11

Targeting Production Systems in the Small Ruminant CRSP: A Typology Using Cluster Analysis

Keith A. Jamtgaard

Agricultural R&D programs that propose to alter production practices in some fashion are faced with the prior task of identifying the potential beneficiaries of their efforts. This typically involves choices within three criterial areas: broad policy questions; socioorganizational structures; and production systems. An example of the first might be whether research on a given topic (e.g., small ruminants) is needed in the first place, and, if so, in which countries. Within the countries selected, political-economic, as well as scientific, criteria may be considered in targeting populations and regions. Even after these policy choices have been made, much of the work of targeting still remains, however.

The second step centers on diversity in the social organization of agricultural production systems within the R&D area. This requires choosing among different types of producers of a commodity, or, at the very least, being aware that different social relations of production may limit the usefulness of given technologies. In Peru, for example, systems with very different social relations of production include independent commodity producers, cooperatives, plantations, and peasant communities.

The third step is to target beneficiaries by production systems. Commodity-oriented R&D might be presumed to hold an advantage over broader spectrum approaches such as FSR (farming systems research) since they can simply target "the producers of commodity X," but, in fact, commodity programs may encounter *more* difficulties. FSR typically targets a single socioorganizational type of producer, i.e, "peasants." Moreover, FSR recognizes that peasants usually manage risk by raising a variety of plant and animal species. Thus, from the outset, FSR is sensitive to the complexity of peasant production systems. (From this standpoint, perhaps one of FSR's shortcomings is that the simplicity gained by targeting production systems is traded for increased technical complexity since the whole system must be addressed—not just one commodity within it.) Even so, FSR projects still must choose among production systems (Bernsten et al. 1984).

Commodity-oriented programs face an analogous problem. A single commodity can fit into many different production systems. The question is which of the many systems incorporating the commodity to target. This chapter describes and evaluates a set of empirical procedures devised by SR-CRSP sociologists that helped answer this question for the SR-CRSP/Peru. This case is instructive for other agricultural R&D initiatives faced with difficulties in defining target populations.

A TARGET POPULATION FOR THE SR-CRSP/PERU

Diversity in the Social Organization of Production

Peru manifests enormous socioorganizational and environmental diversity in production, even within a single category such as "peasants." Small-scale independent farmers work irrigated river basins in the coastal desert. Only a few hours away, peasant communities (*comunidades campesinas*, or CCs) cultivate mountain slopes at over 3,600 m in the high Andes. Farther to the east, medium-sized farmers in the Amazon basin pursue a thoroughly distinct tropical agriculture. Large cooperative enterprises created by the agrarian reform of 1968–1980 also operate throughout the major agroecological zones of the country.

Each of these forms of production is embedded in a fundamentally different social structure, with distinct relations of production, legal structures, linkages with the state, and scales of operation. For instance, the cooperative sector is an assortment of entities constructed primarily from the large haciendas expropriated by the central government during the agrarian reform. They are still closely affiliated with the state. Private producers, whom the government perceives as being among the most productive farmers, have also benefited from government policies aimed at increasing agricultural outputs.

Peru's peasant communities, however, are the most numerous of the rural sector. From the beginning of the SR-CRSP/Peru in 1980, it was clear that CCs were significant producers of livestock, holding an estimated 52% of the nation's sheep; another 15% of the national flock are owned by cooperative institutions, and the remaining 33% by independent producers (DCCN 1980).[1] As much as 80% of Peru's alpaca herds are in the hands of peasant producers (Vidal and Grados 1974, cited in Flores Ochoa 1977:41). Moreover, about 44% of all alpaca are raised within officially recognized CCs[2] (DCCN 1980). Peasant communities also play a commanding role in producing Peru's major plant food staples, notably potatoes, barley, and maize (DCCN 1980).

Diversity in Production Systems

Despite their numerical and economic importance, peasant communities have been historically disfavored by development projects, agrarian policymakers, and credit institutions. Given the SR-CRSP mandate to assist the "poorest of the poor," however, such communities constituted the program's logical target group. Yet, even after narrowing its socioorganizational choices to CCs, the SR-CRSP still faced difficulties in specifying its target population. Two problems often arise when generalizing about cropping and animal husbandry in Peruvian CCs; both result from the tremendous environmental variation that exists from one end of the country to the other—or even within a single community, from its highland pastures over 4,000 m to the valley floor 1,000 m below.

This variation obfuscates comparisons of data from one community or region with basic production parameters from the larger population of all CCs. Moreover, when designing development programs with applicability to some subset of CCs, it is exceedingly difficult to distinguish even the most general production differences among communities. The tendency has therefore been to view Andean peasant communities as impossibly diverse and to confine observations to individual communities or small regions, or, conversely, to make monolithic generalizations about all CCs.

Nevertheless, to target its R&D population, the SR-CRSP/Peru still needed to answer two questions. The first was: How important are small ruminants in the economy of different types of peasant communities? From the very beginning of program activities in Peru, two general types of CC production systems were evident: pastoral and agropastoral.

Peruvian peasants everywhere value small ruminants for their ability to utilize high-altitude grasslands and other areas not under cultivation. In many highland CCs in the central Andes, people's livelihood primarily depends on their herds of alpaca, llama, and sheep; these communities may be characterized as "pastoral." However, small ruminants are also important for agropastoral CCs. While many such communities likewise utilize highland pastures, they often follow a rotational fallowing system (Custred and Orlove 1974; Orlove and Godoy 1986) in which fallow fields are grazed and manured by herds, and crop residues are a critical dry-season feed resource for herds (Jamtgaard 1984). In fact, small ruminants and the manure they provide are criterial to the continued functioning of this production system (Winterhalder et al. 1974).

Animal husbandry is subject to quite different constraints under these two production systems. For example, since agropastoral households actively engage in both cultivation and herding, their labor needs are very different from those of households pursuing only one or the other (Orlove 1977; Vincze 1980). This presents both opportunities and costs. As noted above, plant and animal crops enjoy some mutual benefits in agropastoralism. At the same time, however, the two compete for land and labor, thus necessitating complex mechanisms for integrating the two sectors of production (McCorkle 1986, 1987). Awareness of such constraints is critical in designing successful interventions to increase outputs from the CC livestock sector.

The second question the SR-CRSP needed to answer was: Which of these two types of peasant communities controls more small ruminants? In other words, given limited program resources, which group should be targeted? In the absence of any solid information, it was initially assumed that pastoral communities held more small ruminants and should therefore be the primary target group. But SR-CRSP social scientists pointed out that the program could have greater impact if the universe of small ruminant producers could be empirically delineated and the major producer types defined.

Gathering firsthand data on a population as large and diverse as that of all Peruvian peasant communities was manifestly impractical. However, program sociologists located an exceptionally rich data set in Peru's Dirección de Comunidades Campesinas y Nativas (DCCN), which generously made this information available to the SR-CRSP. These data derived from a 1977 survey that recorded important production and other indicators in 2,716 CCs, or 99% of all officially recognized peasant communities at the time (DCCN 1980).[3] For Peru, this is a unique data set, both because its scope is so broad and because its unit of analysis is the peasant community. With this information, SR-CRSP sociologists were able to elaborate a useful typology of CC production systems.

A PRODUCTION SYSTEMS TYPOLOGY

Approaches to typology construction are traditionally classed as heuristic or empirical. In the former, categories are delineated by reference to a theoretical framework, and the researcher essentially specifies the criteria for bounding the categories. In the latter, categories are developed to conform to salient differences within the data themselves, often employing algorithms such as cluster analysis. However, this heuristic/empirical dichotomy is less useful than are approaches that directly consider the need to measure objects and assign them to groups (Bailey 1973). If research includes a stage in which observations will be assigned to categories, and the objects to be classified lack features that conclusively locate them in one or another type, then typology construction should come after measurement. The goal should be to achieve the best fit between the categories needed and the empirical observations.

For SR-CRSP sociologists, analysis of Peruvian CCs began with an image of different theoretical categories: pastoral; agropastoral; and agricultural. However, these served mainly as guideposts for evaluating the results of the empirical analysis. Cluster analysis was selected for this task because of the lack of criteria for clearly delimiting boundaries among these theoretical categories. Two kinds of production indicators from the DCCN study formed the basis for typology construction: CC herd populations by species, and hectares of principal plant crops under cultivation in each CC.[4]

In the vertical ecology of the Andes, production of many of the most common plant and animal species is altitudinally bounded (Custred 1977; Dollfus 1981; Gade 1975). Knowing which species a community raises usually provides some basic information about its ecological resources. For instance, camelids (especially alpaca) are today most often found above 4,100 m. Sheep and potatoes are increasingly important at the lower limits of this zone (about 3,900 m). Barley, wheat, and broadbeans are the chief crops between 3,900 and 3,300 m, and maize dominates the zone between 3,300 and 2,400 m. Cultigens like sugarcane, fruit trees, and coffee are generally grown at lower altitudes.[5] Therefore, certain production figures can sometimes furnish a crude indicator of the ecozones exploited by a community. If a CC primarily produces livestock, its access to arable land is likely to be minimal. Conversely, many maize-growing CCs lack access to the high-altitude rangelands necessary for significant livestock production.

In reality, communities display enormous diversity in their particular combination of ecozone access and utilization. Anthropologists have documented the historic Andean ideal of maintaining vertical control over multiple ecozones (Masuda et al. 1985; Murra 1972). Many contemporary peasant communities still do so (Brush 1977; Masuda 1981; and many others). Hence, the typology presented here is not claimed to represent any absolute or "true" characterization of CC production systems. SR-CRSP sociologists had a specific goal: to reduce the great variation in CC systems to relatively few categories capturing principal differences among them. As Everitt (1980:6, italics his) notes:

> [I]n many fields the research worker is faced with a great bulk of observations which are quite intractable unless classified into manageable groups, which in some sense can be treated as units. Clustering techniques can be used to perform this *data reduction*. . . . In this way it may be possible to give a more concise and understandable account of the observations under consideration. In other words simplification with minimal loss of information is sought.

Procedures

Analysis was performed in four stages: (1) selection of the variables to be analyzed; (2) data preparation, including logarithmic transformation, standardization of variables, and treatment of outliers; (3) factor analysis in order, to collapse the number of variables into frequently occurring combinations; and (4) cluster analysis of the scores derived from the factor analysis.

Selection of variables. Analysis began with the full range of production indicators listed in Table 11.1. The DCCN study incorporated additional data on forests, overall community area, native pastures, and human demographics, but these were omitted in the SR-CRSP analysis because they lacked the same sense of "production." If the goal of this undertaking had been to develop a typology of natural resources, or to classify communities according to overall production potentials, then including these and other measures might have been desirable. But the SR-CRSP's aim was to define and rank production systems in terms of small ruminant husbandry.

Data preparation. Nearly all of the production indicators listed in Table 11.1 had highly skewed distributions. For example, while 97% of CCs raised some sheep, just three communities accounted for over 5% of the total 7,807,851 head. The median number of sheep per community was 1,000, with a mean of 2,875 - also indicating a highly skewed distribution. Initial attempts at clustering suggested that a relatively small proportion of communities were unduly influencing the results. The exact proportion of CCs with high values varied by plant and animal species, averaging about 10% for each species. Since the communities exhibiting extreme values differed from one species to another, too many CCs were involved simply to remove them all from analysis.

This problem was solved with a logarithmic transformation of the variables. In cluster analysis, the "arbitrariness involved in scaling and combining different variables" means that "there is rarely any justification for using the particular values rather than values obtained from some monotonic transformation; for example, their logarithms or square roots" (Everitt 1980:68). Transforming production indicators to their logarithms dramatically reduced the effect of extreme values, while retaining a semblance of their original variation.

Another problem was that the variables displayed widely differing scales. In order to permit joint analysis of such disparate indicators as "hectares of barley" and "head of sheep," these were standardized to a mean of 0 and an SD (standard deviation) of 1. This was also helpful in scoring the variables for cluster analysis, since the Euclidean D dissimilarity measure that was employed in this analysis is sensitive to differences of scale (Everitt 1980).

No attempt was made to standardize the data with respect to size criteria, such as community land area or human population; that is, production indicators were not adjusted to form such ratios as "sheep per hectare of

TABLE 11.1. PRODUCTION INDICATORS COLLECTED IN THE DCCN SURVEY

Livestock (Head)	Crops (Hectares)
Cattle[a]	Potatoes[a]
Sheep	Maize
Goats	Barley
Llama and alpaca (combined)	Wheat
Swine[a]	Alfalfa
Burros, horses, and mules (combined)	Broad beans
	Coffee
	Rice[b]
	Tobacco[b]
	Sugarcane
	Oranges

[a]These indicators had loadings of .40 or above on more than one factor during factor analysis, and were therefore dropped.

[b]These indicators had communality estimates of .15 or lower during factor analysis, and were therefore also dropped.

community land" or "hectares of maize per inhabitant." This might have given a more accurate image of the actual deployment of resources, particularly in smaller CCs, but it would have eliminated the effect of the volume of production itself, which was also important.

Taken together, the foregoing steps permitted comparisons among variables while still signaling whether a community was a large- or small-scale producer. The next step was to exclude outlier cases and CCs with insufficient data. Only eight CCs registered zero on each of the variables of interest and hence were excluded prior to the logarithmic transformation. To identify outliers, a disjoint cluster analysis was performed with 50 clusters specified; clusters consisting of only one observation were then removed. Four CCs were eliminated in this manner. Finally, the variables for the remaining 2,704 CCs were once again standardized.

Factor analysis. A factor analysis was performed prior to clustering[6] in order to determine which variables or groups of variables would best capture differences between production systems and to organize this information in a compact form. In this stage of analysis, many different solutions were iteratively examined, and a number of indicators were eliminated rather quickly (Table 11.1). For example, those for swine, cattle, and potatoes were dropped because they were found in many combinations of production systems, and hence did not characterize any one system. For the opposite reason (i.e., nonco-occurrence with any other indicators), rice and tobacco were also dropped.[7] This operation greatly reduced the number of variables, thus facilitating cluster analysis both in terms of computing resources and in the interpretation of results.

A "varimax" rotation was also performed; this provided a much clearer identification of variables to factors. Since the eigenvalue noticeably dropped from the fourth to the fifth factor, a four-factor solution was chosen. Each of the four factors had an eigenvalue greater than 1 following rotation.

Next, factor-based scores were computed. These were used instead of common factor scores because of the likelihood of measurement error in the data. Also, using all of the information from variables with smaller factor loadings might be misleading (Kim and Mueller 1978). As it turned out, each of the four factors had three variables loading on it (Table 11.2). The observations were assigned factor-based scores by multiplying the standardized values by 1 for each variable with a high loading, and by 0 for the others. The results were then summed for each factor. Each of these factor scores had a mean of 0 and an SD of about 2.3 (Table 11.2).

The factor-based scores also incorporate a sense of production scale. Higher figures indicate greater commitment to the production activities that make up the factor while lower figures point to their absence. However, at this stage of analysis, a community that ranks high on one factor can rank even higher on another. A CC's score on each of these factors simply indicates the relative importance of that kind of production vis-à-vis the population of CCs studied. Zero indicates that a CC scored close to the population mean; a positive or negative number means it scored above or below the mean, respectively.

Given the strong relationship in the Andes between vertical ecozone and production activity, labels were tentatively assigned to the four factors in Table 11.2 based on the production zone best represented by the variables emerging from the factor analysis. Sierran agriculture (I) was assigned its title because three of the principal, nonpotato crops (barley, wheat, and broadbeans) are produced above 3,300 m, often without irrigation. A high score on this factor signals large hectarages planted to these crops, but it may mean either major production of only one crop or minor production of some combination of the three.

Although most of Peru's 2,716 CCs lie in the Andes, some are found on the coast and on the eastern slopes of the mountains.[8] Nonsierran agriculture (II) represents three crops typically raised at lower altitudes—coffee, sugarcane, and oranges. A high score on this factor simply indicates a CC's

TABLE 11.2. CONFIGURATION OF THE FOUR FACTORS USED IN SUBSEQUENT ANALYSES

	Factor Label	Components[a]
I.	Sierran Agriculture	Hectares of barley, wheat, and broad beans (SD = 2.4)
II.	Non-Sierran Agriculture	Hectares of coffee, sugarcane, and orange trees (SD = 2.3)
III.	Intermontane Valley	Hectares of maize, alfalfa, and head of goats (SD = 2.2)
IV.	Livestock	Head of sheep, camelids, horses, and burros (SD = 2.2)

[a]Factor scores were computed by summing the multiplication of the standardized scores of each of the variables identified with the factor by 1, and for the variables not identified with the factor, by zero. They each have a mean of zero. Standard deviations (SD) varied as indicated.

substantial commitment to these crops relative to the total population of predominantly Andean CCs.

Probably the most difficult factor to label was III. A key distinction among CCs was the presence of maize fields. Alfalfa and goats were often associated with maize.[9] All three of these crops are frequently raised in the Andean mountain valleys; hence the name intermontane valley.

The livestock factor (IV) likewise implied access to a particular altitudinal zone. Since most sierran communities primarily rely on extensive grazing, and since mountain rangelands are the principal feed source for their herds, a high score on this factor suggested access to native grasslands, usually located above the limits of cultivation.

Cluster analysis. In this stage, the four factors were used to generalize about CCs' involvement in different production sectors by developing a typology of the combinations of factor-based scores across all of the sample CCs. From a technical perspective, a challenging feature of this undertaking was the large number of observations to be classified. Cluster analysis is not a single technique, but rather a family of algorithms that group observations according to criteria of similarity or difference. However, analytic alternatives rapidly shrink when numerous observations are to be classified. This practically necessitated the use of a nonhierarchical clustering algorithm. The procedure selected was based on the k-means algorithm (MacQueen 1967),[10] employing Anderberg's (1973) centroid sorting method as implemented in FASTCLUS of SAS version 82.3. Euclidean distance was the measure of dissimilarity.

A major uncertainty in this or any cluster analysis is how many groups to accept since this is equivalent to determining how many categories the typology will have. This decision must therefore be carefully considered. After testing numerous possibilities, including solutions ranging between four and 20 groups, a 14-group solution was accepted[11] (Table 11.3); but as in many statistical techniques, objective criteria offer little "proof" of one typology's superiority over any other. The final decision is largely subjective. In this analysis, solutions with fewer groups seemed to mask important differences among production systems, while those with more groups seemed to dwell on minor variations in scales of production rather than on new combinations of systems or substantial scale differences within already defined systems.

The 14 clusters can themselves be used as "building blocks" for higher-level generalizations. Indeed, some sort of generalization is necessary to answer the SR-CRSP's initial question about the importance of agropastoral communities for small ruminant production in Peru; hence, Table 11.3's aggregation of the clusters into four broader categories: Lowland, Agropastoral, Pastoral, and Agricultural.

Perhaps the most distinctive feature of this typology (and of the alternative solutions examined) is the numerous clusters for lowland CC production systems relative to the small number (123) of CCs involved. Of the 14 clusters identified by the algorithm, six had noticeably high scores on factor **II**. This is neither an important finding nor a problem for understanding the other categories. It is merely a consequence of including an entire factor just to distinguish a few CCs.

Eight clusters emerged for the numerically more important highland CCs. From Table 11.3, clusters 7, 8, and 9 were typed as Agropastoral. Compared to the other clusters, they had important activities in both plant and animal agriculture. CCs in cluster 7 had major commitments to factors **III** and **IV**, and a lesser one to **I**. This contrasts moderately with cluster 8's strong emphasis on **I**, diminished involvement in **IV**, and nonparticipation in **III**. Cluster 9 clearly represents the largest highland CCs, with major investments in all sierran production sectors—factors **I**, **III**, and **IV**.

Two clusters were classed as Pastoral. The first (10) is a fairly clear-cut case of CCs with substantial livestock activities and little more. CCs in cluster 11 simply appeared to be more involved with livestock than anything else. Note that size of production is a consideration here; cluster 11 appears to be primarily composed of small highland CCs.

The three remaining clusters (12, 13, 14) were categorized as Agricultural because of their low scores on factor **IV**. Cluster 12 represented CCs with large investments in **III**, but little else. Cluster 13 also scored high on **III**, but even higher on **I**. CCs in cluster 14 paralleled those in cluster 11 in their low scores on all factors. Discounting cluster 14's score on

TABLE 11.3. MEAN SCORES ON FOUR MEASURES FOR 14-CLUSTER SOLUTION, GROUPED BY GENERAL CATEGORIES[a]

Category Label	Cluster	N	%[b]	Factor I Sierran Agri- culture	Factor II Non-Sierran Agri- culture	Factor III Inter-Montane Valley	Factor IV Livestock
Lowland	1	9	.3	-1.95344	24.96425	1.20431	-0.18355
	2	19	.7	-0.84408	8.91746	1.88506	0.74285
	3	38	1.4	-2.14259	3.53655	0.42143	-1.42240
	4	24	.9	-2.09161	14.07012	1.17883	-1.03576
	5	14	.5	-2.15002	8.65796	-0.06523	-4.73965
	6	19	.7	2.85802	5.47319	2.63845	0.43129
		123	4.5				
Agropastoral	7	273	10.1	0.58379	-0.41116	2.54995	1.98740
	8	296	10.9	2.77679	-0.43071	-1.64558	0.47271
	9	148	5.5	5.29509	-0.37591	3.51572	2.03488
		717	26.5				
Pastoral	10	350	12.9	-1.82401	-0.43258	-1.70847	2.87303
	11	539	19.9	-1.12328	-0.43220	-1.82030	-0.21976
		889	32.8				
Agricultural	12	338	12.5	-1.52349	-0.41930	1.77389	-0.77548
	13	288	10.7	2.13457	-0.43058	1.15632	-1.21898
	14	349	12.9	-1.37510	-0.41812	-0.63908	-3.24633
		975	36.1				

[a]The 14 categories derived from the cluster analysis have been reordered under the labels provided to reflect the interpretation given here.

[b]Percents do not always sum to 100 due to rounding.

II, which is already at its minimum, its next highest score was on III. Thus, cluster 14 might best be described as very small CCs with some production emphasis in maize, alfalfa, and goats.

Discussion

Table 11.3 indicates that of the 2,704 CCs analyzed, the largest number were Agricultural (975, or 36%). The second largest type consisted of Pastoral communities (a third of the total). Agropastoral CCs accounted for 717, or

27% of the population. Finally, 123 communities were categorized as Lowland.

SR-CRSP social scientists' original question concerned the distribution of plant and animal resources across different types of production systems. Table 11.3 is suggestive in this regard, but not conclusive. Since we already know that many of the CCs typed as Pastoral or Agricultural are small (clusters 11 and 14, respectively), simply knowing numbers of CCs may not be particularly helpful. More conclusive information may be obtained by examining the values of the original crop and livestock population figures for the four categories.

Table 11.4 shows that Pastoral communities are of primary importance in camelid production. They hold nearly three-fourths of the llama and alpaca found in the 2,704 CCs. The remaining fourth is held by Agropastoral CCs. However, Pastoral and Agropastoral communities are equally important in terms of sheep population, with 45% and 44%, respectively, of the flocks in the sample. Cattle are more evenly distributed across different production systems. But even here, Agropastoral CCs hold a dominant position, with 47% of all cattle.

Agropastoral communities are important actors in plant crops, too. Across the three key crops (potatoes, barley, and maize), Agropastoral CCs are outstripped by Agricultural CCs only in maize. Agropastoralists control about half of potato and over two-thirds of barley production. Moreover, Agropastoralists make up over a third of all inhabitants in the sample CCs (Table 11.5),[12] thus representing the most important production system in terms of human subsistence production as well.

For other R&D programs wishing to duplicate these procedures, a question arises as to what constitutes suitable data and whether such data are likely to be available. For the case described here, it would be difficult to imagine a better information source. The DCCN study addressed the same unit of analysis as did the SR-CRSP; it gathered the kind of production data needed; and it was relatively current. But if these data had not been available, how useful might alternative sources have been?

Even though data may not be available according to the desired unit of analysis (whether peasant communities, individual farmers, cooperatives, or other units), they can still be useful. When a data set mixes different socioorganizational types of producers, additional information on the degree to which each type controls production in the aggregate unit would be required. One possibility would be to include units with a minimum predetermined level of participation in the production variable of interest. Alternatively, the procedures described here could be applied, but with careful examination of each cluster for the degree to which the socioorganizational type of interest is present.[13]

TABLE 11.4. AGRICULTURAL PRODUCTION INDICATORS BY PRODUCTION SYSTEM TYPE

A. Animal Crops

	Sheep		Cattle		Camelids	
Production System	Head	%[a]	Head	%	Head	%
Lowland	178,436	2.3	170,733	6.5	1,450	0.1
Agropastoral	3,502,251	45.1	1,230,090	46.6	368,864	26.8
Pastoral	3,416,596	44.0	729,207	27.6	989,428	72.0
Agricultural	659,968	8.5	507,686	19.2	15,228	1.1
Total	7,757,251	99.9	2,637,716	99.9	1,374,970	100.0

B. Plant Crops

	Potatoes		Maize		Barley	
Production System	Ha	%	Ha	%	Ha	%
Lowland	8,175	2.6	34,320	15.7	1,555	1.3
Agropastoral	157,792	50.4	88,794	40.6	83,882	68.0
Pastoral	94,189	30.1	6,059	2.8	16,601	13.5
Agricultural	52,874	16.9	89,436	40.9	21,381	17.3
Total	313,030	100.0	218,609	100.0	123,419	100.1

[a]Percents do not always sum to 100 due to rounding.

TABLE 11.5. HUMAN POPULATION[a] BY PRODUCTION SYSTEM TYPE

	Population	
Production System	N	%
Lowland	263,137	10.2
Agropastoral	895,583	34.6
Pastoral	654,690	25.3
Agricultural	773,826	29.9
Total	2,587,236	100.0

[a]Population data were obtained from the 1972 census as published in DGOR 1977, and then integrated with the production typology discussed in the text.

Other problems concern the content of the data gathered. Even in the absence of desired production indicators, valuable insights can be gleaned. For instance, data on camelids disaggregated by alpaca and llama would have been useful for the SR-CRSP since these species are often raised in somewhat different ecozones. Such information might have clarified the factor-based scores and otherwise enhanced the analysis. Even so, the simple inclusion of aggregate data on camelids significantly contributed to typology development.

CONCLUSION

The identification and enumeration of major producer types helps target limited research resources to those beneficiaries who best match the goals of a project. On the SR-CRSP/Peru, it was initially assumed that pastoral communities owned most of the livestock held by Peruvian peasants. Through careful statistical analysis of empirical data, however, SR-CRSP sociologists demonstrated that this supposition was in error. Peruvian agropastoralists are nearly equally important producers of livestock. Hence, they needed to be included in the program as well.

Based on these and other findings, the program focused its efforts to validate livestock technologies for peasant communities on the dual character of small ruminant production in the Andes: pastoral and agropastoral. Sites for field research were therefore selected to represent these two very different groups of producers. Recommendations for interventions to improve small ruminant production in Peruvian peasant communities now draw upon field research and experimentation in these sites.

Such findings might be taken to mean that scarce R&D resources must be thinly spread across very different kinds of producers, but, in fact, this kind of analysis can conserve limited resources since it allows projects to more tightly target their efforts on a reduced set of like producers. Other R&D programs can apply the procedures described here to do the same.

The usefulness of such analyses lies not only in the typology generated, but also in the identification of producer units falling into each of the categories. This makes sampling from a large population easier, more accurate, and more cost-effective.[14] Added benefits are increased understanding of the characteristics of the target population; greater awareness of the limits to generalizing from research results; and a set of parameters that can serve as benchmarks for monitoring and evaluating changes in production. These represent just a few kinds of contributions that social scientists can and do make to the sensitive design and successful implementation of international agricultural research and development.

NOTES

This study was conducted as part of the USAID Title XII SR-CRSP under grant numbers AID/DSAN/XII-G-0049 and AID/DAN/1328-G-SS-4093-00 in collaboration with the Instituto Nacional de Investigación y Promoción Agropecuaria (INIPA). Additional support was provided by the University of Missouri–Columbia. The author gratefully acknowledges the contribution of DCCN members José Portugal, Victoriano Cáceres, Ivan Pardo Figueroa, and Juana Jeri. Thanks are also due Mario Tapia and Jorge Flores for encouragement in locating the data source.

1. Production data disaggregated by socioorganizational criteria are rare. These rough estimates were obtained by combining figures on livestock transferred to the associative sector toward the end of the agrarian reform (Caballero and Álvarez 1980) with figures on livestock owned by officially recognized peasant communities (DCCN 1980). The remainder was attributed to independent producers.

2. Likewise, these estimates are confounded by the fact that many alpaca producers reside in peasant communities unrecognized officially.

3. The DCCN study sought to evaluate the effects of the agrarian reform, when the central government expropriated most of the large, privately held haciendas in Peru, formed cooperative enterprises on these lands, and in some cases distributed land to neighboring peasant communities.

4. One question in this approach is: what relevance do production indicators have across communities? To give an example, all areas planted to barley are not equal. Soil quality, management practices, water availability, and still other variables can account for great production differences. Likewise for livestock; many factors combine to determine the yield from different herds of the same size and species. Still, certain basic tasks in raising a given plant or animal species impose some similar constraints upon its producers regardless of ecozone. As in FSR, the truly critical part of analysis is understanding the particular array of plants and animals exploited, along with their relative importance within the production system as a whole.

5. These altitudinal boundaries represent the upper limits for Andean cultigens, with livestock occupying the nonarable lands above. There appear to be no effective lower ecological limits for many plant or animal crops, perhaps including alpaca (Flores Ochoa 1982). Most small ruminants can be produced on land suitable for maize, although Andean peasant common sense and, indeed, agroecological rationality dictate against this. Opportunity costs, of which peasants are keenly aware, may serve as more effective limits.

6. Either principal components or common factor analysis is often used prior to cluster analysis (Dowling 1987). Factor analysis was chosen in this case because of its greater flexibility in handling measurement error.

7. Interestingly, these results suggest an approach to distinguishing monocultural production systems, though this alternative was not pursued since monocultural community production systems are few in Peru and are largely located at lower altitudes.

8. The numerous indigenous settlements of the Amazon Basin (*comunidades nativas*) differ from CCs in both socioorganizational structure and legal status. However, some CCS are located at the edge of the jungle region, as well as along the coast.

9. This does *not* mean that numerous CCs in Peru supplement caprine

diets with maize and alfalfa, but simply that the three activities co-occur with sufficient frequency to be considered together. The label attached to the factor is less important for this analysis than is the usefulness of the factor for distinguishing production systems.

10. The k-means algorithm is sensitive to the ordering of the data (Milligan 1980), particularly for data sets with less than a hundred observations (SAS Institute 1982). However, it provides satisfactory results when compared to other iterative and hierarchical cluster techniques.

11. After 18 iterations, no observations shifted to new clusters, thus terminating the procedure.

12. In previous publications (DGOR 1977), data from Peru's 1972 population census were organized by peasant community. This analysis shows how the 1972 population was distributed across the production system categories discussed here.

13. A danger with this kind of aggregate data is the "ecological fallacy" (Robinson 1950), although proper specification of the analysis can greatly reduce this problem, too (Langbein and Lichtman 1978).

14. A template has been developed for use with spreadsheet programs that essentially performs this function by incorporating the key features of the procedures described here. After entering production data from a real or hypothetical observation (e.g., a CC), one can quickly learn which typological category most closely matches the observation. By slightly varying the different indices, one can also detect how near the boundary of a category an observation is located.

REFERENCES

Anderberg, M. 1973. *Cluster Analysis for Applications.* New York: Academic Press.

Bailey, K. D. 1973. Monothetic and Polythetic Typologies and their Relation to Conceptualization, Measurement and Scaling. *American Sociological Review* 38:18–33.

Bernsten, R. H., H. A. Fitzhugh, and H. C. Knipscheer. 1984. Livestock in Farming Systems Research. In Proceedings of Kansas State University's 1983 FSR Symposium. Cornelia Butler Flora, ed., pp. 64–109. Manhattan, KS: Kansas State University.

Brush, Stephen B. 1977. *Mountain, Field, and Family.* Philadelphia: University of Pennsylvania Press.

Caballero, José María, and Elena Álvarez. 1980. Aspectos cuantitativos de la reforma agraria (1969–1979). Lima: Instituto de Estudios Peruanos.

Custred, Glynn. 1977. Las punas de los Andes centrales. In Pastores de puna: Uywamichiq punarunakuna. Jorge Flores Ochoa, ed., pp. 55–85. Lima: Instituto de Estudios Peruanos.

Custred, Glynn, and Benjamin Orlove. 1974. Sectorial Fallowing and Crop Rotation Systems in the Peruvian Highlands. Paper presented to the 41st International Congress of Americanists, Mexico.

DCCN. 1980. Comunidades campesinas del Perú: Información básica. Lima: Ministerio de Agricultura y Alimentación.

DGOR. 1977. Comunidades campesinas del Perú: Información censal población y vivienda 1972 (19 vols.). Lima: SINAMOS.

Dollfus, Olivier. 1981. *El reto del espacio Andino.* Lima: Instituto de Estudios Peruanos.

Dowling, Grahame R. 1987. Dialogue on Systems as Clusters: Identifying Systems. *Behavioral Science* 32:149–152.

Everitt, Brian. 1980. *Cluster Analysis.* New York: Halsted.

Flores Ochoa, Jorge. 1977. Pastores de puna: Uywamichiq punarunakuna. Lima: Instituto de Estudios Peruanos.

———. 1982. Causas que originaron la actual distribución espacial de las alpacas y llamas. In Senri Ethnological Studies 10. Luís Millones and Hiroyasu Tomoeda, eds., pp. 63–92. Osaka: National Museum of Ethnology.

Gade, Daniel. 1975. *Plants, Man and the Land in the Vilcanota Valley of Peru.* The Hague: Dr. W. Junk. B.V.

Jamtgaard, Keith. 1984. Limits on Common Pasture Use in an Agro-Pastoral Community: The Case of Toqra, Peru. SR-CRSP Technical Report No. 42. Columbia: Department of Rural Sociology, University of Missouri.

———. 1986. Agro-Pastoral Production Systems in Peruvian Peasant Communities. In Selected Proceedings of Kansas State University's 1986 FSR Symposium: Farming System Research & Extension—Food and Feed. Cornelia Butler Flora and Martha Tomecek, eds., pp. 751–765. Manhattan: Kansas State University.

Kim, J., and C. W. Mueller. 1978. Factor Analysis: Statistical Methods and Practical Issues. Sage University Paper Series on Quantitative Applications in the Social Sciences, Series No. 07-014. Beverly Hills and London: Sage.

Langbein, Laura I., and Allan J. Lichtman. 1978. Ecological Inference. Sage University Paper Series on Quantitative Applications in the Social Sciences, Series No. 07-010. Beverly Hills and London: Sage.

MacQueen, J. 1967. Some Methods for Classification and Analysis of Multivariate Observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability 1:281–297.

Masuda, Shozo (ed.). 1981. *Estudios etnográficos del Perú meridional.* Tokyo: University of Tokyo Press.

Masuda, Shozo, Izumi Shimada, and Craig Morris (eds.). 1985. *Andean Ecology and Civilization: An Interdisciplinary Perspective on Andean Ecological Complementarity.* Tokyo: University of Tokyo Press.

McCorkle, Constance M. 1986. Integrative Strategies of Labor Organization for Crop-Livestock Production in an Indigenous Andean Community. In Selected Proceedings of Kansas State University's 1986 FSR Symposium: Farming System Research & Extension—Food and Feed. Cornelia Butler Flora and Martha Tomecek, eds., pp. 513–531. Manhattan: Kansas State University.

———. 1987. Punas, Pastures, and Fields: Grazing Strategies and the Agropastoral Dialectic in an Indigenous Andean Community. In *Arid Land Use Strategies and Risk Management in the Andes: A Regional Anthropological Perspective.* David L. Browman, ed., pp. 57–79. Boulder: Westview.

Milligan, G. W. 1980. An Examination of the Effect of Six Types of Error Perturbation of Fifteen Clustering Algorithms. *Psychometrika* 45:325–342.

Murra, John V. 1972. El "control vertical" de un máximo de pisosóecol gicos en la economía de las sociedades Andinas. In *Visita de la Provincia de*

León de Huánuco (1562). Iñigo Ortiz de Zúñiga, visitador. Vol. 2, pp. 429–476. Huánuco: Universidad Hermilio Valdizan.

Orlove, Benjamin S. 1977. *Alpacas, Sheep and Men: The Wool Export Economy and Regional Society in Southern Peru*. New York: Academic Press.

Orlove, Benjamin S., and Ricardo Godoy. 1986. Sectoral Fallowing Systems in the Central Andes. *Journal of Ethnobiology* 6(1):169–204.

Robinson, W. S. 1950. Ecological Correlations and the Behavior of Individuals. *American Sociological Review* 15:351–357.

SAS Institute. 1982. SAS User's Guide: Statistics. Cary, NC: SAS Institute.

Vidal, Orlando, and Eduardo Grados. 1974. La alpaca, el vellón y la esquila. Boletín de Divulgación, octubre. Lima: Asociación de Criadores de Alpacas del Perú.

Vincze, Lajos. 1980. Peasant Animal Husbandry: A Dialectic Model of Techno-Environmental Integration in Agro-Pastoral Societies. *Ethnology* 19:387–401.

Winterhalder, Bruce, Robert Larsen, and R. Brooke Thomas. 1974. Dung as an Essential Resource in a Highland Peruvian Community. *Human Ecology* 2(2):89–104.