reservations about the comparability of the years, the income shares to the bottom 30% of the families declined throughout the entire 1950-1963 period, and the share of income received by the "middle" classes in the 51 to 95% groups increased. Both Weisskoff\textsuperscript{41} and Urbina\textsuperscript{42} use various measures of concentration to estimate the trends in income distribution between 1950 and 1963 (Urbina also uses 1969-70 data), and although the measures do not show a consistent trend, it
does appear likely that concentration did not decrease significantly between 1950 and 1963, or even 1969-70.

But it should be emphasized that the 1957 and 1958 Dirección General de Estadística samples have serious methodological problems, and the 1950 census data and the 1963 Banco de Mexico survey data are not strictly comparable. For that reason, great care should be taken in the use of the existing estimates of personal and family income distributions.

For a more recent period, it may be possible to make more accurate estimates of changes in income distribution. The Banco de Mexico has recently published the results of its 1968 family budget study. This survey is national in scope, and the data is representative of the farm and non-farm population. It is estimated that the margin of error of the estimates is less than 3%, and 5,608 families were covered. That survey is comparable to that carried out by the same institution in 1963, and therefore the comparison of the two distributions would seem to be justified. Although some experts have doubts about the quality of these surveys (for example, it has been found that expenditures on food do not always increase as income decreases), Leopoldo Solís has made substantial corrections to the 1968 survey data, and has offered to make available the tapes with the information to investigators interested in the study of income distribution. Work on the 1968 survey and comparison with previous distributions, especially that for 1963, would therefore appear to be a fruitful avenue for research.

In addition to the national income distribution studies already discussed, Mexico appears to have a great wealth of sectoral and regional statistics. From national accounts GNIP per sector and per capita can be
derived, and Navarrete has used this data and that from social security, regional government expenditures and consumption of electricity to elaborate indices of regional income inequality. The Comision Nacional de los Salarios Minimos has also collected detailed statistics on wages, production, and productivity per hectare in agriculture for a great number of regions for 1967-1970. It may be possible to obtain similar information for earlier periods.

There is also an extremely interesting and detailed labor force study, which includes high quality income distribution data carried out under the direction of Jesus Puente Leyva for the city of Monterrey. That publication also includes interesting data on minimum subsistence incomes and other information on the standard of living of the very poor, such as who benefits from the city's social services.

The information existing in Mexico for wages and income of the commercial and service sector is of particular interest, since these are the least understood sectors in economic development literature. The data from the Comision Nacional de los Salarios Minimos should definitely be used for a study of the characteristics and conditions of life of the labor force of these sectors.

Research on the determinants of income distribution

The best published material on the determinants of income distribution can be found in Leopoldo Solis. He attempts to explain the changes in functional distribution. He emphasizes the role of import substitution policies in the deterioration of the participation of labor income, and of the large investments in irrigation on the concentration of income in agriculture. In addition he considers that the fiscal
did not contribute to equality and that technological change was an important determinant of changes in income distribution. Although the Solis hypotheses are interesting, and some are supported by econometric work and input-output estimates, there is much room for further research. For example, did the increases in profits occur in industries that became protected in the import substitution phase? If the large increases in profits occurred in commerce can this be explained by the monopoly profits of those importers who obtained scarce import licenses? Profits should be disaggregated to see if those sectors who should have gained from import substitution are the ones responsible for the increase of the capital share.

It would also be of interest to determine who benefited most from infrastructure investment, and how the policy for amortizing irrigation works led to income concentration.

Mrs. Navarrete⁴⁹ has also discussed the determinants of income concentration in Mexico. She emphasizes labor saving technological change as the major cause of concentration, but also blames the growth oriented strategy of the government which has been biased toward high profit and interest rates in order to facilitate savings. Recently she has become convinced that the Mexican capital market, with its high interest rates and tax exemptions, is a major cause of income concentration. However, as yet, little empirical testing of these hypothesis has been carried out.

Data availability for new estimates of income distribution

The first priority is to analyze the great wealth of information on income obtained by the IX Population Census of 1970, which refers to
incomes for 1969. The Comision Nacional de los Salarios Mínimos, has published tables, by region, of the income of the economically active population that declared income by economic activity and by principal occupation. The income concept includes income from all sources (including profits) before the payment of taxes and social security. The tabulations include 8 income groups. The last group, which is open, includes all workers with more than 10,000 pesos a month, which makes the top group too large, but still manageable.

Although this information is ideal, since it makes possible the study of income distribution by region, by occupation and by activity, and makes possible the decomposition of total income variance into component parts attributable to sector, region, age and education, it is necessary to see if such Census data (which is not usually good for income distribution studies) is of good quality. The Census data should therefore be checked for consistency with the 1968 Bank of Mexico survey and the income aggregates derived from national accounts.

If adjustments can be made to the census data, much research can be carried out with it in the area of income distribution. It may also be possible to make some comparisons with the 1950 census data. The Census also includes data on minimum consumption levels.

As has already been mentioned, the 1968 Bank of Mexico survey can also be used to derive a new distribution, and this is simplified by the fact that Leopoldo Solis has put all the information on tapes. Such a distribution can probably be compared with one derived from the 1963 Bank of Mexico survey, and this would also be useful for identifying recent changes in income distribution. The Bank of Mexico has already published an income distribution for families, divided in farm and
non-farm categories. 52.

The other major area in which new work can be done in the area of income distribution is to utilize data from the agricultural censuses to describe the distribution of wealth and income in rural areas. There is data from such censuses for 1950, 1960, and 1970 on land distribution and production. Following the method used by Albert Berry for Colombia, 53 it should be possible to estimate an agricultural distribution of income for each of these three years, and thus obtain a very good profile of changing income distribution in agriculture.

For example, there is already some information published from 1960 & the 1970 census54. From that census one can already obtain data on productivity and average prices of each product by region. This coupled with land distribution data and number of laborers (and their salaries) makes possible the calculation of an agricultural distribution.

The questionnaire included questions for each production unit on:

I. General characteristics of the production unit, of the producer and the owner.

II. Classification (quality) of the land.

III. Types of crop.

IV. Use during the census year of the land (type of crops)

V. Area of land with lost crops.

VI. Products obtained in land not under crop.

VII. Employed persons.

VIII. Fruit trees and plantations.

IX. Forests and forest products.

X. Livestock, chickens, bees, etc.

XI. Animal production.
The rural distribution derived from production statistics by size of production unit could then be checked for consistency with the census data and the information from the Bank of Mexico survey for 1968.

Other sources which can be used to improve income distribution estimates are those on salaries derived from income tax data. Only the part on salaries can easily be tabulated, since employers send the complete information on withholding to the Secretaría de Hacienda. This information would be useful to correct the last decile of the salary structure, where salaries are usually underestimated in surveys or in the census.

Additional data on salaries can be obtained from Social Security Statistics and from INFONAVIT. In both cases a substantial amount of programming and computation work would be necessary, since these institutions do not at present produce statistics useful for our purposes. Furthermore, the coverage is less than complete and limited to the modern sector.

Finally, the most interesting new data available for income distribution studies is that produced by the Comisión Nacional de Salarios Mínimos. The government has established minimum salaries only in those professions and occupations in which the majority of the workers are not protected by collective bargaining agreements. To set these minimum wages, the government has studied carefully these professions, and has collected detailed statistics on income and education of these laborers. The interest of this work lies in the fact that these are precisely the workers in the "informal sector" (according to the ILO definition) for which most countries have no statistics.
In 1966 only 12 occupations were included, but the number increased to 25 in 1968, and 36 in 1970. By 1972, 49 occupations were included in 87 regions. The Comision has carried out surveys on incomes for these 49 detailed occupations. But to choose these occupations, in 1971 surveys were carried out for 80 occupations which included particularly laborers in commerce and the services whose characteristics are usually unknown. In addition to income data, information was obtained on their education and training, their age, and the detailed characteristics of the jobs. All this information provides, by region, invaluable information on the income of laborers in the lowest deciles.

It should also be mentioned that the Comision has published detailed statistics on wages of the industrial sector, based on the monthly wage survey of industry carried out by the Dirección General de Estadística. This monthly survey includes data on wages, salaries, and fringe benefits (prestachiones), as well as hours worked by industrial sector and region. However, the data does not cover small establishments.

There is therefore very plentiful data on wages, which makes possible the study of recent trends in wages and on the distribution of labor income. Longer wage series can be found in Singer, as well as Everett, Reynolds also has a long series (1934-1965) on minimum wages, as well as data on rural/urban income shares and discussion of regional inequality.

In summary, there is much recent data in Mexico which can be used for new estimates of income distribution. This data also makes it possible to identify poverty groups by region and economic activity, and may make possible serious empirical work on the determinants of income.
concentration. For this reason, it is felt that Mexico, like Colombia, would be a good country to choose for a pilot study on income distribution.

**Interesting subjects for future research**

The most interesting research that can be carried out in Mexico involves the study of the redistributive impact of the nation's fiscal system. The following specific subjects would seem to be worth studying:

a) The tax burden by income deciles. It would be interesting to see who contributes tax revenue, and the possibilities of making the tax system more progressive.

b) The impact of government expenditure on distribution. For example, given the methods used to finance irrigation works and the structure of user taxes, who benefited from those investments? Also who benefits from the large transport investments? It would also be pertinent to study the regional distribution of expenditures, as well as the distributional impact of expenditures in education, health, housing, social security, etc.

c) What has been the distributional trade off of financing public investment through foreign indebtedness instead of through taxes? In order to assure the flow of foreign capital has it been necessary to maintain very high profits rates (and therefore interest rates), thus insuring the concentration of income? Has this dependence on foreign credit tied the economy to a growingly overvalued exchange rate, with the usual negative impact that this has on the labor intensity of technological change?

d) What is the distributional impact of tourism development?
e) What has been the distributional impact of minimum wage legislation is also an interesting subject for research. It would be interesting to see what trade off there has been between unemployment and higher salaries for the lower deciles due to this legislation.

f) Finally, Leopoldo Solis has emphasized the role of import substitution policy on income concentration. A study should be carried out to see whether in fact profits in the protected industries have grown rapidly as protection increased. Some serious work should also be done on the commerce sector to try to determine the causes of high profits in that sector in the early postwar years.

VENEZUELA

Existing income distribution studies

Of the three countries studied here, Venezuela is probably the one in which there has been less analytical work done in the area of income distribution. Although there does exist data on incomes for various years, no one has yet published a complete income distribution estimate based on the existing primary data.

As far as I was able to ascertain, the first work on income distribution was carried out by the fiscal mission headed by Carl Shoup. The estimate of income distribution for 1957 made by the mission was based mainly on material collected by the Banco Central in connection with its estimates of national income and expenditure; on the results of various sample surveys; on the statistics of the Ministry of Finance from income tax returns; on the Encuesta Agropecuaria Nacional of 1956; and on payrolls of a number of establishments.
Later on, and maybe on the occasion of a foreign mission, that of IRFED, the data from a national survey on incomes was used to calculate an income distribution. The data published covers the year 1962, but the results show a more equal distribution than was to be expected. CEPAL has also used this same source to estimate an income distribution, although it has made some adjustments based on complementary statistics from the Oficina de Muestreo and national accounts. CEPAL attempts a short explanation of the particularities of the Venezuelan distribution which although very tentative, is of interest.

But probably the best empirical study of distribution in Venezuela is being finished by Dr. Lourdes de Ferran of the Banco Central de Venezuela who has estimated distributions for 1960 and 1967. The 1960 estimate is based on the income data from the 1961 census, and that for 1967 on the household survey for that year. The author considers that, with the adjustments made, the data on wages and salaries is acceptable, but she has many doubts concerning the incomes of the top decile, where there is usually a tendency to underestimate incomes in census and survey declarations.

One of the problems of the 1961 census is that it presents three types of income: daily, weekly and monthly, and there are some problems of aggregation of the data. The census data appears by occupation and economic activity, and these are sufficiently detailed so that it is possible to define the characteristics of the poorest sectors of society. The census also has income data by levels of education.
Mrs. Ferran also uses the 1967 household survey to estimate a distribution which she thinks is roughly comparable with that of 1960. The 'encuestas de hogares' present a great deal of income information on blue and white collar workers, and they have been carried out periodically. I have seen income data published for 1967, 1968, 1969 and 1970. For example, the 1967 survey covered 10,112 households, and all of the surveys have weekly income of the economically active population classified by occupation and economic activity, although it only covers wage earners in non-agricultural occupations. The income data is not for the reference week, but "usual" weekly income.

As far as I know, the above data has not yet been used for analytical purposes, and only Mrs. Ferran has tried to use it to produce a distribution. Antonio Fernandez of the Banco Central de Venezuela has also recently used this data and that of the 1971 census to obtain a distribution for private employees.

Except for some of the publications of the ECIEL program which analyze some results of the household surveys carried out in various urban areas, there are in Venezuela few studies published on income distribution. One of these is by Tokman, who treats the possible impact on employment generation in industry of the changing consumption pattern brought about by income redistribution, and the possible impact on labor demand of a technology strategy. He finds that the change in consumption patterns caused by income distribution would have a negligible effect on employment, while a strategy to encourage production with labor intensive processes in industries where production with different technologies is possible would increase significantly the demand for labor and thus contribute to a better distribution of income.
Another piece of research which should be mentioned is the attempt by CENDE to estimate an urban income distribution based on household surveys.

Data Availability

Although in Venezuela little research appears to have been done in the area of income distribution, there is relatively plentiful data on the income of families and the labor force. Unfortunately, before someone works with the data and does consistency checks, it is difficult to determine of the quality of the information. For example, the economic characteristics of the population for the 1971 census were obtained from a 25% sample of census units, and there is some evidence that the sample was biased towards small families. In addition, as a rule census information on income is of low quality, and the last open income category of the census includes a rather large proportion of the population, thus making it difficult to calculate a global distribution in which obviously the upper decile weighs heavily.

In addition to the sources already mentioned, such as the 1961 and 1971 censuses, which contain income information for the labor force, much income material can be obtained from household surveys.

There are first of all the periodic labor force surveys of the Ministerio de Fomento, which have income data of salaried urban non-agricultural workers. Probably of more interest are the four family budget studies carried out by the Banco Central de Venezuela. The one for Caracas was done in two phases. The first covered 4,137 families and produced information on income from all sources for these families. The second part of the survey, which produced detailed information
on consumption of 1,100 families was carried out in the same year (1966). In the second half of 1968 a household survey in the area Puerto La Cruz - Barcelona was carried out in two stages as is the Caracas case, covering 855 and 350 families respectively.\footnote{73} Maracaibo was covered by a similar set of surveys (2,539 and 650 families respectively) in 1967-69,\footnote{74} and the district of Valencia in 1969.\footnote{75} These four publications include monthly income data for these families.

Very probably the data from these four surveys is much superior to that obtainable from the 1971 census, especially for the upper income groups. These surveys clearly should make possible the calculation of an urban distribution for 1967-68. For consistency checks on the incomes of the upper deciles, an attempt should be made to compare the survey results with the income tax data on salaries. Unfortunately, in Venezuela it would be very difficult to obtain income data from the tax returns of non-salaried workers. Another interesting source of income data is the public employee salary tables published periodically by the government. It appears that a distribution of public employee incomes was published in the early sixties (I believe for 1961) and there is a new set of tables for 1973.\footnote{76}

There is another good statistical source for urban income data. The Banco Nacional de Ahorro y Prestamo recently carried out a survey in all cities of more than 10,000 inhabitants which includes data on family income by size of family.\footnote{77} This data can obviously be used to check urban income data from other sources and to identify the housing situation of families in different income levels.
Concerning sectoral income data, Venezuela appears to have rather good national income statistics. The first national income series goes from 1957 to 1969, and the new system of national accounts covers 1968-1973. The functional income distribution data in the national accounts unfortunately has the same problem found in most developing countries, which is that the returns of family businesses are included in profits, making it difficult to separate out the labor income of farmers and independent workers.

Various sources can also be used to obtain historic series on wages, and Paez published a heroic effort at estimating wage trends between 1891 and 1971.

Finally, some interesting information on agricultural incomes can be obtained from the Censo Agropecuario of 1971 and from CIDAOEA publications. Parts of the Censo Agropecuario have been published, and there is information available on types of production by farms of different sizes, production consumed on the farm, and land use according to size of the production unit. All this data may be used to obtain a farm income distribution which could complement the urban distribution derived from the household surveys.

**Interesting subjects for future research**

As can be gathered from the above description, very little work has been done in the area of income distribution in Venezuela, but there is some primary data which can probably be utilized for such research. The first priority is clearly to make two income distribution estimates, one for the sixties and one for the seventies with the existing data.
The second most obvious area for research is that of tax incidence and the incidence of public expenditure. Due to oil revenues the Venezuelan government has had large revenues, but it is not clear that these funds have been spent for the benefit of the lower deciles of the distribution. The relative ease with which the government has obtained oil revenues may also have led historically to a weak tax system which makes no attempt at all to redistribute income. The general problem of the role that the government can play to improve income redistribution in a country with an easily taxable natural resource is well worth investigating.

Another general problem which it would be most interesting to study is the impact on labor demand of an exchange rate valued highly due to the impact on the balance of payments of oil exports. Such an exchange rate encourages capital intensive industry, and maybe a low labor share. To correct this, import substitution policies are followed, and these may in turn concentrate income through high profits of protected industries and low demand for labor due to the adoption of labor saving technology. Furthermore, the high prices of protected industry may harm, in a particular manner, low income consumers.

Other studies could include the effect on income distribution of the educational and agricultural policies followed in the last decade, and the impact on labor income of the large migration of Colombians toward Venezuela.

Comparative Studies - Colombia, Mexico, Venezuela

Finally, there are a series of research projects which can most fruitfully be carried out by comparing the impact of certain policies on
the income distribution of different countries. For example, it would be very illustrative to study the impact of financial policy on the distributions of Mexico and Colombia.

Mexico is characterized by an open capital market, with large inflow of capital. Colombia, on the other hand, has for many decades had exchange control and limitations on the free movement of capital. This has resulted in a relatively realistic interest rate in Mexico, while Colombia has had controlled and low interest rates. Thus the foreign exchange policies of the two countries have helped to shape the structure of the capital and money markets, and these are quite different. It would be interesting to try to see which capital market has a less inequitable impact on income distribution. The recent theories of McKinnon and Shaw would suggest that the Mexican model is superior, but the case has not been proven. The comparison of the Colombian and Mexican experiences would therefore be of interest.

A related subject would be to study the distributional impact of export promotion policies in Colombia and Mexico. Contrary to the McKinnon, Ranis and Little, Scitovsky theories, the Colombian and Mexican experiences may show that certain export promotion efforts tend to concentrate income. This is worth studying in a comparative framework. For similar reasons, it would be of interest to analyze the impact on distribution of the different brands of import substitution policies adopted by Mexico, Colombia and Venezuela in the last decades.

Finally, the three countries have followed diverse agricultural policies. Mexico has had land reform in a first phase with large investments in irrigation and new land in a second phase. Colombia has had
little land reform, but a large effort to channel credit to agriculture. The Venezuelan farm policy also has some unique features. But in all three countries the farmers are the poorest group and the sector which has benefited least from development. It is crucial to know why this is so.

With respect to comparative statistical work it may also be possible to produce income distribution data according to the Webb model, in which the modern and traditional sectors are separated out. Non-agriculture data can be found in all three countries, for the years in which household surveys were carried out. Data on government salaries, although sporadic in the case of Venezuela, can be obtained in all three countries. Wage data by size of firm is available in Colombia and Mexico, but may not be easily available in Venezuela (this would have to be investigated). For the rural sector, it would probably be possible to obtain information on net output by size of production units. In Venezuela and Mexico the agricultural censuses would be used, and in the case of Colombia the Banco de la Republica and DANE surveys of 1973, as well as the 1960 agricultural census would be used. In summary, in Colombia and Mexico it would be possible to use the modern-traditional type of analysis used by Webb, while it is not clear whether in Venezuela there would be sufficiently detailed data for analysis of this type. In Colombia such analysis could be carried out for the period 1964-1973 and in Mexico probably for the period 1963-68.
1. Banco Internacional de Reconstrucción y Fomento, Programa de Fomento para Colombia, (Bogota: Banco de la República, 1951).


7. Albert Berry, and Miguel Urrutia, Income Distribution in Colombia, (mimeo, 1974).


12. CEDE, Estudio de Empleo, op. cit.


23. Copies of the main tabulations of the CEDE's unemployment surveys can be obtained from that institution. The tabulations of the DANE, "Encuesta de Hogares" are also available on request from DANE.


27. Data on income tax are available at the Ministry of Finance, Oficina Asesoría Económica. The 1965 tabulations were carried out by C. Taylor who later published some results, (mimeo, 1968).


31. See, for example, Instituto Colombiano de Bienestas Familiar, Hojas de Alimentos Colombianos (Bogota: 1967-70; 1971; 1972); and, La Producción Agropecuaria y las Necesidades Alimentarias de la Población Colombiana (Bogota, 1973).


35. Banco de México, Cuenta Nacionales, op. cit.

36. Ifigenia M. de Navarrete, La Distribución del Ingreso y el Desarrollo Económico de México, (México: Instituto de Investigaciones Económicas, 1960).


41. Ibid.


45. Ifigenia M. de Navarrete, "La Distribución del Ingreso en México," op. cit.


48. Leopoldo Solís, La Realidad Económica, op. cit.

49. Ifigenia M. de Navarrete, La Distribución del Ingreso y el Desarrollo, and "La Distribución del Ingreso en México," op. cit.

50. Comisión Nacional de los Salarios Mínimos, Memoria de los Trabajos, op. cit., Vol. II.

51. Banco de México, Cuentas Nacionales, op. cit.

52. Banco de México, Encuesta sobre Ingresos y Gastos, op. cit.

53. Berry and Padilla, La Distribución de Ingresos, op. cit.


55. Institute for Workers' Housing.

56. Comisión Nacional de los Salarios Mínimos, Memoria de los Trabajos, op. cit.

57. According to the I.L.O. definition, the "informal sector" includes small scale industries, self-employees, skilled craftsmen, service workers, etc.


62. Leopoldo Solís, La Realidad Económica, op. cit.


64. The results appear in Julio Páez Celis, Ensayo sobre Demografía Económica de Venezuela, (Caracas: Ministerio de Fomento, Dirección General de Estadística y Censos Nacionales, 1974). They are based


68. The ECIEL group is one concerned with Latin American economic integration. Latin American research institutes and the Brookings Institution collaborate on studies related to the problems of and prospects for economic integration.


70. Centro de Estudios de Desarrollo (CENDE), Estilos de Desarrollo, (Caracas: Universidad Central de Venezuela, 1971), Vol. III.

71. Ministerio de Pomento de Venezuela, Dirección General de Estadísticas, Encuesta de Hogares por Muestra, op. cit.


73. Banco Central de Venezuela, Estudio sobre Presupuestos Familiares y Índices de Costo de Vida para el Área Puerto La Cruz-Barcelona, (Caracas, 1971).

74. Banco Central de Venezuela, and Universidad del Zulia, Estudio sobre Presupuestos Familiares y Índices de Costo de Vida para el Área Metropolitana de Maracaibo, (Caracas, Maracaibo, 1972).


76. Unpublished material available upon request at the Banco Central de Venezuela, Caracas.
77. Unpublished tabulations.

78. This new system of national accounts is drawn up by Banco Central de Venezuela.


84. For Colombia, see DANE, *Anuario General de Estadística* (various years). For Mexico, see Comisión Nacional de los Salarios Mínimos, *Memoria de los Trabajos*, op. cit.

85. For Venezuela, see Dirección General de Estadísticas, *IV Censo Agropecuario*, op. cit.; for Mexico, see Secretaría de Industria y Comercio, *IV and V Censo Agrícola-Ganadero y Ejidal*, op. cit.

86. DANE, *Directorio Nacional de Explotaciones Agropecuarias*, (Censo Agropecuario, 1960); *Resumen Nacional*, (Two Parts); *Resultados por Departamentos*, (18 Volumes).