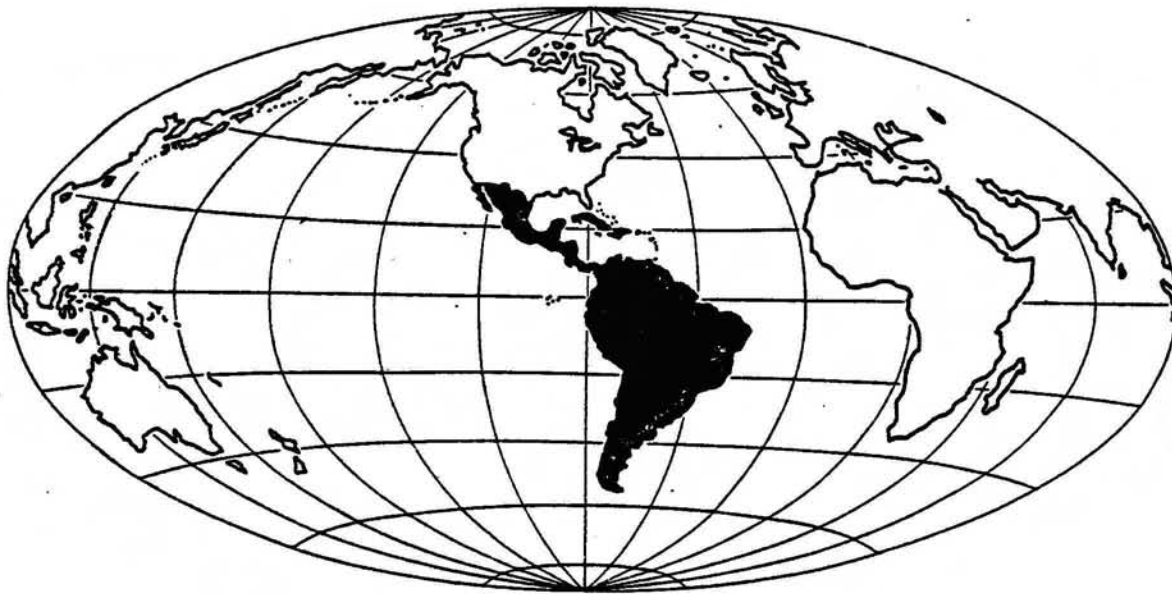


**STATISTICAL  
ABSTRACT  
OF  
LATIN  
AMERICA  
Volume 19**



EDITED BY  
**JAMES W. WILKIE**  
Co-editor **PETER REICH**

**UCLA Latin American Center Publications**

UNIVERSITY OF CALIFORNIA • LOS ANGELES • 1978

## On the Accuracy of Statistics and Development of Time-Series Data

How accurate are the data in a volume on Latin American statistics? Are they "poetry" as the Peruvian Luis Alberto Sánchez would have us believe?<sup>1</sup> Or can we accept the view of Alejandro Portes of the United States that Latin American data do not seem to "be any worse than comparable material collected in the United States"?<sup>2</sup> There is no simple answer to these questions but we can consider the nature of data.

On the one hand, it can be argued convincingly that all statistical data are erroneous. Researchers have not yet developed the capacity to measure economic, social, and political phenomena with much accuracy because we do not yet know the true dimensions of what we seek to gauge. This problem is complicated, moreover, by what may be termed the factor of fickleness: what we sought to measure yesterday is not always of the same interest today because new problems require new data and because criteria and methods of measurement may change suddenly.

Social goals of the 1961 Alliance for Progress, for example, originally were to be measured according to the ability to achieve dramatic growth in number of schools and teachers and hospitals and doctors. Economic goals were to be tested by the ability of Latin America to achieve a gain in per capita GNP of 2.5 percent per year. Since the method for meeting these goals was to involve the amount of \$20 billion in U.S. aid over a ten-year period, presumably this figure provided still another test for the Alliance. True to form, however, not only were the original measures soon discarded by critics of the Alliance even as they were being met,<sup>3</sup> but also the definition of what constituted U.S. aid was revised to exclude funds helping to resolve Latin America's balance of payments crisis, and a new hastily devised goal of income redistribution was imposed. (Immense problems in even beginning to measure income distribution are indicated in table 1406, which samples statistics on variations in existing data.) Thus the Alliance could not succeed even if it did attain many of its own goals. The ultimate goal of wiping out poverty in ten years was so

unrealistic, of course, that the Alliance was doomed from the start, just as was the midcourse revised goal imposed by the critics of redistributing income in less than ten years.

On the other hand, it can be argued with equal conviction that only naïve readers expect data to be fully accurate. Data only reveal the concerns of what a generation deems important to measure. Also, given improvement in methods as researchers become more sophisticated, revised ideas about the dimensions of what is to be measured require changes in definition. Old data series may not necessarily be discarded but "spliced" to the new data series with a note that data are not fully comparable before and after linkage. It is only within such stated limitations (often overlooked) that the data are valid.

A government often uses statistics to deliberately mislead readers for propaganda purposes, yet technically it remains "honest." For example, it does not fully disclose information that only highly sophisticated readers could be expected to know, as when it publishes percentages of the budget projected to be spent on popular educational programs without noting that figures of actual expenditures also exist, and that the difference between the two sets of figures may be as much as 50 percent.<sup>4</sup> Or a government may give an incomplete definition of change in series so that the unsophisticated reader will be unaware that data are not comparable before and after the linkage of differently defined concepts. It has been traditional for the latest Mexican presidency, for example, to show itself off favorably by publishing its plans for the future (with inflated figures) to distribute land to the masses and give actual data of land reform implementation for previous presidents (thus low figures).<sup>5</sup>

These data problems are solved here when possible by presenting alternative data with expanded definitions so that complexity in changes over time may be better understood. Thus in this volume we offer some differing estimates of the same phenomena, such as infant mortality (tables 708 and 709) and time-series population change (tables 600, 601-620, 622, and 623). Also, for example, Marshall C. Eakin

<sup>1</sup> Quoted by Luigi Einaudi, "The Statistics of Political Institutions," in Kenneth Ruddle and Philip Gillette, eds., *Latin American Political Statistics*, Statistical Abstract of Latin America Supplement 2 (Los Angeles: UCLA Latin American Center Publications, 1972), p. 3.

<sup>2</sup> Alejandro Portes, "Sociology and the Case of Secondary Data," in Robert S. Byars and Joseph L. Love, eds., *Quantitative Social Science Research on Latin America* (Urbana: University of Illinois Press, 1973), p. 214. The Byars and Love volume is reviewed here in chapter 37, below.

<sup>3</sup> See James W. Wilkie "The Alliance for Progress and Latin American Development," in *SNP*, pp. 407-431.

<sup>4</sup> Compare presentation of data on educational budgets in México, NAFIN. *Statistics on the Mexican Economy [1977]* (México, D.F.: 1977), p. 415; and James W. Wilkie, *The Mexican Revolution: Federal Expenditure and Social Change Since 1910* (2d ed.; Berkeley: University of California Press, 1970), pp. 160-161.

<sup>5</sup> For discussion of data on land reform, see James W. Wilkie, *La Revolución Mexicana, 1910-76: Gasto Federal y Cambio Social* (México, D.F.: Fondo de Cultura Económica, forthcoming), pp. 221-226.

analyzes problems of defining and giving alternate data for estimating the size of the largest city in each Latin American country since 1900 (chap. 35).

To show how alternative interpretations can be developed from one set of data, Daniel I. Geffner counterposes a "dependency model" of analysis against a "free-trade model" of explanation. Utilizing figures compiled to reveal patterns in movie releases, for example, he analyzes the development of the Latin American film industry (chap. 36).

To illustrate how data may be tested for internal logic, Peter Reich has compiled time-series figures on the U.S. Department of State yearly assessment of the number of Communist party members in Latin America (chap. 34). When the data are assembled over time, the resulting series is so illogical as to suggest strongly that the figures were faked by officials who failed to realize that figures would one day be compared on a year-to-year basis. Yet, from the data one could also contend, less convincingly, that the series is correct in relation to external events, communists going underground and surfacing depending upon the political climate. Or one could argue from the same data that the Communist parties are weak groups in most countries, that they only look strong at opportune times because of the "bandwagon effect" which does not last during inopportune times.

Summarizing the discussion so far, then, we can say that specific data for one year can indeed be quite erroneous. Research into definitions for compilation and discussion of limitations helps immensely to understand data, but also to be recognized is that reliability of figures for any one year is not as important as the statistical trend over time. It is difficult for officials to consistently fudge government data for many years owing to the fact that officials with different interests come into power and anomalies show up in statistics. And if strange patterns in the data appear to be legitimate, it suggests the need for research into those years or periods. The method of identifying epochs for research through patterns stands in contrast to that of some social scientists who set out to study a historical period without knowing how it fits into data trends.<sup>6</sup>

In order to offer increasingly useful data on Latin America, beginning with Volume 17 time series are presented whenever feasible. Series are staggered by volume so that each one contains research on some new areas. It is our goal to give time series systematically according to resources and travel time available. During the period 1975-1977 I visited the following statistical agencies in Latin America:

- Departamento Administrativo Nacional de Estadística, Bogotá
- Dirección de Estadística, Panama City
- Dirección General de Estadística, Guatemala City
- Dirección General de Estadística, Mexico City

- Dirección General de Estadística y Censos, Montevideo
- Dirección General de Estadística y Censos, San José de Costa Rica
- Dirección General de Estadística y Censos, San Salvador
- Dirección General de Estadística y Censos Nacionales, Caracas
- Institute Haitien de Statistique, Port-au-Prince
- Instituto Nacional de Estadística, La Paz
- Instituto Nacional de Estadística, Quito
- Instituto Nacional de Estadística, Santiago de Chile
- Instituto Nacional de Estadística y Censos, Buenos Aires
- Oficina Nacional de Estadística, Santo Domingo

At each agency it was possible to assess data holdings and investigate plans for publication of data. In Mexico, for example, it was determined that the time series on price changes dating back to 1889 is no longer being published because it reveals more inflation than the government wants to admit. Therefore, Mexican unpublished data through the end of the presidency of Luis Echeverría are printed here to complete the series for the years 1889-1976; it is compared to the alternative series now published since 1939 (table 2507).

Articles dealing with specific historical statistics for all countries of Latin America appear in the Statistical Abstract of Latin America series as follows:

Volume 17 (1976)

- Exchange Rate History, 1937-1974, by Bridget Reynolds
- Measuring the Scholarly Image of Latin American Democracy, 1945-1970, by Kenneth F. Johnson

Volume 18 (1977)

- Educational Enrollment History 1880-1929, by José Casimiro Ortal
- Religious Data History, by Peter Reich
- Labor's Real Wages in Latin America since 1940, by John L. Martin

Volume 19 (1978)

- Measuring U.S. Government Perception of the "Communist Menace" in Latin America, 1947-1976, by Peter L. Reich
- Determining the Population in the Largest City of Each Latin American Country, 1900-1970, by Marshall C. Eakin
- Alternative Interpretations of Time-Series Data on the Growth of the Latin American Film Industry, 1926-1970, by Daniel I. Geffner.

Analysis of time-series data also has been developed in two recent volumes of Statistical Abstract of Latin America Supplement series:

Volume 6 (1977)

- Employment and Lack of Employment in Mexico, 1900-1970, by Donald B. Keesing

<sup>6</sup>For an example identifying Mexican years of interest involving the study of election data, see James W. Wilkie, "New Hypotheses for Statistical Research in Recent Mexican History," in *SNP*, pp. 27-37.

Losers in Mexican Politics: A Comparative Study of Official Party Precandidates for Gubernatorial Elections, 1970-1975, by Roderic A. Camp

An Index of Cuban Industrial Output, 1930-1958, by Jorge F. Pérez-López

Research Perspectives on the Revised Fitzgibbon-Johnson Index of the Image of Political Democracy in Latin America, 1945-1975, by Kenneth F. Johnson

Projecting the HEC (Health, Education, and Communication) Index for Latin America Back to 1940, by James W. Wilkie and Maj-Britt Nilsson

Volume 7 (1977)

Democratic versus Dictatorial Budgeting: The Case of Cuba with Reference to Venezuela and Mexico, by Enrique A. Baloyra

Federal Expenditures and "Personalism" in the Mexican "Institutional" Revolution, by James A. Hanson

Financing Industrial Corporate Development in the Aftermath of the First Perón Period, by David K. Eiteman

Time-series data presented in this volume without interpretation implicitly raise a question with which it is necessary to deal. In developing table 2731 to show data at five-year intervals since 1955 for the two leading exports of each Latin American country as a share of total country FOB export values, what impression, for example, should the figures represent? If we want to know what the statistics meant to people during past years, we should give preliminary figures that influenced thinking at the time, as in the case of 1965 Chilean copper exports which appeared to have been about 60 percent of all exports. The revised percentage based upon Chile's recalculations in the 1970s shows the copper share to have been nearly 79 percent, the country being much more dependent upon that one export item for 1965 than it thought.<sup>7</sup> To complicate matters, the question of who knows about the data is also important. International reporting on the value of two leading exports can take time to catch up with revised national sources, as in the case of the International Monetary Fund which in 1976 listed coffee as Mexico's leading export (6.3 percent) for 1975 instead of petroleum which had superseded it (15.8 percent).<sup>8</sup>

In one of the longest yearly time series yet published in the *Statistical Abstract for Latin America*, data on demographic change since 1900 for each country given in tables 601 to 620 illustrate the problem of meshing estimates prepared in different epochs. These figures revise and update the series originally presented in *Statistics and National Policy* (pp. 173-185) but the caution there (pp. 137-139) still holds: given the

problems in estimates for the early 1900s, data for some countries are extremely erratic. Demographic change for Nicaragua (table 615) shows the following percentile growth: 1900-10, 29.3; 1910-20, 17.5; 1920-30, 7.1; 1930-40, 20.8; 1940-50, 28.5; 1950-60, 39.3; and 1960-70, 41.0 percent. Is the high growth between 1900 and 1910 probable, especially compared with the decreases of the 1910s and drastic drop of the 1920s?

And data for Guatemala in the early decades of this century, for example, vary more widely than for any other country in Latin America. Estimates for Guatemala in 1900 range from 885,000 to 1,627,000. If the figure accepted in "standard" sources (1,430,000 persons) is accepted as accurate,<sup>9</sup> then Guatemala population grew about 75 percent faster per year between 1850 and 1900 than between 1900 and 1935.<sup>10</sup> Clearly, population data need to be investigated in relation to (a) national conditions of politics, society, and economy of a given epoch and (b) international opinion pushing governments to overstate their populations at the beginning of the century (in order to make a country important in size) and to understate them by the 1960s (in order to show a country's ability to control its population growth).

Population data for 1970 also appear to be problematic for Cuba and Mexico. Figures for the former seem to involve an overcount because the voluntary exiles that fled the Cuban Revolution of 1959 have never been subtracted by the Cuban statistical agency. Figures for Mexico seem to involve an undercount because of the difficulty of conducting a census in the thousands of rural places with very few inhabitants.

Finally, countries with a large number of jungle Indians have not generally included them in the data unless they were near means of communication. As those uncounted Indians enter "national" social and economic life, the figures may show a percentage spurt in population, especially in the smaller countries.

The long series on political data adapted from the research of Tatu Vanhanen (who himself drew heavily upon *Statistical Abstract of Latin America Supplement 2*) gives estimates to show changing populations incorporated into "national" political life (chap. 33). The Vanhanen political data give the percentage of the electorate eligible to vote and the percentage of the total population that voted in presidential elections since the midnineteenth century. These electoral data, subject to manipulation by government as they are, are quite revealing when put into time-series format. Vanhanen's book from which the voting data are abstracted is one of the books that I review in "The Status of Quantitative Research on Latin America" (chap. 37).

Also of note is my time-series formulation of economic data on national accounts newly presented in chapter 22. Tables given present the most complete historical series

<sup>7</sup> Compare data in Chile, DGEC, *Boletín... Sinopsis 1966*, pp. 476, 552, and data in table 2731, below.

<sup>8</sup> Compare annual time-series data in IFS, May 1976, 1977, and 1978.

<sup>9</sup> The low figure rather than the middle or high estimate is accepted here in table 611 as being the most logical.

<sup>10</sup> See discussion in *SNP*, p. 138.



available, but this compilation was not developed without problems. National sources often differ from international sources that are adjusted to make figures comparable between nations. Also, the United Nations and its regional arm the Economic Commission for Latin America differ in their breakdowns of categories for analysis, and adjustments are made here so that the industrial category is defined as consistently involving manufacturing, mining, electricity, gas, and water types of activities. The Economic Commission for Latin America (ECLA) does not give an industrial total, distinguishing instead between a subtotal for goods — agriculture, mining, manufacturing, and construction — and a subtotal for basic services — including electricity, gas, and water. Where data are not comparably subdivided by category for some countries, differences are noted. Since no agency offers a full series for all countries since 1950, the series developed here draws upon various sources and must only be considered the best longitudinal estimates available.

Latest figures in change of GDP in constant terms are revised and brought up to date here in table 2203. Since U.S. AID source data are no longer published and U.N. source data are delayed until gross product figures for countries outside of Latin America are available, percentage change in GDP for 1974, 1975, and 1976 are taken from latest ECLA calculations. (ECLA GDP data are also used to update GNP data in table 2200, on the assumption that rates of change in GNP and GDP would be essentially the same — for comparison of GNP and GDP totals in national currency, see table 2202.) Because ECLA data are given only as totals, data in per capita terms are calculated from the population trends in table 623. Gross product figures going back to 1950 given in tables 200 and 201 are based on latest IMF historical revisions and on population data given in tables 601 to 621. Data on Cuban GNP given in tables 200, 201, and 2203 need special comment. Figures for years since 1959 should be used with certain caution because they are based here upon recalculation of Cuban and CIA data in order to convert gross social product to GNP; current pesos are converted to constant pesos using the U.S. export price index given in table 2506. Although Cuba does not trade directly with the United States, the cost of goods that it does purchase may be indirectly related to the cost of U.S. goods in the world, higher costs of trade through third countries being offset by some subsidies from the Soviet Union. With regard to pre-Castro economic change, data on GNP in table 2207 may be correlated with political activities as is suggested by the Cuban figure for 1953. In the year when GNP declined by 11.7 percent, a veritable depression, Fidel Castro launched his attack on the Moncada military barracks.

Perception of data is of crucial import in interpreting events in Latin America, witness views of Chilean data treated in constant prices. Preliminary figures for change in GDP for 1970 gave a figure (+2.5 percent) that helped elect Salvador Allende president.<sup>11</sup> But, after Allende took power, 1970

GNP data were revised upward to a figure (+9.2 percent) that could have changed the outcome of the election had it been known at the time.<sup>12</sup> Revised IMF data show, however, a different view of economic change in Chile's benchmark year of 1970 (+3.6 percent).<sup>13</sup> Finally, index data for GDP yield yet another figure for change (+7.0 percent). These varying data reveal the magnitude of the problem the reader faces in choosing between alternative evaluations of the rate of change in gross product.

To help understand outside funding to aid in Latin American economic growth, time-series data on World Bank (IBRD) loans are given in table 3009.<sup>14</sup> The World Bank has supplied unpublished data yearly back to inception of Latin American programs in 1949. It has formally published only its projected loans, and its finance officials had to dig into their records to find out what the Bank had actually expended. In conversation, officials noted that since these data do indeed represent the "bottom line" explaining what the Bank is about, the figures should have been released earlier, but since no one had ever asked for the data either from within or outside of the Bank, it had not occurred to officials to assemble the series for release.

Three related tables among other time-series data developed in this volume merit comment. First, statistics on Latin America's export position vis-à-vis the world is shown in table 2703 for five-year intervals since 1950. Latin America's importance in world export trade declined almost 60 percent between 1950 and 1975, dropping steadily at each interval. Cuba's decline in export importance was the most dramatic of all countries, falling by about 68 percent in the same period, 38 percent of which occurred between 1950 and 1955 in a collapse helping, no doubt, to trigger the Cuban Revolution of 1959.<sup>15</sup> In contrast, U.S. export decreased about 27 percent between 1950 and 1975, the decline coming mainly in the early 1970s. (Japan and West Germany gained in exports at the expense of Latin America and the United States, as suggested in table 2705.)

Second, some of the reasons for Latin America's export decline can be found in table 2206 which shows a general decline in agriculture's share of GDP throughout Latin America, agricultural exports previously having been much more crucial to the economic activity of most countries. In Brazil, Mexico, and Peru, for example, the agricultural share of GDP was more than halved between 1950 and 1976. Only in Argentina and Venezuela has the agricultural share not changed much but both were already low in 1950, 14 and 8 percent, respectively.

<sup>12</sup> See data in table 2203, below.

<sup>13</sup> See IFS, May 1977.

<sup>14</sup> Complementary data for U.S. Export-Import Bank actual loans are given in *SNP*, pp. 339-359, updated here in table 3105.

<sup>15</sup> Economic gains also may "cause" political problems as suggested by data in table 2205: the growth rate of manufacturing in the Dominican Republic and Nicaragua has constituted a kind of "economic miracle" and in both countries new or traditional dictators have faced political challenges in the 1970s.

<sup>11</sup> ECLA-S2, 1970, p. 187.

Third, decline in the agricultural share of GDP has meant that Mexico and Peru exports as a share of GDP fell between 1950 and 1975 from 11 to 4 and 18 to 9, respectively. But as table 2213 also shows, while for most other countries the export share of GDP remained somewhat constant, percentages increased significantly for Chile, Ecuador, El Salvador, and Nicaragua.

Given the varying patterns of twenty Latin American countries seen in the complex relationships of data presented in tables such as the ones dealing with exports, the reader can appreciate the problem of making sense out of what happens in Latin America as a whole. Definitions for time series lack complete consistency not only between countries but also within countries. For scholars accustomed to dealing with only one country such as the United States where time-series data are relatively well developed for interpretation, research in comparative Latin American statistics appears to be quite frustrating.

Recently Latin Americanist scholars have expressed concern about the region's conflicting data, noncomparability of figures, or nonreliability of statistics.<sup>16</sup> For example, remarking on wide variations in data that compare military expenditures from country to country in Latin America, Martin Needler almost laments the fact that alternative estimates exist to confuse the scholar.<sup>17</sup> But could investigation

into data yield a singular valid definition for the needs of all researchers? Do not alternative estimates help us to see dimensions in the problem of achieving "truth" in data?

In an era when the search for "information" is expected to lead to "truth," it is necessary to recall that truth is multi-dimensional and, the more information we gain, is becoming ever the more complex. Data as information cannot lead to truth but perhaps, like poetry, they can give insight into some of its dimensions. To answer a question posed at the opening of this preface, if data for the United States appear to be more reliable than for Latin America, we should note that the U.S. data may at times be more simplistic when one compares its scope with that of the Mexican population census. The Mexican census is intended to be administered to all of the population; the U.S. census questionnaires for 1960 and 1970 were intended only to sample the national population mainly by mail.<sup>18</sup> The former thus offers a much richer source for interpretation of social change; the latter fails to capture or even to identify "pockets" of variations from the "norm."

More important than "truth in data" and absolute accuracy is realizing that information is "valid" only insofar as our interpretations and perceptions of it show complexity in meaning. Statistical data, like all information, yield circumstantial truths, not absolute ones. The ultimate goal of the *Statistical Abstract of Latin America* and its Supplement Series is threefold: to offer a growing core of recent data, especially presenting alternative views where feasible; to develop time-series analyses showing limits and meanings of data; and to guide the reader to an increasing number of sources.

J.W.W.

Malibu, California  
June 1978

<sup>16</sup> See, for example, Philippe Schmitter, "New Strategies for the Comparative Analysis of Latin American Politics," *Latin American Research Review* 4:2 (1969), pp. 83-110; Martin C. Needler, "United States Government Figures on Latin American Military Expenditures," *Latin American Research Review* 8:2 (1973), pp. 101-103; and Thomas I. Dickson, "The Contribution of the Inter-American Development Bank to the Latin American Statistics Muddle," *Inter-American Economic Affairs* 28:3 (1974), pp. 79-85. On data problems for individual countries, for example, see my survey of the literature in *SNP*, pp. 16-22; and Laura Randall, ed., "Special Section on Historical Statistics," *Latin American Research Review* 13:2 (1978), pp. 69-221.

<sup>17</sup> See Needler, "United States Government Figures on Latin American Military Expenditures," p. 103.

<sup>18</sup> See James W. Wilkie, John C. Super, and Edna Monzón de Wilkie, "A Social Census Questionnaire for Latin American Countries," p. 438 in *Statistical Abstract of Latin America*, vol. 18 (Los Angeles: UCLA Latin American Center Publications, 1977).