DOCUMENT RESUME

ED 025 661 VT 007 611

By- Morton, J.E.
On Manpower Forecasting. Methods for Manpower Analysis, No.2.
Upjohn (W.E.) Inst. for Employment Research, Kalamazoo, Mich.
Pub Date Sep 68
Note- 57p.
Available from- W.E. Upjohn Institute for Employment Research, 1101 Seventeenth Street, N.W. Washington, D.C. 20036
EDRS Price MF-$0.25 HC-$2.95
Descriptors- Economic Research, *Employment Projections, *Manpower Needs, Mathematical Models, *Measurement Techniques, Prediction, *Research Criteria, *Research Methodology, Research Problems

Some of the problems and techniques involved in manpower forecasting are discussed. This non-technical introduction to the field aims at reducing fears of data manipulation methods and at increasing respect for conceptual, logical, and analytical issues. The major approaches to manpower forecasting are explicated and evaluated under the headings: (1) Some Curve-Fitting Techniques, involving essentially the methods of population forecasting, (2) Direct Manpower Forecasts, which make use of manpower variables only, (3) Derived Manpower Forecasts, which rely on safely predictable variables (population or economic) which are associated with manpower variables, and (4) Econometric Models, which mathematically depict relationships of single or multiple variables. An introductory section discusses the role of manpower forecasting and its historical development. The concluding section reviews the forecasting techniques in terms of the following dichotomies: (1) short-term versus long-term forecasting, (2) stochastic versus the deterministic approach, (3) point versus interval forecasts, (4) unconditional versus conditional forecasts, and (5) first-order versus higher-order forecasts. (ET)

Methods for Manpower Analysis No. 2

# On Manpower Forecasting



J. E. Morton

THE W. E. UPJOHN INSTITUTE FOR EMPLOYMENT RESEARCH
ESTABLISHED 1945

# ON MANPOWER FORECASTING

By

J. E. MORTON



*September 1968*

*The W. E. Upjohn ~~Institute~~ for Employment Research*
*300 South Westnedge Avenue*
*Kalamazoo, Mich~~igan~~ 49007*

## *The Staff of the Institute*

HAROLD C. TAYLOR, PH.D.
*Director*

HERBERT E. STRINER, PH.D.
*Director of Program Development*

### *Kalamazoo Office*

RONALD J. ALDERTON, B.S.
SAMUEL V. BENNETT, M.A.
EUGENE C. MCKEAN, B.A.
HAROLD C. TAYLOR, PH.D.
HENRY C. THOLE, M.A.
W. PAUL WRAY

### *Washington Office*

E. WIGHT BAKKE, PH.D.
A. HARVEY BELITSKY, PH.D.
SAUL J. BLAUSTEIN, B.A.
ROSANNE E. BURCH, B.A.
SAMUEL M. BURT, M.A.
SIDNEY A. FINE, PH.D.
HENRY E. HOLMQUIST, M.A.
J. E. MORTON, JUR.D., D.SC.
MERRILL G. MURRAY, PH.D.
HAROLD L. SHEPPARD, PH.D.
IRVING H. SIEGEL, PH.D.
HERBERT E. STRINER, PH.D.

# Preface

This paper is intended to serve as a non-technical introduction for the general manpower analyst who, not himself an expert in the use of forecasting techniques, is desirous to form an opinion about some of the problems and techniques involved; in particular, it aims at reducing fears of data manipulation methods and at increasing respect for conceptual, logical and analytical issues.

Since the field of manpower forecasting is in a state of flux, and includes methods borrowed from early demographic writings as well as from the most recent econometric literature, it should be pointed out that the manuscript was completed early in Spring 1967 and that, accordingly it reflects the state of the arts as of approximately that time.

*The W. E. Upjohn Institute*
*for Employment Research*

THE INSTITUTE, a privately sponsored nonprofit research organization, was established on July 1, 1945. It is an activity of the W. E. Upjohn Unemployment Trustee Corporation, which was formed in 1932 to administer a fund set aside by the late Dr. W. E. Upjohn for the purpose of carrying on "research into the causes and effects of unemployment and measures for the alleviation of unemployment."

One copy of this bulletin may be obtained without charge from the Institute in Kalamazoo or from the Washington office, 1101 Seventeenth Street, N.W., Washington, D.C., 20036. Additional copies may be obtained from the Kalamazoo office at a cost of 50 cents per copy.

## Contents

| | Page |
|---|---|
| I. THE ROLE OF MANPOWER FORECASTING | 1 |
| II. SOME CURVE-FITTING TECHNIQUES | 5 |
| III. DIRECT MANPOWER FORECASTS | 20 |
| IV. DERIVED MANPOWER FORECASTS | 25 |
| V. ECONOMETRIC MODELS | 32 |
| VI. RETROSPECT AND PROSPECT | 46 |

# On Manpower Forecasting

## *I. The Role of Manpower Forecasting*

The forecasting of manpower and of manpower components has a long and distinguished history. From the time of the early demographers, numerical outlook estimates of the population and related variables have influenced the opinions of social philosophers, statisticians, politicians, and administrators, as well as the public at large. In the past few years, a period during which the term "manpower" has acquired technical status, forecasts of labor force and employment have become increasingly familiar.[1] They are nowadays used often as a basis for forecasting other major economic measures, e.g., the Gross National Product (GNP).

The President's Committee on Manpower, established to assist in carrying out federal responsibilities under the Manpower Development and Training Act of 1962, has shown a strong interest in the design and construction of consistent govermental forecasts of the nation's labor resources and needs. This interest is reflected in a 1967 report of a Working Group established by the Bureau of the Budget.[2] The group's activities, in turn, have stimulated the preparation of a comprehensive and thought-provoking commentary on long-range forecasting of the labor force.[3]

The objective of forecasting is to reduce our uncertainty about future events.[4] Forecasting, thus, is the art of forming expectations and antici-

[1] See, e.g., H. Goldstein, "Projections of the Labor Force of the United States, "*in* U. S. Congress, Senate Committee on Labor and Public Welfare, Hearings, *Nation's Manpower Revolution*, pt. 5 (1963) and "Projections of Manpower Requirements and Supply," *Industrial Relations*, Vol. 5 (1966); S. L. Wolfbein, "Manpower in the United States with Projections to 1970," *in* U. S. Congress, House Committee on the Judiciary, Special Series No. 3 (1962); B. Grais, *Forecasting of the Active Population by Occupation and Level of Skill* (Paris: Organization for Economic Cooperation and Development, 1966); U. S. Bureau of Labor Statistics, *Projection 1970*, Bulletin No. 1536, 1966; U. S. Department of Labor, *Manpower Projections*, a bibliography, 1966.

[2] U. S. Department of Labor, Manpower Administration, *Manpower Projections: An Appraisal and a Plan for Action*, August 1966. In this publication there also can be found examples of the application of manpower forecasting to a wide spectrum of matters relevant to national policy. For bibliographical references, see pp. 34 ff.

[3] U. S. Bureau of Labor Statistics, *Long-Range Projection of Labor Force*, by Dennis F. Johnston, September 1967; mimeographed.

[4] Two terms are available: "forecasting" and "prediction." Both terms are in general use. Sometimes "prediction" is used for the more formal and "forecasting" for the more informal procedures—see, e.g., Jay W. Forrester, *Industrial Dynamics* (Cambridge, Mass.: Technology Press, MIT, 1961); sometimes "prediction" is used for vague conjecture and "forecasting" for quantitative procedures based on specific observation and information—see Robert G. Brown, *Smoothing Forecasting and Prediction of Discrete Time Series* (Englewood Cliffs, N. J.: Prentice-Hall Inc., 1963). The term "forecasting" seems to have been more widely used in book titles—C. W. F. Butler and Robert A. Kavesh, *How Business Economists Forecast* (New York: Prentice-Hall, Inc., 1966); E. Bratt, *Business Cycles and Forecasting* (5th ed.; Homewood, Ill.: Richard D. Irwin, Inc., 1961); and Milton

pations in which we have greater confidence than in the unguided and unsupported guess. Whether because of the urge of idle curiosity, or the necessity to arrive at an important decision, man has since ancient times sought to unveil the future; in this attempt to discount the tomorrow, and to strengthen confidence in his own expectations, he has been looking to oracles and prophecy, to astrologers and numerologists, and later to more nearly objective and scientific devices for clues and revelation.

In the behavioral sciences the importance of the problem of forecasting in the modern sense was early recognized by demographers and actuaries, who were in many respects predecessors of the students of manpower problems. The construction of life tables, one of the first major forecasting devices, can be traced to Graunt and Halley, both of the seventeenth century.[5] Later, in 1783, Richard Price translated the findings into actuarial practice. During the nineteenth century the first documented attempts were undertaken to make projections of population totals as well as of vital rates on the basis of mathematical functions.[6]

In the area of economics during the second half of the last century and the first quarter or so of the twentieth, interest in prediction and forecasting techniques originated with business cycle and time series analyses.

The study of repetitive "commercial crises" resulted in the identification of apparent oscillatory movements or "cycles." Anticipating turning points of the wavelike movement assumed considerable practical importance, and the business cycle became a basic concept for analysis and interpretation of these recurrent economic movements.[7]

By the first quarter of this century the problem of the economic cycles

---

H. Spencer, Collin J. Clark, and Peter W. Hoguet, *Business and Economic Forecasting* (Homewood, Ill.: Richard D. Irwin, Inc., 1961). H. Theil's recent book, *Applied Economic Forecasting* (Amsterdam: North-Holland Publishing Company, 1966), refers in the index under "forecasting" to "prediction." C. F. Christ's recently published *Economic Models and Method* (New York: John Wiley & Sons, Inc., 1966) uses the two terms interchangeably.

[5] John Graunt, *Natural and Political Observations Made Upon The Mortality* (London: 1662; reprinted in Baltimore: The Johns Hopkins Press, 1939); and E. Halley, *An Estimate of the Degrees of Mortality ....*, Philosophical Transactions of The Royal Society of London (London: 1693; reprinted in Baltimore: The Johns Hopkins Press, 1942).

[6] See especially P. F. Verhulst, "Notice sur la loi que la population suit dans son accroissement," *in Correspondance mathématique et physique* publiée par A. Quêtelet, Tome X (Bruxelles: M. Hayez, 1838) and "Recherches mathématiques sur la loi d'accroissement de la population," *Nouveaux mémoirs de l'Académie Royale des sciences et belles lettres* de Bruxelles, XVIII, 1 (1845); and H. Gylden *Försäkringsföreningens Tidskrift* (Stockholm, 1878); see also H. Cramer and H. Wold, "Mortality Variations in Sweden," *Skandinavisk Aktuarietidskrift*, Uppsala 161 (1935).

[7] See, e. g., Clément Juglar, *Des crises commerciales et leur retour périodique* (Paris: 1862; reprinted in New York: Augustus M. Kelley, 1967).

had become one of the central topics of economic theory and analysis.[8] Once the cyclical nature of economic change was believed to have been established, growing interest manifested itself in the practical question of how to forecast the behavior of this dynamic system. The careful description and decomposition of economic time series into various components—the long waves,[9] the shorter cycle, and the still shorter seasonal component—occupied leading economists of the time.[10]

In this country the new insight generated, at an early stage, devices and schemes for use in practical business forecasting which later (1917) led to the famous barometer system of the Harvard Economic Service.[11] From the United States these forecasting services and institutions spread over most of Europe, the best known being the London and Cambridge Economic Service, established in 1923, and the Institut fuer Konjunkturforschung in Berlin established two years later.

This early enthusiasm was damped by those who raised the fundamental question of forecastability—on the continent, among others by Kondratieff and Morgenstern—and shortly thereafter by the unforeseen disaster of 1929.[12] But by the mid-thirties the scene was set for the reappearance of forecasting and the study of economic time series as a legitimate subject deserving most serious attention in economic and statistical theory and analysis.

Special concern about the particular problem of manpower prediction is relatively recent. But its late appearance has made manpower forecasting the beneficiary and captive of earlier experiences and efforts elsewhere, notably in demography and in the empirical analysis of business cycles. This venture into the domain of prognostication—at first very slowly and only gradually expanding—received a potent fillip during the depression of the thirties. The unemployment problem, and the resulting attempts by government to cope with it, required decisions, manpower policies, and strategies predicated on "early warning systems" and on more or less clearly visualized effects on the future supply-demand

[8] See, for instance, Henry L. Moore, *Economic Cycles* (New York: The MacMillan Company, 1914) and *Forecasting Yield and Price of Cotton* (New York: The Macmillan Company, 1917).

[9] N. D. Kondratieff, "Die Langen Wellen der Konjunktur," *Archiv fur Sozialwissenschaft und Sozialpolitik*, LVI (1926).

[10] Wesley Mitchell, *Business Cycle* (New York: National Bureau of Economic Research, 1927); and much of the literature produced by the staff of the National Bureau of Economic Research.

[11] For earlier attempts in this direction see, e.g., *R. W. Babson Reports* since 1904 and Thomas Gibson's *Stock Market Forecasts* since 1907. The Harvard barometer system is best described in the works of Warren M. Persons.

[12] N. D. Kondratieff, "Das Problem der Prognose" *Annalen der Betriebswirtschaft*, Vol. I (1927); and Oskar Morgenstern, *Wirtschaftsprognose, eine Untersuchung ihrer Voraussetzungen und Nuetzlichkeit* (Vienna: J. Springer, 1928).

situation. A similar contingency, produced by a reversal in the labor supply-demand discrepancy during the war and postwar years of the forties, further accentuated the need for better and more detailed manpower forecasts.

Among the present problems, to the solution of which manpower forecasts are supposed to contribute, are: the future demand for workers; the expected composition of the labor force with reference to educational and training need; the feasibility of new programs and ventures in terms of availability of required skills; early warning systems foreshadowing disruptive effects of automation; and the development of vocational guidance and occupational outlook information. Labor force forecasts also play an important role in developing general economic projections.

Then there are the manpower forecasts which are not directed to a specific decision problem and which, by analogy with government statistics, might be termed general-purpose forecasts. There are also the forecasting objectives that are best attributed to the "fears and fascination of the future,"[13] and the broad interests illustrated by the French group of "futuribles" and their principal exponent, Bertrand de Jouvenel.[14]

These are macroeconomic objectives, but manpower forecasting has also found its way into microeconomic practice. Many smaller but specialized firms, large corporations, and of course, individuals are aware of the uncertainties of future labor markets. Consequently, it has become important for the forecaster to form an opinion about how best to cope with the future in terms of optimum utilization of employees, and other personnel policy matters aimed at the long run. Because of uncertainties inherent in forming an opinion, decisions about such matters are sometimes predicated on the earlier mentioned type of macro forecasts; but often special micro predictions must also be prepared, tailormade for the particular purposes at hand. Among them should be mentioned forecasts of the need by a given company for highly trained, specialized personnel; anticipation of labor turnover rates; and process analysis applied to absenteeism, accidents, and sickness.[15]

In view of this vast array of diverging objectives and uses of manpower forecasts, it is not surprising to find widely differing methods and schemes of forecasting, ranging from quite primitive to highly sophisticated ones.

[13] Daniel Bell, "The Study of the Future," *in The Public Interest* (1965), as quoted by Garth L. Mangum and Arnold L. Nemore, "The Nature and Functions of Manpower Projections," *Industrial Relations*, v. 5, n. 3 (May 1966).

[14] See, for instance, the stimulating tract by Bertrand de Jouvenal, *The Art of Conjecture* (New York: Basic Books, Inc., Publishers, 1967).

[15] See, e.g., Ore Lunberg, *Random Processes and Their Application to Sickness and Accident Statistics* (Uppsala: Almquist Wiksells, 1940), and the copious recent literature in operations research and the management sciences. Methods other than those used in macro prediction will not be discussed in this paper but be deferred to a later occasion.

But none of them is fully satisfactory, and much criticism and doubt have been expressed with respect to methods as well as concepts.[16] Nevertheless, forecasting, in one form or another, has become a basic and necessary constituent of decision making in modern society. It is an essential ingredient in the choice of man's responses to a dynamic environment. All forecasting efforts, by whichever method, have one element in common: the search for relative stabilities and invariances of relationship. These suspected relationships may be simple ones between the phenomenon to be forecast and the data (not necessarily quantitative) to be used as "predictors." With increasing acceptance of factualism and quantification, it was perhaps to be expected that sooner or later the quest would be directed toward discovering such stable dependencies of "predictand" on "predictors"—in terms of original values of variables, rates of change, transformations of variables, and even entire systems of equations.

However complex the resulting methods and models, their forecasting powers stand and fall with the validity of the assumed or implied invariances. No procedure for *a priori* validation of such presuppositions has yet been found. Although comparison of realization with forecast can help refute the soundness of the method that produced the particular forecast, the reverse generally cannot be inferred with certainty. Even the more recent application of the "science of uncertain inference" to the study of economic time series has not, by and large, yielded satisfactory results with respect to forecasting of economic time series. But there are better and worse forecasting methods, and there are more and less useful techniques. All in all, however, socioeconomic forecasting still remains an art.

## *II. Some Curve-Fitting Techniques*

As suggested above, manpower forecasting has developed in close historical and logical affinity with population forecasting. The manpower forecaster therefore needs to be aware of population forecasting methods and techniques.

Among the simplest expressions used in demographic forecasting are the polynomials,[1] beginning with the polynomial of degree one, i.e., the

[16] For a recent and succinct discussion of some of the issues, see Lee Hansen and Herbert E. Striner, "The U. S Manpower Future," *Proceedings of the Eighteenth Annual Winter Meeting of the Industrial Relations Research Association* (Madison: The Association, 1966)

[1] A polynomial can be expressed by the equation $P = a + bT = cT^2 + dT^3 + \cdots$, where $P$ might stand for population, $T$ for time (say, calendar years, and $a$, $b$, $c$, $d$, $\cdots$, are coefficients. The highest power of $T$ whose coefficient is not zero determines the degree of the polynomial; thus $P = a + bT$ is a polynomial of degree one, or the equation of a

straight line. An early example in this country is Abraham Lincoln's projection by straight-line extrapolation of the population growth between 1790 and 1860 to 1925;[2] even after a downward revision, he thus obtained an estimate of 250 million. It is of course easy to criticize *ex post facto* such simple procedures; nevertheless, there are many instances in which linear projections are of interest to the manpower analyst because of simplicity, plausibility for short-term forecasting, or lack of a better hypothesis. For instance, a linear estimating equation may suffice when changing labor force participation rates are to be projected over the short run. The linear extrapolation may be considered an approximation to the unknown nonlinearity—an approximation that is better the shorter the forecasting horizon.

Several techniques for fitting a straight line to empirical data are available, among them the method of selected points and the method of least squares.

A linear trend characterizes a relationship between the particular observed variable, say, population ($P$) and time ($T$) so that both $P$ and $T$ progress by constant absolute differences. In other words, both $P$ and $T$ form an arithmetic progression. Since time is conventionally measured in equal intervals, e.g., years, quarters, or months, our attention is centered only on $P$. Setting up a difference table, we observe the first differences in the successive $P$ observations. If they are constant, or nearly so, the second differences tend to vanish. Using a hypothetical example (see table 1), we find the second differences ($\Delta^2$) to be "small" indeed. We may therefore wish to fit a straight line.[3]

In the method of selected points, it is useful to plot the data (see Diagram 1). The selection of points is greatly facilitated by first fitting a line by inspection. The closer the points actually do cluster about such a line, the easier it will obviously be to agree on such a line. Once the points are plotted, we select any two (preferably two distant ones) and read off their coordinates. In the example, points $P^*_{52}$ and $P^*_{58}$, for the years 1952 and 1958, respectively, were selected. (The asterisks indicate "estimated" quantities, i.e., those on the trend line, not the actually observed population figures.) The yearly increment is $\dfrac{P^*_{58} - P^*_{52}}{1958 - 1952}$ or $\dfrac{55 - 44}{6} = \dfrac{11}{6}$. Thus, the constant annual difference (the slope of the line in the diagram) is $\dfrac{11}{6} = 1.83$.

---

straight line, where $a$ indicates the value of $P$ at $T = 0$, and $b$ is the "slope," i.e., the increment, which is constant, or decrement in $P$ per unit increase in $T$.

[2] J. J. Spengler, "Population Prediction in Nineteenth Century America," *American Sociological Review*, 1 (1936).

[3] See also F. S. Acton, *Analysis of Straight-Line Data* (New York: John Wiley & Sons, Inc., 1959).

*Table 1*
*A Hypothetical 1950–1960 Population*
*(in millions)*

| *Year* | *P* | $\Delta^1$ | $\Delta^2$ | *t* | $t^2$ | *tP* | $t^4$ | $t^2P$ |
|---|---|---|---|---|---|---|---|---|
| 1950 | 38 | 5 | | −5 | 25 | −190 | 625 | 950 |
| 1951 | 43 | 1 | 4 | −4 | 16 | −172 | 256 | 688 |
| 1952 | 44 | 2 | 1 | −3 | 9 | −132 | 81 | 396 |
| 1953 | 46 | 0 | 2 | −2 | 4 | −92 | 16 | 184 |
| 1954 | 46 | 2 | 2 | −1 | 1 | −46 | 1 | 46 |
| 1955 | 48 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1956 | 50 | 2 | 0 | +1 | 1 | 50 | 1 | 50 |
| 1957 | 52 | 2 | 0 | +2 | 4 | 104 | 16 | 208 |
| 1958 | 55 | 3 | 1 | +3 | 9 | 165 | 81 | 495 |
| 1959 | 58 | 3 | 0 | +4 | 16 | 232 | 256 | 928 |
| 1960 | 59 | 1 | 2 | +5 | 25 | 295 | 625 | 1,475 |
| | 539 | | | | 110 | 214 | 1,958 | 5,420 |

If one prefers a more formal criterion of best fit, he may choose the so-called least-squares principle.[1] Accordingly, a best fitting line is the one about which the squared deviations of the observations are smaller than about any other such line. To simplify the mechanics, the "origin" of the observations will be shifted to the arithmetic mean of $T$. Put differently, instead of measuring time in actual units, say, calendar years, we measure it in terms of deviation from the middle of the period, $\bar{T}$, so that $t = T - \bar{T}$. Under these conditions it turns out that the slope or annual increment, equals $\frac{\Sigma tP}{\Sigma t^2}$.

Taking the same example, we find that $\frac{\Sigma tP}{\Sigma t^2} = \frac{214}{110} = 1.95$.

This value differs, but not very much, from the slope or annual increment found earlier, 1.83.

To write the equation of the linear trend using the slope-intercept form, $P = a + bt$, we need to know, in addition to the slope of the line ($b$), the coordinates of one point on the line. Such a point is $a$ in the above equation. In particular, $a$ is the value of $P^*$ at the "origin" of the series. If the origin is taken at $\bar{T}$ (for which $t = 0$), the least-squares principle

[1] Irving H. Siegel, "Productivity Measures and Forecasts for Employment and Stabilization Policy," *in* Sar A. Levitan and Irving H. Siegel, eds., *Dimensions of Manpower Policy: Programs & Research* (Baltimore: The Johns Hopkins Press, 1966), pp. 269–88.

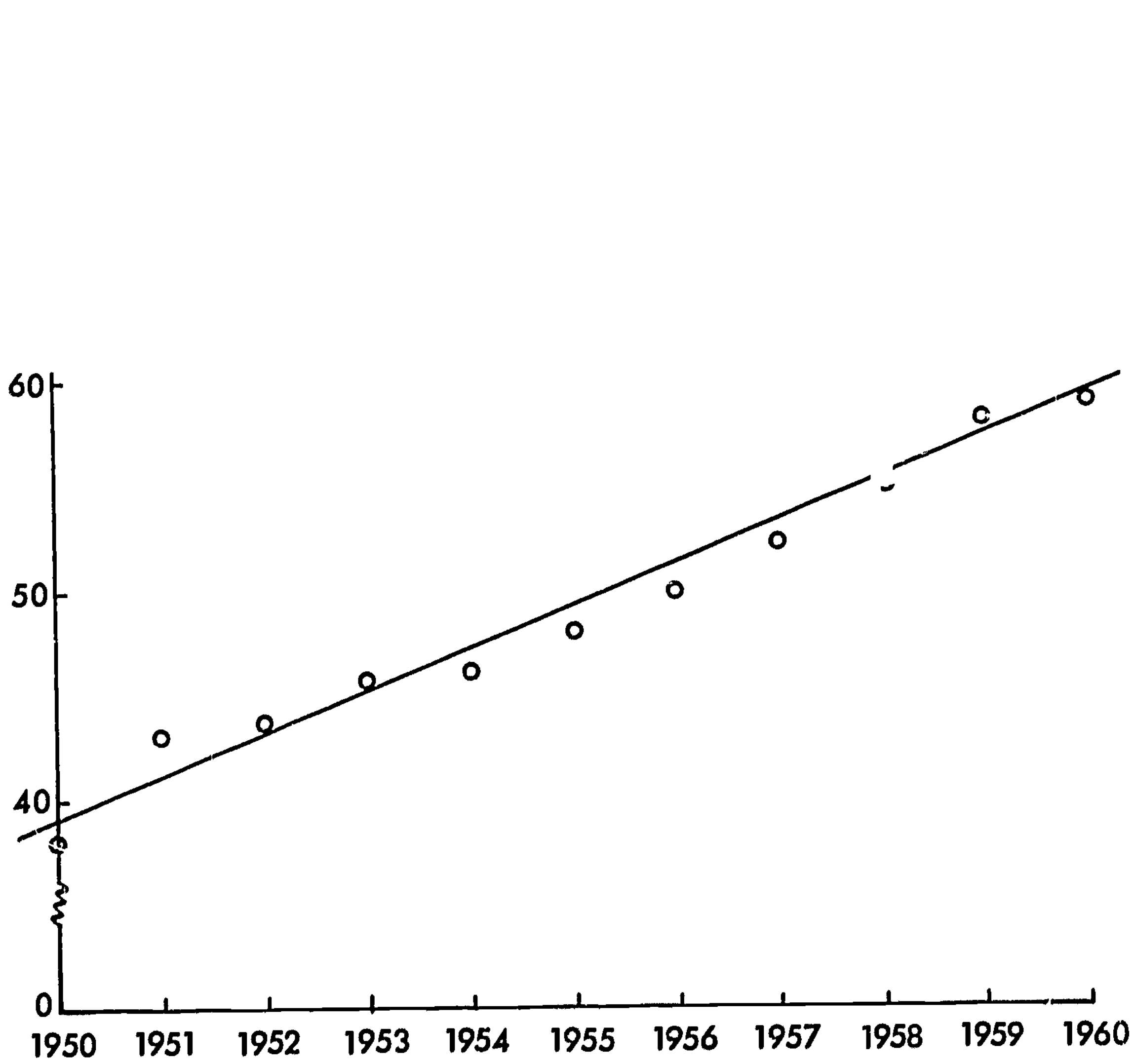


*Diagram 1*
*A Hypothetical 1950–1960 Population*
*(in millions)*

yields $a = \frac{\Sigma P}{n}$. That is, the least-squares line passes through the arithmetic mean of the observations.

Since in our example $\frac{\Sigma P}{n} = \frac{539}{11} = 49$, the equation of the least-squares line is $P^* = 49 + 1.95t$ (with origin at $t = 0$, i.e., at 1955).

Had we used the selected-points method, for which we had found the slope $b = 1.83$, the value of $a$ would have to be read off Diagram 1. Using again $\bar{T}$ as the origin, we find that $a = 49.5$, and that the equation of the trend is $P^* = 49.5 + 1.83t$, which is similar to, but not identical with, the one found by the method of least squares.

As already mentioned the implication of the linear trend is that the variable $P$ is a linear function of time. In other words, its average increase or decrease per unit of time is assumed constant over time.

It was soon observed that, in considering longer stretches, the population increment per unit of time did not remain constant, especially for a dynamic nation such as the United States. In the quest for a flexible but simple descriptive device, polynomials of higher than first degree soon attracted the attention of forecasters.

Early interest in fitting polynomials of higher degree to population changes was aroused by the mathematician and astronomer Pritchett, who fitted a third-degree polynomial to the population of the United

States. On the basis of this equation he made forecasts for 10 centuries hence, arriving at 222 million for 1960, and at 41 billion for 2900.[5]

Although nowadays one would hardly use polynomials for projecting population totals over long periods, the method is of interest for fitting curves over shorter periods to selected manpower variables. Polynomials of varying degree provide a very flexible means for fitting nonlinear functions to observational data.

The $(n + 1^{th})$ differences of an $n$ degree polynomial vanish a useful property for experimenting with functions to fit time-series. Polynomials thus can be regarded as a flexible, curved ruler; a polynomial of degree $n$ has $n - 1$ points of inflection. In fitting a polynomial of degree two to the data in our example by the method of least-squares, we obtain a trend equation of the form:

$P^* = a + bt + ct^2$, in which $a = \frac{\Sigma P - c\Sigma t^2}{n}$

$$b = \frac{\Sigma tP}{\Sigma t^2}$$

$$c = \frac{n\Sigma t^2 P - \Sigma t^2 \Sigma P}{n\Sigma t^4 - (\Sigma t^2)^2}$$

Substituting in these expressions, we find:

$$a = \frac{539 - 3\ 85}{11} = 48.65$$

$$b = \frac{214}{110} = 1.95, \quad \text{and}$$

$$c = \frac{59{,}620 - 59{,}290 - 59{,}290}{21{,}538 - 12{,}100} = \frac{330}{9438} = .035$$

Hence, the sought trend equation is: $P^* = 48.65 + 1.95t + .035t^2$.

The greater flexibility of polynomials is not necessarily an advantage from the manpower analyst's point of view. The closer fit of functions to more or less erratic oscillations of economic time series obscures the structural meaning which is necessary for forecasting. In other words, a function that closely fits irregular fluctuations of empirical manpower data for the past should hardly be assumed to follow the same complex path in the future unless good theoretical reasons suggest such a course. In socioeconomic situations, we rarely ever are justified, *a priori*, to expect such behavior.[6] Where time series are to be decomposed into components that separate systematic from more nearly accidental

[5] H. S. Pritchett, "A Formula for Predicting the Population of the United States," *Transactions of the Academy of Sciences of St. Louis* (1890).

[6] Hence the problem of overfitting, as it is called by H. O. A. Wold in his *Study Week on The Econometric Approach to Development Planning* (Amsterdam: North-Holland Publishing Company, 1965), Chapter 2; he properly relegates this pitfall to the "nonsense department of statistical method."

movements, other smoothing devices are available. Should the reader wish to apply higher-degree polynomials to fit observational data forming a time series, he may use the method of *orthogonal polynomials*. The principle of fit involved is again the least-squares criterion, but the technique is considerably simpler than that for the classical least-squares solution if a set of appropriate tables is available.[7] Among the advantages of the method is that the necessary coefficients for higher polynomials may be completed in successive steps; thus a polynomial of degree $n$ can be changed to one of the $n + 1^{th}$ degree by computation of just one additional coefficient. This is a significant advantage in experimentation with polynomials of uncertain degree.

Exponential functions are most widely used in fitting curves to a growth phenomenon in the social sciences. Various polynomials are characterized by the constancy of their first or higher order differences in the dependent variable. If the independent variable (time) proceeds by equal steps, exponential functions describe a similar constancy in the pattern of change in relative rather than absolute terms. The exponential is therefore a useful device for describing a time series for which the additions are in turn a function of the magnitude of the variable; in other words, large values of the variable are associated with large increments, and conversely. Accordingly a series that increases at a constant percent rate can be conveniently presented by the simple exponential function or its graph.

The most common example for exponential patterns of growth is the so-called *organic law*,[8] (also called compound interest law). Unlike the linear function for which both variables proceed as an arithmetic series, one variable (say, time) in the exponential case progresses arithmetically; the other (say, $P$) geometrically. Such a relationship between two variables, e.g., $P$ and $T$ can be expressed by $P = ab^T$. Taking logarithms, we obtain $\log P = \log a + T \log b$, which gives $\log P$ as a linear function of time. In other words, if we plot logs of $P$ instead of the original values against time, we shall obtain the graph of a straight line. To simplify the plotting we may, rather than look up the logs of $P$, use graph paper ruled along the vertical axis in logarithms—the well known semilogarithmic paper. This simple transformation not only facilitates the diagnosis of

[7] Tables with instructions for use in fitting by the method of Orthogonal Polynomials can be found in R. A. Fisher and Frank Yates, *Statistical Tables for Biological, Agricultural, and Medical Research* (2nd ed.; London: Oliver & Boyd, 1943); W. Beyer, ed., *Handbook of Tables for Probability and Statistics* (Cleveland: The Chemical Publishing Company, 1966); and R. L. Anderson and E. E. Houseman, *Tables of Orthogonal Polynomial Values Extended to* $N = 104$, Research Bulletin 297 (Ames, Iowa: Agricultural Experiment Station, Iowa State College of Agriculture and Mechanic Arts, 1942). For a not-too-technical description of the method, see also W. E. Milne, *Numerical Calculus* (Princeton: Princeton University Press, 1949), Chapter IX.

[8] So-called because many populations in biology follow this pattern of growth, e.g., cultures of bacteria, populations of animals, and the like.

an exponential pattern of growth (or decline) but also simplfies the fitting procedure.

Plotting our data on semilogarithmic graph paper, we first assure ourselves that the resulting cluster of points may be reasonably well represented by a straight line. If such line appears reasonable, we fit a function, linear in the logarithms of $P$, by two selected "representative" points or by least-squares.

Let us consider first the determination of the formula from two points drawing a line fitted with the help of a ruler. We read from the graph the coordinates of two selected points (see Diagram 2) on the fitted line. One such point might be the population on the trend line corresponding to 1951; the other, to 1958. The $P$-ordinates are logarithms of the original $P$ that is the logarithm of 55 for $P$ in 1958, and the logarithm of approximately 41.5 for 1951. In the derived equation,

$$\log b = \frac{\log P_2 - \log P_1}{1958 - 1951} = \frac{\log 55 - \log 41.5}{7}.$$

Looking up the logarithms, we find

$$\log b = \frac{1.740363 - 1.618048}{7} = .017474.$$

Taking again the origin at $T = 1955$, we find that log 48 equals 1.681241; hence the equation of the exponential trend is $\log P^* = 1.681241 + .017474t$ with the origin in 1955.

If we prefer a least-squares solution, we shall first transform the successive $P$'s in Table 1 into logarithms (see Table 2), and proceed as earlier, obtaining

$$\log b = \frac{\Sigma(t \log P)}{\Sigma t^2} = \frac{1.908612}{110} = .017351 \quad \text{and}$$

$$\log a = \frac{\Sigma(\log P)}{n} = \frac{18.553078}{11} = 1.686643, \quad \text{so that}$$

$\log P^* = 1.686643 + .017351t.$

The two sets of coefficients are not identical but differ little.

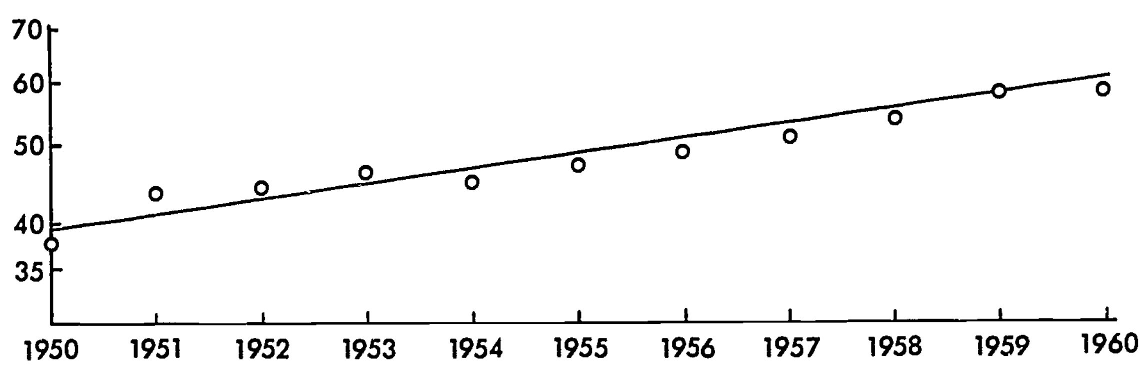


*Diagram 2*
*A Hypothetical 1950 1960 Population*
*(in logarithms of millions)*

*Table 2*

*A Hypothetical 1950 1960 Population*
*(in logarithms of millions)*

| *Year* | *t* | *Log P* | *t log P* |
|---|---|---|---|
| 1950 | −5 | 1.579784 | −7.898920 |
| 1951 | −4 | 1.633468 | −6.533872 |
| 1952 | −3 | 1.643453 | −4.930359 |
| 1953 | −2 | 1.662758 | −3.235516 |
| 1954 | −1 | 1.662758 | −1.662758 |
| 1955 | 0 | 1.681241 | 0.000000 |
| 1956 | 1 | 1.698970 | 1.698970 |
| 1957 | 2 | 1.716003 | 3.432006 |
| 1958 | 3 | 1.740363 | 5.221089 |
| 1959 | 4 | 1.763428 | 7.053712 |
| 1960 | 5 | 1.770852 | 8.854260 |
| | | 18.553078 | 1.908612 |

Translating these logarithmic equations of the exponential into their conventional, nonlogarithmic form, $P = ab^t$, we obtain (after looking up the appropriate antilogs) $P^* = 48\ (1.04)^t$ and $48.6\ (1.04)^t$, respectively, where $a$ is the value of $P^*$ for $t = 0$, i.e., 1955. Here $b = 1.04$ is the growth ratio, or $r$, in the geometric series $P, Pr, Pr^2 \ldots Pr^n$. Letting $r = 1 + i$, where $i$ is the rate of increase of the population expressed as a decimal increment, we recognize the compound interest formula of the accountant and may write the above exponential as $P_n = P_0\ (1 + i)^t$. In our example, subscript $n$ refers to the year to be estimated, subscript 0 to the base year, and $i$ to the percent rate of growth,[9] i.e., approximately 4 percent per annum.

The technique of fitting curves to manpower observations is evidently

[9] If interest is compounded $k$ times year, so that

$$P_n = P_0(1 + \frac{i}{k})^{kt},$$

it could be shown that as $k$ increases beyond all bounds, (i.e., as compounding takes place more and more often) the expression $(1 + \frac{1}{k}{}^k)$ reaches a limit called $e$ (approximately 2.7183). Hence we may also write $P_n = P_0e^{it}$, a formula representing the so-called compound interest law or the law of organic growth and reflecting continuous exponential change.

a simple enough matter.[10] A quite different matter, however, is the choice of an appropriate function or line.

To illustrate this point, we project the several lines obtained above, all of which appear to fit the observed data quite well, to the year 1975. Substituting 20 for $t$, in the several equations (since the origin was taken at 1955, $t = 20$ for 1975), we obtain:

> linear-trend values of 86.1 and 87.8 million for the visual-selected-points method and the least-squares method, respectively;
> exponential-trend values of 107.4 and 108 million for the two methods, respectively; and a second-degree polynomial value of 101.7 million.

Thus, although different fitting techniques may give sufficiently similar estimates, the discrepancies resulting from the choice of different functions are commonly quite pronounced. Unfortunately, no simple automatic procedure exists for answering the question as to the "proper" function to use. A thorough understanding of the nature of the underlying process a knowledge of the fundamental relationships to be described—is needed. The manpower analyst needs a valid theory rather than mere access to empirical data; and it is at this point that he may envy the occasional scientist whose theory enables explanation of the very mechanism or process he wishes to predict. This, of course, does not diminish the importance of manpower forecasting, which is more of an art rather than a science.

For the purposes at hand, the lack of a ceiling diminishes the plausibility of simple exponential patterns of growth. The exponential function here considered increases without limit. It would be easy enough, however, to provide for such a ceiling in the formula for exponential growth. One particularly simple possibility is to introduce a constant, say, $L$, to serve as an upper limit toward which the function tends without ever quite reaching it. Consequently, one might use the function $Y = L + ab^T$, where $L$ is the just-mentioned upper "asymptote"; $a$ is negative (if it were positive, the curve would approach $L$ as a lower limit), and $b$ is positive, but less than 1. The curve is concave downward and gradually tapers off in the right-hand upper corner as it approaches 1, the upper limit.[11]

For manpower projections this modification of the exponential has not found favor. It might be of interest, however, in developing economies

[10] In practice the technique is even simpler to apply than the above example suggests because any desk calculator with reasonable capacity will permit skipping the intermediate steps (shown here for expository purposes) and obtaining the needed totals immediately.

[11] A treatment in nontechnical language can be found in F. E. Croxton and D. J. Cowden, *Applied General Statistics* (New York: Prentice-Hall, Inc., 1939 and later).

when a target for employment is to be reached gradually after an initial spurt.

Another pattern of growth, an *S*-curve, also approaches an upper limit and can be represented by a modification of the simple exponential. This curve is of more direct concern to population and manpower analysts. The growth of certain phenomena advances slowly at first, then rapidly, and thereafter tapers off toward an upper asymptote. This type of change caught the fancy of nineteenth century social scientists, especially during the earlier part, when scientific patterns of thought became prominent. Quételet, the outstanding Belgian social statistician and astronomer was deeply impressed by the suggestion of stability and invariance in the social sphere, long thought of as an area not amenable to scientific approaches. Under his influence the Belgian mathematician, P. F. Verhulst, addressed himself to the problem of population growth. In his quest for an explanatory function to help trace population growth, he soon realized the limitation of simple exponential functions and began to modify them. His ideas attracted attention in the United States about two and a half generations later, when Reed and Pearl independently developed their law of population.[12]

From the point of view of prediction, curve-fitting techniques do not assure discovery of rational trends. The latter are intended as statements of explanatory, analytical structures rather than mere historical description.[13]

In his quest for a mathematical function giving a logical explanation of population growth, Verhulst was deeply influenced by the then-current Malthusian ideas.[14] He was impressed that populations have a tendency to grow geometrically at first and that at some point serious obstacles begin to manifest themselves, so that growth is slowed as the process continues. The devised function therefore should express a "law of growth," as in the biological sciences—growth acceleration during the first half of the process and deceleration but continuing growth during

[12] L. J. Reed and Raymond Pearl, "On the Summation of the Logistic Curve," *Journal of the Royal Statistical Society*, Vol. 90 (1927); and Raymond Pearl, *Studies in Human Biology* (Baltimore: The Williams & Wilkins Co., 1925). See also G. Udny Yule, "The Growth of Population and the Factors Which Control it," *Journal of the Royal Statistical Society*, Vol. 88 (1925). A few years earlier, unknown to Reed and Pearl, Verhulst had been rediscovered by L. G. duPasquier—see his article, "Esquisse d'une Nouvelle Théorie de la Population," *Vierteljahrsschrift der Naturforschungs Gesellschaft*, 63 (Zurich: 1918).

[13] The earlier much stressed methodological discrepancy between idiography (description) and nomothesis or nomology (the postulation of laws and stable relationships), though of considerable practical importance in the days of Wilhelm Windelband—see his *Die Lehre vom Zufall* (1870)—is probably less acute today. In prediction problems, however, the predominance of idiography and pragmatic optimism has probably contributed much to some of the slipshod approaches that can still be found.

[14] It was Verhulst who gave the particular function the name, "logistic," which since has been generally accepted.

the second half with population asymptotically approaching some upper limit.

The logistic curve describes expected change during a complete growth cycle. The growth rate begins near 0, rises to a maximum, and declines again toward 0.

This pattern was found to fit some animal and human populations extremely well. Pearl successfully applied the function to the growth of experimental populations of drosophila, of saccharomyces, and of yeast cells; to the individual growth (in terms of body weight) of the male white rat; and to the growth of populations in a great many western countries, including the United States (1790 1920) and Sweden (1750 1920).[15]

The logistic "model," the first deserving the name in the field of population and human resources, encompasses two, possibly quite different, growth-process types. As Davis observed,[16] it seems to express better and more cogently the growth pattern of populations and colonies of independent elements rather than the growth patterns of the organisms themselves—e.g., of dynamic structures subject to a "central mechanism," such as body weight. Are patterns of manpower change in a particular socioeconomic setting like the biological growth of a group of individuals, or are they like the growth of an organism, of an ordered structure or system subject to autonomous and interdependent responses? As social structure develops and loses some of its random aspects, long-range forecasting via the logistic would seem a hazardous procedure. For shorter periods, however, the logistic form of growth may at least be a worthwhile device to ascertain the discrepancy between the actual (perhaps "directed") and the theoretical (free) growth of a population of organisms.

In essence, the logistic formula is a reciprocal of the earlier mentioned modification of the exponential, i.e., $P = \frac{1}{L + ab^T}$ · In particular, $P = \frac{L}{1 + e^{fT}}$, where $f(T) = a + bT$ is a first degree polynomial in $T$, and $L$ again is the upper limit. A higher degree polynomial can, of course, be considered for a better fit,[17] but the underlying rationale rapidly becomes involved and obscure. The first differences of logistic function vary; and when plotted, form a symmetrical, bell-shaped (not necessarily

[15] Pearl, *op. cit.* For other biological examples see, e.g., T. Carson, "Über Geschwindigkeit und Grösse der Hefevermehrung in Würze," *Biochemische Zeitschrift*, Vol. 57 (1913); or T. B. Robertson, *The Chemical Basis of Growth and Senescence* (Philadelphia: J. B. Lippincott Co., 1923).

[16] Harold T. Davis, *The Analysis of Economic Time Series* (Bloomington, Indiana: Principia Press, 1941).

[17] Second- and third-degree polynomials were used to advantage in smoothing observations, reflecting the presence of some "central mechanism" as alluded to earlier.

normal) curve with the mode at the point of inflection of the fitted curve.

To fit the logistic $\frac{L}{1 + e^{a+bT}}$ to series of reasonable length is not too bothersome with usual desk calculators. As in the earlier illustrations, we may, for example, select equally spaced points and let the function pass through these points. Three points are needed, since there are three constants, $L$, $a$ and $b$.

Let us assume that annual observations ($P_i$) are available for the years 1953 to 1967. For convenience we code the years into $T$ units so that only the last digit of the calendar year is used during the fifties and the last digit plus 10 for the years during the sixties. We select three equidistant points of time, say $T_1 = 5$ (that is, 1955), at $T_2 = 10$ (that is, 1960), and $T_3 = 15$ (that is, 1965). Next we plot the observations (that is, $P_i$) and fit a freehand S-shaped curve or any part of such a curve. From the curve thus fitted by inspection to the observational data, we read the values of the "synthetic" $P$'s corresponding to the years $T_1$, $T_2$, and $T_3$. We label the synthetic points $\Pi$.

To obtain the coefficients in $\frac{L}{1 + e^{a+bT}}$ we use

$$L = \frac{2\Pi_0\Pi_1\Pi_2 - \Pi_1^2(\Pi_0 + \Pi_2)}{\Pi_0\Pi_2 - \Pi_1^2}$$

$$a = \log_e \frac{L - \Pi_0}{\Pi_0}, \text{ and}$$

$$b = \frac{1}{N} \log_e \frac{\Pi_0(L - \Pi_1)}{\Pi_1(L - \Pi_0)}$$

These expressions may be written in terms of the base $P$, rather than $c$. That is, $e^{a+btT} = \log_{10} e(a + bT)$; and, since $\log_{10} e = 0.4343$, we have $e^{a+btT} = 0.4343\,(a + bT)$.

Although only an approximation, the method commonly gives acceptable results. The different constants do vary however with the choice of the three points.

Other methods are available, among which are Hotelling's least-squares method[18] and its adaptation by Davis (see footnote 16). Non-

---

[18] Harold Hotelling, "Differential Equations Subject to Error, and Population Estimates," *Journal of the American Statistical Association*, 22 (1927); see also G. Tintner, *Econometrics* (New York: John Wiley & Sons, Inc., 1952).

technical descriptions can also be found.[19] Extensive use of the logistic curve was made by Kuznets[20] to describe economic growth phenomena.

Before leaving the topic of modified exponentials, we call attention to one more modification which resembles the logistic. It too depicts an S-shaped pattern of growth; however, unlike the logistic, the first differences are not symmetrically distributed about a point of inflection. Named after an actuary of the early nineteenth century, the Gompertz curve was invented primarily with a view to describing certain mortality phenomena. Gompertz justified the function by reference to the inability of organisms, increasing with age, to avoid destruction—this inability accounts for the asymmetry of the distribution of first differences.[21]

The equation of the Gompertz function is $P = ab^{c^T}$; in log form, this becomes $\log P = \text{Log } a + c^T (\log b)$, where $a$ is the upper limit and $b$ the difference between $P$ and the ceiling at the successive points in time, $a - P$. $P$ thus increases with $T$, but in decreasing relative amounts. To fit a Gompertz function, we may again select three equidistant time points and proceed similarly as in the prescription for fitting a logistic.

Discussion of modified exponentials with reference to the underlying mechanisms of growth helped to instill a more cautious attitude towards extrapolation of past observations. The ensuing quest for rational rather than past empirical relationships marked an important step beyond the purely empirical descriptive curve-fitting stage. The next forward stride was taken with the emerging realization that broad aggregates are perhaps too heterogeneous to yield clearly visible and rationalizable patterns of change. The quest for improvement then turned toward the decomposition of complex phenomena into simpler ones. The search for components proceeded along two lines: components of the population and their effect on the changing population structure; and components of the growth process, of the time series themselves. Both approaches attempt to understand in the small what is not disclosed when phenomena are viewed in the large. But micro-information requirements are usually more severe and more difficult to satisfy than macro needs, so the shift toward disaggregation and decomposition has also shifted the emphasis in needs for statistical data and other information.

More general compulsory vital registration increased the ensuing

[19] Croxton and Cowden, *op. cit.*; F. C. Mills, *Introduction to Statistics* (New York: Holt, Rinehart, & Winston, Inc., 1956), Appendix; and J. D. Smith and A. J. Duncan, *Elementary Statistics and Applications*, Part I (New York: McGraw-Hill Book Company, 1944).

[20] S. S. Kuznets, Secular Movements in Production and Prices (Boston: 1930; reprinted in New York: Augustus M. Kelley, 1967).

[21] For simple fitting techniques for the Gompertz curve, see Mills, *op. cit.*

availability of data on births and deaths and on other vital events, and this trend no doubt contributed much to population forecasting by components. This approach is of particular interest also to the manpower analyst.

In this country the responsibility for vital statistics laws has generally been left with the states; so nationwide statistics have tended to be weaker than in most Western countries; at first, so did the forecasting via components—the vital statistics method.[22]

As forecasting via components captured the attention of vital statisticians and actuaries, it found its way into manpower analysis with the appearance of analytical interest in work-history patterns, in specific age and sex participation rates, and in working-life tables.

This shift in emphasis from totals to components has also emphasized the view of population change as a renewal process.[23] This view has in turn emphasized attempts to apply concepts of stochastic processes to population and manpower. More will be said later about stochastic processes when we consider simulation in manpower forecasting.

The essential formula for the renewal of a closed population (i.e., a population for which inward or outward migration can be neglected) is $P_n = P_0 + {}_0B_n - {}_0D_n$. Here, $P_n$ is the population of the $n^{\text{th}}$ year (the year for which a forecast is to be made); $P_0$, the population of the base year: and ${}_0B_n$, ${}_0D_n$, the number of births and deaths, respectively, during the period from 0 to $n$. For an "open" population, i.e., one into and out of which migration takes place, another term has to be added, say ${}_0M_n$, for net inward migration during the period from 0 to $n$. This addition makes it possible to analyze the projected $B$, $D$ and $M$ separately. The pattern of change of $B$ and of $D$ is likely to be more stable than that of population totals. The least tractable component is $M$, especially for estimates for smaller geographic subareas.

The procedure just outlined depends heavily on two assumptions: (1) expected mortality and fertility and (2) anticipated inward and outward migrations.[24] In other words, this forecasting method for totals is based on projections of age-specific mortality rates, of distributions of females by reproductive age groups, and of corresponding age-specific fertility

[22] It may be mentioned that an outstanding early contribution by an American, A. J. Lotka, *Théorie analytique des associations biologiques*, especially Part II, *Analyse démographique avec applications particulières à l'espèce humaine* (Paris: Hermann & Cie., 1934 39), has never been made available in English.

[23] On the general topic of renewal processes see, e.g., W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. I (2nd ed.: New York: John Wiley & Sons, Inc. 1957). The problem is related to the replacement problem in operations research; see e.g., C. West Churchman, *et al.*, *Introduction to Operations Research* (New York: John Wiley & Sons, Inc., 1957).

[24] For details see, e.g., G. W. Barkley, *Technique of Population Analysis* (New York: John Wiley & Sons, Inc., 1958); and M. Spiegelman, *The Society of Actuaries' Textbook* (Chicago: The Society, 1955).

rates.[25] The effects of other factors can, of course, be explicitly introduced; for instance, if major changes in marriage patterns are expected, estimates of marriage-duration and specific fertility rates may yield a better measure of at least the "legitimate" component of ${}_0B_n$. The greatest indefiniteness as already suggested, relates to the migration component. If migration is strongly affected by government policy, this factor is "exogenous"—one to be treated gingerly, and often arbitrarily as will be pointed out later in a discussion of econometric models in manpower forecasting. However, even the use of $B$ and $D$ involves assumptions as to their future course; and these assumptions, although perhaps more convincing than those made with respect to population totals, could prove unwarranted.

Conceptually related to this component method of forecasting are recent attempts to apply the mathematically powerful tools of matrix algebra to summarize fertility and survivorship patterns.[26] One such approach, which a quarter of a century ago was explored by Bernadelli in a study of populations of beetles,[27] should be mentioned here because of its potential relevance to manpower analysis. The virtues of this method are compactness and speed; the use of electronic data processing now permits quick and detailed application of survival and fertility rates to a population with a given initial size, sex composition, and age structure. Efficient and fast tracing of the effect on future populations of an entire set of fertility and survival rate structures, whether actually observed or merely assumed, becomes feasible. Although not a prediction device *per se*, the method enables the researcher to make a great many forecasts for different future conditions. Bernadelli's approach may be regarded as a particular variety of simulation, a technique treated more fully later.

So far, the methods described have been concerned with population forecasting. Since population is the overall frame for the various manpower categories and subcategories, its taxonomic role for manpower prediction can hardly be exaggerated. Nevertheless, the forecasting of manpower and its subdivisions poses some special problems and has led to the development of prediction techniques *sui generis*.

For the sake of orderly exposition we classify manpower forecasts as: explicit; derived via general population forecasts; derived via economic forecasts; and "techni-cultural" expectations and conjectures.

[25] The fertility rate is the number of live births per time unit (say, year), usually per 1,000 women in the childbearing ages. This "crude" rate can be refined by introducing consideration of specific age groups, socioeconomic classes, ethnic characteristics, and the like.

[26] N. Keyfitz, "Matrix Multiplication as a Technique of Population Analysis," *The Milbank Memorial Fund Quarterly*, Vol. XLIX (1964) and "Reconciliation of Population Models: Matrix, Integral Equation and Partial Fraction," *Journal of the Royal Statistical Society*, Series A, Vol. CXXX, Part I 1967).

[27] H. Bernadelli, "Population Waves," *Journal of The Burma Research Society*, Vol. XXXI (1941).

## *III. Direct Manpower Forecasts*

Among the explicit and direct forecasting schemes (direct in the sense that they do not have recourse to "intermediate" predictions of other than manpower variables) must be mentioned the expectation and opinion survey of employers. This predictive device is one particularly popular during periods of national defense pressures, to obtain a consolidated picture of the future labor market from information supplied by a group of employers on their own expectations of future manpower needs.

Apart from difficulties in the way of collecting and interpreting anticipation data of any kind, this survey is handicapped by the frequent lack of knowledge about future manpower requirements of an establishment or company. Unless longer range personnel and production policies have been firmly established for the respondent organization, it is hard for an organization to arrive at a projection of its own manpower needs. Few organizations usually make such projections. Furthermore, respondent's anticipations seem to reflect seasonal variations, more than other components of change.[1] Estimates of future vacancies require allowances for expected separations and other attritions. If, furthermore, the information collecting organization is identified with a major contracting or manpower-regulating agency, the respondent may have to overestimate his future manpower needs. The bias is similar to that associated with a capacity-oriented survey when the respondent suspects a connection between this capacity estimate and the chance of obtaining a contract. There is also the problem of how to define the universe of establishments from which to select the sample, considering the variability over time of this universe.

In short, the company forecast method would seem to be only a last resort. It might be justifiable in estimating the short-term demand of critical special skills by a well-defined group of companies, able and willing to provide answers to the posed questions.[2] This method, therefore, can hardly be suggested as a major manpower forecasting device.

More interesting in their applicability to manpower forecasting are the various attempts to decompose time series, say, of employment. The object is to elucidate patterns not immediately visible to the naked eye.

[1] R. Ferber, *et al.*, "Chicago Area Employers' Labor Force Anticipations," *Journal of Business*, 1961; and R. Ferber, *Employers' Forecasts of Manpower Requirements, A Case Study* (University of Illinois, Bureau of Economic and Business Research, 1958). For an example of an expectation-type survey, see National Science Foundation. *The Long-Range Demand for Scientific and Technical Personnel*, NSF 61 65 (Washington)

[2] See American Statistical Association, *1960 Proceedings of the Business and Economics Section*, and the paper by S. Lebergott, "Government as a Source of Anticipatory Data" in the *1959 Proceedings*; see also National Bureau of Economic Research, Special Conference Series X, *The Quality and Economic Significance of Anticipation Data* (Princeton: Princeton University Press, 1960).

These patterns are important for analysis of the underlying time process for assessment of stability and future change.

Efforts to decomposition of economic time series which by now have a long and respectable history, have been mentioned. Although the primary impulse for study came from the concern with economic crises and the so-called business cycle, the scope of time series analysis broadened considerably. Furthermore, it is not restricted to forecasting.

The conceptual image of a time series and the purpose of analysis determine the choice of technique. As a rule, the separation of a particular time series into components aims either at the study of the components as such or at their elimination; in other words, it typically aims at "computing out" the effect of the components. A conventional approach—by no means the only one—is to conceive of the empirical time series (e.g., the employment series for a given economic sector) as composed of four parts: a long-range component ($T$, the "trend"); a shorter, wavelike one ($C$, the "cycle"); an intra-annual one ($S$, the "seasonal"); and a random element ($\mathcal{E}$) the "error", or noise, in the language of the communcations engineer).

Other frameworks may be devised. For instance, if the series under consideration is one of annual observations, $S$, the intra-annual component, vanishes. Instead of a single trend, major and minor ones may be distinguished.[3] For instance, the employment growth pattern may be characterized over long periods as indicating a major upward trend on which are superimposed S-shaped minor trends (analogous to the Reed-Pearl curve) representing shorter-term influences. Thus, a particular growth industry may eventually reach an upper asymptote; but a new S-shaped pattern may emerge at the new, higher level, perhaps as a result of technological breakthrough. Each "takeoff" leads, after an upsurge, to a new plateau, from which a second "takeoff" occurs, and so on. If the manpower impact of widely spaced major technical innovations is to be investigated, the complex model of superimposed trend components may prove useful. Similarly, there is not necessarily only one composite business cycle, but there may appear superimposed industry-specific cycles (e.g., in the textile industry, in hog breeding, etc.).

The "theoretical" framework within which a given time series is to be decomposed must identify not only the several components of interest, but also the postulated relationship among them. Simple models have often proved quite adequate: an additive model, such as $T + C + S + \mathcal{E}$; a multiplicative one, $T \cdot C \cdot S \cdot \mathcal{E}$; combinations as $TCS + S$; and such. In the first of these models, the effect of each component is assumed to be independent of the effects of the others. The second model

[3] See, for instance, S. S. Kuznets, *Secular Movements in Production and Prices* (Boston: 1930; reprinted in New York: Augustus M. Kelley, 1967).

implies that the effects of the components are interdependent and the relationship is one of direct proportionality. A large value of a component entering a model as a multiplier affects the system more strongly than if it exerted merely an independent additive influence.

In manpower analysis, seasonal fluctations command attention because of their importance for the study of short-term employment changes. For seasonal analysis, a multiplicative model appears more acceptable than an additive one; the technical question is to identify the trend and other components to "eliminate" them from the empirical data. If the observed series is regarded as a composite $T \cdot C \cdot S$, attempts should be made to ascertain the $C \cdot S$ components and exclude their effect by division; i.e. $\frac{T \cdot C \cdot S}{T \cdot C} = S$. If an additive model is used, the effects of $T$ and $C$ would, of course, have to be excluded by subtraction.

The arsenal of techniques for the analysis of oscillatory series is by now impressively large and advanced.[4] Furthermore, one component, the long-term trend, can be addressed by a classical technique that is closely related to the curve-fitting methods described earlier.

A well-established method for the "smoothing" of major irregularities in a series of time-ordered observations is the so-called method of moving averages. Although not yielding a mathematical expression for the resulting smooth series, which is the analogue to the trend, this method has the advantage of great flexibility. Moving averages are based on the idea of the arithmetic mean as a descriptor of a collection of magnitudes; times to be averaged are $n$ successive observations in a longer time series. We may form successive, time-ordered sequences of observations from $t_0$ to $t_n$, $t_1$ to $t_{n+1}$, $t_2$ to $t_{n+2}$, and so forth and characterize each set by its arithmetic mean value, centered at year $\frac{t_0 + t_n}{2}$, $\frac{t_1 + t_{n+1}}{2}$ and so forth. If an oscillatory series of constant period $n$ is smoothed by the just-described procedure, a straight line is obtained after connection of the successive moving arithmetic means; the oscillations will have been completely "eliminated." Otherwise, rather than making the oscillation disappear altogether, the method does iron out some of the wavelike movements. Here appears the conceptual affinity with curve-fitting methods discussed earlier. In effect, the moving average fits polynomials over successive portions of a longer time series, and lets some mid-value of each successive polynomial constitute the "smooth" line.

It is possible to provide a series by weighting observations that precede those for successive points. Such a smoothing method adjusts suc-

[4] See M. G. Kendall, *Contributions to the Study of Oscillatory Time-Series* (Cambridge, England: Cambridge University Press, 1946); and M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics*, Vol. III (New York: Hafner Publishing Co., Inc., 1966), with ample literature references. See also Robert G. Brown, *Smoothing Forecasting and Prediction of Discrete Time Series* (Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1963).

cessive time series data by means of preceding observations usually attributing most weight to the immediately preceding one and decreasing weight to older observations. Thus, one may conceive of the time-ordered process as having a memory that gradually weakens and finally fails completely. Looked at in this fashion, the procedure displays a built-in self-correcting feature.[5] The technique resembles smoothing by moving averages, but is sufficiently different to warrant some attention by persons other than systems engineers, who have been the primary users.

Indeed, it would seem eminently suitable for the analysis of the trend in observed manpower data. New estimates are obtained from earlier ones, by means of the formula $[\bar{M}]_T = cM_T + (1 - c)[\bar{M}]_{T-1}$. Here $M$ is a particular manpower variable, say, employment, in a given industry; $T$ refers to a particular time (year, quarter, or month); $[M]$ is the moving average; and $c$, a percentage, is a smoothing constant. More specifically, $[\bar{M}]_T$ stands for the moving average over the interval (of years, quarters, or months) ending with period $T$. Since $cM_T + (1 - c)\ [\bar{M}]_{T-1}$ can be written more fully as

$$cM_T + c(1 - c)\ M_{T-1} + c(1 - c)^2 M_{T-2} + \cdots + (1 - c)^T M_0,$$

it becomes apparent that $[\bar{M}]_T$ is a function of all previously observed $M$. Each M enters the expression with geometrically declining weights as its distance from time $T$ (the present) increases—hence the term exponential smoothing. If $c$ has a small value, it will decline only slowly. If $c$ is large, the weights will decline rapidly, quickly approaching zero; consequently, the moving average will be affected by relatively few past observations, the process has only a "poor memory," and estimates react quickly to recent changes.

This pattern contrasts with the conventional moving average which implicitly assigns equal weights to all observations entering the particular average. A relationship can be established between the number of observations on which a given moving average is based and the corresponding $c$ or smoothing constant, under the assumption that the weight is zero or a negligible quantity at the time the first observation $(t - n)$ enters the conventional moving average based on $n$ observations. Thus a smoothing constant of 0.5 corresponds to 3 observations in an individual moving average; of 0.4 to 4, of 0.33 to 5, of 0.2 to 9, of 0.1 to 19, and of 0.10 to 199 observations. When the number of observations is 12, $c$ = 0.154, a constant of interest when moving averages are used in manipulation of seasonal adjustments.[6]

One of the idiosyncrasies of the method of moving averages as a

[5] See Robert G. Brown, *Statistical Forecasting for Inventory Control* (New York: McGraw-Hill Book Company, 1959), who introduced the term "exponential smoothing" for the particular method, recommending it for its advantages if electronic data processing equipment is to be used.

[6] The above conversions were taken from Brown, *Smoothing Forecasting . . . , op. cit.*

smoothing device is that the process of summation involved may introduce cycles where in reality there are none.[7] In general, the user of the moving-average method should be aware that, if the length of the moving average (the $n$) does not equal the constant wavelength of the original series, the result will not be a straight but a fluctuating one. It is likely, however, that the new wave will have a smaller amplitude than the original series. In the decomposition of economic time series, which are not strictly periodic, the danger of introducing new fluctuations is great. Nevertheless, when considering alternative choices, the moving-average method remains attractive because it is simple. It is widely used for the "elimination" of the $T$ or $T$ and $C$ components together.

We return now to the elimination of seasonal variations—a problem relevant to the analysis of many manpower series. The method of moving averages is attractive because of the constancy, by definition, of the seasonal pattern's wavelength (12 months). On the other hand, the seasonal pattern may actually change, not only in form, but in length; it may gradually be compressed or stretched along the time axis.

A 12-month moving average, thought to represent the components $T \cdot C$, is first used to eliminate interannual fluctuations (the seasonal and "error" components). Then, if a multiplicative model is assumed, as it usually is, the ratio of $S\mathcal{E}$ to the original manpower series $M$ is found, and we compute $S\mathcal{E} = \frac{T \cdot C \cdot S \cdot \mathcal{E}}{T \cdot C}$ (In an additive model, $T$ and $C$ would of course be subtracted from $M$ in order to yield $S + \mathcal{E}$). The resulting ratios are now averaged for all Januaries, Februaries, Marches, etc. In averaging, the arithmetic mean of only the central values, thus excluding the extremes, is sometimes used to weaken the influence of irregular and extraneous factors. The resulting monthly averages of ratios, perhaps because of the presence of a long-term upward or downward component, may not add to 1,200 (i.e., 12 times 100 percent); they are usually adjusted to do so. The resulting adjusted ratios, frequently called the seasonal indexes, are then considered the "typical" seasonal pattern of the particular series (with respect to trend and cycle) and, in this sense, they may provide a useful short-term forecasting device. When the original data are to be deseasonalized, the seasonal index for month $m$, that is, $I_m$, is divided into that month's observation, $M_m$.[8]

[7] This effect of the method of moving averages, or of any other method based on sums of successive time-ordered observations subject to random fluctuations, is referred to in the literature as the Slutsky or Slutsky-Yule effect after the statisticians who (independently) investigated this property. See E. Slutsky, "The Summation of Random Causes as a Source of Cyclic Processes," *Econometrica*, Vol. 5 (1937); and G. U. Yule, "On a Method of Investigating Periodicities in Disturbed Series, With Special Reference to Wolfer's Sunspot Numbers," *Philosophical Transactions*, Series A, 226.

[8] See S. S. Kuznets, *Seasonal Variations in Industry and Trade* (New York: Columbia University Press, 1933); Kendall and Stuart, *op. cit.*; Ya-lun Chou, *Applied Business and Economic Statistics* (New York: Holt, Rinehart, & Winston, Inc., 1963).

For various reasons, seasonal patterns cannot be assumed to remain consta[illegible]e. The careful forecaster, therefore, will plot seasonal ratio[illegible]sive months as a function of time, to decide whether a new index [illegible]uld be computed or whether the pattern of changing seasonal ratios for a given month displays a sufficiently consistent trend to warrant projection of the given seasonal index.[9]

The Bureau of Labor Statistics (BLS) has, since [illegible], engaged in a major program of studying the problem of identifying the seasonal component in manpower series. It has found a multiplicative model preferable to an additive one for manpower series. Experiments with mixed additive-multiplicative models did not yield significantly different results. The BLS method, too, is in essence based on a ratio to moving-average. The observations are weighted before averaging, however, by what is called a "credence factor" to adjust for the influence of the $\mathcal{E}$ component. The BLS technique also provides an explicit method for restoring the $TC$ residuals which may have been "overlooked" by the moving average and hence absorbed into the $\mathcal{E}$ component.[10]

## *IV. Derived Manpower Forecasts*

The prediction techniques discussed in the foregoing section are applicable to attempts to project manpower series directly—"from within" the series itself, so to speak. Other approaches attempt forecasting "from without"; here the manpower predictions are based on some other variable or variables with which the manpower series are assumed to be associated and which are deemed to be more readily or more safely predictable. Following historical development and present practice, we have divided the presentation of such "derived" prediction methods into two groups:

[9] See, for instance, Board of Governors of the Federal Reserve System, *Industrial Production: 1959 Revision* (Washington: 1960). For general references to the problem see also J. Shiskin and H. Eisenpress, *Seasonal Adjustments by Electronic Computer*, Technical Paper No. 12 (New York: National Bureau of Economic Research, 1958); H. M. Rosenberg, "Seasonal Adjustment of Vital Statistics by Electronic Computer," *Public Health Reports*, Vol. 80 (Washington: 1965); Organization for Economic Cooperation and Development, *Seasonal Adjustments on Electronic Computers* (Paris: The Organization, 1961); H. M. Rosenblatt, "Spectral Analysis and Parametric Methods for Seasonal Adjustment of Economic Time Series," American Statistical Association, *1963 Proceedings of the Business and Economic Statistics Section*.

[10] The computation, based on three successive iterations, transcends manipulation by desk calculator but is readily amenable to electronic data processing, for which it was designed. The program, available from the Bureau of Labor Statistics, is intended for direct use on an IBM 1401 or 1460 which, in conjunction with a 729 II magnetic tape unit, will process the data about twice as fast as the 1401. The program results in direct printout of tables in a printing unit, in a single run. See the *BLS Seasonal Factor Method (1966)* (Washington) and Chapter VI (and literature there quoted) of *Measuring Employment and Unemployment* (Washington: President's Committee to Appraise Employment and Unemployment Statistics, 1962).

those leading to prediction via the intermediary of population forecasts; and those founded on forecasts of economic variables other than manpower itself.

*Based on Population*

If it is desired to predict merely the major dimensions of manpower resources,[1] and if such broad measures as "population of working ages" are adequate for the purpose, the problem reduces to that of demographic forecasting. If, correspondingly, prediction of the dependent population is desired (e.g., to form an opinion of the future course of dependency ratios), then only the population total ($P_T$) and the population of working age, say, from 15 to 64 years (${}_{15}P_{64}$), need be predicted in order to obtain a forecast of the dependency ratios[2]

$$\frac{P_T}{{}_{15}P_{64}} - 1$$

As soon as more complex concepts and measures are to be forecast, intermediate links have to be introduced to translate a population forecast into a forecast of the desired manpower category. The usual procedure is to estimate the relevant economically active component, according to sex and age, of the total population. If the whole labor force is the desired concept, it is necessary to estimate or forecast the percentages of the total population, by sex and age groups, represented by members of the labor force. These fractions are labor force participation rates. What is required is not only a forecast of population by age but also of age-specific labor force participation rates separately for each sex.

The additional burden of estimating and forecasting specific labor force participation rates by age and sex is not minor. Labor force participation rates result from the operation of a complex structure of a multiplicity of factors—economic, social, psychological (motivational), cultural, and others. Some features such as the cultural ones, may be relatively stable; others, such as the motivational ones, may be extremely variable, sensitive to still other influences about which hardly anything is known and which are not amenable to forecasting.

Attempts to forecast manpower categories by labor force participation rates[3] may be based on the assumption, explicit or implicit, that over

[1] Although manpower and similar concepts are typically defined in terms of number of persons, other measures are conceivable and might, for certain purposes, be preferable.

[2] Dependent population divided by working-age population equals total population ($P_T$) less working-age population (${}_{15}P_{64}$).

[3] $\frac{L}{P} \cdot 100$, where $P$ is the total population in a particular category and $L$ is the corresponding number of those in the labor force (i.e., the employed and the unemployed).

the period under consideration the age and sex-specific labor force participation rates are relatively stable or display a simple pattern of change. Then, from the latest labor force participation rates or their trend over a period of time and from population forecasts, we may obtain future labor force estimates.[4] The assumptions, of course, are tenuous, especially for periods of varying rates of employment and of structural change. Accordingly, considerable effort has also gone into the prediction of labor force participation rates that need not remain constant or even nearly so for the length of the forecasting interval.[5]

Fortunately the bulk of the labor force remains concentrated over time within certain age groups; the highest and most stable participation rate has been observed for married men, in the age group from 25 to 54. The rates for women too are relatively high for this particular age group, and might be expected higher for unmarried women than for married ones. Forecasting participation rates is accordingly, more difficult for some categories than for others. Older men have a variable retirement pattern; the circumstances of married women make their propensity to join or to rejoin the labor force uncertain; younger persons have dissimilar and changing patterns of school attendance. The prediction of labor force participation rates requires various strategies. It is feasible to fit curves when regular patterns of change have been observed; it may be necessary first to analyze the behavior of such ancillary variables as marriage rates and marriage age, childbearing pattern reflecting age at time of each birth, child spacing, and family size. It may be necessary to form an opinion about future government action influencing school attendance and military service or other economic conditions relevant to unemployment, wage rates or hours of work. Finally conjecture may be needed about the development of sociocultural values likely to change attitudes toward women's work, leisure time or employment of different ethnic groups.[6]

[4] See, e.g., the earlier labor force estimates by the Census Bureau in *Current Population Reports*, Series P-50, No. 42 (1952); and Harold Wool, "Long-Term Projections of the Labor Force," *in* National Bureau of Economic Research, *Long-Range Economic Projections, Studies in Income and Wealth*, Vol. 16 (Princeton: Princeton University Press. 1954).

[5] Excellent recent presentations of the complex nature of such efforts are Sophia Cooper and Stuart Garfinkle, *Population and Labor Force Projections for the United States, 1960 to 1975*, BLS Bulletin 1242 (Washington: U. S. Department of Labor, Bureau of Labor Statistics, 1959); and Sophia Cooper and Dennis F. Johnston, "Labor Force Projections for 1970 to 1980," Monthly Labor Review, 88 (1965).

[6] See, e.g., the testimony by Harold Goldstein on September 26, 1963, before the Subcommittee on Employment and Manpower of the Senate Committee on Labor and Public Welfare *Projections of the Labor Force of the United States*, Hearings, Pt. 5, pp. 1529 ff. See also the simple and useful presentation of general manpower forecasting techniques in *Demographic Techniques for Manpower Planning in Developing Countries* (1966); U. S.

*Based on Production*

Population estimates are of no avail when detailed manpower categories and subcategories need to be predicted—e.g., employment in a particular industry or economic subdivision. Here specific forecasts of economic activity are commonly made the basis for forecasts of the related manpower variable.

For projecting an industry's activity and employment, we need various building blocks. At the proper level of detail we need estimates of production trends and outlook, and anticipated changes in the ratio of output to per-worker input (so-called productivity). This task may require estimates of future demand for an industry's output; of the expected inflow of capital investment; of technical changes likely to affect labor intensity and productivity; of the effect of domestic and foreign competitive factors; and of the impact of government.[7] Such intensive study is in the best tradition of the historical school. It is, of course, a major undertaking yet a very subjective exercise in conjecture, and possibly more nearly objective where quantitative data are available for use in curve fitting, regression analysis, and the like.

Activity estimates are available on various levels of aggregation, beginning with the highest one—the national economy. They include official forecasts of GNP and its components, and unofficial ones (e.g., those prepared by the National Planning Association).[8] It is tempting to utilize such forecasts for manpower predictions, but the danger of circular reasoning exists, since manpower is sometimes used as a variable in estimating GNP.

Once an acceptable GNP estimate for a future year is available, translation into a manpower figure requires some assumption as to relationship with the investigated manpower variable. Since the ratio is unlikely to remain stable over time, explicit assumptions need to be made about changes, e.g., in productivity and technology. Study of the behavior of the ratio over a past period may disclose a linear or curvilinear pattern that is projectible.

If we use only two variables, relating say, manpower to an independent

---

Department of State, Agency for International Development; T. F. Dernburg in Discussion of Lee Mansen's paper, *Proceedings of the Eighteenth Annual Winter Meeting of the Industrial Relations Research Association* (Madison: The Association, 1965); and S. Cooper, "Notes on Labor Force Projections," *in* R. A. Gordon, ed., *Long-Term Manpower Projections*, (Berkeley: 1965).

[7] See *Long-Term Manpower Projections*, "Industry Employment Projections," by S. Swerdloff, regarding the use of this approach by the Bureau of Labor Statistics; also H. Goldstein, "Projections of Manpower Requirements and Supply," *Industrial Relations*, Vol. 5 (1966), and literature there referred to.

[8] See particularly Joel Darmstadter, "Manpower in a Long-Term Projection Model," *Industrial Relations*, Vol. 5 (1966), and literature there mentioned.

one (estimate of output or some other appropriate measure), we may determine the constants of the regression line as made in fitting a line by the least-squares method. If a linear regression function is not appropriate, a polynomial of higher order, or perhaps an exponential regression may prove useful.

The basic problem facing the analyst is the choice of the function, i.e., the particular type of curve. This choice must be based on some preknowledge or theory about the underlying structure governing the process. No mechanical device can make this choice for the forecaster.

If a simple linear estimating function were to be found, relating say a manpower variable ($M$) with GNP ($G$), the equation of the fitted line would be $M = a + bG$. We find that $a = \overline{M} - b\overline{G}$, ($\overline{M}$ and $\overline{G}$ are again the respective arithmetic means); and $b = \dfrac{n\Sigma MG - \Sigma M \Sigma G}{n\Sigma G^2 - (\Sigma G)^2}$, where $n$ is again the number of observations. The derived $a$ is the value of $M$ for $G = 0$; and $b$ is the slope of the line.

If a second-degree polynomial is used to describe the regression[9] or estimating line, the equation is as before, $M = a + bG + cG^2$. The coefficients may be found by solving simultaneously the three equations:

$$\Sigma M = na + b\Sigma G + c\Sigma G^2$$

$$\Sigma MG = a\Sigma G + b\Sigma G^2 + c\Sigma G^3$$

$$\Sigma MG^2 = a\Sigma G^2 + b\Sigma G^3 + c\Sigma G^4$$

A simple technique, especially satisfactory if errors are expected in both variables, is to divide the observations into thirds and to compute $\overline{M}$ and $\overline{G}$ for each extreme set of observations. Then $b$ is obtained from

$$\frac{\overline{M}\text{ (upper third)} - \overline{M}\text{ (lower).}}{\overline{G}\text{ (upper third)} - \overline{G}\text{ (lower}}$$

The line $M = a + bG$ passes through the common mean of all observations.[10]

When an exponen[illegible] first into logarithmic form: log $M$ [illegible] g $b$ are

---

[9] For a com[illegible] Correlation Anal[illegible] York: John Wi[illegible] 4 [illegible] sion function" or "regre[illegible] rally traced to Sir Francis Galton[illegible] Stature," *Journal of the Anthropo*[illegible] ency of the filial type to depart [illegible] average ancestral type.

[10] See M. S. B[illegible] Variables Are Subject to Error," *Biometrics*, [illegible]

found as in the usual case of linear regression with log $G$ substituted for $G$, log $a$ for $a$, and log $b$ for $b$.

Once the desired regression equation is available, the appropriate value of $M$ may be found that corresponds to a predicted value of $G$.

A regression based on a single predictor, however, may not be satisfactory if $M$ fluctuates considerably about the fitted regression line; then $G$ may not alone account for enough of the variability in $M$, so that a sizeable "unexplained" residual remains. Additional explanatory series or predictors may be sought for a better forecast of $M$.[11]

When multiple regression is involved, more than one predictor is related to the predictand ($M$). The conceptual framework is the same as before, but the equation of a linear regression function may look like $M = a + bG + cX$, the geometric representation is a plane, instead of a line as in the two-variable case, fitted to the $M$-$G$-$X$ observations (one such triplet per year, quarter, etc., located in three-space). As earlier, the least-squares principle yields three equations to be solved simultaneously for the three unknown coefficients $a$, $b$ and $c$.

For a prediction of $M$ we must now substitute in the regression equation two forecasts, one for the anticipated value of $G$, and another for the corresponding future value of $X$. Hopefully, the residual variability of the points (that is, the $M$-$G$-$X$ points) about the regression plane will have diminished. That is, $G$ and $X$ should "explain" more of the variability of $M$ than did $G$ alone. However, as models become more complex, more care needs to be taken in interpretation and analysis.[12] It may be helpful to study an empirical pattern by graphic methods when the underlying process structure is not fully understood, as is so often the case in socioeconomic problems.[13] Even a simple plot on the same graph of the several time series to be included in the analysis may reveal aspects of mutual behavior which are not apparent when only a numerical regression analysis is carried out.[14] When more than three constants need to be

[11] Note, for instance, the Canadian attempt to forecast employment and unemployment from interrelations of changes in output, productivity, and labor force, as described *in* Organization for Economic Cooperation and Development, *Techniques of Economic Forecasting* (Paris: The Organization, 1965).

[12] See, G. U. Yule, "Why Do We Sometimes Get Nonsense Correlations Between Time Series?" *Journal of The Royal Statistical Society*, Vol. 89 (1926).

[13] On graphic methods see, e.g., Ezekiel, *op. cit.*

[14] See, e.g., Milton H. Spencer, Colin J. Clark, and Peter W. Hoguet, *Business and Economic Forecasting* (Homewood, Illinois: Richard D. Irwin, Inc., 1961), Chapter 3; and Jan Tinbergen, *Statistical Testing of Business Cycle Theory*, Vols. I and II (Geneva: League of Nations, 1939).

evaluated, more than three equations have to be solved simultaneously, and recourse to matrix notation may be desirable.[15]

Graphic analysis of time series may also disclose the presence in the time series of lagged relationships.[16] The discovery that one series anticipates another is of considerable interest to the forecaster, especially if historical observation has a valid theoretical explanation.

Closely related to this approach are various early warning systems and the use of economic barometers.[17] They highlight a particular forecasting objective, that is narrow but nevertheless important. Rather than predict the future value of the variable under consideration, or the interval within which that future value is likely to be found, they concentrate on anticipation of major turning points.

Similar considerations underlie the National Bureau of Economic Research's well-known cyclical indicators. A careful analysis of over 800 economic time series[18] has disclosed leading and lagging indications and other coincident with the business cycle. Four indicator lists have been published in 1938, 1950, 1960 and 1967.[19] The last one includes, among the many series, 14 that reflect the employment and unemployment situation. Only one of the 14, measuring long-duration unemployment, lags (eight of the remaining are roughly coincident with the cycle) and hence is a *prima facie* candidate for forecasting by many other indicators. With less assurance two roughly coincident indicators (job vacancies and composite employment) might be anticipated by generally leading series.

[15] None-too-technical introductions into matrix algebra can nowadays be found in many standard works in econometrics. Among these are Gerhard Tintner *Econometrics* (New York: John Wiley & Sons, Inc., 1952); L. R. Klein, *A Textbook of Econometrics* (Evanston, Illinois: Row, Peterson & Co., 1953); and A. S. Goldberger, *Econometric Theory* (New York: John Wiley & Sons, Inc., 1964). See also S. Valavanis, *Econometrics* (New York: McGraw-Hill Book Company, 1959); and J. Johnston, *Econometric Method* (New York: McGraw-Hill Book Company, 1963). For a lucid recent presentation see Clopper Almon, Jr., *Matrix Methods in Economics* (Reading, Mass: Addison-Wesley, 1967).

[16] More elegant analytical techniques are available, such as the correlogram, lag correlation, and periodogram analysis, but they have limited usefulness to the manpower practitioner. Some of these trace their ancestry back to early meteorology, weather forecasting and astronomy. For a general description see, e.g., Harold T. Davis, (*The Analysis of Economic Time Series* (Bloomington, Indiana: Pricipia Press, 1941); Kendall and Stuart, (footnote 4, p. 22), Vol. II; G. U. Yule and M. G. Kendall, *Introduction to the Theory of Statistics* (London: Charles Griffin & Co., 1950); and Valavanis, *op. cit.*

[17] See earlier cited efforts by the Harvard Committee on Economic Research and its business barometers of questionable fame, and the life work of Warren M. Persons.

[18] See W. C. Mitchell and A. F. Burns, *Statistical Indicators of Cyclical Revivals* (New York: National Bureau of Economic Research, 1938); G. H. Moore, *Business Cycle Indicators* (Princeton: Princeton University Press, 1961).

[19] G. H. Moore and J. Shiskin, *Indicators of Business Expansion and Contraction*, (New York: National Bureau of Economic Research, 1967), provide a detailed description of the construction and qualifications of these indicators.

Since 1961 the Bureau of the Census has been publishing indicators monthly in a report named *Business Cycle Developments*.[20] Similar measures are used in other countries.[21]

The indicators have serious limitations as practical forecasting devices. It is difficult to distinguish major turning points from what appears in retrospect to have been an accidental oscillation. A corrective is sought in so-called diffusion indexes, which measure the percentage of leading indicators moving in unison during a particular month and thus convey some idea about the pervasiveness of the observed change.

## *V. Econometric Models*

In the search for forecasting techniques in more recent times, interest has gradually shifted toward econometric methods. The purpose of constructing a particular econometric "model" may be to deepen the understanding of a process; to trace the economic effects of a policy or managerial decision; or to forecast. These objectives are, of course, not mutually exclusive, but the primary intent of the problem affects the nature and use of a model.

From the forecaster's point of view, econometric models may be conveniently divided into those based on single relationship (usually one equation) and those involving entire systems of relationships.[1] The first kind was first in the historical development of econometrics. In its classical form, the single-relationship model emerged in connection with the study of the market mechanism. It is exemplified by mathematical demand, supply, cost and production functions. Perhaps the most directly relevant function for manpower forecasting is the production function—a macroeconomic cost function, so to speak.

The production function associates changes in aggregate output (say,

---

[20] See in J. Shiskin, *Signals of Recession and Recovery* (New York: Columbia University Press, 1961).

[21] See, e.g., Organization for Economic Cooperation and Development, *Techniques of Economic Forecasting* (Paris: The Organization, 1965).

[1] Another classification of particular interest in the theory of forecasting has been proposed by H. O. A. Wold, in *The Econometric Approach to Development Planning* (Amsterdam: North Holland Publishing Co., 1965). For a more extensive treatment, see *Econometric Model Building*, edited by Wold (Amsterdam: North Holland Publishing Co., 1964). Wold distinguishes between causal chains (CC), or recursive models and interdependent (ID) systems. In forecasts by the CC principle, each result is used for the next forecast as if it were an actual observation. In spite of the theoretical merits of the recursive model (CC), its application to manpower forecasting is limited by inadequacy of information. Most data refer to a period much too long (say a year) to permit satisfactory decomposition of the underlying adjustment processes, step by step. See E. Malinvaud, *Statistical Methods of Econometrics* (Amsterdam: North Holland Publishing Co., 1966).

in manufacturing or in the private sector) with changes in categories of aggregate input. The pioneering Cobb-Douglas function[2] has been used to connect production ($P$) with labor ($L$) and capital ($C$) inputs in an exponential expression $P = aL^rC^s$. This corresponds to a linear equation involving logarithms, $\log P = \log a + r \log L + S \log C$. The exponents, in addition to their usual meaning as the slopes, may also be interpreted as partial elasticities indicating the relative changes in $P$ corresponding to small changes in $L$ and $C$.

Although the exponents need not add to unity, they often do. Thus Cobb and Douglas derived for manufacturing in the United States for the years 1900 to 1922, the formula $P = 1.01\ L^{.75} \times C^{.25}$ The partial elasticity of production with regard to labor was 0.75; of production on capital, 0.25. The sum of exponents, unity, indicates constant return to scale. Although the constancy of returns assumption has been challenged, the Cobb-Douglas model remains attractive to manpower analysts; it has also been used extensively at the microeconomic level, under the name of cost functions.[3] The measurement of elasticities has been developed primarily in connection with the analysis of commodity demand function.[4]

Once a specific relationship involving labor input, output, and other relevant variables has been established, the function can be interpreted for forecasting purposes. Thus, if in the equation $P = aL^r \times C^s$, future estimates or assumed values are substituted for $P$ and $C$, the appropriate value of $L$ can be found.

Even within the Cobb-Douglas framework, production functions with more than two inputs might, of course, be considered. For instance, we may set $P = f(L, C_1, C_2, A_1, A_2, A_3,)$ where $P$, as before, stands for product, $L$ for labor, and $C_1$ for physical capital. But we add $C_2$ for capital in education, and $A$ for changes in technological level, which

[2] P. H. Douglas, *Theory of Wages* (New York: The Macmillan Company, 1934); C. W. Cobb and P. H. Douglas, "A Theory of Production," *American Economic Review*, Supplement, Vol. 18 (1934); and many subsequent papers by Douglas and collaborators investigating the function for various countries and for various periods. See, for example, P. H. Douglas, "Are there Laws of Production?", *American Economic Review*, Vol. 38 (1948). The original formulation of the Cobb-Douglas production function goes back to Knut Wicksell, *Lectures on Political Economy* (London: G. Routledge & Sons, 1934).

[3] A good technical review of this topic can be found in J. Johnston, *Statistical Cost Analysis* (New York: McGraw-Hill Book Company, 1960). See also earlier, and the mathematically more demanding presentation by R. W. Shephard, *Cost and Production Functions* (Princeton: Princeton University Press, 1953).

[4] See the monumental work of H. Schultz, *The Theory and Measurement of Demand* (Chicago: Chicago University Press, 1938). Appendix C of this work is still one of the most complete presentations of the techniques of fitting regression lines appropriate to such functions. For an up-to-date treatment, see also H. O. A. Wold and L. Jureen, *Demand Analysis* (New York: John Wiley & Sons, Inc., 1953).

might be subdivided into labor saving ($A_1$), capital savings ($A_2$), and neutral ($A_3$)[5]

More ambitious attempts to investigate economic interdependence make use of several, sometimes very many, simultaneous relationships.[6] The computational aspects need not deter forecasters. The logical basis is the same as that underlying the diminutive system of two equations which had to be solved in fitting a straight line to empirical data.

As more equations are to be solved simultaneously for more unknowns, the computing burden increases rapidly, and efficient methods need to be devised.[7] For really large systems of equations, the use of electronic data processing equipment becomes imperative; although the forecaster can leave to computer experts the details of the numerical solution, he does need to be sure of the rationale of his model and, above all, he should be conversant with the empirical data underlying the system of equations, including their limitations.[8]

As an illustration of a large system, we refer to a simplified model relating a nation's economic growth, labor force, and educational requirements through six linear equations.[9]

The number of labor force participants with a high school education is assumed to increase in constant proportion with production. Labor force participants with only a grade school education are neglected since such education is generally available and obligatory. In order to permit for flexibility, the labor force participants are divided into two groups: the new entrants during the preceding year and all those who had entered

[5] See, e.g., J. E. Meade, *The Neo-Classical Theory of Economic Growth* (London: G. Allen & Unwin, 1959). Discussion of some technical aspects of production functions may be found in G. Tintner, *Econometrics* (New York: John Wiley & Sons, Inc., 1952); Harold T. Davis, *The Theory of Econometrics* (Blcomington, Indiana: Principia Press, 1941); R. G. D. Allen, *Mathematical Analysis for Economists* (London: MacMillan and Co., Inc., 1938); and R. G. D. Allen, *Mathematical Economics* (London: St. Martin's Press, 1957). See, also Shephard, *op. cit.*; and R. G. Bodkin and L. R. Klein, Non-linear Estimations of Aggregate Production Functions, in *Review of Economics and Statistics* (February 1967).

[6] The concept is not new, but its operational formulation and application have been recently developed by econometricians. Among earlier contributions to the approach L. Walras should be mentioned especially.

[7] For desk calculators, the Doolittle method is probably the most widely used. A description appears in various elementary statistics textbooks, such as F. E. Croxton and D. J. Cowden, *Applied General Statistics* (New York: Prentice-Hall, Inc., many editions since the first in 1939); and in such old authoritative handbooks as T. W. Wright and J. F. Hayford, *The Adjustment of Observation* (Princeton: D. Van Nostrand Co., Inc., 1906).

[8] On the use of computers for problems involving many variables in the social sciences see, for example, W. W. Cooley and P. R. Lohnes, *Multivariate Procedures for the Behavioral Sciences* (New York: John Wiley & Sons, Inc., 1962).

[9] "Linear" here refers to the fact that the influences of the several terms are assumed to be additive and independent of each other, thus the possible presence of interaction and joint effects is disregarded.

during the past six years (that is, the average period assumed necessary for post-elementary education). Explicit consideration is also intended for the fraction of labor force leavers (because of death or for other reasons). New entrants with high school education are equal to the number of high school students a year earlier, excepting those who are going on to colleges and universities. New entrants with a college or university education go either into teaching (in high schools, colleges, or universities) or other productive activities.

Before exhibiting the equations we describe the symbols used therein. The volume of production (say, in billions of constant dollars) is designated by $P$; the members of the labor force by $L$, of whom $E$ have entered the labor force during the past six years; and the number of students in millions by $S$. Subscripts $h$ and $c$ refer to high school and college (or university) respectively. The coefficients to be used in the equations are designated by $b$, with appropriate subscripts. Thus, $b_{Ph}$ and $b_{Pc}$ stand for ratios of labor force to production, the first one for persons with high school training and the second for persons with college and university training: $b_{Lh}$ and $b_{Lc}$ represent fractions of labor force leavers with high school and higher education, respectively; subscript ($t$) indicates the time of reference.

With the above symbols, the postulated relationships form the following six linear equations;[10] (from Tinbergen, see footnote 11 below).

(1) $$L_{h\,t} = b_{Ph}P_{t}$$

(2a) $$L_{h\,t} = (1 - b_{Lh})L_{h\,t-1} + E_{h\,t}$$

(2b) $$L_{c\,t} = (1 - b_{Lc})L_{c\,t-1} + E_{c\,t}$$

(3) $$E_{h\,t} = S_{h\,t-1} + S_{c\,t}$$

(4) $$E_{c\,t} = S_{c\,t}$$

(5) $$L_{c\,t} = b_{Pc}P_{t} + b_{Rh}S_{t} + b_{Rc}S_{c\,T}$$

Numerical values of the coefficients may represent, *a priori*, policy-determined quantities (e.g., a teacher student ratio considered desirable) or estimates from cross section data or time series. Cross section data might refer to a set of regions or subregions of a particular country, or to several countries which have a similar structure. Time series data might supply useful arithmetic mean values from observations of a decade by estimates of trend values.

For the sake of illustation, let us assume that the estimated coefficients are 0.2 for $b_{Ph}$, 0.02 for $b_{Pc}$, 0.1 for $b_{Lh} = b_{Lc}$, 0.04 for $b_{Rh}$, and 0.08 for $b_{Rc}$, $R$ being the postulated teacher–student ratio.

[10] See I. H. Correa and J. Tinbergen, "Quantitative Adaptation of Education to Acceleration of Growth," *Kyklos*, Vol. XV (1962).

Time ($t$) will be measured arbitrarily in units of six years since the average period for the kind of education here considered is six years; correspondingly, $t - 1$ will stand for the last year before the end of period $t$.

Let us suppose a forecasting problem in which this question is to be answered: Is the particular educational structure compatible with demands six years hence, if the target increase in production for which manpower is to be supplied is 4.5 percent per annum or about 30 percent in six years? We express this expected increase of production and the correspondingly required increases in the labor force and education variables from 1 to 1.3 by the end of the first six-year period, by letting $P_1 = 1.3P_0$, $L_1 = 1.3L_0$, $E_1 = 1.3E_0$, and $S_1 = 1.3S_0$; for the last three variables, both the high school and college components are to be considered. By substitution in the above system of equations and solving, we obtain:

$$L_{h0} = 0.2P_0 ;\ L_{c0} = 0.02450_0 ;\ S_{h0} = 0.094P_0 ;$$

$$S_{c0} = 0.0098P_0 ;\ E_{h0} = 0.062P_0 ;\text{ and } E_{c0} = 0.0076P_0$$

If initial production stands at \$260 billion the educational structure required to support, six years later, a volume of \$338 billion ($1.3 \times 260$) includes (0.094) (260) = 24.5 million high school students and (0.0098) (260) = 2.5 million college and university students. If initially, there are, say, 30 million students in high schools and 2 million in colleges and universities, we should expect the future manpower requirement for high-school-trained-personnel to be met easily, but there would be a shortage (approximately 20 percent) of college-and-university-trained people unless appropriate remedial action is taken.[11]

Well-known monumental models descriptive of the entire economy are also available—e.g., the Brookings model and the Klein-Goldberger model. In spite of their comprehensiveness, the usefulness of such models for manpower forecasting is limited, up to this time, to conditional prediction.

This paper cannot even attempt to scan the outstanding peaks of these efforts or pretend to give an introduction to the econometrics of large national models.[12] For the purpose of this brief introduction, it may suffice to recall the classification of econometric variables into

[11] Slight modifications have been made in the more realistic, hence considerably more complex, model presented *in* Organization for Economic Cooperation and Development, *Econometric Models of Education* (Paris: The Organization, 1965).

[12] In addition to the several well-known textbooks in econometrics, the following contributions should be mentioned: G. Cooper, "The Role of Econometric Models in Economic Theory," *Journal of Farm Economics*, Vol. 30 (1948); E. G. Bennion, "Cowles Commission's Simultaneous Equation Approach: A Simplified Explanation," *Review of Economics and Statistics*, Vol. 34 (1952); W. C. Hood and T. C. Koopmans, *Studies in Econometric Method* (New York: John Wiley & Sons, Inc., 1953); E. Malinvaud, *op. cit.*; and C. F. Christ, *Econometric Models and Methods* (New York: John Wiley & Sons, Inc., 1966).

endogenous ones (whose values it may be desired to predict), and exogenous ones postulated or assumed to be known *a priori*, and into jointly dependent (current endogenous) and predetermined (lagged endogenous and the exogenous) ones.

Many of the variables of interest to the manpower forecaster are usually treated as exogenous and thus are not convenient candidates for model-based forecasting. Future government policies, union policies, propensities of individuals to enter or leave the labor force, family formation, and attitudes toward ethnic groups are examples of variables usually treated as exogenous.[13] Some aspects relevant to manpower forecasting are often absent altogether. Among these are "production functions" for labor itself—that is, the specification of variables and relationships describing education, training, and other forms of skill transfer.

Most of all, some of the major aggregative variables included in models may have been estimated via manpower forecasts. Thus we expose ourselves, as earlier suggested, to the danger of circular reasoning when we use these variables in turn in manpower forecasting.

On the other hand, a closely related procedure, also structural in nature, at this time seems to have much greater appeal to manpower forecasters. The method alluded to goes under the name input-output or interindustry analysis. In essence it is a form of multiple-entry national accounting, or accounting for large geographic aggregates. This family of models has been of considerable interest to manpower analysts from its origin. At first glance the approach may resemble that of structural models just-sketched. It, too, is based on interrelations presented in a system of equations.

Subdividing the economy into many activities, e.g., industries or sectors, the input-output analysis presents an orderly tableau,[14] an *n*-by-*n* grid. The rows are the components of output (measured in dollars of transactions) of a given industry or activity channeled into other industries or activities, to which they become inputs and which are represented by the columns. Each row-column intersection, or cell, lists the amount of the row activity entering the column activity as an input. Also attached to the table are several columns, including final demand (the so-called autonomous sectors), so that the tableau gives at one glance a picture of the dollar transaction among industries according to input and output and to final demand.[15]

---

[13] See, however, the succinct contribution by S. Lebergott *in* J. S. Duesenberry, *et al.*, *The Brookings Quarterly Econometric Model of The United States* (Amsterdam: North-Holland Publishing Co., 1965).

[14] In this aspect input-output traces back to the famous *tableau économique* of Quesney.

[15] On the use of input-output analysis for forecasting, see the comprehensive presentation by Clopper Almon, Jr., *The American Economy to 1975* (New York: Harper & Row, Publishers, 1967).

By appropriate manipulation of this so-called transaction matrix, derived matrices or tables can be obtained that are of particular interest.

For example, the entries in the table may be expressed as percentages of the respective column totals (the sector's output totals), indicating the purchasing sector's direct transfers or purchases from the producing sector's per unit of output, say, per dollar (the so-called coefficient matrix); from this matrix can be obtained a further Table showing the direct as well as the indirect requirements placed on all industries to support the delivery to final demand of a unit of output (say, one dollar), from each industry (the so-called inverse). Thus it becomes possible to "compute" rather than literally to trace step by step, the primary, secondary, and higher order transactions which result from, say, an increase in output of automobile production. This increase creates a direct demand for steel—and many indirect ones for steel too—e.g., for amounts needed to support the very increase of steel production.

Once a table of direct and indirect input requirements has been computed (by solution of systems of simultaneous equations) it is easy enough to translate these dollar amounts into manpower units. The resulting entries in the cells of the table may then be interpreted as reflecting direct and indirect manpower requirements associated with a given output for a particular activity. This last-mentioned aspect of input-output analysis is of primary interest for manpower analysis and forecasting. A stipulated change somewhere in the economic structure (perhaps an increase in a specific demand) may be traced in its manpower aspect throughout the system.

To convert the entries in a transaction matrix into a corresponding manpower matrix, we must relate manpower and dollar variables in the several industries or activities. From this manpower matrix an inverse can be computed that shows direct and indirect manpower requirements, say, per $1 of final demand. This inverse makes it possible to analyze manpower implications of proposed unemployment policies, of given export and import patterns, and of changes in the several components of final demand.[16] In addition to prediction, with the help of the inverse, we should recognize the forecasting of the input-output coefficients themselves as a challenge to manpower analysis.

So long as the economic structure as described by the coefficient matrix remains reasonably stable, conditional prediction with its help may be safe enough. However, if the forecasting horizon is long, the coefficients themselves may change and a matrix more appropriate for a specified future date must first be provided.

Since the input-output coefficients of the manpower matrix indicate

[16] See, for instance, National Bureau of Economic Research, *Input-Output Analysis: An Appraisal* (Princeton: Princeton University Press, 1955), particularly the paper by W. Duane Evans and M. Hoffenberg.

the manpower inputs (say, in terms of number of jobs) per unit of output, the factors likely to change the numerical values of these coefficients need to be explored. Among such factors, man-hours, productivity, and hours of work (since the measurement unit is a job) are particularly important. They must be predicted for forecasting the economic structure, the manpower inverse, and such.

Early work on employment oriented input-output analysis was undertaken under the sponsorship of the Bureau of Labor Statistics;[17] the most recent contribution in this direction was also made by the Bureau, although the basic input-output information used here was predicated on the Department of Commerce's general input-output matrix for 1958.[18]

The manpower-projection scheme involves these phases when a basic input-output table is already available:[19]

1. Total GNP prediction.
2. Prediction of distribution of GNP over the various major components of final demand.
3. Prediction of the distribution of major demand components over specific demand items.
4. Application of these estimated dollar amounts to the inverse, that is, multiplication of the inverse by each of the anticipated levels of demand. The resulting table gives in its cells the dollar amounts of direct and indirect requirements of the column industries that are to be satisfied by the row industries. (Intersection of column 15 with row 6, for instance, would show the 15$^{th}$ industry's input from industry 6). Now these predicted output requirements need to be translated into corresponding manpower requirements.
5. The original input-output table is converted to an employment-requirement table by multiplication of the entries by per-dollar employment estimates for each of the activities. Since per-dollar output requirements for manpower change with technology, factor substitution, etc.,[20] the table obtained in this step must be amended to incorporate such anticipated changes in its cells.
6. The Bureau of Labor Statistics projects the employment-requirement table, taking account of anticipated developments in labor

[17] See J. Cornfield, W. Duane Evans, and M. Hoffenberg, "Full Employment Patterns, 1950," *Monthly Labor Review* (February and March 1947).

[18] See *Survey of Current Business* (November 1964 and September 1965)—papers by M. R. Goldman, M. L. Marimont, B. N. Vaccara, and the National Income Division staff.

[19] For details see U. S. Bureau of Labor Statistics, *Projections 1970*, BLS Bulletin 1536 (Washington: 1967), and literature cited therein.

[20] On measurement of labor input see W. Duane Evans, "Indexes of Labor Productivity As a Partial Measure of Technological Change," *Input-Output Relations, Proceedings of a Conference on Inter-Industrial Relations Held at Driebergen, Holland* (Leiden: H. E. Stenfert Kroese, N. V., 1953).

productivity[21] and man hours. Then predictions of the coefficients in the table are made on the basis of intensive industry studies as well as conjectures about technological changes, the length of the future working day, etc.[22] Here and there, the coefficients observed for leading establishments are imputed to an entire industry, on the assumption that new technologies and factor-use patterns will eventually permeate the industry.

The forecasts of employment requirements derived by the multiplication of the anticipated bill of goods with the corresponding entries in the employment-requirement inverse (after the inverse, in turn, has been projected to the date of the desired forecast) are subject to various restrictions. In addition to forecasting difficulties, there are difficulties inherent in the input-output model itself. A fundamental identity is assumed for example, between product and activity. A related assumption is product or out-put homogeneity (i.e., all outputs of a particular activity or industry are essentially identical and interchangeable, dollar for dollar). For a given time the coefficients of input to output are assumed constant; in other words, input-output ratios do not alter with changing volume of output.

Thus far no attention has been given to the all-important question of the validity of information. In manpower or other socioeconomic forecasting, the information available to the analyst is usually deficient in some degree.

Incompleteness of information is a typical situation, and much, if not all, of statistical method relates to the very problem of drawing inferences from partial information. Modern sampling theory, for example, is used for estimating unknown characteristics of a population (the parameters) from statistics obtained from samples. If proper procedures are followed in gathering such sample data, then probability theory will point the way to estimates of sampling error and how to proceed from sample measures to population values for the particular population characteristics or the range within which this value is likely to be found. Other things being equal, the wider the acceptable range the higher our confidence that this range will contain the value of the characteristic. In bivariate and multivariate situation, theory provides means for estimating from sample data the "true" but unknown regression coefficients for the population.

[21] See Siegel (footnote 4, p. 7 in Chapter II) *in* Sar A. Levitan and Irving H. Siegel, *Dimensions of Manpower Policy.*

[22] A procedure relying perhaps more heavily on formal techniques in the projection of input-output coefficients is described in a paper developed by the Department of Applied Economics, University of Cambridge, England, entitled "Input-Output Relationships, 1954–1966" (distributed in the United States by MIT press, Cambridge, Mass., 1963).

The problem of prediction from time series is more complex. The standard (parametric) approaches are based on assumptions hardly ever realized in typical manpower situations[23] e.g., independence of successive observations, normality of distribution of the residuals about a line of regression, and unchanging variances over the observed time domain (so-called homoscedasticity).

Errors often result from incompleteness of the forecasting model itself. Not all the relevant variables may be included the very purpose of model building. The form of the relationship may be oversimplified beyond the permissible limit implicitly sought in a model. Errors of measurement, the so-called errors of observations are presently the kind toward which classical statistical theory is directed, but error in time series observations is still not easy to handle. Heroic assumptions may be needed e.g., that our time series is in turn a sample from a hyperpopulation of all possible relevant time series. Still, it is important to know the variances of the estimates where they are available, to compare the performance of different models and procedures and to provide a check on likely usefulness and meaning of forecasts.

Of course, the specification of the model may itself be faulty. A tendency to include in a model as many relationships as can be obtained is difficult to avoid; but if two or more of the predicting "independent" variables are actually interdependent, the validity of the system is limited. Thus two predictors that move in unison, that are highly in-

[23] A glimpse into this problem is provided by Davis (footnote 17), Kendall and Stuart (footnote 4), Tintner, Goldberger, Johnston, and Yule and Kendall (footnotes 15 and 16) with ample reference to more advanced literature. In addition to parametric methods that assume knowledge of the underlying probability distributions, distribution-free (so-called nonparametric) methods have been developed for the analyais of economic time series. Mar and Hotelling were important early contributors; a not-too-technical presentation may be found in M. G. Kendall, *Rank Correlation Methods* (London: Charles Griffin & Co., 1948) and in G. A. Ferguson, *Nonparametric Trend Analysis* (Montreal: McGill University Press, 1965).

At this time still of minor importance for the manpower forecaster are harmonic analysis and spectral analysis. They are not discussed here. On the first, see Davis (footnote 18 and M. G. Kendall, *Contributions to the Study of Oscillatory Time Series* (Cambridge, England: Cambridge University Press, 1946); on the latter, see J. Cunningham, "The Spectral Analysis in Economic Time Series," Census Bureau Working Paper No. 14 (Washington: 1963) and, for a more complete and more technical treatment, see W. J. Granger and M. Hatanaka, *Spectral Analysis of Economic Time Series* (Princeton: Princeton University Press, 1964).

The application of stochastic process models to economic time series has been discussed by H. O. A. Wold, *A Study in the Analysis of Stationary Time Series* (Uppsala: Almqvist & Wiksells, 1938), and in the abundant recent literature see, for example, J. G. Kemeny and J. L. Snell, *Finite Markov Chains* (Princeton: D. Van Nostrand Co., Inc., 1960); and Wold, *Bibliography on Time Series and Stochastic Processes* (Edinburgh: Oliver & Boyd, 1965).

tercorrelated, amount to only one effective predictor. The macro variables to be used in manpower forecasting (e.g., GNP, its major components and population) may all be strongly intercorrelated during sustained periods of expansion. The resulting problem, called "multicollinearity" by Frisch,[24] introduces complexities and weakens the forecasting capability of the model. Its presence should, at least, be recognized by the manpower forecaster.[25]

The difficult problems arising from the use of macro approaches that depend on highly aggregated information have recently evoked much attention. Thiel's effective illucidation[26] has encouraged the exploration of other avenues. This quest has also been stimulated by practitioners' dissatisfaction with the state of statistical inference for the analysis of economic time series. In the meantime, the investigation of the behavior of complex systems has also received a potent fillip in the engineering and related fields from a computer application known as simulation. This technique was applied with reported success to a miscellany of problems ranging from general management to transportation, traffic control, and military logistics.[27]

The essentially stochastic nature of economic and behavioral relationships affecting manpower can be turned to advantage in simulation, which involves two operations:[28] construction of a model representing the "real" process being considered; and experimentation with the model via the computer, with a view to forecasting, testing, or design. The model is stated in mathematical terms but not, as earlier, to specify a system of contemporaneous equations to be solved together in order obtain, say, predicted values of the variables.

The essential constituent elements of the model are the particular de-

[24] Ragnar Frisch, *Statistical Confluence Analysis by Means of Complete Regression Systems* (Oslo: Universitets Økonomiske Institutt, 1934) proposed a useful graphic method (now unfashionable) called bunchmap analysis. See also O. Reiersöl, *Confluence Analysis by Means of Instrumental Sets of Variables* (Uppsala: Almqvist & Wiksells, 1945).

[25] On important problems facing the model builder, including the problems of statistical inference in multiple regression, see E. F. Beach, *Economic Models* (New York: John Wiley & Sons, Inc., 1957).

[26] H. Theil, *Linear Aggregation of Economic Relations* (Amsterdam: North-Holland Publishing Company, 1954), addresses the conceptual obstacles to aggregation and shows the difficulties that exist even when "perfect" aggregation is theoretically possible.

[27] See, e.g., *Reports of the System Simulation Symposium, American Institute of Industrial Engineers* (New York: The Institute, 1957 and 1959).

[28] See G. W. Morgenthaler, "The Theory and Application of Simulation in Operations Research," *in* R. L. Ackoff, ed., *Progress in Operations Research*, Vol. I (New York: John Wiley & Sons, Inc., 1961), and an extensive bibliography on applications to other than the behavioral sciences. For some examples of the application of simulation to business see *Proceedings on Simulation* (New York: American Statistical Association, New York Area Chapter, 1966).

cision-units—say, individuals or firms—whose behavior is to be studied, not the large economic aggregates underlying the economic macro models earlier described. Another departure from macro-modeling is in the primary use of the computer to generate responses of the individual components.

Still in its early stages of development, simulation has been applied to forecasting in the behavioral fields—especially in the general area of demography, which has many parallels to manpower analysis.[29] A brief listing of ingredients found in typical simulation indicates its nature and data requirements:[30]

1. The "elements or components"—e.g., individuals, households, firms, or goods.
2. The "attributes or variables" which may be divided into
   a. "Responses" or "output variables"—e.g., joining or not joining the labor force, having a job, not earning an income, having three dependents, having children.
   b. Factors influencing the behavior of the decision units other than their direct responses—input invariables such as the season and the responses (outputs) of other elements.
   c. Status characteristics of the decision units—e.g., age, sex, educational achievement, income group.
3. Relationships necessary to prediction divisible into—
   a. Definitions or identities (e.g., the statement that income consists of straight-time wage plus overtime wage plus nonwage income).
   b. The so-called operating characteristics (relationships which, for a given element such as a person, tie item 2a to 2b and 2c); empirical or postulated, they are fed into the simulation to specify how elements act or respond to stimuli.

Operating characteristics usually have to be stated in probabilistic rather than deterministic terms. For example, the probability that, in a specific situation, a person of given sex, age and educational achievement (the status variables) will, during winter (an input variable reflecting the effect of the seasonal component) join the construction industry (an output variable) is an operating characteristic. If age-and-sex-specific labor

[29] See, e.g., M. C. Sheps and J. C. Riley, eds., *Public Health and Population Change: Current Research Issues* (Pittsburgh: University of Pittsburgh Press, 1965); Riley and Sheps, "An Analytic Simulation Model of Human Reproduction with Demographic and Biological Components," *Population Studies*, Vol. 19 (1966); and M. Shubik, "Bibliography on Simulation, Gaming, Artificial Intelligence and Allied Topics," *Journal of the American Statistical Association*, Vol. 55 (1960). Of particular interest to the manpower student is J. Korbel's chapter on "A Labor-Force Model" *in* G. H. Orcutt *et al.*, *Micro Analysis of Socioeconomic Systems: A Simulation Study* (New York: Harper & Row, Publishers, 1961).

[30] The presentation here follows closely Orcutt, ibid, which also provides specific examples. See especially pp. 15 ff.

force participation rates are available, estimating the corresponding probabilities of joining the labor force may be relatively simple. In other instances, however, research may first have to be undertaken to provide a reasonable probability estimate.

Since the model is descriptive of a historical process, it may be viewed as a sequence of equispaced cross sections or states, through time, perhaps one for each month. A state can be described by an appropriate tableau of probabilities corresponding to the operating characteristics. Applying the probabilities (often called transition probabilities) to the tableau of one state produces the tableau for the next.[31] What characterizes simulation is the selection, according to these probabilities, of individual elements (say, persons) for assignment to new locations or destination in the next following state. Thus, given the probability of a particular person's moving from unemployment in state $t$ to employment in state $t + 1$, his movement from $t$ to $t + 1$ can be simulated by a selection mechanism that places a particular element chosen at random into a new place with the desired probability. This selection, on the basis of random choice (sometimes called the Monte Carlo method, a name given by von Neumann and S. Ulam) is undertaken by the computer at such an impressive speed that processes which would take years or generations in real time can be collapsed to hours.

Starting with a sample of individuals or households, or with a cohort of women of, say, age 22, we may simulate the process of forward change. We may desire and attempt to modify the probabilities themselves as the simulated process proceeds. The computer reads the original data input for the initial state, updates with specific probabilities of transition to state two, updates again, and so forth.[32]

Problems of sampling or statistical inference are difficult enough for empirical time series; they are, as yet little explored for input-output analysis and simulation. The manpower forecaster looking for some way to improve his forecasts, to attach to them measures of reliability, should consider if the forecasting method itself can be improved? As in the proof of Cervantes' pudding, comparison of forecast and actual outcome (the "realization") plays an important role. For such comparative evaluation Theil has proposed an intriguing yet simple technique.[33]

This method, applicable to manpower forecasts, systematically compares a predicted percentage change, $P_T$ (where $T$ refers to the particu-

[31] For an application of transition probabilities to a labor force problem, see I. Blumen, M. Kogen and P. J. McCarthy, *The Industrial Mobility of Labor as a Probability Process* (Ithaca: Cornell University, 1955).

[32] To facilitate communication with the computer, special-purpose languages have been devised, e.g., Simscript. Of particular interest to labor force type of simulation is *Program Simulate, Revision AB*, by C. C. Holt *et al.* (Madison: University of Wisconsin, 1965).

[33] H. Theil, *Applied Economic Forecasting* (Amsterdam: North-Holland Publishing Co., 1966), and "Who Forecasts Best?" *International Economic Papers*, No. 5 (1955).

lar time, say, year), and the actually realized change, $A_T$. The expression $\frac{\Sigma(P_T - A_T)^2}{N}$ is an average measure of the magnitude of total discrepancy. Relating this term to the average of squared realized changes, $\frac{\Sigma A_T^2}{N}$, Theil obtained $U^2 = \frac{\Sigma(P_T - A_T)^2}{\Sigma A_T^2}$. A measure of discrepancies, $U$ is called the inequality coefficient. A valuable property of this measure is that, like variance, it can be decomposed into three components, each indicative of a particular source of errors. The three components are:

1. $(\bar{P} - \bar{A})^2$, the error in central tendencies;
2. $(\sigma_p - \sigma_A)^2$, the error of unequal variations; and
3. $2(1 - r)\sigma_P\sigma_A$, where $r$ measures the correlation between predicted and actually observed changes, and $\sigma$ is the standard deviation.

Expressing these three components as fractions of the total discrepancy $\left(\frac{\Sigma(P_T - A_T)^2}{N}\right)$ we obtain three fractions: $U^m$, $U^s$, and $U^c$. $U^m$, the bias fraction indicates the degree of discrepancy between average predicted and average actually observes values. $U^s$, the variance fraction, suggests the presence of residual cycles and fluctuations in the deviations of predictions from realizations, (the forecasting procedure may not take adequate account of business cycles or other cyclical features of the process). $U^m$ vanishes only if average predicted changes equal actually observed changes—a property expected from a good forecasting procedure, since the forecast should be correct "at least on the average." Thus, if $U^m$ is relatively large, the forecasting procedure would seem deficient; and adjustments and improvements would be called for. A relatively large $U^s$ should alert the forecaster to the possibility that cyclical forces have been overlooked.

$U^c$, the "covariance" fraction, characterizes a more abstract aspect of the relation between forecast and actually realized changes. Since $U^m + U^s + U^c = 1$, $U^c$ is the residual component: it represents errors due to what Theil calls incomplete covariation. The forecaster can do little, if anything about them.[34]

Another test criterion for prediction schemes is Wold's Janus coefficient.[35] This measure, analogous to the $F$ ratio (named after R. A. Fisher) is defined as the ratio of (a) the mean squared discrepancy be-

[34] On inequality coefficients, see Theil, *Applied Economic Forecasting op. cit.*, and his *Economic Forecasts and Policy* (1st ed.; Amsterdam: North-Holland Publishing Co., 1938). Also of interest is Thiel's simple graphic method (the so-called prediction-realization diagrams) which at a glance reveals the extent to which the forecasting procedures have resulted in overestimation or underestimation, and their sensitivity in anticipating turning points.

[35] A. Gadd and H. O. A. Wold, "The Janus Coefficient: A Measure for the Accuracy of Prediction," *in* H. O. A. Wold, ed., *Econometric Model Building* (Amsterdam: North-Holland Publishing Co., 1964). The reference to Janus indicates that the measure looks forward over the prediction horizon and backward over the sample period.

tween projected and (later) observed values to (b) the mean squared discrepancy between estimates that would have been obtained by application of the model to the sample period itself and the corresponding actual observations. In algebraic terms,

$$J^2 = \frac{\sum_{T=N+1}^{T=M} (P_T - A_T)^2}{\sum_{T=1}^{T=N} (E_T - A_T)^2}$$

Here $P$ again refers to the predicted values, $E$ to the estimates which would have been obtained by application of the model to past data, $A$ to the actual values, and $T = 1, 2, \cdots N, N + 1, \cdots N + M$, to time, say, in years or months. The first $N$ observations refer to the sample (the observed past) on which the prediction model is based, and the next $M$ intervals refer to the prediction horizon. The measure can of course be obtained only after accumulation of experience with the forecasting model during the $M$ periods.

As Wold himself points out, $J$ is an indicator of the relative size of the discrepancy between $P$ and $A$ It thus does not appraise the particular forecasting procedure for sensitivity to turning points or for ability to anticipate them accurately.

## *VI. Retrospect and Prospect*

Manpower forecasts are, like other statistical measures and estimates, answers to particular questions,[1] so there are many varieties of forecasts obtainable by many techniques. The alternatives would appear even greater had we included the methods adaptable to manpower forecasting but not actually so employed. Thus, the manpower forecaster really works from an embarrassment of riches. He therefore must often choose among different methods and interpret competitive forecasts. We now review the gamut of forecasting approaches, attempting to sharpen the reader's appreciation and comprehension of them by an exposition in terms of dichotomies, which in reality represent only the extreme points of continua.

### *Short-Term Versus Long-Term Forecasting*

Although the distinction between short-term and long-term forecasts is common, it is among the less informative ones. Manpower forecasting when concerned with a variable that is a simple function of population, frequently focusses on a longer forecasting range or horizon than does the

[1] See R. A. Lester, *Manpower Planning in a Free Society* (Princeton: Princeton University Press, 1966).

prediction of other important economic variables. If, for instance, a manpower component is to be forecast which strongly depends on the "closed active population," i.e., the population that is between ages $A$ and $A_u$ and not subject to change by migration, a relatively long forecasting horizon is generally permissible. For example, if the closed active population includes all persons between the ages of 15 and 65, a forecast looking 15 years hence would seem to be safer than for nearly any other economic variable independent of or less dependent on, population size. But even such a forecast involving persons already alive, involves restrictive assumptions, a major one being the stability of the age-specific death rates over the next $A$ years.

In general, however, long-term forecasts are less certain of realization than short-term ones. For one thing, assumptions of stability in trend, in ratios to other variables, etc., become more and more tenuous the longer the period anticipated. Furthermore, the chance of major structural alterations or of a sudden change in an exogenous variable[2] increases with the time horizon. If a forecast is made for a year 10 or 15 years hence, the secular trend need be considered, but not the variations and cycles shorter than 10 or 15 years.

A short-term forecast, say, for six months hence, does require knowledge about short-term (in this case intra-annual) fluctuations. The analysis of seasonal variations and their inclusion in the projection is vital.

### *Stochastic Versus Deterministic Approach*

The term "stochastic," originated with James Bernoulli in 1713.[3] In its present meaning it has been in wider use only for a relatively short time, having been reintroduced by von Borthiewicz and Tschuprow.[4] It is now frequently employed in economics and elsewhere to characterize errors of observation and other sources of variability about the underlying "true" value that is to be estimated. In a stochastic relationship between two or more variables an unequivocal sharply defined relationship is not postulated as in the "deterministic" case. A stochastic variable may be described in terms of a probability distribution. Empirical economic relations are, as a rule, stochastic relations.

Nonstochastic, i.e., deterministic, approaches are typically used in manpower forecasting because of their relative simplicity or for lack of sufficient and appropriate estimates. We usually assume a functional re-

[2] Exogenous variables, as opposed to endogenous ones, are variables independent of the economy; many acts by governments are of this nature. Which specific variables are to be treated as exogenous and which as endogenous depends of course on the forecasting method or model used.

[3] *Ars Conjectandi* (Basel: Impensis Thurnisiorum, fratrum, 1713).

[4] A. A. Tschuprow, "Ziele und Wege der Stochastischen Grundlegung der Statistischan Theorie, "*Nordisk Statistisk Tidskrift*," Vol. 3 (Stockholm: 1924).

lationship for what we know to be a stochastic one. Although no practicable solution to the stochastic problem may be available, it is important to keep in mind the inherently stochastic character of forecasting from empirical data. Neglect may result in forecasting-blunders of considerable magnitude.

### *Point Versus Interval Forecasts*

Forecasting of an empirical stochastic process will result, not in a single value of the predicted variable, but in a distribution of predictions with varying probabilities. As a rule, therefore, we can forecast with greater assurance in terms of a range of values than in terms of a single point.

Even in nonstochastic procedures, the validity of the forecast can presumably be improved if a range is substituted for a point. Instead of a simple forecast we may with greater assurance provide an upper and lower limit or provide several forecasts with slightly different assumptions. The Census Bureau's alternative population forecasts illustrate this approach.

It is the user of predictions rather than the forecasting model that often has to be blamed for apparent shortcomings of the latter. Enormous pressure is placed on forecasters to provide point estimates, to make one forecast only. The elementary urge of the pragmatic man of action is to obtain a unique figure. An aversion to interval estimates, however, does not speak well for the sophistication of the statistics user.

### *Unconditional Versus Conditional Forecasts*

An unconditional forecast is not contingent on the realization of a specific condition. Forecast of the size of the male labor force by means of a trend fitted to, say, the male labor force figures for the past six years is unconditional if no additional qualifications are introduced.

A great many manpower forecasts are conditional, although the user does not always ascertain what the conditions are. Furthermore, the maker of forecasts does not always bother to spell out the major implicit conditions that he really accepts for his procedure. Examples of well-qualified conditional forecasts are the Census Bureau's population forecasts.[5] Practically all economic forecasts are conditional in that they assume certain stabilities on which there could be considerable disagreement. Few forecasts, however, are adequately qualified,[6] and full qualification is exceedingly difficult if not impossible.

[5] For other examples, see H. Theil, *Applied Economic Forecasting* (Amsterdam: North-Holland Publishing Co., 1966).

[6] This fact was recently stressed by W. Lee Hansen, "Labor Force and Occupational Projections," *in* Gerald G. Somers, ed., *Proceedings of the Eighteenth Annual Winter Meeting of the Industrial Relations Research Association* (Madison: The Association, 1966).

Conditional forecasts based on clearly stated major assumptions are very useful. They help to realize that important premises may not actually be met; they help us to trace the implications of a policy under consideration. Alternative conditional forecasts permit a comparison of likely results of competitive assumptions or decisions.

*First-Order Versus Higher-Order Forecasts*

A forecast is usually treated as a first-order forecast—as if it were "neutral" with respect to the future course of the event predicted. The forecast itself, however, might affect the future value of the predicted variable. In many situations, in the social sciences, as elsewhere, a forecast may be "adaptive" and the process under consideration may be subject to prediction-feedback.[7] It is also conceivable that higher than second-order effects may have to be considered, for the expected or estimated reaction to a forecast could further modify the ultimate outcome. Crop estimations and stock market forecasts have early been recognized as giving rise to such higher-order effects.

In microeconomic situations, second-order forecasts in the above sense are frequently encountered.[8] On macro levels they will be encountered when the variable in question is a subject of active and deliberate policy. Much manpower forecasting, therefore, should be treated as of a higher order, especially when the adaptive period, e.g., the period for training skilled and specialized manpower, is a long one. Unless higher order considerations are incorporated in the forecasting procedure, first-order predictions by themselves may actually frustrate the intended policy; they may, for instance, result in more violent fluctuations and distortions than in the absence of prediction.

Logically related to this issue is so-called "teleological" forecasting.[9] Such forecasting relates not to what is likely to happen in the future but to targets that should be achieved. Teleological macro prediction poses a problem that is pertinent to the United States, as well as to centrally planned economies and the so-called developing countries. In the evaluation of structural unemployment, for example, it is useful to predict future skill needs to meet the "target" of full employment. On the micro level, teleological forecasting is encountered, for example, in the application of linear programming techniques.

If manpower forecasting is used as a diagnostic device, and if diagnosis and therapeutics are responsibilities of the same person or or-

---

In more recent sociological literature one finds the term "self-fulfilling prediction" to describe a similar effect.

[8] Theil, *op. cit.*, proposes the term "anticipations" if the forecaster has no direct control over what he predicts, and the term "plan" if the forecast variable "is directly controlled by the forecaster himself."

[9] See R. K. Srivastava *ecting Manpower Demand* (Government of India, Directorate of Manpower, Ministry of Home Affairs, 1964).

ganization, the anticipated effect of therapeutics cannot be ignored. Some of the implications are only gradually being recognized, and we are far from finding satisfactory solutions.[10] In the manpower field, as in others, the problem is complicated by our limited ability to influence the future. Even where government policy may be expected to alter effectively the results based on a "neutral" forecast, only a part of what is predicted may be subject to corrective therapeutic action. Forecasting in such instances would have to proceed along two lines, distinguishing what is amenable to control via feedback from what is autonomous and uncontrollable, yet not necessarily random.

Manpower forecasting schemes typically assume the relative stability of the underlying socioeconomic structure. For the usual prediction horizon, this premise does not seem unreasonable. However, the expanding scope of manpower policies and the momentum of technological change compress the scale of process time as compared with historical time. What once was assumed fixed will itself have to be forecast. Thus, we shall have ensuing need to forecast the changing future shape of the skill and occupational structure—on both the "supply" side (as affected by education and training) and the "demand" side (as affected by alternatives of skill requirements).[11]

Formal prediction usually avoids the difficult issue of interstructural versus intrastructural change. No structural change is the usually stipulated condition, or such change is relegated to the exogenous part of the model. Yet the manpower forecaster must face the issue in one way or another. In the physical and life sciences, it is possible on occasion to modelize extensive structural change,[12] but much too little is known about socioeconomic subject matter to procede in comfort along "objective" lines.

We come now to the much vaguer and highly subjective areas of conjecture and guess, if not augury.

[10] See, for instance, Bertrand de Jouvenel, *The Art of Conjecture* (New York: Basic Books, Inc., Publishers, 1967); also H. O. A. Wold, in *Study Week on the Econometric Approach to Development Planning* (Amsterdam: North-Holland Publishing Co., 1965).

[11] For projections of scientific and engineering manpower aspects in this country, where the problem of anticipating technological change is fundamental, see the Bureau of Labor Statistics study, *The Long-Range Demand for Scientists and Technical Personnel*, NSF 61-65; D. M. Black and G. J. Stigler, *The Demand and Supply of Scientific Personnel* (New York: National Bureau of Economic Research, 1957). For the importance of technological factors in estimating future occupational composition in general, see Bureau of Labor Statistics, *Special Labor Force Report No. 28* (1963); and R. A. Gordon, ed., *Long-Term Manpower Projections* (Berkeley: Institute of Industrial Relations, University of California, 1965).

[12] For an imaginative and not-too-technical exposition in a particular instance, see D'Arcy Wentworth Thompson, *On Growth and Form* (2nd ed.; Cambridge, England: Cambridge University Press, 1959).

During the last few years efforts have been made to sharpen our insights into the future:[13]

1. The most broadly gauged approach is that of the French organization, the Futuribles, under the direction of Bertrand de Jouvenel.[14]
2. In the United States the Department of Defense and the Rand Corporation[15] have sought to anticipate technological developments; the latter organization has employed the [illegible] Delphi method for establishing a concensus of expert opinion.[16]
3. Industry has followed these develop[illegible]perimenting with additional forecasting techniques [illegible] major alteration of the technological structure.[16]

These attacks on a most significant problem [illegible] impressive, not because of their methodological results but [illegible] their courage in facing a very difficult and important problem. It seems [illegible] to conclude that these endeavors have as yet not extended beyond the range of the structural certainties, beyond the range in which "our future certainties are features inherent in an order in which we have confidence".[17]

The preplexing difficulties of the task of solving what appears a [illegible]itable riddle of the sphynx reach deep into the epistemological foundations of economics and the behavioral sciences. Hence it would seem that the more conventional manpower forecasting methods, sometimes termed naive, should be improved rather than discarded; they are all we really have. Bold attempts, however, should be encouraged. They will, if nothing else, generate new hypotheses that require confrontation and testing.

---

[13] Among individuals see, e.g., J. von Neumann "Can We Survive Technology?" *Fortune*, LI (1955); Harrison Brown, *et al.*, *The Next Hundred Years* (New York: The Viking Press, 1958); T. H. Gordon, *The Future* (New York: St. Martin's Press, Inc., 1965); and I. J. Good, ed., *The Scientist Speculates* (New York: Basic Books, Inc., Publishers, 1962).

[14] See the series Futuribles of the bulletin *Sedeis* (Paris).

[15] Army Research Office, *Long-Range Technological Forecast* (Washington: 1963); and RAND Corporation, publications P-3063 and P-2986 (Santa Monica). For a broadgauged synopsis see Erich Jantsch, *Technological Forecasting in Perspective* (Paris: Organization for Economic Co-operation and Development, 1967).

[16] In addition to the just-cited RAND papers, see also O. Helmer, *Social Technology* (New York: Basic Books, Inc., Publishers, 1966), especially his contribution jointly with T. H. Gordon discussing the specific application of the Delphi method. Among the industries are General Electric Corporation, Westinghouse Electric Corporation, TRW Incorporated (whose forecasting technique, Probe, is reminiscent of the Delphi method) and the Hudson Institute (whose concern is in the general area of international economic relations).

[17] de Jouvenel, *op. cit.*