

Managing not-so-small Numbers *Between Comparative and Statistical Methods*¹

GAYL D. NESS

University of Michigan, Ann Arbor, U.S.A.

ABSTRACT

The problem here is what to do with an N of 20? Comparative case studies effectively deal with up to five or so observations. Aggregate statistical studies can easily with hundreds and thousands of observations. But with an N of 20 is too large for detailed case comparisons and too small for the use of powerful statistical analyses. This article proposes a middle ground that weaves an argument from a combination of multiple regression equations and case histories. Multivariate outliers identify cases for historical analyses; and the exposure of data in bivariate scattergrams permits useful validity testing. The procedure is illustrated with analyses of modern Asian population policy changes.

I. The Issue

NOT-SO-SMALL NUMBERS stand in an ambiguous position at the intersection of a number of sociological approaches. We have little trouble understanding a comparative study of two to six countries. (Tocqueville, 1863, Bendix, 1956; Moore, 1966; Skocpol, 1979) Nor does a study of five families (Lewis, 1959), 15 suicides (Douglas, 1967), a handful of inmates in a mental hospital (Goffman, 1961) or a street corner gang (Whyte, 1943) seem unfamiliar or out of place. At the other extreme we are also in familiar territory as we approach political behavior through voting studies of 1500 respondents or reproductive behavior through surveys of 6,000.

Between these extremes, however, we meet a number of intractable problems. What do we do with an N of 20 or 40? With small numbers we can make reasonable interpretations through our capacity to grasp a great deal of historical or individual detail. For large numbers we can permit assumptions of randomization to wash out much of the noise of detail and to make statistical inferences either about system characteristics or variable relations. It is the not-so-small numbers that give us trouble. With 20 to 40 observations we have too many cases to permit our entry into the rich detail of the quality of behavior, and too few for the use of our powerful, number-hungry new analytical technology.

II. Comparative and Statistical Methods

The problems of not-so-small numbers are found in both micro-level “qualitative” research and in macro-level cross national research, but it is in the latter that the problems may be said to be particularly acute today. Theories of modernization made cross-national research highly fashionable three decades ago, and the emergence of conflicting theories in Dependency and World Systems perspectives have kept their popularity very much alive. But the popularity of cross-national research derives from forces more mundane and possibly more insidious than those expressed in theoretical developments.

Much of the popularity derives, I believe, from the combination of three salutary methodological developments. One is the rapid growth of data collection by national governments and international agencies. The resulting archives of “data” (or better, numbers) cry out for manipulation simply because, like mountains, they are there. Second is the creation of an electronic technology for manipulating numbers. Third is the parallel development of a statistical technology that provides a logic for the electronic technology.

The convergence of these three developments makes it possible to spin out theories about societal dynamics and to test their propositions with great ease. Does urbanization-industrialization lead to a reduction of fertility, to produce a massive social change in reproductive behavior? Do different political systems lead to different rates of economic growth? Does political-economic “dependency” retard economic growth and equality, to hold together a world economic system in understandable and predictable ways? These are heady questions, and the capacity to address them with answers to three decimal points provides both delight and comfort, especially to academics whose rewards in life come from publication.

The delight and comfort are hardly deserved, of course, regardless of the material benefits that accrue. Cross-national statistical comparison presents a wide array of beguiling problems. Ragin points out three of central importance. (1981) First is the question of the unit of analysis and observation. Because nation states collect data, we typically use them as they are and assume that we are dealing with societies, or with systems whose basic parameters we understand. These assumptions are rarely critically examined and are often simply ignored even if recognized. Second is the issue of validity. We typically use indirect measures for highly complex patterns of behavior: per capita gross domestic product for example, is used to indicate the wealth or productivity of a society. Voting patterns, the existence of an opposition, or the manner of selecting an executive, come to stand for some vaguely defined quality of the political system. Foreign investment, trade or aid indicate degrees of dependency, or a “nation’s” position in the international division of labor. In the urge to generate numbers for our hungry analytical technology, we seldom pause to question the validity of these indicators.²

The third problem Ragin identifies goes to the heart of the problem of not-so-small numbers. The number of cases we have for cross-national analysis is

small even in the best of circumstances. There are today only about 150 nation states that can be used as units of observation, and for many of these, problems of missing data abound.³ Further, the greater the theoretical specificity we generate, or the longer the time span we use, the smaller is the number of cases we usually have. We might start with 100 cases, but soon find ourselves dealing with less than 50—too many for close comparative case analysis and too few for effective statistical analysis.

Ragin's solution to this problem is not of much help. In effect, he asserts that we must distinguish between two methods. The *comparative* method is limited to the logical examination of similarities and differences between whole cases, necessarily limited to perhaps less than ten. It differs from the *statistical* method, which "...requires the investigator to focus only on *relationships* between variables, ... and not on the differences and similarities between cases." (1981, p 113, emphasis added) Comparative and statistical methods are qualitatively different and suited to different types of problems.

Let me propose an alternative solution that is both eclectic and modest. It permits us to exploit both methods together, gaining the detail of the case analysis and the power of the statistical method. Its rhetoric consists, first, in weaving an argument with regression coefficients and the detail of case histories, and second, in exposing the argument with a liberal use of bivariate scattergrams. The former addresses the necessity of theory construction and hypothesis testing. The latter addresses the issue of validity.

The steps are simple and not at all uncommon. First is the definition of a problem. This both implies and requires the elaboration of a theoretical framework that specifies the units of observation and analysis and provides a general statement of dynamic conditions, or variable relations in those units. This is followed by the specification of variables and their operational definitions. Indicators are obtained for the variables and the general model is tested. In the process the cases are exposed in simple bivariate scattergrams, which permit knowledgeable observers of the individual cases to validate the indicators. With even as few as 20 cases, the model can be tested with multiple regression techniques. Following this, however, both influential and deviant cases (or multivariate outliers) are identified for more specific historical comparative analysis. In Ragin's terms, this approach uses the *statistical* method to identify those cases whose *comparative* analysis will best illuminate or test the dynamics proposed in the general model.

III. An Example: The Political Ecology of Modern Population Change in Asia

Here an example is required to clothe the bare bones of the steps in this process. I shall focus on the problem of rapid population growth in Asia during the last half of the 20th century. I begin with the observation that modern Asia is distinguished by the configuration of three conditions. The first is rapid population growth. The second is the rapid spread of official antinatalist

policies, and the third is the construction of public programs that are designed to intervene directly into and to reduce marital fertility.⁴

The observation that these conditions constitute a distinct configuration derives from a particular perspective, which is best identified as *ecological* in nature. It thus derives from a theoretical framework that is marked by a number of specifiable elements. The ecological perspective leads us to see the phenomenon of population growth and fertility change as a pervasive adjustment, a population adapting to a changing environment. Further, the perspective assumes that such adaptations always take place through some form of organization and involve a specifiable technology. For theoretical reasons we see political forms of organization as important, though not exclusive, forms through which modern populations adapt to their environments.⁵ Like all forms of organization, political organizations vary in their capacity to mobilize populations and to direct the adaptive process in a concerted manner.

We must now return to comparative case histories to suggest a model of the dynamics by which political organizations emerge, gain a specific character, adopt specific policies and have a specific impact on the adaptive process. The content of the model we propose derives originally from a comparison of Malaysia and the Philippines, and the historical processes by which they adopted antinatalist policies, constructed the organizational instruments of fertility-limitation, and thus actually had an impact on fertility. (Ness and Ando, 1971) From demographic theory and observations of the distribution of antinatalist policies, two other sets of conditions are included: the level of social and economic development⁶, and the level of population density.⁷

With these sets of conditions, or variables, the model proposes the following process: Ecological pressure (population density) combines with the strength of the political and administrative organizations to determine the timing of the antinatalist policy decision. The greater the density and the stronger the political-administrative organizations, the earlier the state will make an antinatalist policy decision.⁸ Second, the earlier the antinatalist policy decision and the higher the level of social and economic development, the stronger will be the implementing organization, or the family planning program.⁹ Finally, the stronger the family planning program and the higher the level of social and economic development, the more rapid will be the rate of fertility decline.

With these specified variables and their operational definitions, we can obtain quantitative measures for 21 of the major states of Asia. (Our study omitted India and China, which were submitted to a comparative case analysis; and Japan, which had completed the demographic transition by the time the other Asian nations were entering it.) As the model is stated, it can be estimated by three multiple regression equations. The dependent variables are, respectively, the time of the antinatalist policy decision, the strength of the family planning program, and the rate of fertility decline between 1965 and 1975.¹⁰

These data can be arranged in a variety of ways to exhibit their relationships, and thus to make statements about the similarities and differences of

Asian states (to use the rhetoric of both comparative and statistical methods). Exploratory Data Analysis (Tuckey 1980) provides a new and highly simplified mechanism for both displaying and analysing distributions and relationships. We choose here, however, to use the more familiar multiple regression and path analytic techniques, largely because they are so widely used in cross national research. We believe, however, that most of the users of this technique rely far too much on coefficients and far too little on qualitative, or comparative, historical detail. Thus we believe it is especially useful to develop a strategy that exposes our measures to the careful scrutiny necessary for more effective validation.

We expose the statistical estimations in three sets of scattergrams and path diagrams (Figures 1-6). The scattergrams permit knowledgeable observers to engage in at least crude validity checks on our measures.¹¹ The first scattergram (Figure 1) shows the scores of political-administrative system strength

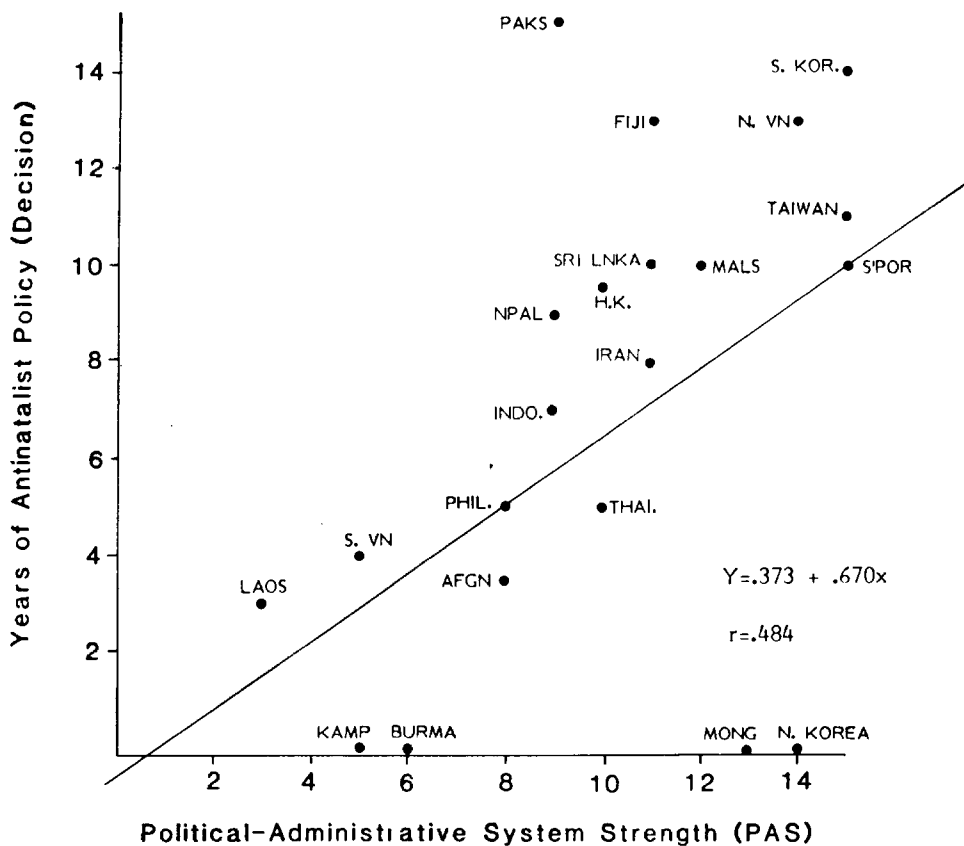
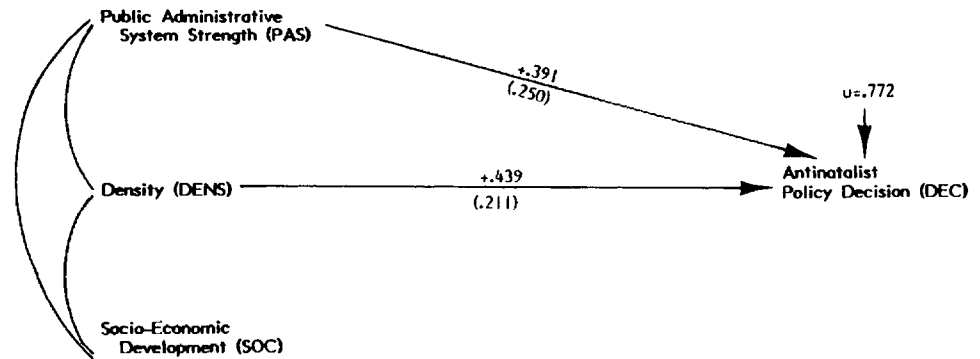


Figure 1. Scattergram of Political-Administrative System Strength and Years of Antinatalist Policy.

on the horizontal axis, and the timing of the antinatalist decisions on the vertical axis. The accompanying path diagram (Figure 2) shows the impact of the two independent variables together on the timing of the antinatalist decision. Each subsequent pair of figures follows this process. The scattergram presents the immediate organizational score on the horizontal axis and the score of the dependent variable on the vertical axis. Each scattergram is followed by a path diagram that shows the full range of independent variables in relation to the dependent variable.



$$Y(\text{DEC}) = +.391\text{PAS} + .439\text{DENS} - .0395\text{SOC} \quad (1)$$

(.250)
(.211)
(.271)

$$F=3.8358 \text{ Sig.}=.0289 \text{ R}^2=.404$$

Figure 2. Path Diagram of the Determinants of The Timing of the Antinatalist Policy Decision.

In these figures, the coefficients and standard errors (shown in parentheses) in the path models provide some support for the general theory. Political and administrative conditions plus ecological pressures drive policy decisions. Together with the level of social and economic development, the policy decisions determine the strength of the family planning program; and program strength together with the level of social and economic development largely determine the rate of fertility decline.

The final step in our analytical strategy leads us back to a qualitative assessment of individual cases, or back to comparative case histories. We identify the cases to be examined by the size of their residuals.¹² It is the deviant cases, or the multivariate outliers that are most interesting.

In the first scattergram (Figure 1), North Korea and Pakistan are, respectively, negative and positive outliers. In the case of North Korea the question is

what led it to maintain a pronatalist policy despite its high levels of both population density and political-administrative strength? Our explanation draws on qualitative historical data and focuses on the hostile relations between North and South Korea, and the fact that the South has twice the population of the North. These political and demographic conditions can adequately explain why and how North Korea has retained a pronatalist policy.

Pakistan is the other deviant, with a much earlier policy decision than predicted by its relatively low level of population density and its weak political-administrative system. Pakistan's history shows clearly, however, that when it made its antinatalist policy decision, it had just experienced a surge of what we call political-administrative system strength. Ayub Khan had just come to power with new policies to strengthen the center. The country also moved forward rapidly in centralized national economic planning, and it was precisely the planning activity, combined with a new knowledge of the size and growth rate of the population (a monitoring of ecologically relevant conditions) that led to the policy decision. All of this is quite clear from a reading of the political, organizational and economic history of the period.

The "strength" of the administrative system did not last out the decade of the 1960s, of course, giving it an overall low score, but the early decision remained as a part of policy. It was not effectively implemented, and has not yet been to this day, but the time of the policy decision stands. This not only clarifies the dynamic relationship, it also amounts to a qualification, or partial invalidation, of the indicator of policy decision. Our measure was the time of the decision, which, as we see here, is not by any means a fully valid measure of the strength of the policy decision.

The scattergram (Figure 1) shows Mongolia as a significant negative outlier, but with its exceptionally low population density, it performs according to expectations in the multiple regression equation. It can also be readily demonstrated that Mongolia's political elite perceive its small and scattered population as an obstacle to economic development. Thus, its population policy is deliberately pronatalist. It is, in effect, designed to promote national economic development, just as is the antinatalist population policy of South Korea, India, or Indonesia.

Let me pass over the deviant case analysis of the second step (Figures 3 and 4) and go to the final one (Figures 5 and 6) where the rate of fertility decline is explained. Here the major deviant case, in both the scattergram and the equation, is Iran, with virtually no decline in fertility, despite its alleged strong family planning program, its early policy decision and its apparent high level of social and economic development. To understand Iran's failure in fertility reduction we can engage in a number of comparisons. In any exercise, we find that Iran failed to produce a human service system in general and a family planning program in particular that was capable of moving resources to the poor rural population. It was a system in which attention, interest and resources were concentrated at the top. Iran's wealth lay in oil, discovered, exploited, paid for and consumed by other societies. Its great wealth thus did not

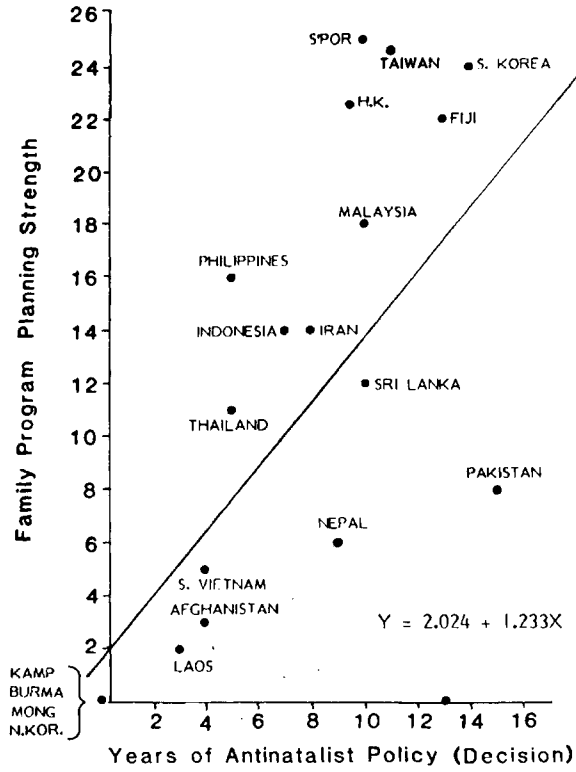


Figure 3. Scattergram of Years of Antinatalist Policy and Family Planning Program Strength.

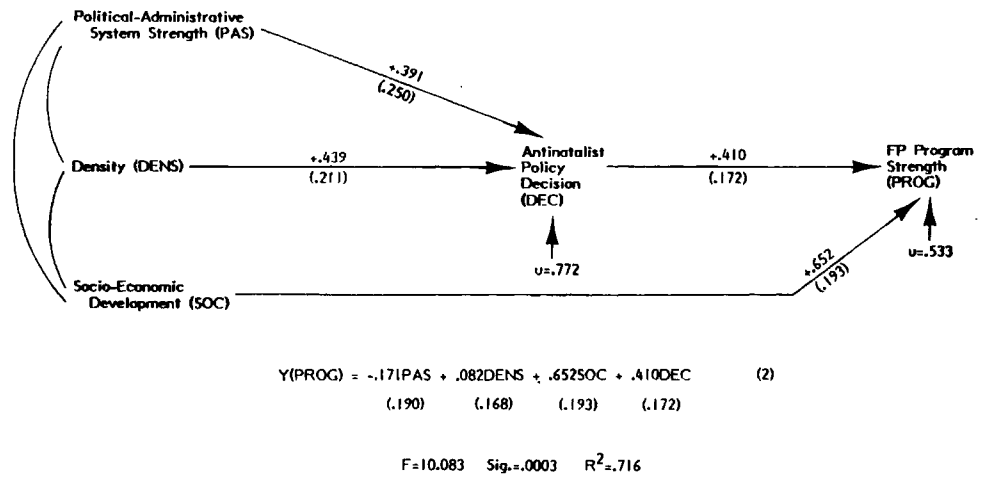


Figure 4. Path Diagram of the Determinants of Family Planning Program Strength.

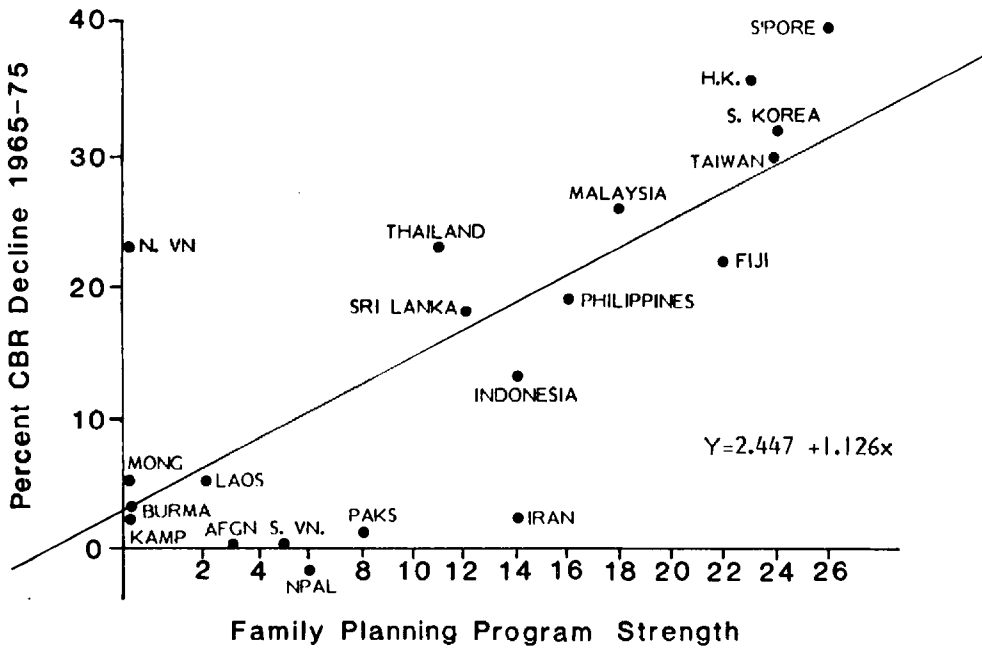


Figure 5. Scattergram of Family Planning Program Strength and Rate of Fertility Decline 1965-75.

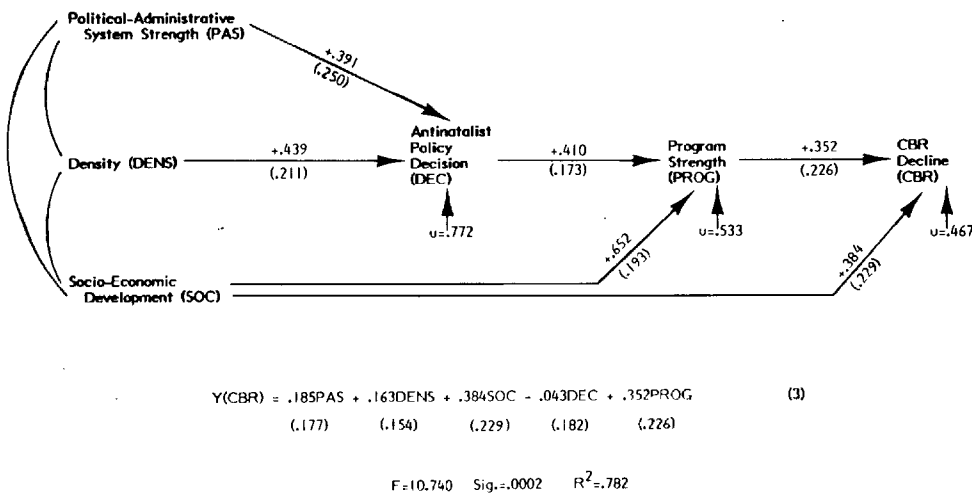


Figure 6. Path Diagram of the Determinants of the Rate of Fertility Decline 1965-75.

imply an effective mobilization of its own human resources, as was true, for example, in South Korea, Taiwan, Malaysia, Singapore or Sri Lanka. Its family planning program did not, as was very true in much poorer Indonesia, move out of the urban centers to the villages and to the rural poor. In fact, in explaining the failure of Iran's fertility-limitation program, the content of the discussion sounds similar to the one used to explain the fall of the Shah.

There are many similar examples of the combination of statistical and comparative analyses that can be generated in this exercise, but space precludes their exposition here. One central issue must be mentioned, even if only in passing. To this point our theoretical framework has permitted us to work with nation states as the units of both observation and analysis. This is justified by our specification of the political organizations of the state as at least some of the major organizational forms through which the adaptive processes are working. It is not necessary to stop with this unit of analysis, however, and extended theoretical development would lead us to examine the states in the larger setting of a world community or a world economic system. We can move to this level of analysis and ask to what extent the external world intruded into these state decisions and programmatic implementations.

The same strategic process can be used. We can develop a theoretical statement that specifies the condition of the larger setting and the mechanisms by which it affects state level actions. In this case, we found these in the technical and organizational operations of the United Nations regional organizations, and of other international agencies that promoted the organization, technology, and ideology of national economic planning. This analysis employs largely a narrative historical form, with a focus on organizations that span national boundaries. We can also quantify certain elements of this boundary spanning by noting the per capita level of foreign assistance received by each state for its modern fertility-limiting activities. The resulting measures can be entered into the same three series of regression equations used for the states alone, both to test the general theory and to specify those cases that should be subjected to individual comparisons. We do not show the data and findings, since the technical details require more space than is warranted by what is, after all, merely an illustration.

Our interpretation from these analyses, however, is that in Asia, the processes of policy decision and program implementation came more from within than from without the states. There was, to be sure, external influence from both the emerging Asian community and from the western nations. All of this, however, constituted an important interactive process that supported, but could not counter, processes whose political and social roots lay within the states. The exposition of this argument is woven out of historical treatment of international organizations, and specific foreign assistance programs, regression coefficients of the impact of foreign assistance on national population planning programs, and individual country analyses that examine the determinants of both effective and ineffective foreign assistance programs.

Space precludes a full exposition of this story, but this should at least make clear the steps in the strategy that permit us to weave together comparative and statistical methods of analysis. Most cross-national analysis focuses exclusively on one method or the other. The systematic comparative illustration of a few cases provides rich detail but is weak in the generation of more general principles of social change. Statistical crossnational studies seldom permit the reader to check the validity of the indirect measures that are used to indicate highly complex behavior patterns, nor do they provide the historical detail by which we can gain a better understanding of the complex processes of social change. The rhetoric of statistical cross-national analyses is typically the rhetoric of numbers and equations, often lacking even the most cursory mention of the fact that what we are attempting to understand is collective human behavior rather than simply the relationships between numbers.

It is not necessary, however, to make the severe choice between these two strategies. Cross-national studies can exploit both comparative and statistical methods, and we would argue that it is imperative that they do so. Our indicators are far too indirect, and their validity far too suspect to permit the extensive statistical treatment we often see in studies of larger numbers of cases. And our need to understand the general processes by which many different states or societies move and change with and against one another is too great to permit us to be content with comparisons of very small numbers alone. It is in the not-so-small numbers that we can find both the rigor of statistical analyses and the rich detail of individual cases.

NOTES

- 1 Naming this class of numbers presents an interesting challenge. We thought of calling them RANs, for Reproductive Age Numbers, or MANs, for Middle-Aged Numbers. They could also, of course be called Not-So-Large Numbers, but a creative lapse has forced us to rely on this somewhat more prosaic term.
- 2 I shall not attempt to deal here with the equally intractable problem of reliability. Few people read Oskar Morgenstern (1950) any longer, but his criticism of both the reliability and validity of economic measurements are still well worth reading, as is his argument against their frequent uncritical use.
- 3 The problems are compounded as missing data also suggest less than certain validity for those data that are available.
- 4 Elsewhere, I apply these observations to the whole of the currently high fertility world, which is also almost coterminous with the Third World. (Ness 1983) The focus on Asia is based on a further observation that this configuration emerges and is most fully developed in Asia, and this for understandable historical reasons whose exposition is precluded here due to limitations of space. (Ness and Ando 1984 provide the full explanation.) The first condition has occurred before in other parts of the world, but not with as great power as we find today in the Third World. The second and third conditions are quite unique today. Most governments throughout most of human history have been pronatalist. Thus the rise of official antinatalism today represents a significant revolutionary condition. Finally, throughout most of human history, the marital bond has typically carried with it both the right and the obligation to reproduce. Today the public intervention to reduce *marital fertility* constitutes another revolutionary condition.

- 5 The theoretical base of this observation lies in the Weberian perspective on the bureaucratization of the world. (Weber 1968, Coleman 1972, and Jacoby 1968)
- 6 Many measures are available for the level of social and economic development, and most are highly intercorrelated. We used a combination of Infant Mortality Rates, GDP per capita, and female school enrollment. This may be said to indicate the extent to which a society is healthy, wealthy, and wise.
- 7 We have observed that states with antinatalist policies in 1973 had higher average levels of population density than those that had not made such decisions. This was true for the Third World as a whole, and for each of the four major regions within the Third World: Africa, Asia, Latin America, and the Near East-North Africa. This observation is intuitively acceptable in an ecological perspective and leads to the use of population density as a general and indirect measure of ecological pressure. See also Galle (1978) for a more general theoretical and empirical treatment of the relation between population density and a population's adaptive processes.
- 8 For the strength of the political and administrative organizations we coded three dimensions, whose contents were derived from the Malaysia-Philippines comparison. These were the strength of the political center, the commitment to modern national economic planning, and the capacity to monitor social and economic change. To code individual states, we used a narrative five point scale with multiple coders, whose results showed high intercoder reliability. (Ness and Ando, 1984)
- 9 The measure of family planning program strength was developed by Lapham and Mauldin (1972) and updated by Freedman and Berelson (1976), and Mauldin and Berelson (1978). It codes and scores a variety of organizational and resource conditions that are typically found in varying degree in modern family planning programs.
- 10 The problem of the validity of the measures is a serious one and we cannot claim to have solved it to full satisfaction. There is still considerable uncertainty about the actual rate of fertility decline, for example. (Cavanaugh, 1981) At this time, we can only move on to our exposition of the cases by bivariate scattergrams to permit knowledgeable others to make independent judgements.
- 11 We argue that this should be standard practice in cross national analysis. We are seldom shown specific countries in their critical relationships, and often we are neither told which countries are included or excluded in specific equations, nor what the individual scores are for the countries included.
- 12 We have usually considered outliers to be those cases whose standardized residuals are + or - 2 greater than the mean of the residuals. In some cases the residual of the outlier is less than 2, but the case remains quite alone above or below others.

REFERENCES

- BENDIX, Reinhard,
1956 *Work and Authority in Industry* (New York: John Wiley & Co.).
- COLEMAN, James S.,
1974 *Power and the Structure of Society* (New York: W. W. Norton and Co.).
- DOUGLAS, Jack D.,
1967 *The Social Meanings of Suicide* (Princeton: Princeton University Press).
- FREEDMAN, Ronald, and Bernard Berelson,
1976 "The Record of Family Planning Programs," *Studies in Family Planning*, 7:1.
- GALLE, Omer and Walter Gove,
1978 "Overcrowding, Isolation and Human Behavior: Exploring the Extremes in Population Distribution," in Karl E. Taeuber, Larry Bumpass and James A. Sweet, eds., *Social Demography* (New York: Academic Press), pp. 95-132.
- GOFFMAN, Erving,
1961 *Asylums* (Chicago: Aldine).

- JACOBY, Henri,
1972 *The Bureaucratization of the World* (Berkeley: University of California Press).
- LAPHAM, Robert, and W. Partker Mauldin,
1972 "National Family Planning Programs: Review and Evaluation," *Studies in Family Planning* 3:2 (March, pp. 29-52).
- LEWIS, Oscar,
1959 *Five Families* (New York: Basic Books).
- MOORE, Barrington Jr.,
1966 *The Social Origins of Dictatorship and Democracy*, (Boston: Beacon Press).
- MORGENSTERN, Oskar,
1950 *On the Accuracy of Economic Observations*, (Princeton: Princeton University Press).
- NESS, Gayl D.,
1983 "The Political-Ecology of Population Change in Developing Countries," Paper read at the Population Association of America meetings, Pittsburg, April.
- NESS, Gayl D. and Hirofumi Ando,
1971 "The Politics of Population Planning in Malaysia and the Philippines," *Comparative Administration Review*, pp. 42-54.
- NESS, Gayl D., and Hirofumi Ando,
1984 *The Land is Shrinking: Population Planning in Asia* (Baltimore: The Johns Hopkins University Press).
- RAGIN, Charles,
1981 "Comparative Sociology and the Comparative Method," *International Journal of Comparative Sociology*, XXII:1-2, pp. 102-120.
- SKOCPOL, Theda,
1979 *States and Social Revolution* (Cambridge: Cambridge University Press).
- TOCQUEVILLE, Alexis de,
1863 *Democracy in America* (Cambridge: Sever and Francis).
- TURKEY, John W.,
1977 *Exploratory Data Analysis* (Reading, Mass.: Addison Wesley Publishing Co.).
- WEBER, Max,
1972 *Economy and Society* ed. by Gunther Roth and Klaus Wittich, (Berkeley: University of California Press).
- WHYTE, William F.,
1943 *Street Corner Society* (Chicago: University of Chicago Press).