

RESEARCH

The Village Studies Programme

When are the Ceteris Paribus?

by Anne McManus*

It is surely wise to test a hypothesis empirically before charging full speed ahead applying it as *a priori* and self evidently given. In the case of the Village Studies Project (VSP), reality obstinately refuses, in many areas, to conform to the Liptonian 'universal' definition of a 'village' as a small, settled group of persons, living in and forming all or most of the population of a locality, with two thirds or more of the families getting most of their income from agriculture, and with most of the economic, social and political relationships and hierarchies within the locality.(1)

The model hypothesis of the Project is:

"Villages of type X, in less developed environments, respond better to inputs of type A than villages of type Y, for the reasons R_1, R_2, \dots, R_n ."(2)

It is contended that differences between villages rather than between individuals or social macrocosms (nations, states, regions, classes) explain the varying success of different sorts of developmental efforts. The hope is to show that a village of a given size, cropping pattern, tenure system, soil-climate ecology, man-land ratio and urban contact, will respond similarly to the same inputs whether it is located in India, Mexico or Kenya; in one region or another. The idea is that the data to test particular hypotheses at the micro-level already exist in the form of intensive single village studies which are often

* Anne McManus is an IDS Research Assistant who is working on the Village Studies Project under the overall supervision of Michael Lipton, IDS Fellow.

(1) Michael Lipton, "The Less Developed Village: Some Conceptual Issues", IDS Mimeo, July, 1970

(2) *ibid.*

unpublished, widely scattered, and of varying methodology, 'typicality', orientation and value.

The primary stage of the Project is to collect and file at IDS all the 'village studies' done since 1950 from India, Africa, and Latin America. In the first instance, the aim is to analyse them with a view to publishing a comprehensive annotated bibliography of all village studies, and a guide to their methodology, by October 1971. The bibliography will indicate for each study the coverage of items of physical, socio-economic, and demographic information, which will ultimately be available on punch-cards for widespread subsequent use.

The second stage of the Project consists of classifying the villages according to a number of key variables which are required to carry out a set of correlations for cross-section comparative analysis in the three subprojects: Nutrition, Employment/Migration, and Demography; and to estimate the types of villages that are likely to respond 'best' to, say, alternative bundles of measures to improve nutrition, or to identify types of job situation that can be linked to types of village environment. The ultimate objective is to be able to define and measure specific benefits and to see how they vary with specific input levels in villages of different types, thus making it possible to offer policy-makers investment manuals which corroborate the original hypothesis: 'If you put this type of input into this type of village, you'll get this result', and presumably helping them in resource allocation.

That, in very summarized form, is what VSP is about. In its present form, it has a number of important weaknesses. Apart from the obvious fallacy of attempting to construct a micro theory of change from essentially static data, there are further methodological and empirical defects which cast severe doubt on the validity of VSP within its original frame of reference.

The basic problem is that what needs to be investigated is simply asserted. The methodology assumes away the crucial problem of whether in fact villages in Africa/Latin America/India are comparable; and whether national/regional/international relationships are not the key factors influencing village level behaviour. Whether this is so or not, it is certainly not self-evident. Such a tautologous methodology dictates that those villages fitting the classification may be used without any attempt to answer the prior and

fundamental questions of whether these villages are typical, or whether there are crucially different macro-level factors, which have in the past, and may potentially, in the future, produce a different effect. Obviously an exercise in comparative statistics of this kind, with a methodology which excludes consideration of them at a particular point in time by definition, can never offer even tentative predictions. Even in India, from which the classification seems to have been extrapolated, the peasant economies have a dual orientation, both subsistence and market. They exist in one nation state and are linked with urban areas which they must sustain. These relationships need specific empirical investigation in each case, and it cannot just be assumed that "the level of any social, economic or political variable in the village will not in the short term be significantly affected by changes in the national social, economic and political system". (1)

An orgy of selectivity validates the Project's comparative claims if only 50 Latin American villages appear to fit the classification at the particular time of the survey. Unless external/macro factors are vitally integrated into the model and continually assessed, any comparison of village characteristics could be blown apart at any time by the infiltration of, or changes in, regional/national/international influences. The most cursory and preliminary investigation would have tempered this myopia with the factual observation that in Latin America, it is meaningless to look at small groups of *minifundia* in Chile as behavioural units in abstraction from the presence and effects of the *hacienda*; or to look for a single economic criterion for geographical subdivisions of areas in East and Central Africa, where land settlement patterns are widely different over short distances, and tend to be scattered anyway. The only relevant area for analytical purposes in the latter case would be those demarcated by the agricultural systems, which often cover very large areas. Whether or not it is useful in these cases to conduct inquiry at the community level, or at the level of farm management schemes, or at that embracing whole *municipios/haciendas* etc. is a matter for empirical enquiry first. Failing that, at least a prior pilot survey should be conducted in the areas intended for study. The data should be subordinate to the hypothesis, not the inverted process of looking for hypotheses which will fit into the village level framework once one possesses n-thousand village studies that, once collected, must be used. If it is decided that the data are simply not available or useful for Latin America or East Africa, and that these two

(1) Mick Moore, "Introduction" to *The Methodology of Village Studies*. IDS mimeo.

inconveniently large anomalies are to be abandoned by the Project, this involves dropping the basic comparative framework, and entails a fundamental policy decision. It cannot simply be left until the data, collected on expensive expeditions, speaks for themselves.

The 'village level' of analysis is defended by offering a false dichotomy between the village and the universe, i.e. the argument that if one does not accept the village as a valid economic or social unit, then one must go to the logical limit of studying no unit smaller than the world. Though it is tempting to answer: "Yes, why not?", the point is rather a plea for relevant hypotheses which should dictate the spatial/structural area for study, rather than starting from the wrong end by taking behavioural characteristics as given. All in all, the project seems to have been extended far too much in the direction of village studies as units and ends in themselves, rather than in the direction of developing basic hypotheses and testing them via the village study material, if indeed it is at this level that they should be posed.

A further difficulty is raised by the assumption that the village may be treated as a unitary behavioural unit because, at the very least, it neglects the interpersonal distribution of costs and benefits within the village. Even though this is explicitly not the intention, the aggregated nature of the data available on village inputs and outputs:- income, education, nutrition, owned land, and production, leads inevitably to a mystification of 'village (community) benefits'. It is not necessary to be a raving Naxalite to ask who benefits within the village; or to realise, for example, that higher 'benefits' in the sense of higher crop returns on investment obviously accrue to inputs into an irrigated, rather than a rainfed village. This benefit will tend to be concentrated in the hands of a few in the irrigated village, but will be more widely spread if it goes to an unirrigated village, where people are poorer anyway. Whose cost and whose benefit is at stake: what about growth and development?

The format of the bibliography includes both variables that have been selected as key for the Project's eventual cost-benefit comparisons, and those considered relevant to a wide variety of potential users. It is thus possible to record whether key indicators of benefit are available in each village study catalogued. But apart from the doubtful use of secondary sources for correlation analysis, the feasibility

of this correlation analysis itself depends on the availability of enough studies of villages of the same type and recording the same sort of data. For example, a major difficulty of comparability arises in relation to the nutrition studies since the separate methodology involved in nutrition work has rarely, if ever, been applied to villages for which detailed economic analyses have been made. For these nutrition studies alone there are at the moment very few which contain data on the other key variables necessary for correlation analysis (such as labour input) and the nutrition data in the general village studies are so sparse as to be useless. So the information for correlations between nutrition and other key variables is simply not currently available. The further question of a cost-benefit analysis demands not only correlating enough villages with the same type of input to be statistically significant, but an accurate costing of all 'input' programmes, and a measurement of the effect of benefits within the village, not accounting for any benefits dispersed outside it. This demands at least an adequate number of re-surveys (even allowing time-sequence observations as an adequate assessment of changes) to enable benefits to be estimated. Until now, the Project has been able to acquire only about 30 re-surveys for India. Can it rely on obtaining more? If not, how is it possible to estimate benefits?

So even if the methodological questions are set aside, and *ceteris paribus* conditions beyond the village are assumed, the limitations of the data available are becoming increasingly obvious, at least within the original cost-benefit framework, which is based on cross-section village data. Further, correlation, even if possible and significant, is not, of course, causality, and the highest association of village-level phenomena, plus casual and doubtful evidence about time-sequence, is scanty evidence to induce planners to put resources into Village X rather than Y. Planners will need heavily substantiated causal hypotheses about reasons for associations/correlations, and these are dubiously possible without original research and resort to extra-village-level data.

All this is not to say that there is no point in micro-level studies, or indeed Studies of Studies. It is rather to suggest that studies of villages in isolation beg the key question of whether they are in fact so, and whether for the purpose of analysis they can be isolated. If, for example, the Project were to discard its totally global focus, it could undoubtedly turn up useful information about villages in India. The bibliography will undoubtedly be very valuable - for Village hunters, planners, and students.

Whether the second stage of doing something with the data within a cost-benefit framework is feasible and worthwhile remains to be seen. But the whole affair would appear much more convincing if there had been less of a Grand On-Going Concern, and more of an evident attempt to investigate initial hypotheses and to conduct healthy critical introspection.