

---

## FORUM

---

### APPLYING QUANTITATIVE TECHNIQUES TO QUANTITATIVE HISTORY: POVERTY AND FEDERAL EXPENDITURES IN MEXICO\*

*Felix G. Boni and Mitchell A. Seligson, University of Pittsburgh*

IN RECENT YEARS THERE HAS BEEN AN UPSURGE IN WHAT HAS COME TO BE CALLED "quantitative history."<sup>1</sup> Despite the enthusiasm for this approach, however, some of its critics have validly claimed that its payoffs have all too often failed to live up to its promises of new and more accurate findings. Perhaps a central reason for this criticism is that although the quantitative historian may have been sufficiently thorough in collecting his data, he has often failed to apply to them the sensitive techniques of modern data analysis. Rather, he has continued to rely on more traditional methods of description such as means and percentages. As a result, quantitative works are often long on tables and short on analysis.

One of the best examples of a work reflecting both the virtues and shortcomings of the quantitative school is James W. Wilkie's (1970) extensive investigation of the relationship between federal expenditures and social change in Mexico. This work, owing to its pioneering nature, has attracted wide interest. Our purpose in this note is to illustrate how the use of factor analysis can help overcome some of these shortcomings. In particular, we hope to demonstrate the value of factor analysis in the evaluation of indices such as Wilkie's Poverty Level, by examining (1) his measurement of poverty in the Mexican states and (2) the possible relationships between this measure and federal expenditures.

#### MEASURING POVERTY

Wilkie sets for himself the task of measuring poverty over a fifty year period in

\* Grateful acknowledgment is given to Professors Paul A. Beck, Raymond Corrado, James M. Malloy, Michael Margolis, Reid R. Reading, and Harold D. Sims of the University of Pittsburgh, and to Professor Peter H. Smith of the University of Wisconsin for their helpful comments on an earlier, unpublished version of this note.

*Latin American Research Review*

each of the 32 federal entities (i.e., 29 states, 2 territories, and the Federal District). He does this by averaging the values for each entity of the following seven variables (in percentages): illiteracy; population speaking only an Indian language; persons living in communities of fewer than 2,500 people; barefoot population; population regularly wearing sandals instead of shoes; population regularly eating tortillas instead of wheat bread; and population without sewage disposal. A major problem with this measure is that it implicitly assumes that all seven items are equally a part of the same unidimensional phenomenon, namely poverty. As Skidmore and Smith (1970: 78) query: ". . . one might ask whether the first three [items] do not differ radically from the other four in that they restrict an individual's mobility and therefore his 'life chances' in a much more fundamental way." Through the use of factor analysis we have attempted to examine the relevance of this criticism.

The major advantage of factor analysis lies in its ability to reduce a large number of variables into one or more distinct dimensions.<sup>2</sup> It does this by delimiting a subset of a group of variables which are relatively strongly correlated with each other and relatively weakly correlated with the remaining variables in the group. The greater the extent to which the variables of the subset are intercorrelated, the greater the degree to which they can be said to measure the same phenomenon. In a similar manner, the lesser the degree to which these variables are correlated with other variables, the greater the extent to which this subset measures a dimension distinct from what the remaining variables measure. This dimension is given mathematical specification through the creation of a "factor," which essentially is that "space" in which the variables of the subset share a common variance. The factor loading, then, is the extent to which a given variable shares in that common variance.

The findings presented in Table 1 reveal that two of the seven items used by Wilkie to measure poverty for 1960 have little relationship to the dimension measured

TABLE 1  
*Factor Loadings of Poverty in Mexico (1960)\**  
Factor Solution #

| Variable                                     | Factor 1 | Factor 2 |
|--|----------|----------|
| % Eating Tortillas instead of Wheat Bread    | .857     | .036     |
| % Regularly Wearing Sandals instead of Shoes | .811     | .084     |
| % Rural Population (smaller than 2,500)      | .751     | .531     |
| % Illiterate                                 | .681     | .612     |
| % Without Sewage Disposal                    | .659     | .520     |
| % Barefooted                                 | .134     | .843     |
| % Speaking only an Indian Language           | .109     | .854     |
| Explained variance                           | 58.8%    | 16.4%    |

\* Eigenvalue = 1

# Varimax orthogonally rotated factors

Data Source: Wilkie (1970: 207-230).

#### POVERTY AND FEDERAL EXPENDITURES

by the remaining five.<sup>3</sup> This is demonstrated by the fact that the correlations between the first factor and the individual variables (i.e., the factor loadings) are quite high for the first five (.857 to .659) but extremely low for the remaining two items (.134 and .109). More specifically, the variables of "barefoot population" and "population speaking only an Indian language" lie on a dimension different from that on which the other five variables lie, thus creating a situation in which the seven variables form two factors or dimensions.

The multidimensional nature of the variables used by Wilkie leads to the problem of the determining nature of these two dimensions. We would suggest that the first is "poverty" while the second, because of the high loading of the "speaking only an Indian language" variable, is what could crudely be called "Indianness." Given the extreme poverty of this indigenous population it is not surprising that the "barefoot population" variable (which is probably indicative of a much more rudimentary lifestyle than are the other variables) also loads quite highly on the second factor. It would seem therefore, that the basic thrust of the Skidmore and Smith criticism is substantiated: some of the variables comprising the index are indeed different from the others. What factor analysis has done is to specify precisely which variables those are.

In an ideal factor analysis solution, each variable should demonstrate high (i.e., near unity) on one factor and low (i.e., near zero) loadings on all other factors. As can be seen in Table 1, three of the seven variables (% rural, % illiterate, and % without sewage disposal) have distributed loadings. That is, they load on both factors 1 and 2. This means that in addition to being indicators of poverty, the three variables are to some extent also related to the notion of Indianness. The connection is understandable when one realizes that an important distinction between the two factors is that while the indigenous population in Mexico lives in poverty, only a fraction of those living in poverty are Indians. Thus, although the final two variables loading most highly on the Indianness factor are not highly correlated with the poverty dimension (most poor Mexicans are not Indians), some aspects of poverty are related to Indianness (most Indians are poor). However, because these variables are more highly correlated with the dimensions of poverty than with the Indianness dimension we shall consider them as part of the former.

#### FEDERAL EXPENDITURES

The second shortcoming of Wilkie's discussion of poverty is his failure to provide the reader with a definitive analysis of the relationship between it and the major variable of his study, namely federal expenditures. He does, however, suggest the existence of such a relationship: "Federal investments were largely directed to entities which do not lie within regions of high poverty." (Wilkie, 1970:247). In effect, Wilkie suggests a strong negative correlation between federal expenditures and the degree of poverty in the states.<sup>4</sup> This is indeed a fairly straightforward proposition, and regression analysis, a relatively simple statistical technique, is used to test it.

*Latin American Research Review*

Skidmore and Smith (1970:80) employ a product-moment correlation to measure the relationship between declines in poverty (as measured by Wilkie's Poverty Index) and federal expenditures. Our present interest lies in a similar but somewhat different relationship. What we have examined is not the extent to which reductions in poverty may be attributed to federal expenditures, but rather whether the federal government, in the disbursement of funds, discriminates against the poorer states.

In this part of the analysis we compare the relationships of the distribution of federal expenditures with both poverty and Indianness (the former composed of the first five variables in Table 1 and the latter, of the remaining two). In this way we attempt to determine whether the distinction which we have made on the basis of factor analysis has any significance with regard to the allocations of federal expenditures. In constructing measures for the poverty and Indianness dimensions we use the method employed by Wilkie, that is, averaging separately the values of the variables comprising each of the two dimensions for each federal entity.

In measuring federal expenditures Wilkie uses total federal capital investment for five years of the López Mateos term of office, 1959-63. He uses it in two different ways, however. First, he calculates the difference between the percent of this investment received by a state and the state's population as a percent of Mexico's total population. Second, he subtracts from the percent of federal investment received by a state that state's "poverty characteristics" as a percent of the country's total poverty characteristics.<sup>5</sup> The former variable measures the extent to which a state is getting its "fair" share of federal investment in terms of its share of the country's population, while the latter measures it in terms of the state's poverty population.

The results from this analysis are presented in Table 2. The most striking finding is the difference in the correlation coefficients when the poverty scale as opposed to the Indianness scale is correlated with Wilkie's two measures of federal expenditures. With the former the correlations are  $-.48$  and  $-.62$ , while with the latter

TABLE 2  
*Poverty, Indianness, and Federal Expenditures*

| Federal Expenditures#                        | Five Variable Poverty Scale | Two Variable Indianness Scale |
|--|-----------------------------|-------------------------------|
| Investment—Population Difference             | $-.48^*$                    | $-.16ns.$                     |
| Investment—Poverty Characteristic Difference | $-.62^*$                    | $-.28ns.$                     |

Correlations are Pearson  $r$ 's

\* statistically significant at .05 level or less.

ns.—not statistically significant

(for justification of use of test of significance, see note 6)

# Excludes expenditures by decentralized agencies and mixed public and private enterprises. Investment data is from Wilkie (1970: 247-248).

#### POVERTY AND FEDERAL EXPENDITURES

they are only  $-.16$  and  $-.28$ , respectively. The significance of this finding is clear. The two dimensions delimited by factor analysis refer to real differences in the analysis of federal expenditures. The evidence shows that although the correlations are in the predicted negative direction—which is to say that as the poverty or Indian-ness of a state increases, the amount of federal investment, as operationalized by Wilkie, decreases—only the relationships with the poverty variable are strong and statistically significant.<sup>6</sup> The size of a state's indigenous population has little relevance to the distribution of federal expenditures, while the extent of its poverty does.

In conclusion it should be pointed out that the criticism leveled against Wilkie's work does not negate its value for the study of Mexican history. Wilkie's contribution is that he has collected large amounts of useful data. The task of quantitative history, however, is not completed at the data collection stage. This is only the first step. The data, which often come in large unwieldy proportions, must then be reduced into more manageable forms so that they can be analyzed in an effective manner. In this note we have attempted to demonstrate how factor analysis can be a valuable tool in the process of data reduction and how, as a consequence, its use enables the researcher to construct concepts, test hypotheses, and reach conclusions that otherwise might pass unobserved.

#### NOTES

1. For example, see the pages of the *Historical Methods Newsletter* and Rowney and Graham (1969).
2. For an excellent basic introduction to the use and interpretation of factor analysis see Rummel (1967). For a more thorough work see Rummel (1970).
3. A principal components factor solution with a varimax orthogonal rotation was employed. The eigenvalue was set at 1.0. The computation was completed using the Statistical Package for the Social Science programs (Nie, Bent, and Hull, 1970) as implemented on the University of Pittsburg's IBM 360 and the Digital Equipment Corporation's PDP 10 system. A factor analysis of the same seven variables for the years 1940 and 1950 reveals the same basic structure. A principal components analysis of socio-economic development has been conducted by Luis Unikel and Edmundo Victoria (1970), using twelve variables, only one of which (% illiterate) corresponds to variables used by Wilkie in constructing his Poverty Level. Their interest, however, is different from ours in that they are not directly concerned with the dimensionality problem. Note, however, their comments on p. 316.
4. The question of cause and effect is a complex one in this case. We do not know if low federal expenditures are responsible for poverty in the states or whether the poor states, having less political influence, are unable to win greater federal support and therefore remain poor. Wilkie apparently accepts the former view since he goes under the assumption, "... that government policy creates the climate in which the private sector will operate . . ." (1970:259). This, of course, does not preclude the former view since the motivation for differential levels of federal spending in the states is not at all clear. The real question is whether federal expenditure policies are correlated with decreases in the poverty index over time. Wilkie demonstrates (1970:237-243) that poverty has declined more rapidly in some areas than in others. Because of the lack of spending data, however, the effect of spending on these changes cannot be accurately determined. For a further discussion of this problem of causality see Skidmore and Smith (1970:79-83).

### *Latin American Research Review*

5. A poverty characteristic, essentially, is a poor person. If, for example, a state's poverty level is 30 per cent then its number of poverty characteristics is 30 per cent of its population (Wilkie, 1970:251).
6. Since the question of statistical significance enters this study in a way different from its normal use, an explanation is needed here. Typically, measures of statistical significance are employed when one is attempting to draw inferences from a sample about a population. In this study all of the Mexican entities are included and therefore the data cover the entire population. However, one must also face the fact that any fairly large list of correlation coefficients, even one based on purely random data, is likely to contain a few moderately high correlations. One would not want to infer from these correlations that they are necessarily theoretically significant. For this reason, it is best to treat all of those correlations which might have occurred more than five times out of one hundred (i.e., the .05 level of significance) as theoretically insignificant. For this point see Dye (1966:38-46), who deals with the problem of statistical significance when using data on all fifty U.S. states. See also Morrison and Henkel (1970).

#### REFERENCES

- DYE, THOMAS  
1966 *Politics, Economics and the Public*. Chicago.
- MORRISON, DENTON E., and RAMON E. HENKEL, eds.  
1970 *The Significance Test Controversy: A Reader*. Chicago.
- NIE, NORMAN, DALE H. BENT, and C. HADLAI HULL  
1970 *SPSS: Statistical Package for the Social Sciences*. New York.
- ROWNEY, DON KARL, and JAMES Q. GRAHAM  
1969 *Quantitative History: Selected Readings in the Quantitative Analysis of Historical Data*. Homewood, Illinois.
- RUMMEL, RUDOLPH  
1967 "Understanding Factor Analysis." *Journal of Conflict Resolution*. 11: 444-480.  
1970 *Applied Factor Analysis*. Evanston, Illinois.
- SKIDMORE, THOMAS E., and PETER H. SMITH  
1970 "Notes on Quantitative History: Federal Expenditure and Social Change in Mexico since 1910." *Latin American Research Review*. 5: 1: 71-85.
- UNIKEL, LUIS, and EDMUNDO VICTORIA  
1970 "Medición de algunos aspectos del desarrollo socioeconómico de las entidades federativas de México, 1940-1960." *Demografía y Economía*. 4:292-361.
- WILKIE, JAMES W.  
1970 *The Mexican Revolution: Federal Expenditure and Social Change since 1910*. Berkeley.