
NATURAL RESOURCES JOURNAL

Volume 8
Issue 2 *Spring 1968*

Spring 1968

Clawson and Stewart, Land Use Information

George H. Hall

BOOK REVIEWS

Land Use Information

By
MARION CLAWSON AND CHARLES L. STEWART

Baltimore: The Johns Hopkins Press. 1965.
Pp. xvii, 402, $6.00

The appearance of an analysis of land statistics addressed to the broad community concerned with land use is something of a milestone. It marks the transition to a new era with three significant characteristics. First, there is an increasing concern with patterns of land use and a willingness to influence them through numerous and diverse public programs. Second, there is an increasing interest in basing decisions on quantitative analysis. Third, there is an increasing capability to create and handle quantities of data heretofore infeasible. The first two characteristics may, in economic parlance, be thought of as providing a new and extensive demand for information. The third characteristic represents a technological change in the production function for information. The result is a new "industry"—systematic land use information—that now involves considerable resources in the production process and is obviously a "growth industry." More significantly, the increased output of this industry is sure to have wide ramifications both for understanding land use patterns and for the policymaking that affects them. Indeed, perhaps the most significant message of this book is that the generation of land use information can no longer be left to a few statistical specialists, but is a subject about which all involved with natural or urban resources must be concerned.

This volume is actually two books. The first is a report of the deliberations of the Committee on Land Use Statistics sponsored by Resources for the Future, Inc. The second is the *Standard Land Use Coding Manual,* prepared by the Urban Renewal Administration and the Bureau of Public Roads and reprinted in an appendix. This manual was developed concurrently and in conjunction with the RFF report, and there is a valuable complementarity between the two efforts. The value of *Land Use Information* is also increased by appendixes containing reports from various governmental agencies on their statistical programs.

Land information can be analyzed at three different levels. The first level is that of data elements or the specific statistics produced. The second level is the data system or process by which the data elements are generated, stored, and made available to the user. At the third level are the users and their needs. While each level receives attention in this book, the major emphasis is on data elements. This review, however, will concentrate on the data system and data user levels. Admittedly, dwelling on topics the authors placed beyond the scope of their work violates one of the rules-of-the-game for reviews, but three reasons justify concentrating on the data system and data user dimensions. The first is that the topics discussed in *Land Use Information* are so well covered that further work will begin from this analysis. This is high praise but it also makes criticism largely superfluous. The second reason for the emphasis in this review is that element choices depend upon the system of collecting and handling data and how these data are used. These relationships received less attention than they merit. The third reason for stressing data system and data use problems is that the volume under review is doubtlessly the first of many attempts to improve land use statistics, so it seems appropriate to concentrate on the problems that should have pride of place on future agenda.

In analyzing data elements, both the RFF study and URA-BPR manual are based on the sound, fundamental principle that "pure line" data series are essential. The characteristics or aspects that describe land and its uses should not be combined into a single classification system or index at the enumeration stage. Instead, each characteristic should be separately reported, and combination should take place at higher stages of analysis or decision-making.

This is a vital point. Previously, with the exception of the censuses, most statistics about land have been produced as by-products of special studies or administrative procedures. The result is that available data usually aggregate a number of characteristics. To illustrate, the number of dwelling units (a measure of intensity) is frequently combined with the type of ownership, for example, governmental or private. Later analysis requiring separate data on either ownership or use-intensity becomes impossible. Separate enumeration of characteristics is the only approach consistent with the heterogeneity of the units collecting and utilizing land information.

The URA-BPR classification system meets this need for flexibility and aggregation ability. A four-digit code is used, with each

additional digit place permitting more disaggregaton of activities. A particularly helpful feature of the code is the handling of activities that are separate but functionally linked. By using an auxiliary code attached to the four-digit number, warehouses, parking lots, laboratories, central offices, and the like can be identified and related to other organizational counterparts. The code is similar to the Standard Industrial Classification Manual system and uses SIC terminology for its classes. The URA-BPR categories are not necessarily identical to SIC categories, however. More about this later.

The RFF report provides a history of land use statistics and analyzes present series. Emphasis is given to the compatibility of different series, coverage, and biases. The major attention is devoted to land use data, but other characteristics of land such as natural qualities, intensity of use, tenure, and prices are also considered.

If the numerous organizations and agencies collecting land use information adopt the approach of the RFF and URA-BPR committees, much more compatible, useful and accurate data than now exist will be collected. Like the SIC code, the system proposed in *Land Use Information* could provide a unifying force in an area that desperately needs unification. The choice of data elements, however, is importantly influenced by the data system and the use made of data.

Consider the data system aspect of land use information first. When data were scarce, one did not need to worry much about their handling and presentation. By a cut-and-try system the usefulness of the data could be investigated. One accepted whatever limitations were inherent in the data since it was usually prohibitively costly to obtain better information. The handling of the data could be casual and heuristic because the small amount of data permitted one to go back and correct mistakes or even start over if a better approach was later discovered. When one must deal with the masses of data that a system such as the one advocated in *Land Use Information* will produce, the situation becomes much different, however. A more formal approach to the collection, storage, and presentation of the data is required if biases and serious errors are to be avoided and if usable and timely information is to be made available. With the great quantities of statistical data that it is now feasible to produce, there is a serious danger that information collected at great expense will be useless.

There are a number of reasons why large data collection efforts may yield few usable results. One problem is that even if data exist it

may be a laborious and costly process to obtain them. If data are stored at different geographic locations, if different offices use different systems of punch card and computer tape storage, if the data are intermixed with other data on tapes and cards, if different reporting periods are used by different agencies—in short, if different data handling processes are used—then even if the data elements are the same, obtaining the data may be laborious, time-consuming and costly. Frequently the result is the same as though the data had never been collected.

A related problem is that data handling, data destruction, and disclosure policies frequently differ among collecting agencies. The analyst trying to construct a statistical series may discover gaps and serious biases. Not only does the old proverb that "a chain is only as strong as its weakest link" apply here, but also Murphy's Law that "if a thing can go wrong it will."

The answer to these, and most of the other relevant problems that cannot be discussed here, lies in creating a data system sufficiently simple that there is a high probability that collection, storage, and retrieval will be reliable and inexpensive. Simplicity requires several things. Foremost are data elements that fit in with the missions of the collection agencies and that can be applied uniformly among agencies. Here the RFF and URA-BPR reports are extremely helpful. Also important, however, are simple methods of collecting the data and getting them to users promptly. The more numerous the data elements and the more detailed the reporting system, the more complex are storage and retrieval problems. Thus, it is hard to separate data element questions from data system questions.

Pessimism about our ability now, or in the near future, to create large well-functioning data system seems called for. The RFF and URA-BPR reports sweep these problems aside by blithe statements that modern computers make storage and combinations of many data elements possible. True, but they can also make retrieval and use impossible. This problem can be illustrated by asking the question: If we are going to generate pure data series about many intensity, tenure, product, and other land characteristics, can we afford the time and expense necessary to publish these in report form? Previously a few pages of journals or special reports have sufficed to make data available. It is doubtful that this will be possible for the large masses of land use data that will soon become

available. Perhaps only a description of the data need be published, with an offer to provid users with copies of computer tapes or punched cards. The Bureau of the Census has been moving in this direction, but I have no idea whether it is a feasible policy for land use information. I do know that we are in danger of producing more data than can be made usefully available to consumers. Instead of being unable to answer questions because so few data are available, in a short time we may be unable to answer questions because so many data are available.

Land Use Information focuses on data production rather than data consumption. The argument is that if the data elements are detailed enough the reader can "roll his own." The RFF study particularly stresses the need to record data by parcel, defined as the smallest identifiable tract of land for a given purpose. There is a good reason for the attitude expressed in this book. The uses to which any series will be put is difficult to forecast. Further, the individuals, institutions, and public agencies utilizing land use statistics vary widely in size, goals ,and interests. No all-purpose set of numbers is possible.

Granting this, there is still need for more attention to how the data that Clawson and Stewart and the URA-BPR propose gathering will be used. To mention only one problem, the comparability of land use data with other statistical series needs more attention. A big step towards comparability was made when the URA-BPR based their system on the SIC, used for most governmental series. Unfortunately, there are important differences between the SIC and the URA-BPR systems, particularly when mulitiple-product units are involved. Most SIC series classify activities on the basis of the product that yields the most revenue. For land use, classification is on the basis of the product that requires the most space. Also, the "establishment," the basic unit for SIC-based series, is inappropriate for land use studies. Thus, the user who wants to match up land use data with other economic indicators is going to have a difficult analytical and statistical task. In short, the adequacy of data must be evaluated in terms of the use to which it is or could be put.

Compromise is a way of life when one is designing statistical series. The compromises that the RFF or URA-BPR committees made probably were the best choices. It would have been helpful, however, to have had more discussion of the many uses to which land use data are put—particularly because, as Clawson and Stewart

point out, once a statistical series has been established there is a disinclination to change it since changes lessen the usefulness of the data for historical analyses.

As previously mentioned, the primary reason for dwelling less on the accomplishments of this volume and more on what remains to be done is that the RFF and URA-BPR studies are undoubtedly the first of many examinations of land use information. The RFF study recommends the establishment of a permanent committee or organization to refine and develop land statistics. The URA-BPR system is undergoing changes and improvements. The expanding public programs affecting land use and the growth in our data handling capabilities demand such efforts. RFF, the URA and the BPR have taken an important first step. Much more needs to be accomplished, however, before land use information problems are solved.

GEORGE R. HALL*

* Logistics Department, The Rand Corporation, Santa Monica, California. The views expressed herein are those of the author. They should not be interpreted as reflecting the views of The Rand Corporation or the official opinion or policy of any of its governmental or private research sponsors.