A GRAPHIC TECHNIQUE FOR PROJECTING FAMILY INCOME SIZE DISTRIBUTION

Mitsuo Ono*
U.S. Bureau of the Census

Introduction

Projections of family income distribution are used widely not only by businessmen in developing marketing plans but also by government officials in formulating public programs, e.g., urban development studies. These projections can be calculated in a variety of ways, depending upon the availability of resources, e.g., time, money, skilled manpower, equipment. However, all of these projections, whether simple or complex, have one thing in common: the adequacy of the projections depends ultimately on the validity of the assumptions used to make the projections, and usually these assumptions are no more than calculated historical guesses. Thus, under certain circumstances, the use of relatively simple estimation methods may yield results comparable to that obtained from use of more complicated methods, e.g., regression equations, because use of both of these methods encompasses the same growth factors.

The primary purpose of this paper is to describe a relatively simple graphical technique using logarithmic normal probability (lognormal) graph paper by which projections of family income size distribution can be developed. The discussion is divided into four parts: (1) An outline of the graphical method; (2) an outline of an arithmetical method; (3) comparison of results obtained from using other estimation procedures; and (4) an example of the use of the graphical method.

The Graphical Method

The technique involves essentially the following: (1) Plot historical data to determine the relative "stability" of the cumulative income distribution; (2) if the distribution "shape" appears to be similar historically, obtain the median family income value for the most current distribution available; (3) compute a projected median value, e.g., using historical growth rates of median family income; (4) arithmetically extrapolate the benchmark year income levels to projected year income levels (at given cumulative percentage levels); and (5) plot "new" distribution on graph paper and obtain required information. A specific example of this procedure is outlined below. Lognormal graph paper is used because it tends to reduce the skewness of the income distribution curve. That is, some skewed distributions are transformed into normal probability distributions by taking the logarithm of the variates. If the distribution is lognormal, the ogive generally forms a straight line when it is plotted on lognormal graph paper (in which the ordinates are scaled logarithmically and the abscissa scaled in accordance with the normal distribution function). The slope of the line is related to the standard deviation of the logarithm of the variate. In a lognormal distribution, it is assumed that the magnitude of a variate changes at a rate proportional to the value of the previous variate. (See publication entitled The Lognormal Distribution by J. Aitchison and J.A.C. Brown, Cambridge University, 1957, for further details.) There are other types of graph papers, e.g., arithmetical, log-log, probability, and semi-log, which can be used just as well for the graphic method but it appears that the lognormal graph paper is more convenient to work with because the linearity of the plotted distributions generally extends over a wider range on this type of graph paper than others.

Assume a need exists for a projected (1985) family income distribution expressed in 1967 dollars. First, we plot the cumulative family income distribution (in 1967 dollars) for 1952, 1957, 1962, and 1967 to examine the relative stability of the income distributions over the past 15-year period. We find that the "shapes" of the income distributions are fairly constant (see attachment 1). We then compute the rate of increase of median family income between 1957 and 1967. We then assume that the growth rate of median family incomes between 1957 and 1967 also applies for the period between 1967 and 1985 (in constant dollars). Thus: the 1957-1967 average annual rate comes up to about 3.1 percent per annum (for 10 years). We then apply this growth rate to 18 years, resulting in a total (compounded) increase of 1.73 times the median income in 1967, i.e.,

$$\frac{\text{1967 Median income \$7,974}}{\text{1957 Median income \$5,889}} = 1.354 \text{ or about 3.1\% compounded annually}$$

$$(1.031)^{18} = 1.732$$

* The views expressed here are not necessarily those of the Bureau of the Census.

The next step is to multiply this net increase against the 1967 income levels to obtain the 1985 income levels (in 1967 dollars). Thus:

1967 Cumulative Percentage of Families	1967 Income Intervals	1985 Income Intervals
2.1	Under $1,000	Under $1,730 (1,000 X 1.73)
6.5	Under $2,000	Under $3,460 (2,000 X 1.73)
12.5	Under $3,000	Under $5,190 (3,000 X 1.73)
25.3	Under $5,000	Under $8,650 (5,000 X 1.73)
41.4	Under $7,000	Under $12,110 (7,000 X 1.73)
50.3	Under $8,000	Under $13,840 (8,000 X 1.73)
58.6	Under $9,000	Under $15,570 (9,000 X 1.73)
65.7	Under $10,000	Under $17,300 (10,000 X 1.73)
88.1	Under $15,000	Under $25,950 (15,000 X 1.73)

This procedure, of course, assumes that the income distribution curve has shifted over time with insignificant changes in the slope or "inequality" of the relative distribution. We plot on the lognormal graph paper the projected 1985 income levels (at the 1967 cumulative percentage points) for various income levels (see attachment 1). If we want a more refined fit, we would use more detailed income intervals which would, of course, take into account more of the "curvature" of the cumulative income distribution. The final step is to connect the points graphically and read off the required information. Thus using this method, it is estimated that in 1985 about 9 percent of all family units would have incomes of less than $4,000 (in 1967 dollars). That is, the $4,000 line cuts the projected 1985 income distribution at about the 9 percent point (found at the bottom of attachment 1).

The Arithmetical Method

The same results obtained graphically can be also calculated as shown below. The first step is to calculate the total net increase in median family incomes between the benchmark year and the projected year. We found this increase to be 1.73 between 1967 and 1985 (in constant 1967 dollars).

COMPUTATIONS FOR ARITHMETICAL METHOD

1967 Income levels	1985 Income levels (1967 values X 1.73)	Cumulative percentage of families in 1967	1985 Income levels	1985 Cumulative percentage of families using arithmetical method	1985 Cumulative percentage of families using graphic method
Under $1,000	$1,730	2.1	Under $1,000	$\frac{\$1,000}{\$1,730}$ (2.1)=(.58)(2.1)=1.2	NA
Under $2,000	$3,460	6.5	Under $2,000	$\frac{\$2,000}{\$1,730}$ (2.1)=(1.16)(2.1)=2.4	2.8
Under $3,000	$5,190	12.5	Under $3,000	$\frac{\$3,000}{\$3,460}$ (6.5)=(.87)(6.5)=5.7	5.8
Under $5,000	$8,650	25.3	Under $5,000	$\frac{\$5,000}{\$5,190}$ (12.5)=(.96)(12.5)=12.0	12.0
Under $7,000	$12,110	41.4	Under $7,000	$\frac{\$7,000}{\$8,650}$ (25.3)=(.81)(25.3)=20.5	19.0
Under $8,000	$13,840	50.3	Under $8,000	$\frac{\$8,000}{\$8,650}$ (25.3)=(.92)(25.3)-23.3	23.0
Under $9,000	$15,570	58.6	Under $9,000	$\frac{\$9,000}{\$8,650}$ (25.3)=(1.04)(25.3)=26.3	26.0
Under $10,000	$17,300	65.7	Under $10,000	$\frac{\$10,000}{\$8,650}$ (25.3)=(1.16)(25.3)=29.3	30.0
Under $15,000	$25,950	88.1	Under $15,000	$\frac{\$15,000}{\$15,570}$ (58.6)=(.96)(58.6)=56.3	56.0

NA Not applicable.

The second step is to multiply this rate against the income levels. Thus, we multiply $1,000 by 1.73 to obtain $1,730, etc., as we had done before in the graphic method. The third step is to estimate the cumulative percentage levels in 1985 for the original income levels in 1967. We can also use either extrapolations or interpolations in order to take account of the "curvature" of the income distribution.1/ This adjustment is needed if the income intervals that are being used in the calculation are large. The following table shows the computational method and a comparison of the cumulative percentage rate obtained from both the arithmetical method and the graphic method.

In the computations, six income levels, e.g., the "$5,000 and under" level in 1985 were interpolated while three income levels were extrapolated, e.g., the "$9,000 and under" level in 1985. Overall, the results obtained from using the graphic method and the arithmetical method are similar and either method can be used depending on one's preference.

A Comparison of Projections

Projections of families receiving under $3,000 income were developed for 1975 (in constant 1965 dollars) using two methods:

1. A regression equation to estimate the change in median family income from 1965 to 1975. Thus,

$$Y_F = \underset{(32.62)}{-1,172} + \underset{(0.008)}{0.749}\,(X_{PF}) \quad N = 16$$

where Y_F = Median family income

X_{PF} = Personal income per family

Personal income in 1975 was estimated from GNP projections described in Joint Committee Print, U.S. Economic Growth to 1975: Potential and Problems, 89th Congress, Second Session, U.S. Government Printing Office, Washington, 1966. GNP projections and implicit GNP price indices are shown on page 16 of this reference. Model B (assumes a 4.5 percent annual rate of growth for real GNP, 1966-1975) was used. This model assumed a 5.5 percent average personal savings rate for the period 1961-1965 and a 4 percent unemployment rate (see page 9 of cited reference). The 1975 GNP (in 1965 dollars) was converted to estimated Personal Income by using the PI/GNP ratio shown in current dollars in the cited reference. The projected PI value was divided by the projected number of families to obtain PI per family in 1975. The total increase in median family income using this regression procedure was 1.37 times the median family income in 1965. Using the graphic method, the percentage of families receiving under $3,000 (in 1965 dollars) was estimated at about 10.0 percent in 1975.

2. The average annual rate of increase in median family income between 1955 and 1965 (in 1965 dollars) was calculated to be about 2.9 percent or a total increase in median family income of 1.33 over a 10-year period. This procedure resulted in an estimate of about 11 percent of families receiving under $3,000 (in 1965 dollars). It is interesting to note that these percentage rates are fairly comparable to those derived from using regression methods. For 1975, the following percentage of families receiving under $3,000 was computed:

Reference source	Assumption	Percentage of all families under $3,000	Basic relationship
"Poverty by Color and Residence--Projections to 1975 and 1980," by J. Patrick Madden, paper presented at American Agricultural Economic Association Meeting, Bozeman, Montana	4% Unemployment rate and 1959-1966 data fitted to log-log equation (1964 constant dollars)	10.7	Percentage of families with $3,000 or less is related to median income and unemployment rate

The point of this discussion is that under certain assumptions, relatively simple extrapolation methods may be used to derive approximations which may be almost similar to results derived from use of more complicated methods.

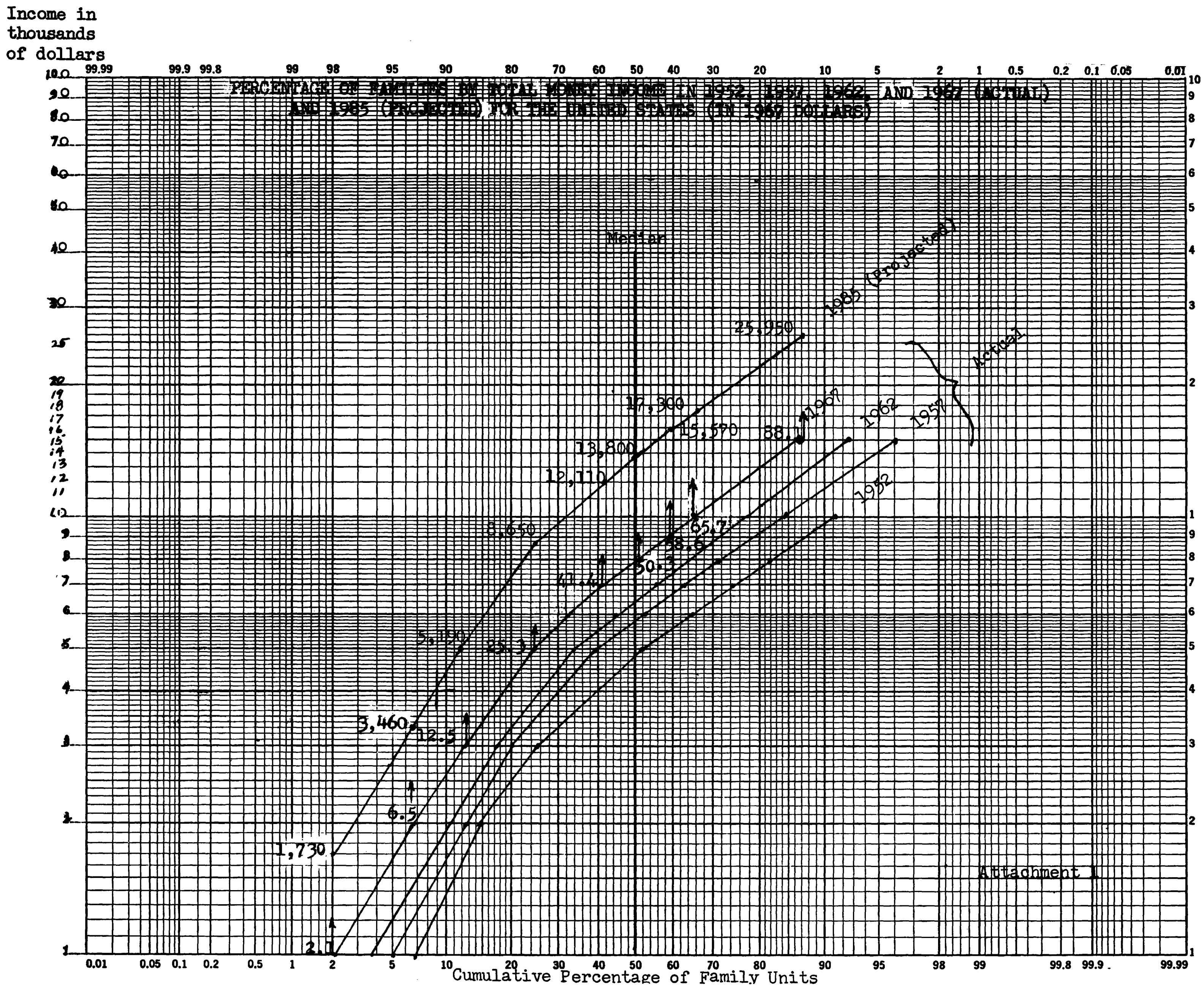
An Example of Use of the Graphical Method

Advantages of the graphical method are the simplicity and flexibility introduced in obtaining the desired information. Thus, we may find that historically the "curvature" of the income distribution has changed over time. Instead of taking a constant percentage change for each income level value, it is possible to apply differential rates of increase to different income levels which would take into consideration the change in relative "inequality" of the distribution over time. Instead of keeping the "slope" constant, changes in both the "slope" and the median income level are taken into account. The "slope" related to the standard deviation in turn is related to measures of inequality, e.g., the Gibrat inequality measure.

As a specific example of this technique, we may relate historical changes in summary measures of inequality to percentage changes in income levels, e.g., quintile values. In turn, these changes in values can be related to changes in socioeconomic characteristics of that group. In this way, relationships are obtained for quintile groups, associating changes in summary measures of inequality with changes in social and economic characteristics of different income groupings.

FOOTNOTE

1/ This comment, attributed to Mr. Albert Mindlin, Chief Statistician, Government of the District of Columbia, is gratefully acknowledged.


PERCENTAGE OF FAMILIES BY TOTAL MONEY INCOME IN 1952, 1957, 1962, AND 1967 (ACTUAL)
AND 1985 (PROJECTED) FOR THE UNITED STATES (IN 1967 DOLLARS)
Income in thousands of dollars
Cumulative Percentage of Family Units
Median
1985 (Projected)
Actual
1967
1962
1957
1952
25,950
17,300
15,570
13,800
12,110
8,650
5,190
3,460
1,730
88.1
65.7
58.6
50.3
41.4
25.3
12.5
6.5
2.1
Attachment 1