APPENDICES TO CHAPTER I

Appendix 1-1: Additional Notes on Possible Biases

The problem of a biased sample is particularly acute in the selection of preindustrial societies. As noted in chapter 1, I selected them from the Standard Cross-Cultural Sample (SCCS), which is by no means a perfect sample. Using a simulation study technique, however, J. Patrick Gray (1996) finds that for all but two of the forty-eight characteristics which he sampled, the SCCS represents a random draw from a population of more than 1,250 preindustrial societies from the Ethnographic Atlas, the most extensive collection of standardized ethnographic data that are currently available.

Selection of the pinpointed date, however, does admit bias and for particular characteristics, such as slavery and serfdom ("unfree labor"), the problem is especially acute. This is because many colonial powers tried to outlaw unfree labor and to suppress slave trading in countries within their jurisdiction before the date pinpointed by Murdock and White (1969) for the SCCS. As a result, in my sample of agricultural societies unfree labor was less prevalent (17.5 percent) at the pinpointed dates than it had been in the preceding centuries (57.5 percent). Of course, in some of the societies where slavery had disappeared, such a change can be traced to internal causes. But in an unknown number of cases, the absence of unfree labor could simply be due to the direct or indirect influence of Western colonial powers. Moreover, other institutions and organizations either closely connected with, or dependent on, unfree labor might also have disappeared, so that these societies in my sample were not "pristine." This does not necessarily mean that the economic systems that I have identified would be different, but certainly the number of societies in each cluster would have changed. Similarly, the pinpointed dates of the societies in my sample are sufficiently recent that wage labor -

particularly people working outside the community - was overrepresented. Again, this may not change the typology of economic systems that I develop, but the number of societies that belong to each type might be different than if the sample had been based on societies at an earlier period in their history.

Another source of possible bias is my focus on the economic system of specific communities, rather than the society as a whole. The SCCS is set up this way for the simple reason that in many cases the ethnographic facts are known for only a few communities in a society. Because communities are ethnographically quite diverse in some societies, it would be unwise to draw general conclusions from just a few. Such a conscious sampling decision could lead to certain biased results if the communities were embedded in a state structure. For instance, particular activities or influences on the economic system by the central government and/or top political leaders might not receive due emphasis in the case studies. Fortunately, this did not usually lead to problems in classifying the economic system, because the central government ordinarily had little direct impact on the economic functioning of individual communities, except through the collection of taxes, the impressment of villagers into the army, and the appointment of officials. I have omitted from my sample several societies for which this generalization does not seem to hold and in other cases have tried in a qualitative fashion to offset the bias resulting from a community focus.

Finally, random elements could have entered into my coding of the relevant data. For each societies I tried to read all of the relevant parts of the major ethnographic descriptions and, in many cases, corresponded with the anthropologists who had conducted the fieldwork. But errors in coding the society for the statistical analysis undoubtedly crept in anyway and these probably lowered the

results of the tests for statistical significance.[1] For the series used in the cluster analysis, as noted above, I coded the data myself, as well as many of the economic variables not used for determining the economic system. For many of the social and political variable, however, I use data coded by others. Their results might be biased as well, although it would be difficult to specify such bias without trying to replicate their coding results.

---

[1] Many societies in my sample were visited and described by only one or two observers, so that the descriptions may not be complete. Further, the observations of the observers are subject to error. For instance, in trying to complete my data gathering for one society, I wrote the two anthropologists who had visited it. Although they were in the same village at the same time and even lived in the same house, their answers to most of the seven questions I had asked were conflicting.

I drew the codes of other scholars from the annual data disk published by the journal World Cultures (Divale and Gray, annual). This disk contains over 1,700 coded series for the SCCS societies and is an invaluable statistical source.

Appendix 1-2: Some Technical Notes on the Cluster Analysis

For the various cluster analyses in this study I employed an unsupervised cluster analysis, which means that I made no attempt to impose my own hypotheses about the types of economic systems on the data. I instructed the computer program (LOICZ, see Maxwell, Pryor, and Smith 2002) to determine the clusters by employing a k-means algorithm and to calculate unweighted multi-dimensional Euclidean scaled distances between sample points in the calculations. Scaled distances means that the cluster program lineally transformers the various codes so that each of the ten institutional variables are recalculated to have a mean of zero and a variance of one. The algorithm used by LOICZ, described originally by MacQueen (1965), incorporates some modifications to improve its robustness.

Because the computer program uses an iterative method to determine a solution, the end results are not always the same. For calculating the MDL (minimum description length) I ran the program 50 times and selected that number of clusters (economic systems) appearing most often as optimal.

After I specified the number of optimal clusters, the LOICZ program then randomly selects the same number of nations, and calculates in an iterative fashion the nations in those clusters. Given the randomness of the initial selection of nations, it was necessary to run the program a number of times and average the results. For this reason I also specified that each run of the program include the average of 100 such calculations and that each calculation be carried out with a maximum of 200 iterations. Finally, I repeated this process ten times (so that a total of 1,000 separate runs were made) and averaged these results for the final identification of the clusters.

In these final ten runs, the clusters usually consisted of roughly the same societies. If, however, a given society fell into one cluster less than 70 percent of the time, I eliminated it from all regression calculations of the properties of the cluster. I did, however, include these societies in the calculation of the development elasticities, which required no knowledge of the economic systems.

Because of my reservations about some of the codes, I also carried out some statistical experiments by rerunning the cluster analyses, but excluding those codes rated "uncertain" or "educated guess." The results were quite similar to those when all codes were included.

BIBLIOGRAPHY FOR APPENDICES TO CHAPTER I

Gray, J. Patrick. 2001. <u>World Cultures</u>. Compact disk. Jamaica, NY.

 ---. 1996. "Is the Standard Cross-Cultural Sample Biased? A Simulation Study." <u>Cross-Cultural Research</u> 30, no. 4 (November): 301-15.

MacQueen, J. 1965. "On Convergence of k-means and Partitions with Minimum Average Variance." <u>Annals of. Mathematical Statistics</u> 36: 1084 ff.

Maxwell, Bruce A., Frederic L. Pryor, and Casey Smith. 2002. "Cluster Analysis in Cross-Cultural Research." <u>World Cultures</u> 13, no. 1 (Spring): 22-39.

Murdock, George P., and Douglas R White. 1969. "The Standard Cross-Cultural Sample and Its Codes." <u>Ethnology</u> 8: 329-69; also pp. 3-45 in Barry and Schlegel, 1980.