

UC Irvine

UC Irvine Previously Published Works

Title

Community of Reference in Rural Stratification Research

Permalink

<https://escholarship.org/uc/item/1df8x8s1>

Journal

Rural Sociology, 46(4)

ISSN

0036-0112

Author

Cancian, F

Publication Date

1981

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at

<https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Community of Reference in Rural Stratification Research¹

Frank Cancian

School of Social Sciences, University of California, Irvine, California

ABSTRACT Local social systems are connected to larger systems, including the world-system, in a variety of complex ways; thus it is often difficult to identify the local stratification system or community of reference. Yet, system specification is crucial in stratification studies, because the phenomena studied are often best understood in terms of *relative position within the local system*. To support this emphasis on local community of reference this paper reviews Cancian's upper middle rank conservatism thesis and Frey *et al.*'s (1979) finding of lower middle rank conservatism in adoption of agricultural innovations in Pakistan, and presents a reanalysis of the Frey *et al.* data. The reanalysis shows that the finding changes when specification of the local system changes. Thus, the system specification problem faced by empirical researchers concerned with rural stratification systems is highlighted.

Introduction

The problem of system definition bedevils rural stratification research in at least two ways. The first comes from the fact that rural people and rural communities are usually involved in many and diverse social, economic and political relationships. Friendship and kinship ties on the one hand, and the world economic system on the other hand, can be shown to affect both the rural individual and the rural community. In the end it is usually the investigator's perspective that determines the approximate boundaries of the specific system that is analysed; and it is frequently true that the system analysed includes and is included in other systems that can be shown to manifest important social, economic and political processes. While it is commonly believed that different perspectives are legitimate, it is also commonly believed, by the partisans of the various positions that their perspective is particularly revealing, more important, and fundamentally closer to basic reality. Neither reductionism nor expansionism can be definitively rejected; and selection of perspective or level of analyses remains a fundamental dilemma in research concerning rural people.

System definition is also a problem at the operational level. When we do empirical work we study complex natural systems. The niceties of philosophy of science aside, the *analytic system* defined by the investigator's perspective is usually meant to approximate an isolatable

¹ I am indebted to Francesca Cancian, Ken Chomitz, Scott Frey, Karen Leonard, Alice Saltzman and anonymous reviewers of this journal for comments on earlier drafts of this paper.

natural or empirical system. Most investigators believe that the natural world is not an undifferentiated blob which indifferently welcomes all perspectives and all definitions of empirical systems. They believe, or behave as if they believe, that natural systems do exist—however open and interconnected they might be. Some “things” are more related than others; and empirical research must identify the patterns that are usefully labeled natural systems. Still, it is not always easy to know where a specific empirical system begins and ends, what it includes, what its boundaries are.

Research on rural stratification systems faces both of these dilemmas: (1) selection of perspective or level of analysis, and, (2) identification of the empirical system once the perspective is set. For the most part, this paper concerns the second of them, but the first is also relevant, if only because the success of the perspective favored here competes with the various reductionist and expansionist alternatives.

In what follows I will concentrate on the problems of identifying² the “community of reference” (Cancian, 1979); that is, the empirical system within which a particular kind of rural stratification system operates. The dilemma involved in selecting a perspective is avoided because the concrete alternative analyses reviewed below share a perspective—that is, they share assumptions about the relevant local system. Neither of them wishes to build rural social structure from individuals and dyads nor to derive it from extralocal (national and international) processes. And, both of them focus on populations with high concentrations of small farmers rather than populations numerically dominated by landless laborers, rural proletarians or tenants. Even given this agreement on perspective and these limitations on the scope of the inquiry, the problems of identifying the concrete system remain.

The paper has three parts: First, I will briefly discuss the notion of a local stratification system or community of reference. Then, I will review a theory that relates risk-taking to rank, and argue that the assumptions and predictions it makes provide a particularly useful “litmus paper” for evaluating specific efforts to identify local stratification systems. Finally, I want to respond to Frey *et al.*'s (1979) contention that my “upper-middle-class” conservatism thesis is not supported by their data from Pakistan, and to discuss their conclusion that a pattern better described as “lower-middle-class” conservatism does exist in Pakistan. I will conclude that the findings derivable from Frey *et al.*'s data vary dramatically depending on how community of reference is specified. This conclusion in turn illustrates the need to attend to local stratification systems.

² I will usually use “specification” to indicate the “definition” or “identification” of a concrete “system” that the researcher believes is a natural system relevant to his/her research.

Local community of reference in stratification research

I want to show that people live in local stratification systems. Larger, extra-local systems are somehow important to social position, but class or rank, and many important social processes seem to be local for most people most of the time. The problem is: how can we conceptualize and study these processes in a productive way?

First we must recognize that such processes are not limited to "simple" or rural societies where people are somewhat isolated from the larger national and international systems. Even cosmopolitan people fully participating in the modern world may have to cope with localized or specialized stratification systems. For example, Rossi (1980:41) suggests that while national occupational prestige ratings show sociologists to be higher than economists, he is still seeking a university setting where his everyday life manifests such a pattern. Whatever the "average" American respondent who defines the larger system thinks of and does with sociologists and economists, ordinary university people and funding agencies commonly give more deference and resources to economists, at least many sociologists think so. Thus, in this case, as in many others, the local system—the one defined by university people and funding agencies—is more relevant to the actor than the "larger" system.

Rossi's remarks in a review of Coleman and Rainwater's *Social Standing in America* help focus the issues in a way that is useful at many levels of analysis. He lists the exploration of the local/global distinction as one of the major problems in stratification research and says,

... there are the many issues raised by how easy it is to confuse local stratification with societal stratification. Our economy is dominated by society-wide processes. There are scarcely any local-producer-local-consumer markets anymore. The fates of our jobs, how much we pay for food, housing, and clothing, and the desirable life experiences within our grasp are all dominated by the national economy, national government, and their connections with the world market and the global power struggle. While our fates may be determined somehow by that complex that scarcely knows distance, we enjoy or suffer the consequences locally in that restricted area encompassing our homes, jobs, family, and kin, and the peculiar distributional organizations that serve that area. The stratification that counts heavily is that range of differences displayed constantly. In that sense, class is a local phenomenon, recognized as such as much among the denizens of Soulside and Talley's Corner as in Crestwood Heights, Jonesville, Kansas City, and Boston (Rossi, 1980:41).

Of course, the phenomenon that Rossi describes is not limited to urban enclaves of industrial societies. In a comparative study of agricultural communities (Cancian, 1979), I have used the term "community of reference" to describe a local unit similar to that evoked by Rossi's rich characterization, and I have argued that rank-seeking behavior must be understood in terms of the local system. Here, I want to review the theoretical status of the community of reference idea before describing relevant empirical studies.

Community of reference may be briefly defined as a naturally occurring, local unit within which most people experience distributional and rank differences. This working definition is a complex assertion that needs immediate elaboration and qualification on a number of points.³ First, the assertion must be qualified to recognize the complexity of communities of reference seen from the standpoint of the individual actor. Most actors may be seen in terms of multiple, complex, overlapping, and contradictory communities of reference, and it is unlikely that any actor can always be fully understood in terms of a single community of reference. On the other hand, any assertion is useful to empirical research only insofar as we can put aside the infinite complexity and interconnectedness of human behavior and concentrate on the search for patterns that stand out from the complexity. Thus, the community of reference idea, like any other abstraction, must be evaluated in practice.

Second, the characterization of communities of reference as natural systems is meant to emphasize the importance of discovery; that is, it is meant to emphasize the need to look at the concrete world before specifying the boundaries of particular systems. It would be almost as apt to substitute the word "empirical" for "natural," for, as the paper's introduction implicitly suggests, I am simply trying to steer a path between those languages of research that see system delimitation as an arbitrary, entirely analytic or logical operation and those naïve realist ones that believe they will find truth full-blown in nature. This level of metatheoretical detail is a bit excessive in the present context, but recent attention to a variety of questions about the scale and reality of socioeconomic systems (see, for example, Barth, 1978 and Kaplan, 1978) make it desirable to be explicit.

Finally, both Rossi's language and mine stress subjective experience. For my part, this aspect of the definition has an "as if" status: that is, actors should be *seen as* behaving in terms of their positions within a community of reference and their desire to attain or avoid other positions in it. However, since I believe that actors are not always fully self-conscious and clearly motivated in their behavior relevant to stratification, their self-reports, status evaluations and oth-

³ The relation of these ideas to "reference group" ideas is discussed in Cancian (1979).

er "subjective" data are not necessarily relevant to evaluating the utility of this conceptualization of local stratification processes.⁴

In sum, community of reference is a concept developed to facilitate work with an old problem—the problem that, in part at least, the structure of specifiable, local socioeconomic systems seems to determine the behavior of actors who are connected in many ways to multiple and larger socioeconomic systems. By exploring the confusions resultant from empirical misspecification of community of reference, I hope in what follows to both demonstrate the importance of local systems and document the need for careful attention to the concrete community of reference from which rural actors are sampled.

A test for community of reference: upper middle rank conservatism in local systems

What is the empirical evidence for the assertion that local systems are important? Since, even at its best, community of reference is a diffuse concept referring to a complex phenomenon, it remains to be shown whether and in what sense natural local systems exist. This will not be easy, because, as Rossi (1980) suggests in the passage quoted, there are many ways to mistakenly identify the "appropriate" system in the study of social stratification. It is not always clear that local systems specified in particular studies are more than arbitrarily delimited parts torn from a larger system, or that individual behavior is more than a local manifestation of a global process. In this situation my theory of "upper-middle-class" conservatism (Cancian, 1967, 1972, 1979) provides a useful test of the importance of community of reference. This is so because the theory makes predictions that will be easily falsified if community of reference is not appropriately specified. Since these predictions are complex, failure to falsify them may be taken as evidence for existence of a natural system at the local level. Thus, while the theory must be supported in the process, the theory itself is only an indicator (a litmus paper) in the more general exploration of the community of reference problem. This logic is more easily explained, I think, in the concrete terms that follow.

The theory of upper middle rank⁵ conservatism predicts that, in the early stages of the spread of an agricultural innovation, upper middle rank farmers will be less likely to adopt than lower middle

⁴ A number of papers in Barlett (1980) provide modern illustrations of the contrast between statistical behavior ("black box") conceptions of the actor (papers by Acheson, Barlett, Cancian, Chibnik, DeWalt and DeWalt, Finkler, and Johnson) and approaches that seek to approximate actual cognitive and evaluative processes (papers by C. Gladwin, H. Gladwin and Murtaugh, and Ortiz).

⁵ I used the word "class" in earlier publications to draw attention to the general (non-rural) applications of the theory, and to provoke discussion of the concept "class," but "rank" is more appropriate in the present context.

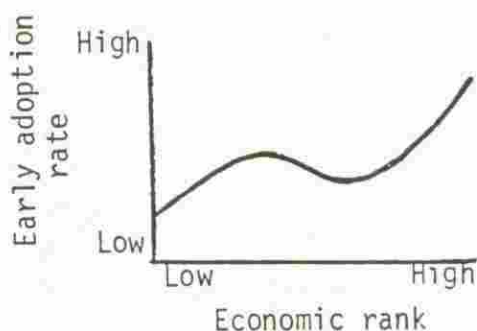


Figure 1. The relation of economic rank and early adoption of new agricultural practices

rank farmers. That is, it predicts that the "dip" illustrated in Figure 1 will appear in the relation of economic rank and the tendency to adopt. A secondary prediction is that the "dip" will disappear in the later stages of the adoption process—resulting in the positive, linear (or at least monotonic) relation that is part of the received wisdom in diffusion studies (Rogers and Shoemaker, 1971:187). An illustration of both predictions for an imaginary sample of 400 farmers is displayed in Table 1 and shown graphically in Figure 2.

As was noted above, the theory itself is not important in the present context. Since it is complex, and since it has been stated, along with evidence for its predictions, in a number of places (Cancian, 1967, 1972, 1979), and has been critiqued, along with presentation of negative evidence, in a number of other places (Frey *et al.*, 1979; Gartrell, 1977; Gartrell *et al.*, 1973; Morrison *et al.*, 1976), it seems pointless to repeat it here. Rather, I will focus on only those aspects of the theory and its predictions that are central to this paper. These flow from the fact that the rank variable is defined as *rank relative to relevant others*—to a community of reference.

Table 1. Predicted distribution of adoption by rank: an illustration, N = 400

	Low	Low middle	High middle	High	
Remainder	80*	55	45	20	200
Stage 2	10	15	35	40	100
Stage 1	10	30	20	40	100
Total	100	100	100	100	400

* Cell entries may be read as counts or as column percentages.

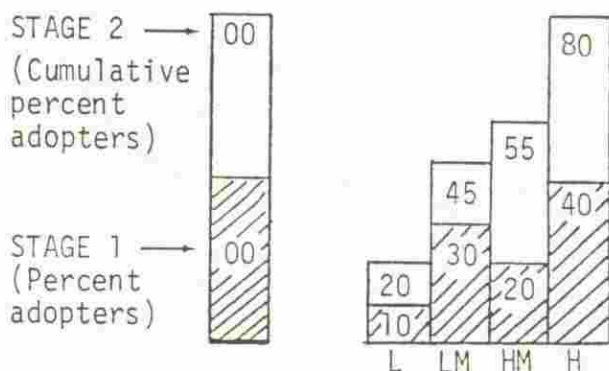


Figure 2. Adoption rate by stage: Table 2 illustrated

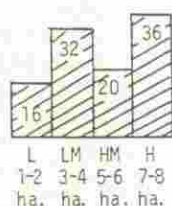
More specifically the "dip" and the remainder of the Stage 1 pattern illustrated in Figures 1 and 2 have particular characteristics that would be obscured if community of reference is not properly identified. Two likely types of misspecification are discussed below. Empirical evidence relevant to one will be briefly reviewed, while a study relevant to the other will be more fully explored. In both cases I will show how misspecification of community of reference may obscure the patterns predicted by the theory.

First, specification of community of reference that is too *inclusive* might homogenize distinct populations to produce a monotonic relation of rank and adoption rate (Figure 3). Concretely, the homogenization might result from the fact that rank is usually indexed by size of farm, value of crop or some other absolute quantity. Since the distribution of farm sizes, for example, might be quite different in different local areas (for a variety of reasons), inclusion of farmers from more than one natural area might obscure the "dip" in the manner shown by Figure 3. This potential misspecification graphically represents the sense in which rank as it is used here is rank, or social position, *in a system*, that is, in the community of reference, as contrasted with relative rank *in a relationship* between two or more individuals. The peculiar shape of the prediction of upper middle rank conservatism in the early stages of the adoption process (Figure 1) requires the former understanding of rank, and therefore requires, conceptually at least, a community of reference.

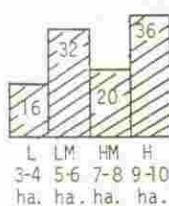
In another place (Cancian, 1979) I have compared empirical studies of almost two dozen populations involving more than 6,000 farmers in eight countries. The results show that the "dip" is more likely to appear as predicted when the sample of farmers considered is drawn from a universe⁶ with less than 100,000 people; and they

⁶ A non-gerrymandered contiguous geographical area. See Cancian (1979:36-38).

The village of Alpha
N=400 (100 per rank)



The village of Beta
N=400 (100 per rank)



The Alphabeta Valley
N=800 (200 per rank)

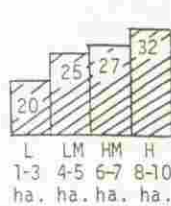


Figure 3. Potential effects of aggregation on curvilinear relationships. The number in each bar indicates the percent of farmers of that rank to adopt in a hypothetical Stage I. The stage is defined as the first 26 percent to adopt. It is assumed that, within ranks within villages, differences in hectares cultivated imply no difference in innovation rate.

suggest that failure to confirm the predictions is, at least in some cases, due to the homogenization of curvilinearity like that illustrated in Figure 3. This is empirical evidence for the existence of natural local systems—communities of reference.⁷ While the total number of cases (identifiable local systems) considered in this comparative research remains small, these results point to the importance of identifying the natural local system in the study of rural social stratification. If its misspecification obscures the particular theory utilized here, it may well have similar effects on other approaches to understanding rural populations.

Before going on to the other likely type of misspecification, I should note that the 100,000 figure stated above is an order of magnitude guess made on the basis of limited personal experience and tested against limited data. Just what size a community of reference will be should be expected to vary for a number of reasons having to do with local physical, geopolitical and socioeconomic conditions. Moreover, the 100,000 figure suggests a limit for an area where the heterogeneity is unknown or known to be moderate. A very heterogeneous area would certainly require a lower figure. At the other extreme, in a homogeneous area in which, say, each village is a community of reference, different conditions might prevail. For the purposes of testing the predictions considered above, for example, in such a homogeneous area there should be no difference between a random sample of one village and a random sample of a population that includes 1,000 villages. Clearly, judgments about local homogeneity and heterogeneity must be made for each area studied.

Second, specification of a community of reference that is too exclu-

⁷The details on a sample of 676 farmers taken in an attempt to represent an especially large "universe" (more than 100 million people) are also discussed in an earlier publication (Cancian, 1976).

sive might truncate the pattern illustrated in Figures 1 and 2. This could lead to another form of improper falsification of the theory. It is easy to imagine, for example, that part-time farmers who form the low rank in a local natural system might be excluded from a sample, or that residential segregation of low ranking people might also lead to their exclusion despite their sociological importance in the local community of reference. Similarly, high ranking people may live apart making it difficult to specify and sample the sociologically significant community of reference. Like the judgments about local homogeneity and heterogeneity discussed above, judgments about the extent of the local community of reference must be made for each area studied.

In the following section, I will argue that Frey *et al.*'s (1979) finding of lower middle rank conservatism in Pakistan results from an overly exclusive specification of community of reference, and I will reinterpret their data to show evidence for upper middle rank conservatism. This outcome lends further support to the thesis that identification of local systems is vital to rural stratification research.

Lower middle-rank conservatism in Pakistan

Frey *et al.* report an "unexpected finding" that is "contrary to the direct and linear relationship reported in the traditional literature" (1979:429). It is also different from the finding of upper middle rank conservatism that I have repeatedly reported (Cancian 1967, 1972, 1979). They say: "Unlike the curvilinear pattern of upper middle rank conservatism reported by Cancian, however, the present findings indicate a curvilinear pattern of lower middle rank conservatism. This finding represents a serious anomaly for Cancian's theory Consequently, we suggest that Cancian's theory is incomplete, for it cannot adequately explain the existence of the pattern of lower middle rank conservatism reported here" (1979:429).

Frey *et al.*'s finding of lower middle rank conservatism is both clear and stable across various measures of the rank variables they use. Thus, it does appear to be a serious anomaly. Yet, its shape (which is illustrated in Figure 4) raises questions about specification of community of reference—at least it does for the investigator whose theory goes unsupported by the finding. In the present context the most important question is: might alternative definitions of community of reference affect the clarity and stability of the finding? And further, of course, are there reasonable alternative definitions of community of reference that lead to findings more consistent with my predictions and less consistent with Frey *et al.*'s "unexpected finding?" What follows seeks to answer these questions.

Specification of community of reference can be, in practice, heavily dependent on holistic natural observation; and few studies include enough detail to permit the secondary analyst to comfortably change

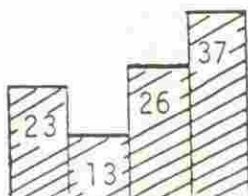


Figure 4. Frey *et al.*, N = 173 illustrated.

the definition adopted by the original investigators. Thus, it may not be easy to find alternatives in concrete situations. Fortunately, in this case, the cooperation of the original investigators was available. They provided the original data, the access to the background publications, and the guidance in their use which made the following reanalysis possible.⁸

The reanalysis of the Pakistan data that follows has three parts. They parallel subsections of the the section on "Data, variables and methods of analysis" in Frey *et al.*: (1) sample, (2) independent variables, and (3) dependent variable. In each case technical details (like the standard procedures used to divide the variables into quartile ranks) and specifics will be provided in the appendix at the end of this paper. The text will concentrate on the overview that is essential to the overall argument of the present paper.

Sample:

Frey *et al.* provide a concise description of their procedures:

The data reported here are taken from a larger study conducted by the Colorado State University Water Management Research Team in 1976 and were gathered through structured interviews with a sample of 387 farmers from 16 villages in the Sind and Punjab provinces of Pakistan.* Since Cancian (1976:1091; 1977a:5; 1979:4-8, 34-38) suggests that an adequate test of the thesis must be based on a sample that approximates a "community of reference," only Punjab cases with complete data on all relevant variables are considered in the present analysis. Eliminating cases on this basis, we present an analysis based on 173 cases representing 10 Punjab villages.† Clearly, such a procedure only allows for a crude approximation of a "community of reference." However, we make the assumption that such a bias does not se-

⁸ I am indebted to Scott Frey, David Freeman and Max Lowdermilk for their cooperation. Scott Frey was particularly helpful in this effort. In the course of reconstructing the tables from Frey *et al.* (1979) we found a typo in the paper: page 426, paragraph 2, last line should read "15 and 21" not "14 and 21."

riously hamper our efforts at providing a general assessment of the two hypotheses (1979:425).⁹

It is important to divide specification of the sample ($N = 173$) analysed by Frey *et al.* into three steps: 1. selection of the villages and the farmers (the original sample, $N = 387$), 2. reduction of the 16 villages studied to the 10 Punjab villages represented in the final sample (making $N = 242$), and 3. elimination of the cases with missing data (making $N = 173$). At each step the relation of the sample to a sociologically meaningful community of reference is, of course, likely to be affected.

In the first step the field investigators sought to represent farmers using irrigation systems in Punjab and Sind provinces. Their estimates indicate that there are more than 2.5 million farms and 25 million rural people in the area covered (Lowdermilk *et al.*, 1978:88). Thus the potential for diversity is tremendous, and, despite the homogeneity brought by the basic production technology common to all farmers included in the sample, it is likely that many communities of reference are involved. There is a dim hope that technological, political and social conditions are constant over this large area. If they are constant, we may imagine that each village is sociologically identical to the others such that a random sample of farmers from a random sample of villages will produce a sample representative of each village's population.

In the second step Frey *et al.* made a deliberate effort to reduce the likely diversity in the universe sampled. They eliminated all four villages from Sind province and two Punjab villages—in order to somewhat homogenize the communities sampled. As they say in the passage quoted above, this "procedure only allows for a crude approximation of a 'community of reference.'" While the universe including the 10 remaining villages is much bigger than the 100,000 people mentioned in the general discussion above, there is at least an increased possibility that diversity is sufficiently reduced to provide a sample that represents actual communities of reference in the natural world. As will become apparent, the outcomes of the Frey *et al.* analysis and the alternative analysis I make support such an assumption.

The third step that I see in the specification of the final Frey *et al.*

⁹The footnotes from the original are as important as the text. They are (1979:425ns): "The purpose of this larger study was to ascertain those factors (physical, social, economic, etc.) which act to undercut the farmer's capacity to realize full crop production potentials. For a discussion of the purposive sampling procedures utilized and related methodological points, see Lowdermilk *et al.* (1978:Volume VI:Appendix 1-A). †We do not present a separate analysis of the Sind data because the available sample of 75 cases is too small for a meaningful analysis."

I have changed the citation of Cancian (1979) in the body of the quote, substituting parallel sections of the published version for the 1977 draft cited by Frey *et al.*

sample involves the elimination of cases with incomplete data on some of the relevant variables (see the extended quote above). This procedure reduces the 242 farmers sampled from the 10 Punjab villages to the 173 cases analysed by Frey *et al.*; and it forms the basis for my response to their analysis. I will try to show that lower-ranking farmers are disproportionately represented among the 69 farmers eliminated; and that restoration of these farmers to the sample will produce results more like my upper middle rank conservatism and less like the Frey *et al.* lower middle rank conservatism based on their $N = 173$ sample of data from Pakistan.

To restore most of the farmers eliminated to the sample I modified Frey *et al.*'s data analysis procedures in two ways. The first (which restores 21 cases making $N = 194$) involves the independent variables. The second (which restores 38 additional cases making $N = 232$) involves the dependent variable. An overview of the procedures is given below, and, again, the details are recorded in the appendix.

Independent variables

Frey *et al.* (1979) used two measures of economic rank, or, two independent variables: "acres of land owned" (Area Owned), and "acres of land cultivatable" (Area Cultivable). Review of their data manipulation procedures with the help of Scott Frey revealed that their efforts to clean up the data included elimination of farmers who owned no land. These 21 farmers are active tenants; though they own no land, they are spread fairly evenly over the Area Cultivable distribution (eleven are above the median). Thus, eliminating them might significantly distort the sample of farm operators.¹⁰ When these 21 farmers are simply tacked on to the bottom of the distribution used by Frey *et al.* for the analysis based on Area Owned the picture suggested by Frey *et al.*'s analysis changes dramatically. Reanalysis of the augmented sample ($N = 194$) based on a new division into four ranks suggests that their finding is very sensitive to how the community of reference is specified.

Table 2 documents these changes. Only the first four rows are relevant to the argument so far. Row A ties the format of the table to the illustration developed in Table 1 and Figure 2. Row B shows the data as analyzed by Frey *et al.* It duplicates the Stage 1 part of the Area Owned, Quartiled distribution displayed in their published table (1979:428); and provides the numbers on which Figure 4 in this paper is based. Row C shows the 21 tenants (who include only one Stage 1 adopter) tacked on to the bottom of the Frey *et al.* distribu-

¹⁰ In some cases in my comparative study I eliminated individuals with a zero score on the independent variable (1979:30-33); but I hope those judgments are, all things considered, more defensible as efforts to define a community of reference than the decision made by Frey *et al.*

Table 2. Standardized percentage rates of adoption by rank, area owned, quartiled, stage 1^a

Sample	Rank					Total
	Added rank	Low	Low middle	High middle	High	
A. N = 400 (Table 1 example)		10.0 10	30.0 30	20.0 20	40.0 40	100.0 100
B. N = 173 (Frey <i>et al.</i>)		23.3 9	13.2 6	26.4 12	37.1 14	100.0 41
C. N = 194 (Frey <i>et al.</i> + 21 cases)	4.7 1	22.2 9	12.6 6	25.1 12	35.4 14	100.0 42
D. N = 194 (Tenants added)		14.5 6	25.8 12	23.4 12	36.3 19	100.0 49
E. N = 232 (New adoption scale)		15.8 10	26.3 15	23.7 15	34.2 21	100.0 61

^a Top figure is percent, bottom figure is the count in the cell. As explained in the text, Area Cultivated data and alternative cutting points on rank are shown in Figure 6. Rank cutting points by row are based on the following Ns A(400); B(173); C(same as B, then bottom rank added); D(242); E(242). For the adoption scale they are A,B,C(same); D(194); E(232). The percentages shown (standardized to sum to 100 to make them more easily comparable when data are lumpy) are based on the original adoption rates, that is, on the percent of the rank who are Stage 1 adopters.

tion. Note that the counts for the regular ranks are identical in Rows B and C. Row D shows the results of reanalyzing the augmented sample (N = 194) with the standard data handling procedures used in both my analyses and Frey *et al.*'s.

The new finding in Row D is illustrated in Figure 5. The low adoption rate of the low rank has reappeared, and the lowest point of the "dip" has moved up the rank distribution. What was clearly lower middle rank conservatism in the Frey *et al.* analysis is now much more like the upper middle rank conservatism I have found in a number of places.

Figure 6 (parts A and B) displays another analysis of the data underlying Rows B and D of Table 2. This alternative method of data analysis produces identical findings for both the Frey *et al.* sample

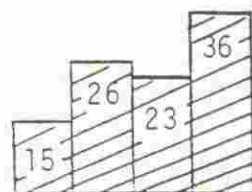


Figure 5. Augmented sample, N = 194 illustrated

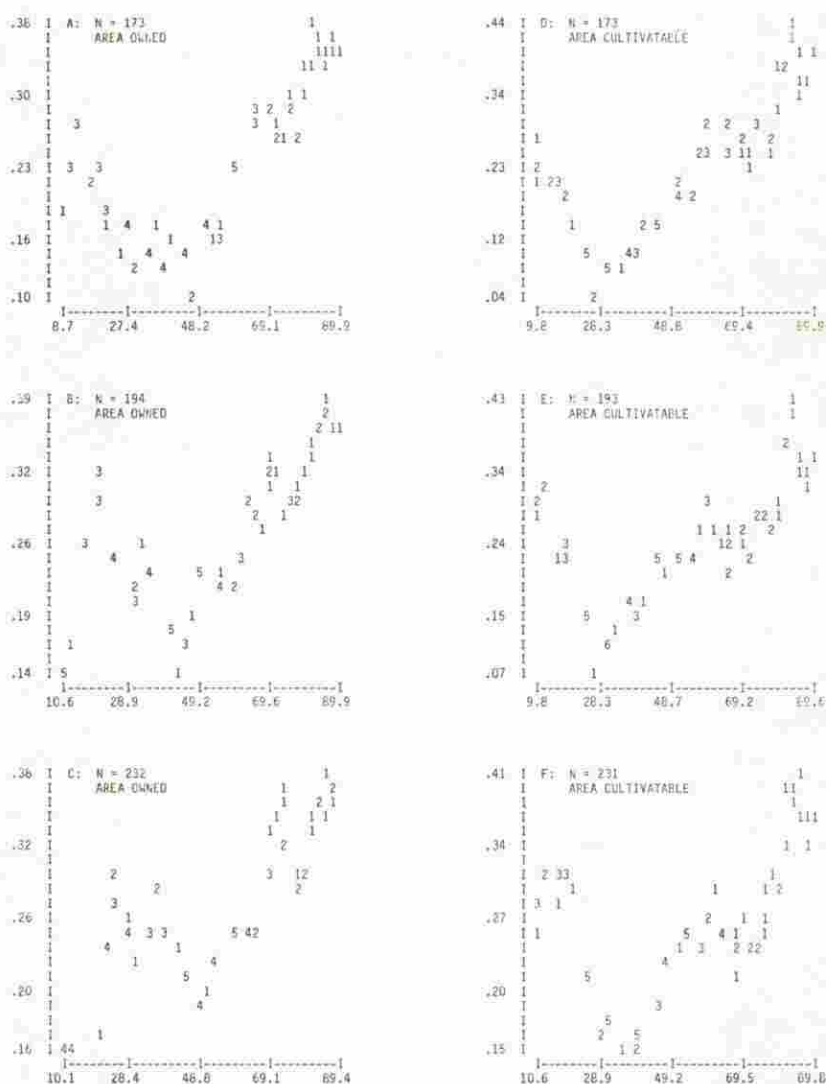


Figure 6.

(N = 173) and the augmented sample (N = 194). The low adoption rate of the low rank reappears and the dip moves up the rank variable (compare Figure 6 part B with part A). The scattergrams in Figure 6 are based on a moving average analysis (see appendix) that avoids some of the deficiencies of arbitrary division into quartiles or other categories appropriate to crosstabular analysis (Cancian, 1979; Morrison *et al.*, 1976). The scattergrams are included here because they:

1. illustrate the stability of the reanalysis based on the Area Owned data, and 2. permit economical presentation of the Area Cultivable analysis that is discussed immediately below.

My presentation of the tables is selective in two ways. One involves technical issues and is discussed in the appendix. The other points to a substantive problem. I have shown results for the Area Owned independent variable only. What about the Area Cultivable variable? In Figure 6 scattergrams for both Area Owned and Area Cultivable are displayed. As comparing these scattergrams (including parts C and E which represent the larger sample discussed below) suggests, crosstabular analysis shows that the finding of lower middle rank conservatism remains unchanged when Area Cultivable is used as the independent variable. I can do no better than *ad hoc* explanations for this difference between Area Owned and Area Cultivable, and, given my lack of direct knowledge of the institutional arrangements in the Punjab, I would rather leave *ad hoc* explanations to others. The pattern revealed in the reanalysis of the Area Owned variable is clear and stable, and that is enough for the purposes of the present paper.

Dependent variables

The sample size can be further increased by some reanalysis of the dependent variable. The addition of 38 more cases previously eliminated for lack of complete information on the dependent (adoption rate) variable was accomplished by changing the definition of the adoption scale used by Frey *et al.* The change permitted me to restore the cases for which very little information was missing (see appendix). Again, the missing cases fall disproportionately into the lower ranges of the Area Owned distribution (not as dramatically of course). The results of the crosstabular analysis with $N = 232$ for the Area Owned independent variable are shown in Row E of Table 2. They include the dip typical of upper middle rank conservatism. While, the crosstabular analysis shown in Row E of Table 2 is not substantially different from that in Row D, comparison of Figures 6B and 6C reveals that with $N = 232$ the pattern in the data is a bit closer to upper middle class conservatism than it was with $N = 194$.

In sum, when rank is measured by area of land owned, the Frey *et al.* finding of lower middle rank conservatism in Pakistan may be seen as an artifact of the community of reference they defined. The alternative analysis outlined above is based on a straightforward reinterpretation of the concrete situation, and it produces alternative results that are closer to previous tests of the theory in question. The pattern characteristic of upper middle rank conservatism is present in the Pakistan data; the anomaly is dispelled. If there is an anomaly that points to the need for further work, I think it is the clear and

stable differences between the pattern of adoption related to rank on Area Owned and Area Cultivable. As long as this difference involves only these Pakistan data, its explanation demands the knowledge of research procedures and local institutional arrangements possessed by the original investigators. Thus, it seems best left to them.

Conclusions

This paper points to the need to identify natural (empirical) social systems. It suggests that, sometimes at least, social behavior must be understood in terms of a locally defined system, that actors seem to eschew the multitudes known to them in the modern world and act in terms of a locally meaningful set of alters.

In substantive terms we have seen yet another demonstration that behavior is better understood when people are seen as occupants of social positions defined relative to a social system. The characteristics attached to people by other people (including social researchers) are signals of their social positions in local systems, and it is the behavior of these positions that we seek to understand. Thus we must recognize the system that gives the positions existence. Deriving this system from the combination of individual actions is not a viable strategy.

Yet, this system is difficult to identify, to "specify" or "operationalize" in concrete settings. The long-standing problem represented by this difficulty has sent anthropologists to remote islands, put sociologists in small rooms behind one-way mirrors and inspired countless clever observations by students of the natural scene. In this paper I have approached a concrete situation with the aid of a theory and comparative findings whose shape tests the validity of the specification chosen by the original investigators. But this particular solution is very limited. Even within the specialized concerns of rural stratification research, it is likely to be useless in a number of situations, including, for example, rural socioeconomic systems dominated by plantation agriculture (see Mintz, 1974). While theory will remain an indispensable guide to system specification, the judgment of the field investigator must not be trivialized merely because it is not subject to rigorous abstraction. Nor can the field investigators abandon responsibility for system definition to irrelevant sampling formulas. Rather, the problem of system definition should evoke a heightened sense of craft among social scientists.

Finally, the problem of system definition drives some to the philosophical extreme so candidly formulated by Wallerstein (1976). Having stated that, "What characterizes a social system in my view is the fact that life within it is largely self-contained, and that the dynamics of its development are largely internal" (1976:229), Wallerstein goes on to argue "... that *the only real social systems* are, on the one hand, those relatively small, highly autonomous subsistence economies not

part of some regular tribute-demanding system and, on the other hand, world-systems" (1976:230, emphasis added). I agree with Wallerstein that the natural world is not likely to provide neatly packaged systems adapted to such demanding analytic constructs. Only part of life is ever effectively "contained" in the local system. If we take most people or most institutions as starting points we will usually find important connections to the larger system. On the other hand, we must not allow analytic conundrums to blind us to empirical patterns. The fact that few empirical (natural) systems are as elegantly closed as the logical systems we define does not mean that we must substitute the entire world for the anthropologist's island. Evidence from studies like the one presented in this paper suggest that we should look at local systems, at communities of reference. These systems will be historically specific manifestations of universally relevant local social processes. The fact that these processes present particularly difficult problems is all the more reason for attention to them.

References

- Anderson, Dennis, and Mark W. Leiserson
 1980 "Rural nonfarm employment in developing countries." *Economic Development and Cultural Change* 28:227-48.
- Barlett, Peggy (ed.)
 1980 *Agricultural Decision Making: Anthropological Contributions to Rural Development*. New York: Academic Press.
- Barth, Fredrik (ed.)
 1978 *Scale and Social Organization*. Oslo, Norway: Universitetsforlaget.
- Cancian, Frank
 1967 "Stratification and risk-taking: a theory tested on agricultural innovation." *American Sociological Review* 32 (December):912-27.
 1972 *Change and Uncertainty in a Peasant Economy: The Maya Corn Farmers of Zinacantan*. Stanford, California: Stanford University Press.
 1976 "Reply to Morrison, Kumar, Rogers and Fliegel." *American Sociological Review* 41:1089-93.
 1977a "Stratification and risk-taking again." Paper presented at the annual meetings of the Rural Sociological Society, Madison, Wisconsin.
 1977b *The Innovator's Situation: Upper Middle Class Conservatism in Agricultural Communities*. Irvine, California: Social Science Working Paper No. 132.
 1979 *The Innovator's Situation: Upper-Middle-Class Conservatism in Agricultural Communities*. Stanford, California: Stanford University Press.
- Early, Alan C., Max K. Lowdermilk, and David M. Freeman
 1978 *Farm Irrigation Constraints and Farmer's Responses: Comprehensive Field Survey in Pakistan, Volume III*. Fort Collins, Colorado: Colorado State University.
- Freeman, David M., Max K. Lowdermilk, and Alan C. Early
 1978 *Farm Irrigation Constraints and Farmer's Responses: Comprehensive Field Survey in Pakistan, Volume VI*. Fort Collins, Colorado: Colorado State University.
- Frey, R. Scott, David M. Freeman, and Max K. Lowdermilk
 1979 "Cancian's 'upper middle class conservatism' thesis: a replication from Pakistan." *Rural Sociology* 44 (Summer):420-30.

- Gartrell, John W.
 1977 "Status, inequality and innovation: the green revolution in Andhra Pradesh, India." *American Sociological Review* 42 (April):318-37.
- Gartrell, John W., E. A. Wilkening, and H. A. Presser
 1973 "Curvilinear and linear models relating status and innovative behavior: a reassessment." *Rural Sociology* 38 (Winter):391-411.
- Kaplan, Barbara H. (ed.)
 1978 *Social Change in the Capitalist World Economy*. Beverly Hills: Sage.
- Lowdermilk, Max K., Alan C. Early, and David M. Freeman
 1978 *Farm Irrigation Constraints and Farmer's Responses: Comprehensive Field Survey in Pakistan, Volume II*. Fort Collins, Colorado: Colorado State University.
- Mintz, Sidney W.
 1974 "The rural proletariat and the problem of rural proletarian consciousness." *Journal of Peasant Studies* 1:291-325.
- Morrison, Denton E., Krishna Kumar, Everett M. Rogers, and Frederick C. Fliegel
 1976 "Stratification and risk taking: a further negative replication of Cancian's theory." *American Sociological Review* 41 (December):1083-89.
- Rogers, Everett M., with F. Floyd Shoemaker
 1971 *Communication of Innovations*. New York: The Free Press.
- Rossi, Peter H.
 1980 "The ups and downs of social class in America." *Contemporary Sociology* 9:40-44.
- Wallerstein, Immanuel
 1974 *The Modern World-System: Capitalist Agriculture and the Origins of the European World-Economy in the Sixteenth Century*. New York: Academic Press.

Appendix

This appendix has four parts: Sample, Independent Variables, Dependent Variables, and Stages.

Sample

The original research concerned farm water management in Pakistan, and the sample ($N = 387$) was selected accordingly. Freeman *et al.* say: "The sampling problem is one of conducting a study of watercourses that are known to be representative of the larger population of over 78,000 watercourses in the Punjab and Sind" (1978:11). Villages were selected for a number of reasons (see especially Freeman *et al.*, 1978:5-11), and within each village watercourses were identified. This produced a sample of 40 watercourses in 16 villages. While the researchers are careful to point out that they had no sampling frame of watercourse types that would justify statistical inference from their results, their intention to select a representative sample is clear. Given their goals, and the small absolute number of villages, we may hope that what is lost to lack of randomness is more than replaced by the considerable field experience of the investigators.

Of the many factors (including agroclimatic zone, crops grown, presence of tube-wells, caste distribution and geopolitical origins of the inhabitants) considered in the holistic judgment made in selection of villages, only one seems directly relevant to the distributions on the variables of concern here. Six of the 16 villages were chosen with the "requirement that there be no exceptionally large landlords present" (Freeman *et al.*, 1978:6). Since my concern is to reduce the heterogeneity of the villages that might be lumped into the same community of reference in order to avoid the mixing of non-commensurable absolute distributions on the independent variables, this sampling bias probably improves the sample.

Within the villages sampling was as follows: "... sample farmers were chosen at random but stratified according to watercourse position—the head, middle and tail thirds" (Freeman *et al.*, 1978:12). There is no reason to think that this criterion for stratification of the sample biases the sampling of farmers in any way of interest here; but, of course, every village has non-farmers who may be important members of the farmers' community of reference. Lowdermilk *et al.* report an estimate that "... 31 percent of village population are landless laborers, shopkeepers, and artisans" (1978:88). This percentage of nonfarmers seems high in comparison with Latin American situations with which I am familiar, but is not very different from many rural populations in which substantial proportions of the people are employed outside of agriculture (Anderson and Leiserson, 1980:229). Clearly, a very high ratio of landless laborers to landlords and/or farm operators would make the farmers a partial system for which the theory involving community of reference may not be appropriate. The size of the "non-farmer" component in the population may be a factor in determining the composition of the community of reference and the location of the "dip" that is of interest here, but I cannot evaluate the effect of small differences in size. They are unlikely to make substantial differences to the outcome of the analysis.

In sum, apart from the major questions raised by the likely heterogeneity of the 16 separate villages selected from a vast area, the sampling appears roughly comparable to many of the samples from various countries that are included in my recent comparative study (Cancian, 1979).

The second step in sample definition (reduction from $N = 387$ to $N = 242$) represents a Frey *et al.* effort to approximate community of reference. They substantially reduced the likely heterogeneity by eliminating all four of the Sind villages. In addition, two Punjab villages (including a total of 17 sampled farmers) which were felt to be inappropriate for a variety of reasons, including their exceptionally high percent of tenants among the farmers in the sample (Lowdermilk, personal communication), were also eliminated at this point, yielding the final Punjab sample of 242 farmers on 21 watercourses in 10 villages (villages 101–104, 106, 107, 109–112 in Table 5, Early *et al.*, 1978:11).

Reduction of the sample from $N = 242$ to $N = 173$ is discussed in the text.

Independent variables

Frey *et al.* replicate data manipulation procedures reported in Cancian (1979: Appendix A): cases are ordered on the independent variable and the distribution divided to best approximate four equal quartiles (25/25/25/25) in one specification of the theory, and four ranks from low to high in the proportion 30/30/20/20 in another specification. Frey *et al.* use both sets of cutting points on both independent variables (Area Owned and Area Cultivable) yielding the four tests of each hypothesis they report in their paper. (Frey *et al.*, 1979 did not follow the cutting point rules stated in Cancian, 1979:95 exactly, but the differences do not prove to be substantial.)

The 25/25 and 30/20 cutting points do not follow from the theory (Cancian, 1967). The 25/25 set was originally adopted for statistical, not theoretical, reasons, to maximize the size of the expected value of the smallest cell when analysis depended on studies with small samples. The 30/20 cutting points were originally added to respond to the observation that stratification systems often have a "triangular" character (Morrison *et al.*, 1976) in which (sociologically meaningful) higher "classes" are smaller than lower "classes." From the point of view of the theory, variations in the size of meaningful societal subdivisions (if any) and consequent prediction of the location of the "dip" are unspecified, and are taken to be dependent on historically specific, local institutional arrangements. Since the "dip" could be obscured by specification of cutting points not appropriate to the local situation, the best specification of the theory is one which avoids commitment to one or another set of cutting points. The "moving average twenty" procedure used in Figure 6 has this desirable characteristic. In it cases are assigned a percentile rank on the distribution of the independent variable and then

an adoption rate (proportion of Stage 1 adopters) is calculated for a moving "window" equal to 20 percent of the sample. The lowest "case" includes percentile ranks 0-20, the next one 1 to 21, and so on until the eighty-first includes percentile ranks 80 to the highest (for details, see Cancian, 1979: Appendix B). Plotting these 81 "cases" in the form illustrated in Figure 6 avoids most cutting point problems likely to be produced by unknown local institutional arrangements, and is therefore a better specification of the theory than the 25/25 or the 30/20 cutting points.

To complete the record: the 30/20 cuts for $N = 194$ and $N = 232$ using Area Owned yield a monotonic pattern which obscures the dip; they show neither the original nor the Frey *et al.* pattern.

Dependent variable

Frey *et al.* define a conventional adoption scale with a maximum total score of 36 (1979:426). It includes eight innovations (each with a possible score ranging from zero to 2, 3, 4, or 8). By eliminating all cases for which there was missing information on any of the eight component variables Frey *et al.* eliminated ten cases that could be rated for 22 or less of the 36 "points" on the scale and 38 cases that could be rated for 28 or more points. In order to include these 38 cases (making $N = 232$) I defined the "new" adoption scale score as the percentage of the maximum possible points given information on the case. This procedure permits the inclusion of the 38 cases with virtually complete information among the 194 cases with complete information—without changing the ordering of the original 194 cases.

Stages

Stage 1 and Stage 2 are defined as the first and second quartiles on the adoption scale respectively (Cancian, 1979; Frey *et al.*, 1979:427). The primary concern of this paper is with the hypothesis that a "dip" will be present in the distribution of the "innovators" (first 25 percent to adopt). The additional prediction that the relation will move towards a monotonic positive one in Stage 2 of the process was also not confirmed by either of the cutting points (25/25 and 30/20) used by Frey *et al.* on the $N = 173$ sample with the Area Owned independent variable. For Area Owned it is confirmed for $N = 194$ and $N = 232$ for the 25/25 cutting points and not confirmed for the 30/20 cutting points. No systematic procedure for evaluating the difference between Stage 1 and Stage 2 is presently available for the "moving average twenty" analysis.