ECONOMIC PROJECTIONS FOR LOCAL AREAS

Robert E. Graham, Jr., Daniel H. Garnick, and Allan L. Olson*
Office of Business Economics

The title of this paper has been changed from that listed in the program -- which emphasizes population projections -- to one that is more generalized. This change reflects the fact that our aim in the Regional Economics Division of the Office of Business Economics is to project three major aggregates -- income, employment, and population. Moreover, our main focus is on the separate and independent projections of income and employment, with population a derivative of the first two.

The decision to make population a derivative of income and employment stems from the assumption that the critical element in a population projection is regional migration and the major factor underlying migration is economic opportunity, or lack thereof. That is, population will move toward expanding economic opportunity and away from a shrinking or static regional economy. Consequently, it would seem that a better population projection can be made by concentrating more directly on the basic motivating factors and then deriving population from the results rather than via a population-to-employment approach.

Economic projections are often classified into 2 major groups: (1) Projections that represent mainly extrapolations of past trends which are usually termed simple or naive, and (2) projections made via an economic accounting model featuring income and product or input-output and which are termed sophisticated.

Without intent to set up a "straw man," it seems useful to point out that the foregoing characterizations of the two types of projections are not really valid.

Reliance on extension of past trends is not a distinguishing feature of the two methods, for both naive and sophisticated projections place equal reliance on past experience. To the extent that the sophisticated methods reflect no past experience, their results must generally be judged less reliable than those of simpler but historically-based series.

Use of past trends in the naive method is well-recognized. Not so obvious, however, is the equal reliance that sophisticated methods place on past developments. An input-output table, for example, to be really useful in making projections requires that at least two exogenous elements -- final demand and technical coefficients be projected, either explicitly or implicitly. To hold technical coefficients constant and project proportionately equal increases in all elements of final demand yield results of the most naive sort. Somewhat more valid is a regional projection based on input-output that relies on a simple projection of past trends to derive a population estimate which, in turn, undergirds a projection of final demand from which is derived a projected economic structure of the area.

Both naive and sophisticated methods, then, rely equally on past trends to the extent the basic data permit. If time series data are not available, and they tend to be scarce in direct proportion to the complexity of the projection framework being used, the technician must either hold all relationships constant or change them through deductive reasoning.

A pragmatic classification of projection methodologies into naive and sophisticated groups reflects more the complexity of the economic measure used in the projection process than the method used to extend that measure into the future. And, <u>ceteris paribus</u> the more complex or detailed the economic measure employed, the more useful will be the results to the extent that the systematic components dominate the random, nonsystematic components of the economic measure. However, in the real world of economic measures other things are seldom equal and in few areas are they more unequal than in regional economic measurement. Indeed, the regional field is characterized by a paucity of economic measures. Here, reference is not to the quantity of data available. What is lacking are time series for major constructs such as income and product tables, input-output accounts, employment and flow-of-funds series disaggregated both industrially and geographically.

In the final analysis the method used, a complex econometric approach versus a simple methodology with a good measure of judgment thrown in, really reflects the type and quality of the input data available. If input-output or income and product tables are available both historically and currently for the geographic areas under study, the so-called sophisticated method of projection would be the choice in nearly every instance. If data availability imposes its usual constraints, a simpler model tends to become the choice.

Our program of regional economic projections calls for the preparation of projections for 165 economic areas initially. Later the number may be increased to as many as 400. Given this very large number of geographic areas, data input becomes a crucial consideration.

---

* The authors are members of the Regional Economics Division.

To prepare input-output tables or income and product accounts that are something more than mirror images of their national counterparts for at least two years (in order to gauge trends) for 165 separate areas would be a task of near-impossible proportions. To measure personal income by local area for 5 selected years has required 2 years and an expenditure of close to $1 million. Preparation of the more detailed economic accounts, if indeed feasible, would require many times the resources needed for measuring personal income.

Given data requirements and data availability, we have chosen to project an economic aggregate that is moderately comprehensive; that can be constructed to show adequate geographic and industrial detail; and for which a time series can be prepared. Personal income meets those requirements more adequately than any alternative.

Specifically, historical estimates of personal income by local area have been prepared for 5 selected years of the span 1929 to 1962. The years include 1929, 1940, 1950, 1959, and 1962. In general, each of these except 1940, represents a roughly comparable point on the business cycle, thereby eliminating the potentially distorting influence of the cycle on economic change. Since area employment estimates statistically comparable to the personal income series have not been completed, we are using employment from the decennial censuses of population. From this a series showing about 36 separate industries by local areas has been assembled.

Because it was desirable to project both employment and income in as much industrial detail as possible, a set of geographic areas in which the various industrial components of income and employment would bear reasonably stable relationships to one another was constructed. These economic areas are based on the nodal-functional area concept. That is, to each urban center are attached the surrounding county units in which economic activity is focused directly or indirectly on the center. Each economic area combines the place of residence and place of work of employees as nearly as possible so that there is a minimum of commuting across economic area boundaries.

Each economic area specializes in the production of certain types of transportable commodities and of nontransportable special services such as education at Cambridge, recreation at Miami, and finance in New York. The production locus of such goods and services is determined not so much by transportation costs as it is by the costs associated with special resources and by the economic benefits derived from economies of scale. Different commodities are associated with production processes requiring different input relationships and the comparative advantage of a region for the production of a commodity is determined by the region's relative endowment of the factors of production. In addition, in many industries the effort to maximize returns to the factors of production leads to expanded production as a means of exploiting the economies of scale. This process, which can be implemented only if trade can be carried on with other areas, further reinforces regional comparative advantage and specialization.

In contrast, each economic area approaches self-sufficiency in its residentiary industry sector; that is, while each area specializes in producing goods and/or services for "export" to other economic areas (and abroad) most of the services (and some goods) required by local residents and businesses are provided within the area.

Thus, the economic areas correspond to the closed trade areas of central place theory in which the number and type of establishments and their size and trade areas are bounded by the relative transportation costs from the hinterland to competing centers. Each area approaches closure with respect to residentiary industries which include general and convenience retail and wholesale trade activities and those other services which are difficult or impossible to transport and are most efficiently consumed in the vicinity of their production.

Application of the foregoing criteria to the U. S. economy yielded 165 areas each of which formed a complete and integrated economic unit characterized by comparative stability in interindustry relationships. Having delineated the 165 economic areas, we then considered alternative projection methodologies.

The first method examined was a naive model, characterized by a complete absence of theoretical underpinnings in its formulation. It was devoid of systematic or interacting components and all projected elements were exogenously determined. It was essentially a "no change" model.

The exogenous determinant or predictor in this naive regional model was the national change in employment or income in a given industry. That is, the base period ratio of regional employment or income to national employment or income in each industry was applied to the projected national level of employment or income for the corresponding industry.

$$(1) \quad E_{ij}^{t} = (E_{ij}^{o}/E_{io}^{o})\, E_{io}^{t}$$

Where the subscripts $\underline{i}$, $\underline{j}$ refer to the $\underline{i}$th industry and the $\underline{j}$th region, the subscript $\underline{o}$ refers to a summation: when in the right hand position, it is the summation of regions (= the Nation), when in the left hand position, it is the summation of industries (= total employment or income); superscripts $\underline{t}$, $\underline{o}$ refer to the projected period and the base period, respectively.

The naive model, though reflecting no more than the national industrial growth rates in each individual industry in each region, does, nonetheless, reflect an aggregate growth rate that differs from that for the Nation when the region's industrial composition differs from the national in the base period. Such a model, however, fails to take account of regional differences in rates of growth among individual industries. To take account of this, we turned to shift-share analysis.

Shift-share analysis is designed to discern regional departures from national industrial growth rates, and while its history goes back to 1943, most of the work using, clarifying and elaborating on the technique appeared only in the late 1950's and in the 1960's. [1] In its simplest form, the shift-share technique distinguishes a proportional growth element and a differential growth element between a region and the Nation in each industry.

$$(2)\quad E^t_{ij} = (E^t_{io}/E^o_{io})\, E^o_{ij} + C^t_{ij}$$

Where C^t_{ij} equals the difference between the level attributable to the national growth rate of the industry and the regional growth rate actually attained in the industry.

It is the attention paid to the differences between regional and national growth rates in each industry that distinguishes the naive share model of equation (1) from the shift-share model of equation (2). Thus, the first term on the right hand side of equation (2) is equal to the entire right hand side of equation (1). The second term on the right hand side of equation (2) is called the share effect (C_{ij}) in shift-share analysis. It is, in fact, the difference between the "hypothetical growth," accounted for by the first term, and the attained level of the left hand side. In basic or export industries the share effect is presumed to be connected with some regional competitive advantage (or disadvantage if the term is negative) in the industry. That is, the region presumably grows faster or slower than the rest of the Nation with respect to the industry in question because of a difference in the marginal productivity of capital in the region relative to all other regions. Thus, the shift-share projection model departs from the naive-share extrapolation model, at least implicitly, insofar as it treats regions as relatively open economies among which capital and labor may flow. In contrast, the naive-share extrapolation model treats each region as a miniature reproduction of the national economy with all national developments occurring proportionally in each region's economy.

The causal economic factors associated with C_{ij} are the essence of industrial location theory. But, over the last 20 years, there has been very little correspondence between developments in industrial location theory and the empirical studies undertaken with respect to locational patterns. [2]

Since industrial location theory has produced so little empirical evidence of the causal factors that determine industrial location patterns, projecting the C_{ij} term is still in an experimental stage. Two approaches have been tested. An econometric model which uses multiple regression to "explain" and project the C_{ij} effect for each of 50 industries has been developed in the Regional Economics Division. In it, the share effect is projected for each industry by a multiple regression analysis. That is, the C_{ij} effect in the most recent period for which data are available is regressed against a number of independent variables that relate to the preceding period or to a preceding point in time. This use of lagged variables obviates the necessity of making separate projections for each independent variable. The most significant of these variables is the C_{ij} effect in the preceding period. Additional independent variables include measures such as the size and rate of growth of the industry, total population, level of income, and the C_{ij} effect in related industries. Inclusion of this last variable makes it possible to establish appropriate interindustry linkages in the regression equations. Regression coefficients are calculated by "cross-sectional" analysis in which the value of the variable in each area forms an observation.

As empirical evidence is gathered, and as regional economic measurement is refined, it will be possible to select independent variables that have a closer and more stable relationship to the C_{ij} effect. At that time the foregoing method would seem to offer the most potential for development. However, in view of the paucity of data with which to measure past changes in the geographic location of industries and the comparative lack of information on factors underlying these changes, the foregoing approach to projecting the C_{ij} element of industrial change with its considerable emphasis on mathematical precision seemed unsatisfactory.

Accordingly, the second approach to projecting the C_{ij} term was a simple one that was less demanding of data and that could make maximum use of available information. For each industry, a simple curve was fitted to each region's share of the national total of income and employment (separately). This curve was then extended into the future and the values of the region's future share read.

This last approach is actually a variation of "shift-share" analysis with regional share effects (C_{ij}) calculated implicitly rather than explicitly. That is, from equation (2) the following relationship between changes in the regional share of the national industry (E_{ij}/E_{io}) and the regional-share effect (C_{ij}) of the shift analysis holds:

$$(2')\quad E^t_{ij}/E^t_{io} = E^o_{ij}/E^o_{io} + C^t_{ij}/E^t_{io}$$

$$(2'')\quad C^t_{ij} = E^t_{io}\left[E^t_{ij}/E^t_{io} - E^o_{io}/E^o_{io}\right]$$

$$= E^t_{io}\, \Delta\, (E_{ij}/E_{io}).$$

Statistical tests were applied to the several models as well as to others not described here. Although results at this stage are inconclusive, indications are that model 3 gave best results. Accordingly, it was chosen for further development. Model 3, it will be recalled, was the curvilinear extension by simple regression of a region's percentage share of the national total of income and employment in each industry. This mechanistic approach was modified in two ways.

First, substantial judgment was used in extending the curves. Such judgment reflected analysis of the numerous erratic observations in the historical time series; the timing of basic developments in a series; the status of the supply of the natural resource on which a particular industry depended; and the shape of the curve fitted to the measured observations. This approach permitted the full utilization of all information that could be assembled on any given industry in any region.

The projections made of the basic industries as outlined above (and specified in equation 2) were considered final. However, analysis of the interindustry relationships that prevailed in both income and employment in the 165 functional economic areas led to the modification of the shift-share projection model (as in model 3) to incorporate some features of an older basic-service model in projecting residentiary industries. This comprised the second of the two modifications referred to above.

Studies of the relationships of local-service or typically residentiary activities to export or basic industries in a region have given rise to an often used basic-service model. The interactions of the exogenous and the local-service industries result in a multiplier effect very similar to a Keynesian consumption multiplier. In the case of the basic-service model, the endogenous or internally determined sector is comprised of local-service activities such as trade, local transportation and other service activities. Since the function of these local-service or residentiary activities is to supply the local businesses and households with commodities and services which do not enter into interregional trade in substantial amounts, the magnitude of these residentiary activities is determined by the size of the population and income of the region. Thus, regional residentiary employment or income is functionally determined by regional total employment or income and hence must be solved simultaneously with the latter two aggregates.

The total employment and total income necessary to solve the residentiary industry equation were obtained by summing the final projections for basic industries and the preliminary projections for the residentiary group. Final projections of residentiary industries were then endogenously determined by functional relationships estimated in cross-section studies and projected forward by means of these relationships together with projected changes in the exogenous sector.

$$(3) \quad E^t_{oj} = \sum_{i=1}^{k} E^t_{ij} + \sum_{i=k+1}^{n} E^t_{ij}$$

$$(3') \quad \sum_{i=1}^{k} E^t_{ij} = \sum_{i=1}^{k} \left[(E^t_{io}/E^o_{io})E^o_{ij} + C^t_{ij} \right] = B^t_{oj}$$

$$(3'') \quad \sum_{i=k+1}^{n} E^t_{ij} = f(E^t_{oj}) = a_{oj} + b_{oj}E^t_{oj} = S^t_{oj}$$

Where E^t_{oj} = the sum of regional employment in industries $(i = 1, \ldots, n)$, of which the exogenous industries $(i = 1, \ldots k)$ are projected by means of the shift-share model as in equation (3) and the endogenous local-service industries $(i = k+1, \ldots, n)$ are jointly determined with total regional employment (here shown as a simple linear relationship where the a + b parameters are estimated by cross-section analysis).

Substituting (3') and (3'') into (3) and simplifying gives the multiplier value similar in structure to the Keynesian multiplier ($1/1$ - marginal propensity to consume locally produced goods and services).

$$(3''') \quad E^t_{oj} = B^t_{oj} + a_{oj} + b_{oj}E^t_{oj} = {}^{1}/_{1-b_{oj}} \, (B^t_{oj} + a_{oj})$$

Indeed, cross-section analysis undertaken by the Regional Economics Division has permitted estimates of industrially disaggregated regional residentiary sector multipliers.

$$(3'''') \quad E^t_{oj} = B^t_{oj} + \sum_i (a_{io} + b_{io})E^t_{oj} = ({}^{1}/_{1} - \sum_i b_{io}) \, (B^t_{oj} + \sum_i a_{io}) \ \underline{3/}$$

Clearly, since the a_{io} and b_{io} parameters represent a national central tendency, they do not necessarily fit the current case for individual regions. The Regional Economics Division is adjusting them for regional use however, by trending the current residentiary mix with respect to total regional employment toward the "national" parameters over the projection period. Thus, the working assumption is that regional local consumption patterns will trend toward national uniformity.

It is a deficiency of basic-service models of the type represented in the (3) series equations as well as of regional input-output models, that regional growth is caused entirely by external stimulation through the growth of the exogenous sector. While this deficiency is not altogether redressed in this methodology, it is diminished to the extent that the relationship between the basic and the local-service sectors is stable. Such stability is, of course, greater, the more successful we are in delineating nodal regional configurations.

Thus, if r_e and r_b are the rates of growth of regional total employment, and regional basic or exogenous employment, respectively, it can be shown that in this model the rate of growth of basic employment determines jointly the rates of growth of total regional and local-service employment.

$$(3''''') \quad \frac{E^t_{oj}}{E^o_{oj}} = (1 + r_e)^t = \frac{1}{E^o_{oj}} \, \frac{1}{1-b_{oj}} \, (a_{oj} + B^t_{oj})$$

$$= \frac{1}{\frac{a_{oj}+B^o_{oj}}{1-b_{oj}}} \, \frac{1}{1-b_{oj}} \left[a_{oj} + B^o_{oj}(1+r_b)^t \right]$$

$$= \frac{a_{oj} + B^o_{oj}\,(1 + r_b)^t}{a_{oj} + B^o_{oj}}$$

This deficiency is all the greater, the greater the variability in the relationship between the exogenous and the endogenous sectors. Conceptually, the relationship between the exogenous and the endogenous sectors, here the basic and the local-service industries respectively, is most stable in a nodal regional delineation scheme and least stable in an arbitrary or administrative delineation scheme. Thus, the relatively closed trade area incorporated in the functional economic area concept would hypothetically permit less variance between exogenous and endogenous sectors than would regions which were identical with county boundaries or delineation based on administrative or homogeneous groupings of county units. Empirical studies performed by the Regional Economics Division with respect to indexes of industrial centralization and of relative regional specialization support the hypothesis that less variation in the basic-service relationship occurs in such nodal regions as OBE Economic Areas than in non-nodal regions comprised of single counties or homogeneous or arbitrary groupings of counties. [4] Hence, the validity for projecting such cross-section relationships as basic-service interactions forward in time decreases as regional delineations depart from the nodal regional concept.

Footnotes

1. For a detailed explanation of this type of analysis, see Growth Patterns in Employment by County, 1940-1950 and 1950-1960, Lowell D. Ashby, Office of Business Economics, U. S. Department of Commerce, 1965.

2. B. H. Stevens and C. A. Brackett argue that this lack of correspondence is in part attributable to the inability of existing theory to generate testable hypotheses. Cf. Industrial Location, A Review and Annotated Bibliography of Theoretical, Empirical and Case Studies, Regional Science Research Institute, Philadelphia, 1967. This dearth of hypothesis testing has also been noted by J. Meyer, "Regional Economics: A Survey," American Economic Review, LIII, No. 1, March 1963.

3. Insofar as the parameters are central tendencies over all the regions of the Nation, they represent, as it were, national coefficients, hence, the index notation a_{io} and b_{io} where the right hand notation position indicates their national character. This deficiency of relying on national or adjusted national coefficients is shared with many regional input-output projection models. Indeed, it can be shown that if the sectors were the same, the parameters in equation (3'''') are algebraically the same as those in the employment multiplier model estimated in the input-output projection study for the New York SMSA. Cf. B. Berman, B. Chinitz and E. Hoover, Projections of a Metropolis, Harvard University Press, Cambridge, 1961, pp. 8-9.

4. It must be borne in mind that for any given industrial sectoralization, increasing the size of the region has the tendency to decrease the variation among regions. Since increasing size actually means aggregating contiguous counties each with their own industrial mix, each county added to a regional configuration implies a discrete and not necessarily monotonic change. Nonetheless, counties lying in the hinterlands of urban centers, on the average, do exhibit the tendency toward reducing regional specialization when added to the urban centers. That is, when location quotients are the means for distributing portions of the industrial sectors among the basic and residentiary sectors we have:

$$\frac{E_{ij}/E_{oj}}{E_{io}/E_{oo}} = L_{ij}$$

$$B_{oj} = \sum_{i,L_{ij}>1} E_{ij}\left(1 - \frac{1}{L_{ij}}\right) = \sum_{i,L_{ij}>1} \left[E_{ij} - E_{io}(E_{oj}/E_{oo}) \right]$$

Therefore, regional specialization, B_{oj}/E_{oj}, tends to zero as regions are summed, $\sum_j (E_{oj}/E_{oo}) = 1$, and $E_{ij} - E_{io}$ tends toward zero. Thus, we have to distinguish between decreases in the regional specialization index resulting purely from the size effects implicit in the mechanical aggregation of counties from those decreases resulting from nodal regional delineation.