AAEA Annual Meetings
Pullman, Washington

July 30, 1979

MODELING RURAL GROWTH

by

Clark Edwards, Senior Economist
Economic Development Division
Economics, Statistics, and Cooperatives Service
U.S. Department of Agriculture

Rural economic growth has been in and out of the spotlight as a public issue. Sometimes the concern is how to spur lagging development. Sometimes it is how to limit an unbalanced surge. Other times interest lags and the issue is neglected. When the issue is in the spotlight, there are always ready answers at hand to the question, "What to do?" During the sixties, we often heard: "Stop rural outmigration or reverse it!" Toward this end, we made loans and grants, trained workers, and located firms. The answer by some persons in areas which were growing too fast during the early seventies was to put up a sign to the effect that "this is a nice place to visit but we don't want you to live here."

Policy makers have identified the issue. And they have many ready answers. What they are frequently missing are the economic reasons which relate the answers to the issues. That is where we, as researchers in agricultural economics, come in: we can help provide the kind of understanding of the issues that comes from research, and which serves as a basis for policy prescription. The point of this paper is that, although we have been performing this function, we are not doing it as well as we might. Our research provides helpful descriptions and useful projections, but our models generally are not reliable for predicting impacts or for calling turns.

This paper addresses three obstacles to modeling rural growth: models, data, and theory. Some of the obstacles are inherent in the way we approach the modeling process; some are attributed to limitations of available data; and others stem from failure to incorporate relevant theory.

Models

The idea of economic modeling brings to many minds visions of mystifying mathematics, dreary data, and cold computers. Of course not all models are mathematical, empirical, and computerized. But let the image stand--let's talk about that kind of model. One important obstacle to building models is that many people simply oppose them. And their grounds for opposition have some validity. They say, for example, that one cannot capture the things that really count in economic progress--the quality of life and a sense of well-being--in a numerical measure or in a computerized system of equations. To many, getting a region to grow is a practical problem unrelated to mathematics. Certainly, negotiations by community leaders to locate a plant employing 500 persons is not a problem in mathematics. And secondary employment induced as a consequence of locating this plant can be calculated with simple counting methods that can hardly be considered mathematical.

"If such problems can be understood in plain English," so the question often goes, "why do we need to treat them mathematically?" The wish to have our models explained in plain language is a reasonable one. It is hard for a researcher to explain that a simple translation of a mathematical model into English is not possible. The mathematical and computerized model is not merely another language, it is also a logical system. The analogy might be that the model translates not into a sequence of simple sentences, but into a syllogism. The model imposes a system of organized reasoning. The reasoning can be understood from ordinary language if one makes a special effort to follow it. But most of us do not think in perfect syllogisms, and the reasoning is likely to be lost when each equation is translated into language.

Maybe if we knew enough about economic growth we would not need to turn to mathematical and computerized models to evaluate alternative policies; we would know by experience what the impacts of each policy would be. But most of us do not know that much. So we make up for it by using models to organize our thoughts and to reason out the likely consequences--either quantitatively or qualitatively--of alternative actions. The gains from organized reasoning need to be compared with costs such as that of reducing to numerical measures certain human values that are properly understood as not numerically measurable.

Let us assume that the trade-off favors modeling and inquire into other obstacles to modeling rural growth that are evident in published research.

Cause or Effect. One common empirical method to identify factors affecting area growth is to sort the data according to the intensity of the problem and then note associated factors. For example, we can rank the 3,000 counties according to rate of growth in population and then notice that slower-growing counties exhibit, say, a lower average level of education than faster growing counties. The beauty of this method is that it always works. There is always a group of slower-growing counties, and such a group can always be found to have certain distinguishing characteristics. This is what makes the method descriptively useful. However, it is not analytically useful to sort on the effect and then notice the cause. If we seek to explain, we should sort on the cause and notice the effect. In this example, we should sort on level of education and test whether counties with lower levels of education are significantly slower-growing.

One reason we persist in doing the thing backwards may be that it requires less thought; in one pass on the computer one may, mindlessly, catch several "explanatory factors" that are statistically significant. To order the data causally, one must think ahead and either do more detailed cross tabulations or make more computer runs. Another reason may be that it is easier to explain anomalies when we do the thing backwards; it is easier to explain as an exception to the rule a slower growing county with a high education level than a county with a low education level which grew rapidly.

Recent statistical tests of causality, based on an operational definition of the concept, may help us to sort out, at least in a statistical sense, what is cause and what is effect in economic growth, and thereby improve our ability to model causal structures.

Simple or Complex. Many rural growth models are exceedingly detailed and complex. The driving idea behind them appears to be to build a model comprehensive enough to be capable of answering any question which might arise. Of course there are certain economies in building models which are large enough and general enough to warrant repeated use. But there is a tendency to go too far and build models which bog down in their own mathematical complexity and their detailed data requirements. Some detailed models, such as input-output, have large data requirements but follow simple logic; they are therefore easy to understand although expensive to construct. Others, such as simulation, may have simple data requirements but follow complex logical relationships that are exceedingly difficult for a user to understand. One ridiculous equation inadvertently embedded in a large simulation model might never be discovered either by the builder or the user. The high probability of such accidents leaves one concerned about building models more complex than necessary.

At the other extreme, some of our models are entirely too simple to capture the complexities and the many facets of economic growth. Almost all single-equation models are suspect because they cannot allow for simultaneous determination of growth variables or for feedback and reciprocal relationships.

Between the extremes of too big and too small, there must be a size which is just right. One approach to determining model size is this: Start with the problem you seek to explain, say rural economic growth, and write down the variables with which you intend to explain it, say population, income, and employment. If it requires three variables to describe the situation, then a minimum of three equations is needed; if the results of a model are to be summarized by these three variables, then fifty or a hundred equations are far too many.

But there is a good argument for using more than three. If you can explain population with one equation, you can usually do it better by explaining a few of its major components with separate equations. For example, two equations, one for working age people and another for dependent oldsters and youngsters, can be expected to give better results than a single population equation. Using this rule, a minimum of six behavioral equations will be needed for explaining two categories each for population, income, and employment. And three identity equations will be needed to form the totals of each variable. A nine-equation model, then, can adequately explain the stated problem. And a model of 90 equations is not likely to do any better job than one of nine or so in explaining the three variables of concern. Worse, unless a priori totals are incorporated to constrain the 90 equations and keep the results "in the ball park," the larger model is likely to behave less well than the smaller one.

Some obstacles to modeling rural growth would be overcome if we gave more thought to: (1) the problem the model is intended to solve, (2) the variables actually needed to describe the problem, and (3) the minimum number of equations needed to adequately explain the essential variables.

Seek Answers or Assume Them. Many models assume answers to questions to which their prospective users are seeking answers. For example, a rural growth model may assume a target population, and perhaps a target level of income, for the year 1990, and then examine the industry mix and resources that would be required to realize the target. There is nothing wrong, of course, in building such models; they are useful. The error arises when the user (and sometimes even the model builder) classifies such a model as a growth model. If a model assumes the level of growth as an input, it is not useful for explaining growth.

Describe or Explain. Many of our mathematical, empirical, computerized models are thought to explain growth when they really only describe it. Descriptions are useful, but they are not explanatory. Stages of growth models describe various stages through which a region is expected to pass without explaining how the region moves from one stage to the next. Shift share models measure the extent to which unique regional factors contributed to growth without identifying what the unique factors were or how they contributed. Projections models describe the likely future based on recent trends. Some of the methods of projection are complex and ingenious but, analytically, they have no more explanatory value than linear extrapolations. Descriptive models can influence decisions by policymakers. But they do not constitute models of growth. They are not capable of forecasting a turnaround. Nor are they capable of being used for evaluating the consequences of policy intervention or for assessing alternative futures.

Data

Major gaps exist in the data base for rural growth models. There are several causes of these gaps. Some data which have been available for a long time, such as for farm employment, are considered obsolete because demands on the data have changed while the supply has not. Some series we would like to use, such as adequate measures of underemployment or hidden unemployment, are not supplied. The reporting units on which data are made available often do not correspond to the analytic units in our models. For example, the data may pertain to establishments (place of work) whereas we want to model families (place of residence). The time detail is often wrong. We have one census for 1970 and will have another in 1980. But families in the 1980 census will not be linked to families in the 1970 census for longitudinal information. And comparable information for intercensal years will not be available to study the turnaround in rural growth exhibited during the late 1960's and early 1970's.

Rurality is a geographic concept and lack of geographic detail in published data is probably the most important source of data gaps with respect to rural growth analysis. Even though many of the geographically specific data needed are collected in various censuses and surveys, this gap continues because such data are generally not released by the data collection agency in sufficient geographic detail.

Data are relatively more or less reliable depending on the number of observations underlying a reported average and on the accuracy with which released data are edited. Several data gaps result from the way the present data system has been conceptualized. For example, much of the industry detail available for rural growth models describes product markets, whereas a growth model may need detail on factor markets instead. Because of increased institutional size, economies of scale, and specialization, the separation of data suppliers from data users has created data gaps which could be narrowed by means of better communications among institutions.

Some data gaps are formidable--their resolution would not be cost-effective and we must learn to live with them. Others can be narrowed at nominal (virtually zero) cost by making small changes in the way some data are collected and reported. Researchers all too often adapt their models to accommodate existing data series. Sometimes a far better solution would be to become more aggressive about insisting that needed data become available.

Some ingenious procedures are being developed by researchers for getting around data gaps without waiting for perfect data and without sacrificing the logical content of their models. Let me give two examples: One is to use primary data insofar as they are available or considered absolutely essential, and to use inexpensive and readily available secondary data sources as default options when primary data are missing. Another is to merge two files (which together have the required information, but which separately do not) into a synthetic, comprehensive file. The direct benefit of these procedures is to maintain operational models. A fringe benefit is that the results can be used to estimate the benefits that would likely follow from collection of more primary information and more comprehensive data sets.

Theory

Theory allows us to explain economic phenomena. To understand and influence the course of rural economic growth, economists first need a growth theory. Alternative futures can be examined, causes can be understood, and intelligent choices can be made. At one extreme stand those whom William James called the tender-minded; they seem to believe that pure theory alone is all that is required to solve problems of economic growth. At the opposite extreme stand those whom he called the tough-minded; they seem to believe that hard facts alone are required. As for the rest of us, we seek to ground empirical models in appropriate theory. The theories we need are scattered through the economics literature. Until these have been merged into a single, comprehensive theory, modelers must chose eclectically. In such an environment, the likelihood that any single model will incorporate all relevant theories must be very low indeed.

The various growth theories can be, and have been, collected and discussed in ordinary language. But integrating them all into a consistent set of equations which function as a growth model is exceedingly difficult. Obstacles to incorporating some of these disparate bases for growth into a single, comprehensive model are the subject of the remainder of this paper.

Increasing Resource Availabilities and Improving Technology. Neoclassical microeconomic theory incorporates two bases for growth. An increase in output results either from using more resources or else from obtaining more output per unit of resource. These two microeconomic ideas so completely dominate our thoughts that when they are presented in the growth chapters toward the back of our macroeconomics textbooks no one appears to notice a contradiction: models which assume that all savings are invested, that supply creates its own demand, that there is full employment, and that money is simply a veil over the economy are introduced in the back of a book dedicated to contradicting these assumptions. Incorporating resources and technology as bases for growth is relatively easy and is frequently done. So let us assume we have a growth model incorporating these two ideas and inquire into the difficulties encountered in adding other bases for growth.

Expanding Markets. The idea that aggregate demand is a basis for growth has three roots. Classical ideas about the importance of foreign trade embrace the concept. And Keynesian theories center on the importance of demand. Economic base theory is at times interpreted as a simplified, one-sided version of either of the above, but is at other times considered to have an independent origin. As far as construction of economic growth models is concerned, economic base theory and Keynesian macroeconomic theory are very much alike. Each model is driven by demand rather than by the neoclassical bases for growth--resources and technology. And each incorporates multipliers which measure secondary impacts. Models for rural growth which are driven by demand abound. The obstacle to growth modeling of concern here is encountered when efforts are made to merge the neoclassical and the Keynesian ideas into a single model. The ideas are seen to be, on the face of it, incompatible. Let me illustrate how this is so.

Consider two ways that an input-output matrix can be used. Using Keynesian logic, we can start with a menu of final demands and inquire into the industry mix and resources required to produce it. If there were idle resources, the economy would grow as demand grew. On the other hand, using neoclassical logic, we can start with a menu of resource availabilities, incorporate the input-output matrix into a linear programming format, and inquire into the industry mix that would maximize final product. In this case, the economy would grow as resource availabilities and technology grew. Now ask yourself how to get a model to behave both ways at once. The answer lies in building in some kind of a flip-flop mechanism so that causation flows in one direction through the input-output matrix when resources are slack and in the opposite direction when they are fully employed. If some resources are slack while others are fully employed, it gets even more complicated. So let us simplify the problem further.

Consider a growing economy which can be described by two variables: an input and an output. There are three equations: a demand for output, a supply of input, and a production function. This system is overidentified: three equations for two variables. The neoclassical solution is tantamount to assuming that the demand relation is an inequality—that there is an effective demand for at least as much as will be produced when resources are fully employed. This is accomplished by Say's Law. The Keynesian solution is tantamount to assuming that the supply relation is an inequality—that production capacity is adequate to produce at least as much as will clear the market. That is, only two equalities are needed to determine the system's growth--one pair if demand is the basis for growth, the other if resource availability and technology are the bases. A consistent clue as to which is which is in the production function. If the model is demand driven, the logic flows through this function from output to input. If the model is supply driven, the logic flows in the opposite direction, from input to output.

In a comprehensive model, there must be inequalities describing demand and supply and a flip-flop mechanism describing the direction of flow of logic through the production function. One obstacle to modeling rural growth is to figure out how to make our models consistently incorporate both Keynesian ideas of demand and neoclassical ideas of supply in a comprehensive and consistent system of equations.

Conquering Space. Rurality is a geographical concept, yet many rural growth models fail to incorporate spatial relationships. We should be able to display the results of a rural growth model on a map. The obstacle here is not only redesigning the logic among variables already in the model as was required above, but also to incorporate overlooked variables. There are four categories of flows over space to examine: people (migration and commuting); goods (imports and exports); capital (balance of trade), and ideas (diffusion of information). Other variables already in the model must be identified by location. The logic of the growth models must be redesigned to incorporate opposing forces which induce centralization and decentralization. Centralizing forces are identified by central place theory. They include agglomerative efficiencies and transportation costs. Decentralizing forces include rent gradients and von Thunen rings. Spatial bases for growth are ignored in most of our rural growth models.

Institution Building. Econometric models have tended to capture short-run phenomena by assuming that needed institutional arrangements were in place. Longer-run issues, which were harder to model, were abandoned to the institutional economists who knew that institutional change does matter and knew how to account for it in economic analysis. Now that we have learned to build dynamic models that cover economic change over time, the importance of integrating institutional economics with econometrics is reemphasized. This situation is giving rise to a new breed of institutional economists who are trying to capture the essence of institutional relationships as equations in econometric models.

Two aspects of institutions are fairly easy to capture. First, one function of an institution is to assess ends or goals. These goals can be written as equations--maximizing gross regional product, minimizing unemployment, and improving the distribution of income, for examples. Institutions help to resolve conflicts among goals; models can be used to exhibit trade-off possibilities among competing goals.

Second, institutions promulgate behavioral rules. These rules can be written as equations. Some examples, such as tax laws and price support levels, are commonly incorporated and recognized as institutional constraints. With imagination, we can capture more institutional rules, for examples, zoning laws, water rights, and licensing requirements.

Other aspects of institution building are harder to capture. Institutions are groups of people acting toward certain ends. How do you capture in an equation a group of area planners acting to promote regional growth? If they change a zoning law or an objective, we can change an equation. But this is *ad hoc* and likely to miss the essence of a planning organization. Some of it might be captured with a zero-one variable which opens certain synapses among economic variables when the institution is present and closes them in its absence.

Each of the several bases for growth discussed above—resources, technology, markets, space, and institutions--can be, and have been, incorporated in rural growth models. Many models incorporate two or three. A few contain four. I have yet to see one that satisfactorily reflects all five. And I think there is good reason for our failure to do so—it isn't easy. But I see it as a challenge which we as a profession need to work on.

Conclusion

A number of models of rural growth have been built and used by agricultural economists. Clients use the results. There is a demand for more and better policy guidance on strategies for rural growth. I have given my reasons for thinking that we are not modeling rural growth as well as we might and that, consequently, we are not providing, as well as we might, the economic reasons which connect policy prescriptions to rural economic issues. I have offered my suggestions about how we, as researchers can improve our rural growth models. The ideas we need are lying about in the literature--but we haven't put them all together yet.