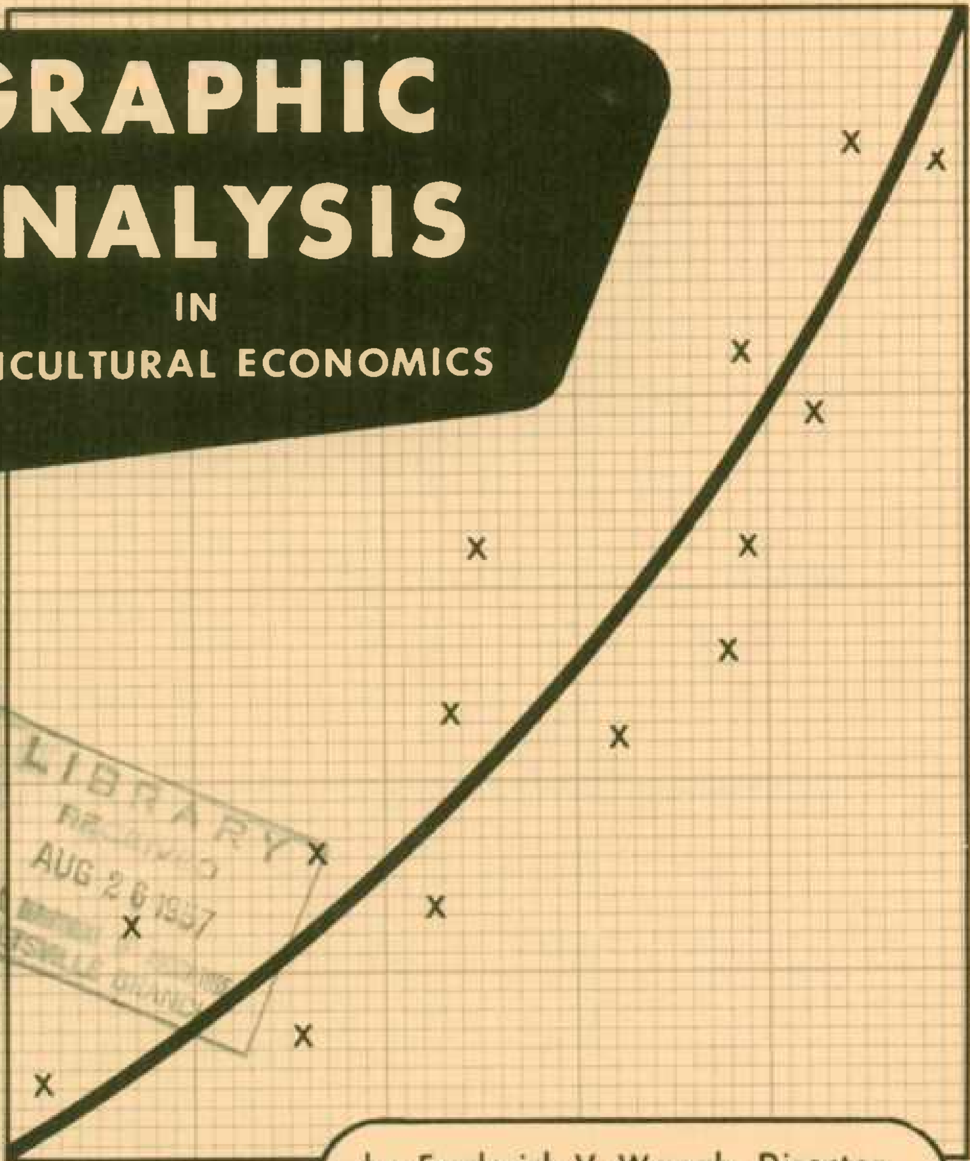
GRAPHIC ANALYSIS IN AGRICULTURAL ECONOMICS



by Frederick V. Waugh, Director
Division of Agricultural Economics

Agriculture Handbook No. 128

UNITED STATES DEPARTMENT OF AGRICULTURE
Agricultural Marketing Service
Washington, D. C.

CONTENTS

Issued July 1957

This publication supersedes Graphic Analysis in Economic Research, Agriculture Handbook No. 84.

For sale by the Superintendent of Documents, U. S. Government Printing Office, Washington 25, D. C. - Price 40 cents

GRAPHIC ANALYSIS IN AGRICULTURAL ECONOMICS

By Frederick V. Waugh, Director
Division of Agricultural Economics,
Agricultural Marketing Service

INTRODUCTION

Since Handbook Number 84, "Graphic Analysis in Economic Research," was issued in June 1955, I have received many suggestions for improvement. Also, I have been pleased to see that a large number of agricultural economists (and also many non-agricultural economists) feel that a handbook of this kind meets a real need.

When our supply of the original handbook became exhausted, we decided to rework it and get out a new handbook--the present volume--rather than simply republishing the old one. I wish I could thank here everyone who has made a useful suggestion. But the list would be too long. I would like especially to thank the staffs of the Departments of Agricultural Economics and Rural Sociology of the land grant colleges. In answer to my request, practically all of these departments sent useful suggestions, many of which have been used. Some of the principal suggestions are acknowledged at various places in this handbook. I am also happy to acknowledge the important help of three members of the Division of Agricultural Economics. They are Richard J. Foote, Hyman Weingarten, and James R. Donald.

Graphic analysis has a long and honorable history in agricultural economics research. Back in the 1920's, it was perhaps the principal research tool used by the former United States Bureau of Agricultural Economics and by the land grant colleges. In recent years less attention has been given to graphics. Agricultural economists and statisticians have become intrigued with new mathematical methods and with improved calculating machines. These are powerful tools and definitely have a place in economic research. Nevertheless, my own view is that graphic analysis is an indispensable tool which should be used right along with the newer and fancier gadgets. In my opinion, there is an unfortunate tendency, especially among so-called "econometricians," to be satisfied with a purely mechanical analysis. Good research is not simply a matter of recording various statistical series on tape, feeding them into an electronic computer, and getting back a lot of numbers computed to 6 or 8 "significant" figures. First, a capable economist must understand the data he uses. He often must work with estimates which are significant to only 2 or 3 digits. Second, he must have a thorough and basic understanding of the nature of the relationships between the variables he is considering. To do this, he must understand economic theory and he must also be able to see the nature of the relationship shown empirically. In many cases the economist cannot assume linear relationships, for example.

As I see it, the greatest value of graphics in economic research is in making a quick, preliminary analysis of a problem to determine which variables to use and the general nature of the relationships. For many practical purposes, the graphic analysis alone is fully satisfactory. In other cases, the economist will want more precise measures. In these cases, he will want to fit some sort of mathematical function to the observations. I believe that graphics is an indispensable tool for choosing the sort of function to fit. I believe that the neglect of graphics has frequently led economists and statisticians to choose inappropriate kinds of functions.

To be sure, graphic analysis has sometimes been misused. So has any kind of statistical analysis that can be named. Whatever tools are used, there is no substitute for sound judgment and common sense. Without this, the economist is going to get into trouble anyway. If he has reasonably good judgment, I believe that his best approach to economic research will be a combination: Using graphics in the preliminary analysis of a problem, then more elaborate mathematical methods to pin down results with greater precision.

FREQUENCY DISTRIBUTION

Number of Dealers Reporting Various Prices Paid by Farmers for Laying Mash, September 1949

The economist often deals with averages. For example, he may be analyzing the average price received by farmers for wheat or the average price paid by farmers for some item used in production or in farm family living.

He must remember that averages often cover up important information. To understand the meaning of the data he uses, the economist needs always to have some understanding about the degree of variation around the average. In some simple cases he knows this in a general way by observation. If he were told that the average height of men in a large group was 5 feet 10 inches, he would not expect many of them to be less than 5 feet or more than 7. But many economists work with data obtained from various sources. They often know little about the kind and amount of dispersion to expect. Probably economists and agricultural statisticians ought to do more work on this subject. And they should publish their findings so others could judge the reliability of averages and the variation to expect around them.

B. Ralph Stauber of the Agricultural Estimates Division, Agricultural Marketing Service, supplied the data used to draw the accompanying diagram. 1/ It is a so-called "frequency distribution" of prices paid by farmers throughout the United States for laying mash in September, 1949. The average price is a little more than $4.50 a hundred pounds. The range is from $3.40 to about $6. This range is, of course, due to many things, including geographical differences that reflect freight rates and differences in the ingredients used in the mash.

Each of the bars in this diagram shows the number of dealers reporting prices paid by farmers within the several ranges shown in the table beneath the diagram. I have drawn a smooth curve representing a judgment as to the general nature of the observed distribution. Note that I have not bent the curve to make it go through the midpoint of each bar. Rather, I have drawn it smooth, to show the general nature of the distribution.

This distribution appears to be fairly near what is commonly called a "normal distribution." When this is true, the arithmetic average, the median, and the mode all come at about the center of the distribution of prices. This is somewhat of an accident. We shall see in the next diagram a case in which the curve is far from normal in the technical statistical sense.

1/ Unless otherwise specified, diagrams referred to in the text are those on the facing page, the data for which are given in the table beneath the chart.

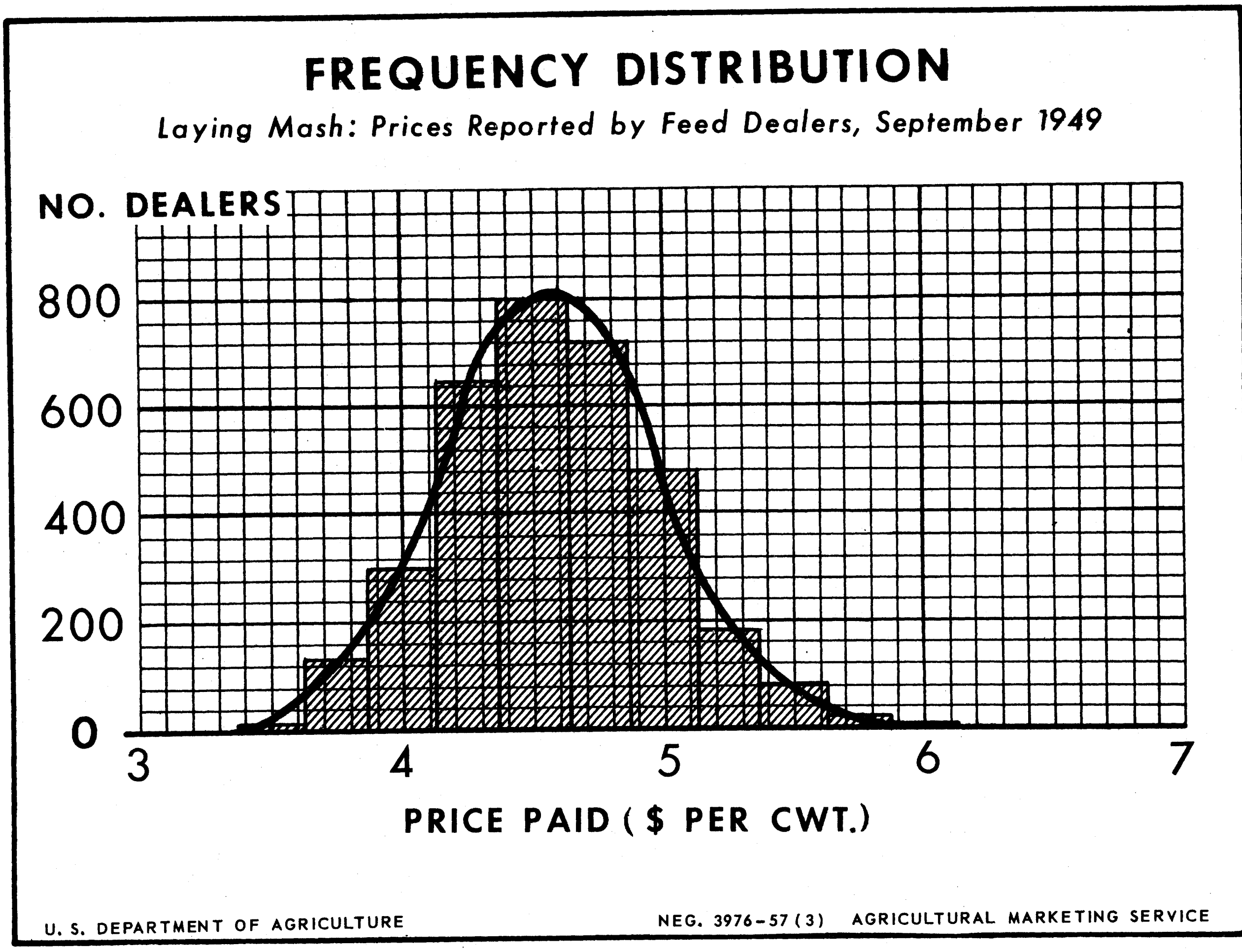


Figure 1

Laying mash: Frequency distribution of prices paid by farmers per hundredweight as reported by feed dealers, September 1949

Price	Dealers reporting
Dollars	Number
3.375 - 3.624	12
3.625 - 3.874	136
3.875 - 4.124	302
4.125 - 4.374	652
4.375 - 4.624	808
4.625 - 4.874	715
4.875 - 5.124	486
5.125 - 5.374	183
5.375 - 5.624	82
5.625 - 5.874	16
5.875 - 6.124	3
Total	3,395

Data supplied by B. R. Stauber, Agricultural Estimates Division, Agricultural Marketing Service.

Corn Acreages in Sample Area Segments Enumerated in 12 Southern States, June 1956

Actually, the so-called "normal curve" is a rather unusual phenomenon in economic research, although it may frequently apply fairly well when we consider residuals unexplained by a statistical analysis. The agricultural economist often works with initial data which are skewed. He may need to know something about the nature and degree of such skewness. One of the best ways to find out is to plot the frequency distribution.

The facing chart is based upon data supplied by Walter Hendricks of the Agricultural Estimates Division, Agricultural Marketing Service. It summarizes the results of a survey of corn acreage made in 12 southern States in 1956. The survey covered 623 segments of a sample area. The figures were tabulated to show how many of these segments reported corn acreage of 0 to 19, how many reported from 20 to 39, from 40 to 59, and so on. The figures are shown in the table below the diagram.

As in the preceding chart, the height of the bars indicates the number of segments reporting corn acreages within the indicated ranges. I again have drawn a smooth curve, attempting to describe the general nature of this distribution. The curve is intended to run approximately through the mid-point of the top of each bar. This smooth curve is an estimate of the actual distribution of corn acreages in the southern States. The variations may be due to errors in sampling and in reporting.

In this case, the average reported acreage of corn was about 49. However, the curve is so badly skewed that 49 acres is far from either the median or the mode. The largest number of segments reported less than 20 acres of corn; many reported no corn acreage at all.

This is an extreme case of skewness, but the agricultural economist often must use data that are noticeably skewed. He should be aware of the kind of distribution he is dealing with. He can inform himself of this in a few minutes by the sort of analysis we have shown.

The shaded area represented by the bars in the diagram is often called a histogram. The histogram is a summary of the observed facts in the sample. The smooth curve is an estimate of the distribution in the statistical universe.

For some purposes the statistician may want to fit some form of mathematical curve to the data in the histogram. However, it is a good idea to draw the histogram and graphic curve first, before deciding what sort of mathematical curve to try. Don't try to fit a normal curve or a Poisson curve to any old data. Look at them first.

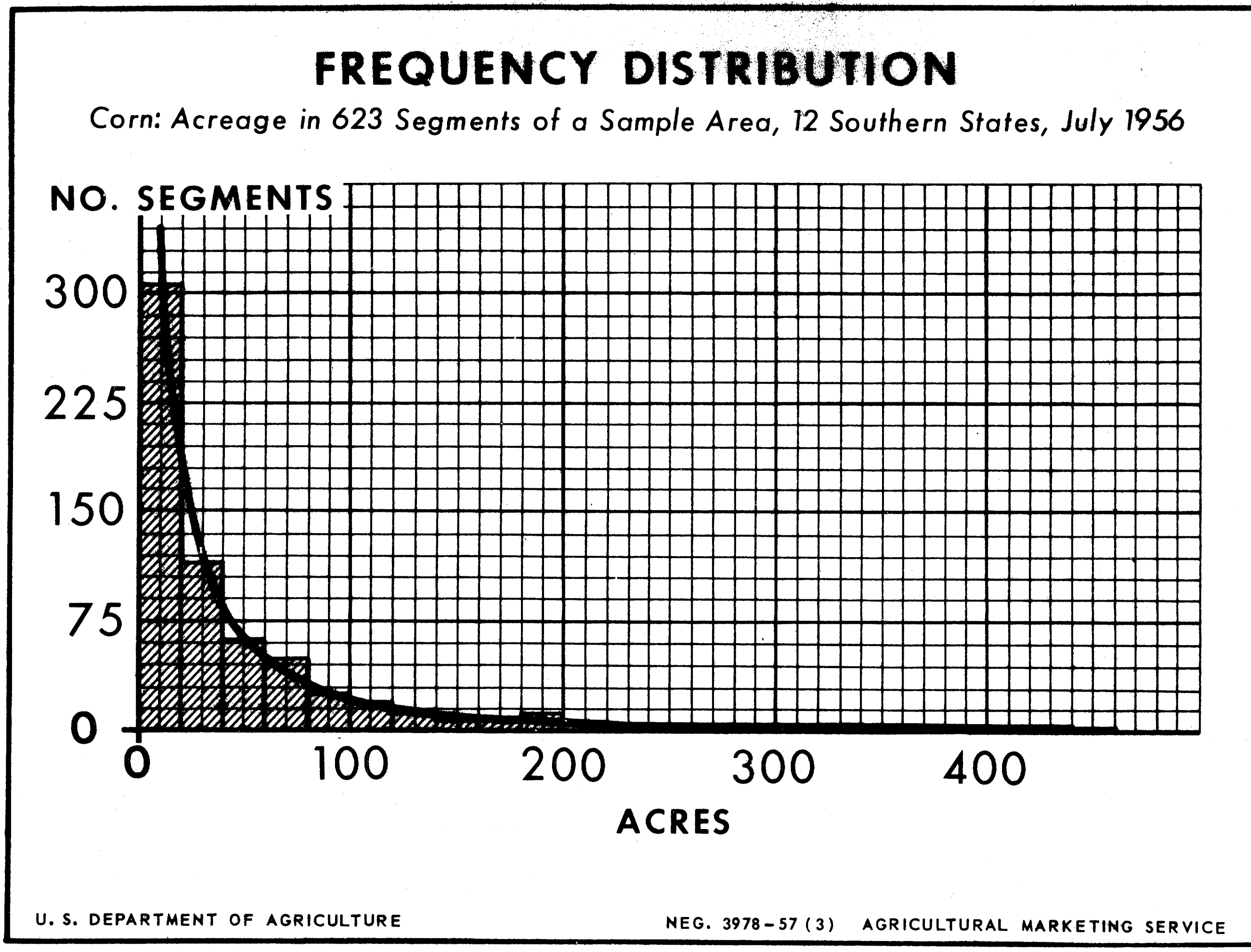


Figure 2

All-corn acreage: Frequency distribution in sample area segments enumerated, 12 southern States, June 1956

Acreage	Segments	Acreage	Segments	Acreage	Segments
Acres	Number	Acres	Number	Acres	Number
0-19	308	120-139 ...	10	240-259	1
20-39	115	140-159 ...	6	260-279	2
40-59	61	160-179 ...	8	280-299	1
60-79	47	180-199 ...	9	300-319	1
80-99	25	200-219 ...	5	320-339	3
100-119 ...	19	220-239 ...	1	340 and over .	1/1
				Total	623

1/ 456 acres.

Data supplied by Walter Hendricks, Agricultural Estimates Division, Agricultural Marketing Service.

Percentages of Families With Incomes Below Stated Levels

Economists and statisticians are concerned with many kinds of frequency distributions. The particular distribution shown on the diagram refers to percentages of families with various incomes. In this case we have shown a cumulative frequency curve. Thus, instead of showing the percentage of families with incomes from 0 to $1,000, from $1,001 to $2,000, and so on, we show the percentage with incomes below $1,000, below $2,000, and so on.

The cumulative frequency curve, or ogive, has some advantages over the more usual noncumulative frequency curve. It can be used whatever the "class intervals" may be. For example, in this case, class intervals of $1,000 were used for the part of the curve from 0 to $6,000. For incomes above $6,000, a larger class interval was used. With unequal class intervals, it is awkward to draw and use the ordinary type of frequency chart, and the cumulative chart is preferred.

In this case there was no problem of drawing a freehand curve to fit the cumulative frequencies. The plotted data all lie almost exactly along the freehand line we have drawn.

Several mathematical functions have been proposed and used to describe the distribution of incomes. Some of these, like the Pareto curve, are purely empirical. Others, like the Gibrat curve, are based upon logical considerations. It is obvious that no mathematical curve could fit the data much better than the freehand curve we have drawn. In fact, the freehand curve probably fits the data on the left hand side of the diagram better than would a mathematically fitted Pareto curve.

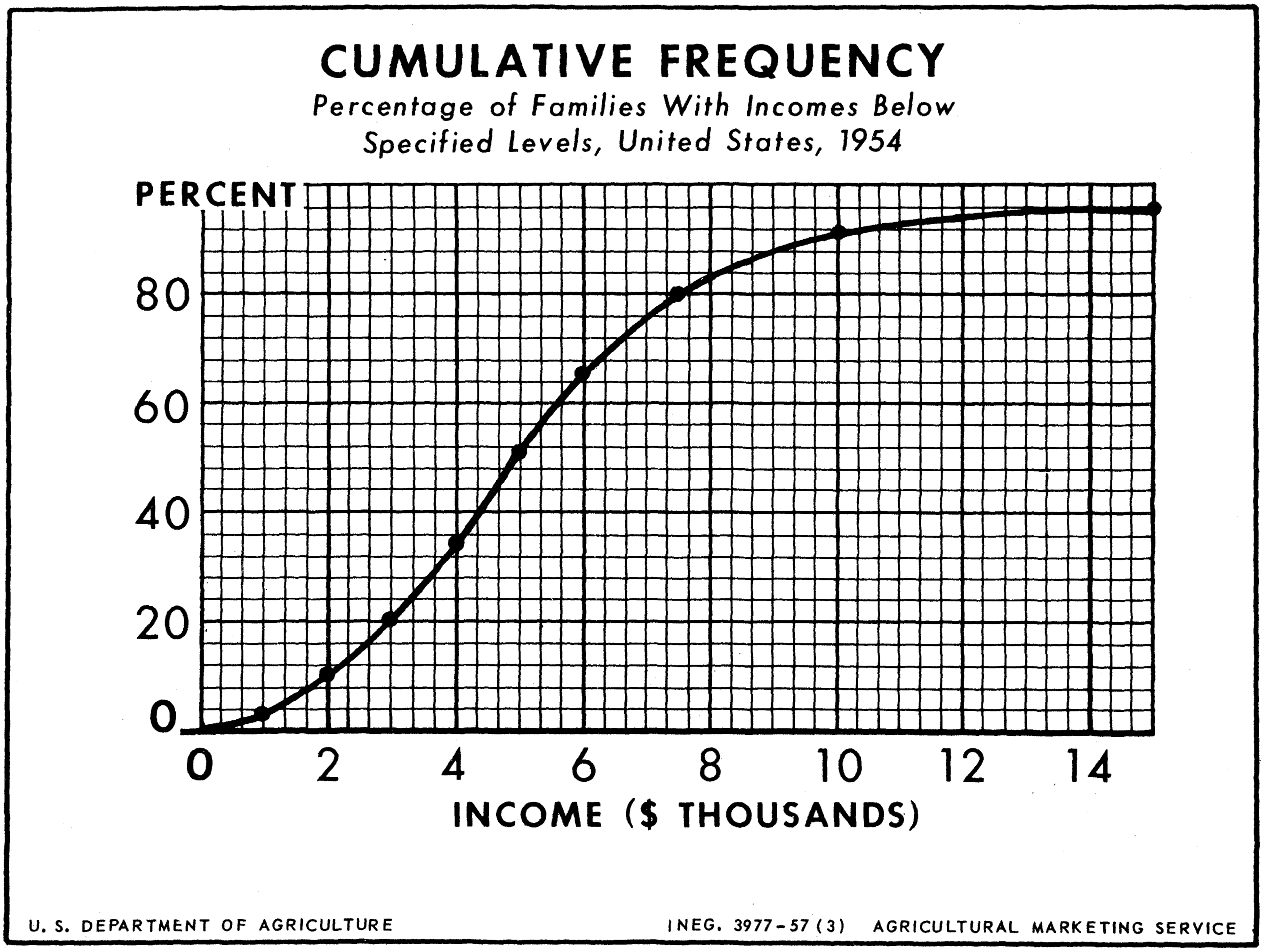


Figure 3

Families: Percentage with personal income below specified levels, United States, 1954

Income level	Distribution
Dollars	Percent
Under:	
1,000	2.7
2,000	9.9
3,000	20.0
4,000	34.2
5,000	50.6
6,000	65.3
7,500	79.9
10,000	91.4
15,000	96.5

Survey of Current Business. U. S. Dept. Commerce. June 1956. p. 15.

Percentage of Families in the United States with Personal Incomes Below Stated Amounts and Percentage of Total Personal Income Obtained by These Families, 1954

Lorenz curves are often used to analyze the distribution of incomes. The chart presented here uses a Lorenz curve for this purpose. Obviously such curves could be used to analyze any kind of distribution, such as those shown in the first two charts in this book.

The distinctive feature of the Lorenz curve is that the data on both the x-axis and the y-axis are plotted as percentages of the total. In this case, for example, the U. S. Department of Commerce figures indicate that in 1954 2.7 percent of the families in the United States had personal incomes of less than $1,000. These families obtained 0.2 percent of the total personal income. The table below the chart shows similar percentages for families with incomes below $2,000, below $3,000, and so on. Each pair of percentages is plotted on the chart and indicated by a dot.

The dotted line on the diagram indicates what the distribution would be if all families got the same incomes. The area between the dotted line and the curve is a measure of income inequality. If we should plot a series of such curves over a period of years, changes in this area would indicate whether this inequality is becoming more or less.

Note that the Lorenz curve can be used to plot data that are grouped by any kind of class interval, whether equal or unequal. As in the preceding chart, the first six groups of families are classified into income ranges of $1,000. Above $6,000, the class intervals are wider.

Another distinctive feature of the Lorenz curve is that it can be read up, down, or sidewise. Reading up, for example, we might estimate that the lowest 40 percent of the families received 19 percent of the income. Reading down we would find that the top 10 percent of the families got 30 percent of the income. Reading from left to right, we find that 20 percent of the income was obtained by the lower 41 percent of the families. Reading from right to left we see that 20 percent of the income was obtained by the top 5 percent of the families. These are just a few illustrations of the many uses of this ingenious form of curve.

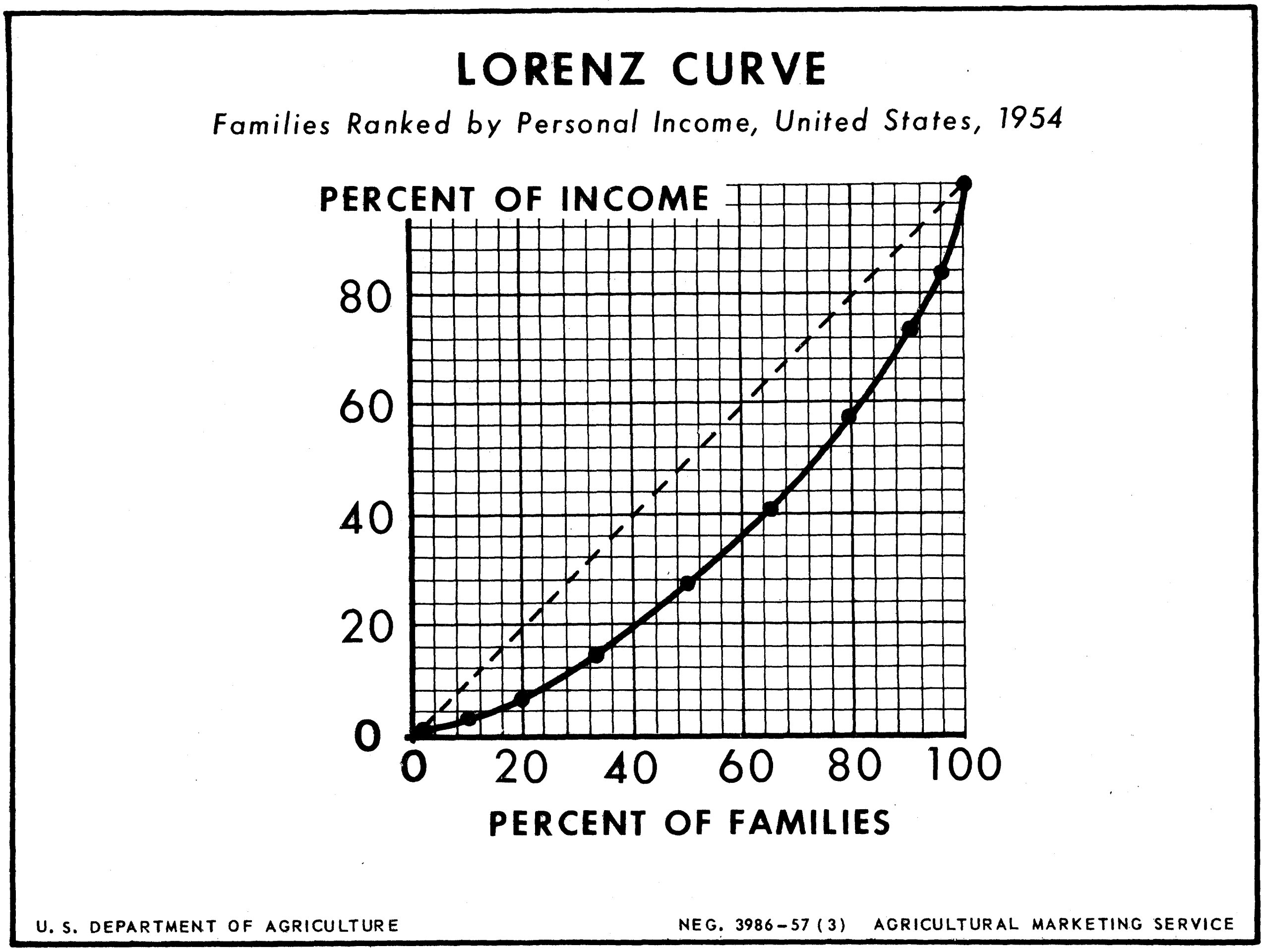


Figure 4

Families: Percentage with personal income below specified amounts and percentage of total personal income obtained by these families, United States, 1954

Incomes less than	Percentage of-- Families	Income
	Percent	Percent
$1,000	2.7	0.2
2,000	9.9	2.1
3,000	20.0	6.4
4,000	34.2	14.8
5,000	50.6	27.2
6,000	65.3	40.7
7,500	79.9	57.0
10,000	91.4	73.3
15,000	96.5	83.7
∞	100.0	100.0

Survey of Current Business. U. S. Dept. Commerce. June 1956. p. 15.

Population Trends in the United States by Decades, 1800-1950

The economist is often concerned with time trends. He wants to find out how some variable has been increasing or decreasing over a period of several years or decades. For example, he may be studying the growth of population in the United States or the rate of decline in the number of farm workers. In such cases he will want to disregard minor fluctuations due to errors in the data or to temporary disturbances. He will also generally want to disregard cycles or other shorter term movements in the data if they exist. He is concerned only with the gradual rate of change in a variable in relation to time.

The chart shows the Bureau of the Census estimates of the population in this country by decades since 1800. It is a simple matter to draw a freehand curve describing the trend. Ordinarily, at least, population does not change abruptly except by major wars, serious epidemics, or a sharp increase in immigration. If we plot the data for each year, or decade, we can usually draw a smooth curve running nearly through the points we have plotted. In this case, departures from the curve could well be due to errors in estimating the population. It should be noted that even official estimates may not warrant the naive faith in their accuracy that sometimes prevails. We, as economists, are probably as much responsible as any other group of users of published data for the insistence upon the publishing of a single number (point estimate) to represent, say, the population of the United States. We are reluctant to accept a lower and upper estimate (interval estimate) of the actual population even though we know that the Bureau of the Census official figure of 150,697,761 persons for 1950 (or that for any other year) may not be exact. All too often we do not even take the trouble to understand what the publisher has to say about the known, or estimated, amount of possible error in his estimates.

Instead of drawing a freehand curve, the statistician could, of course, fit some kind of mathematical function, such as a logistic curve. Our advice would be to draw a freehand curve first. In this case, it is doubtful if any mathematical function would give a better description of the trend than our freehand line. A mathematical curve might have some advantage when comparing trends in population in several different countries. If the same type of function were fitted in each case; results could be summarized in a few statistical measurements.

A practical application of trends is in forecasting. This always involves an extrapolation beyond the range of the data. Extrapolation of trends is dangerous whether it is done from a freehand curve or from a curve that has been fitted mathematically. For example, before the 1950 census data were available (so that we did not have the last observation on the diagram), many population experts drew an S-shaped curve indicating that the rate of growth had started to flatten. When this type of curve was extrapolated it suggested that the population would become stationary, or even decrease, by 1960 or 1970. Such an extrapolation now looks doubtful in view of the census figure for 1950.

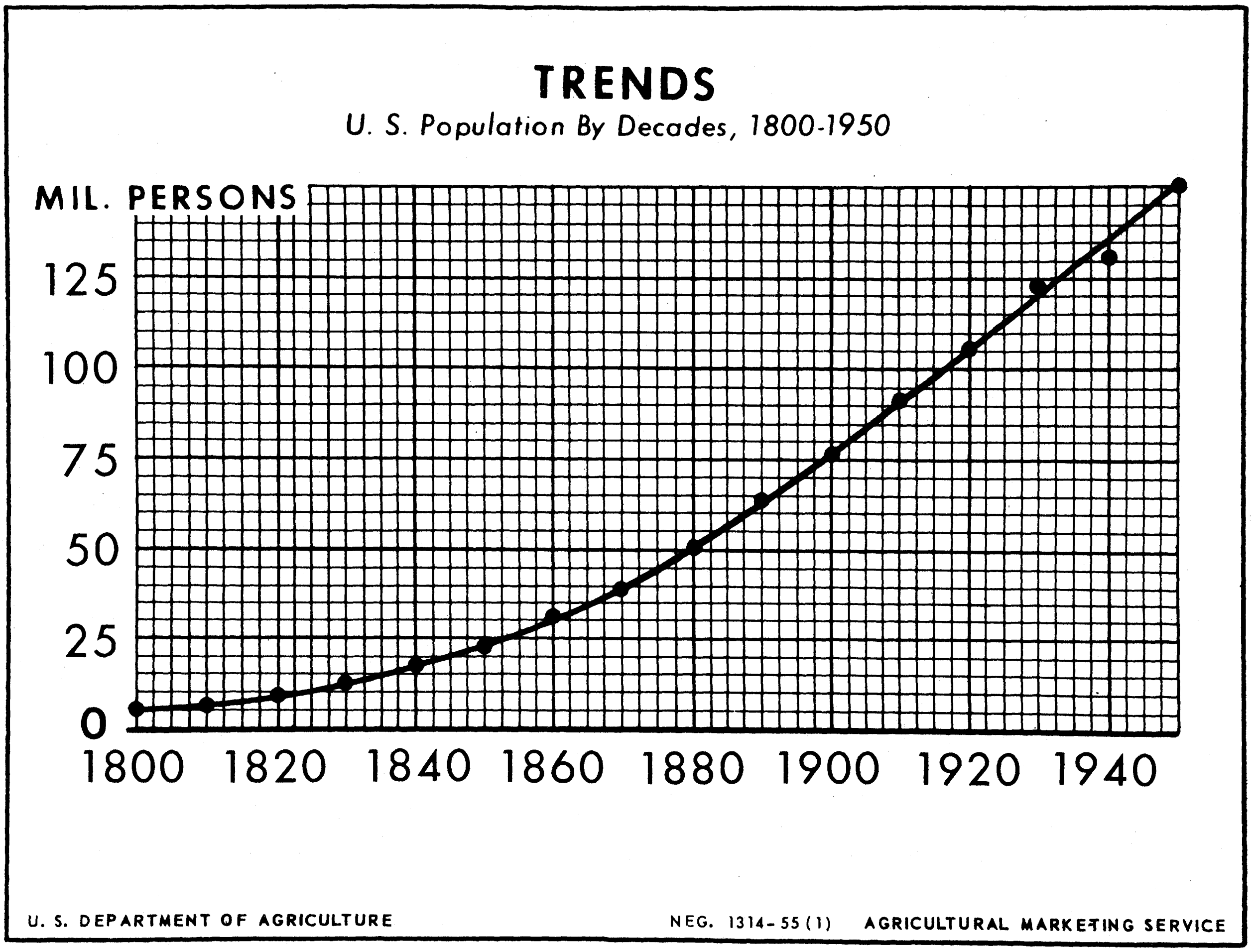


Figure 5

Population: United States, by decades, 1800-1950

Year	Population	Year	Population
	Millions		Millions
1800	5.3	1880	50.2
1810	7.2	1890	62.9
1820	9.6	1900	76.0
1830	12.9	1910	92.0
1840	17.1	1920	105.7
1850	23.2	1930	122.8
1860	31.4	1940	131.7
1870	38.6	1950	150.7

Bureau of the Census.

Volume of Agricultural Marketings

It is easy enough to draw a chart representing the growth of population, because population tends to grow at a steady rate. It may be upset a little sometimes by such things as wars, depressions, and epidemics. But in spite of such factors, estimates of population tend to lie fairly close to a smooth curve. In many cases, however, the agricultural economist must work with data that do not lie along any sort of smooth trend.

A case in point is the index of volume of agricultural marketings. This index is plotted on the accompanying diagram for each year from 1910 through 1956. Quite evidently there has been an upward trend in agricultural marketings. However, this trend has not been steady. For example, the chart shows that from 1925 to 1935 marketings did not increase but fluctuated around an index of about 70.

In a case of this kind the agricultural economist needs to use a good deal of judgment in drawing a trend line. The first edition of this handbook exhibited these same data with a graphic trend. The trend shown in that edition of the handbook was properly criticized by several able economists. I have redrawn it to take account of their criticisms.

In some cases the statistician may want to compute a mathematical trend. Before doing so, he would be well advised to draw a freehand trend to indicate what sort of mathematical function should be used. In this case, for example, a straight line would not adequately describe the trend of the data. A third-degree parabola might give a fair fit but would be an arbitrary sort of trend and an extremely bad one to extrapolate into the future. In general, the economist would do well to avoid a parabolic trend. Logically, they seldom make economic sense.

Of course, one of the main reasons for fitting a trend is to get some idea of the current direction of the series and perhaps its probable direction in the future. However, it is dangerous to extrapolate trends, especially when they exhibit such irregularities as are shown in this chart. In thinking about the future trend of the volume of agricultural marketings, we must also remember that it will be affected by such things as acreage allotments, marketing quotas, and the Soil Bank.

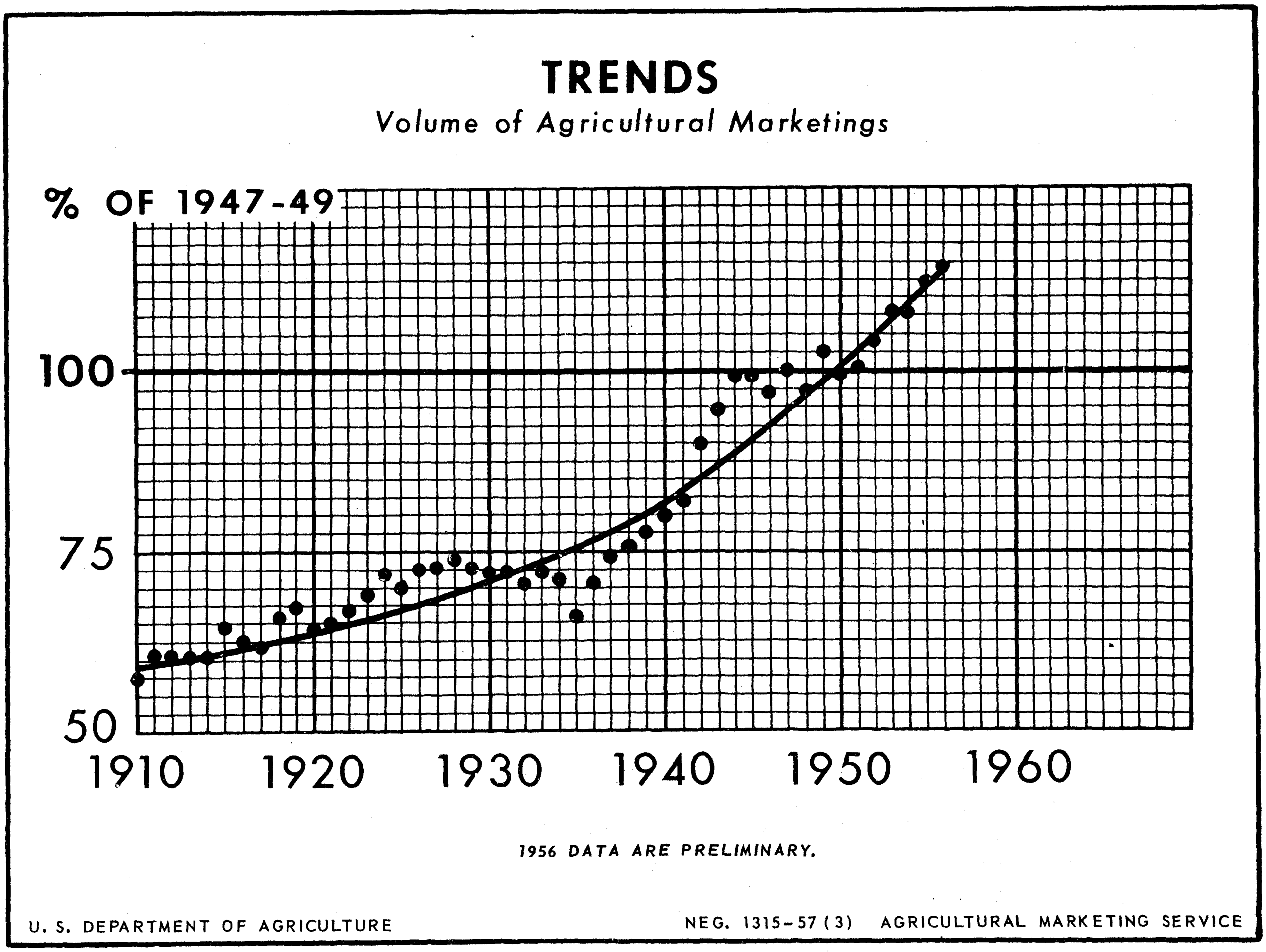


Figure 6

Farm marketings and home consumption: Index numbers of volume, 1910-56

[1947-49=100]

Year	Volume	Year	Volume	Year	Volume
1910	58	1926	73	1942	90
1911	61	1927	73	1943	94
1912	62	1928	74	1944	99
1913	61	1929	74	1945	99
1914	61	1930	72	1946	97
1915	64	1931	73	1947	100
1916	64	1932	71	1948	97
1917	62	1933	72	1949	103
1918	67	1934	71	1950	99
1919	67	1935	66	1951	101
1920	64	1936	71	1952	104
1921	65	1937	74	1953	108
1922	67	1938	76	1954	108
1923	69	1939	79	1955	112
1924	72	1940	80	1956 1/	114
1925	70	1941	82		

1/ Preliminary.

Agricultural Marketing Service.

CYCLES

Cattle on Farms, January 1

Many important economic time series tend to fluctuate more or less regularly up and down around a trend line. Such fluctuations often have important economic implications. For example, the business cycle is of great interest to economists and has been studied in great detail by many competent economic theorists and statisticians.

Cycles are especially important in agriculture. The very nature of agriculture tends to generate cyclical movements. Take cattle, for example. When cattle prices are high, farmers are likely to start breeding for larger herds. It takes several years to increase the herds substantially, and the increase ordinarily continues for some time after prices become unprofitable. Then the reverse happens and herds are gradually decreased.

One of the best ways to forecast the probable behavior of a current cycle from that of previous cycles is to break the total series into individual cycles. In 1956, the most recent cycle in numbers of cattle on farms was beginning a downturn. In the diagram shown here, data for these individual cycles are plotted on the same scale, beginning with the year of the low point in inventories in each instance.

The several cycles of numbers of cattle are remarkably similar. One handicap in this visual scheme is that each cycle is of a different length. Similarity between cycles would appear even closer if the cycles were telescoped into a uniform length.

A good statistician knows that history seldom repeats itself exactly. Cycles vary in length and in amplitude. A knowledge of past trends, and of past cycles, gives some perspective to the present. Often it suggests the general direction of changes in the immediate future. But the wide-awake economist will be looking for factors that may make the current cycle different from the others.

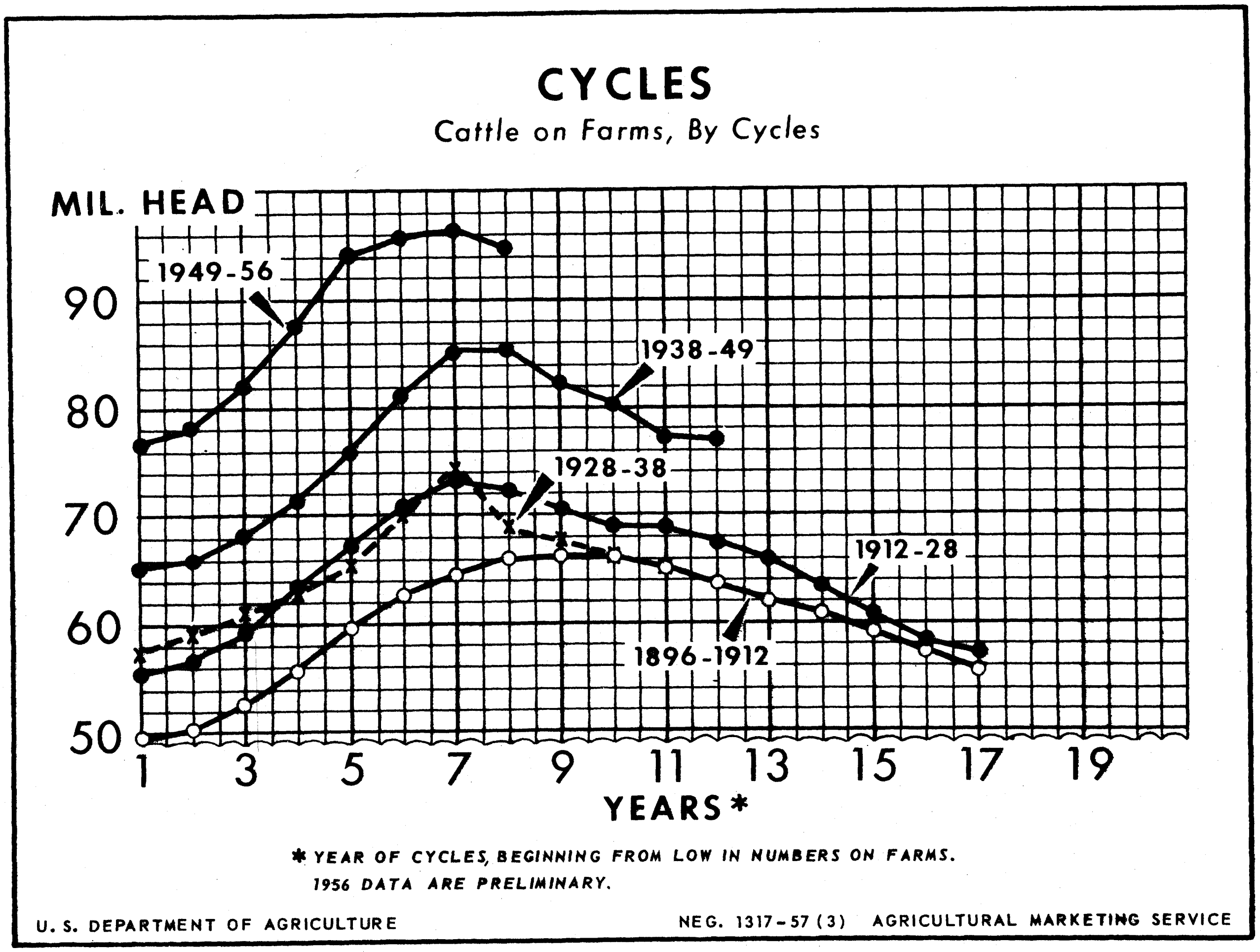


Figure 7

All cattle and calves: Number on farms January 1, 1896-1956

Year	Number	Year	Number	Year	Number	Year	Number
	Millions		Millions		Millions		Millions
1896	49.2	1912	55.7	1928	57.3	1944	85.3
1897	50.4	1913	56.6	1929	58.9	1945	85.6
1898	52.9	1914	59.5	1930	61.0	1946	82.2
1899	55.9	1915	63.8	1931	63.0	1947	80.6
1900	59.7	1916	67.4	1932	65.8	1948	77.2
1901	62.6	1917	71.0	1933	70.3	1949	76.8
1902	64.4	1918	73.0	1934	74.4	1950	78.0
1903	66.0	1919	72.1	1935	68.8	1951	82.1
1904	66.4	1920	70.4	1936	67.8	1952	88.1
1905	66.1	1921	68.7	1937	66.1	1953	94.2
1906	65.0	1922	68.8	1938	65.2	1954	95.7
1907	63.8	1923	67.5	1939	66.0	1955	96.6
1908	62.0	1924	66.0	1940	68.3	1956 1/	95.2
1909	60.8	1925	63.4	1941	71.8		
1910	59.0	1926	60.6	1942	76.0		
1911	57.2	1927	58.2	1943	81.2		

1/ Preliminary.

Agricultural Marketing Service.

Hog-Corn Price Ratio to Hog Slaughter

The agricultural economist usually is not content with simply observing periodic movements in prices or in production. He wants to know what causes the swings. And he especially wants to know how the current cycle is developing--whether, for example, it will be shorter or longer than average.

This chart summarizes the history of the hog-corn price ratio and hog slaughter since 1920. By using a chart of this kind, the agricultural economist cannot only get some understanding of past cycles, he can get a fairly clear idea of the current situation and probable developments in the next year or so.

Note that there is a tendency for hog slaughter and the hog-corn price ratio to move in cycles of about 4 or 5 years in length. The length of the cycle can be measured from one peak to the next or from one trough to the next. Note that there is a lag between changes in the price ratio and changes in hog slaughter. This lag is from 1 to 2 years in length. It is indicated by the dotted lines connecting peak years of price ratio with peak years of slaughter. Thus, information on recent and current hog-corn price ratios gives some indication of what is likely to happen to hog slaughter 1 or 2 years in the future.

Most economic cycles are not regular. They vary in length and in amplitude. Sometimes mathematicians are tempted to fit some type of curve that assumes perfect regularity. For example, a sine curve. Actually, such mathematical curves rarely fit the data well. If for any reason the economist should want to smooth out the irregular variations shown in this chart, he would do well to draw smooth curves that describe the general nature of movement of the two lines. However, for most purposes he would do just as well to leave the chart as it stands.

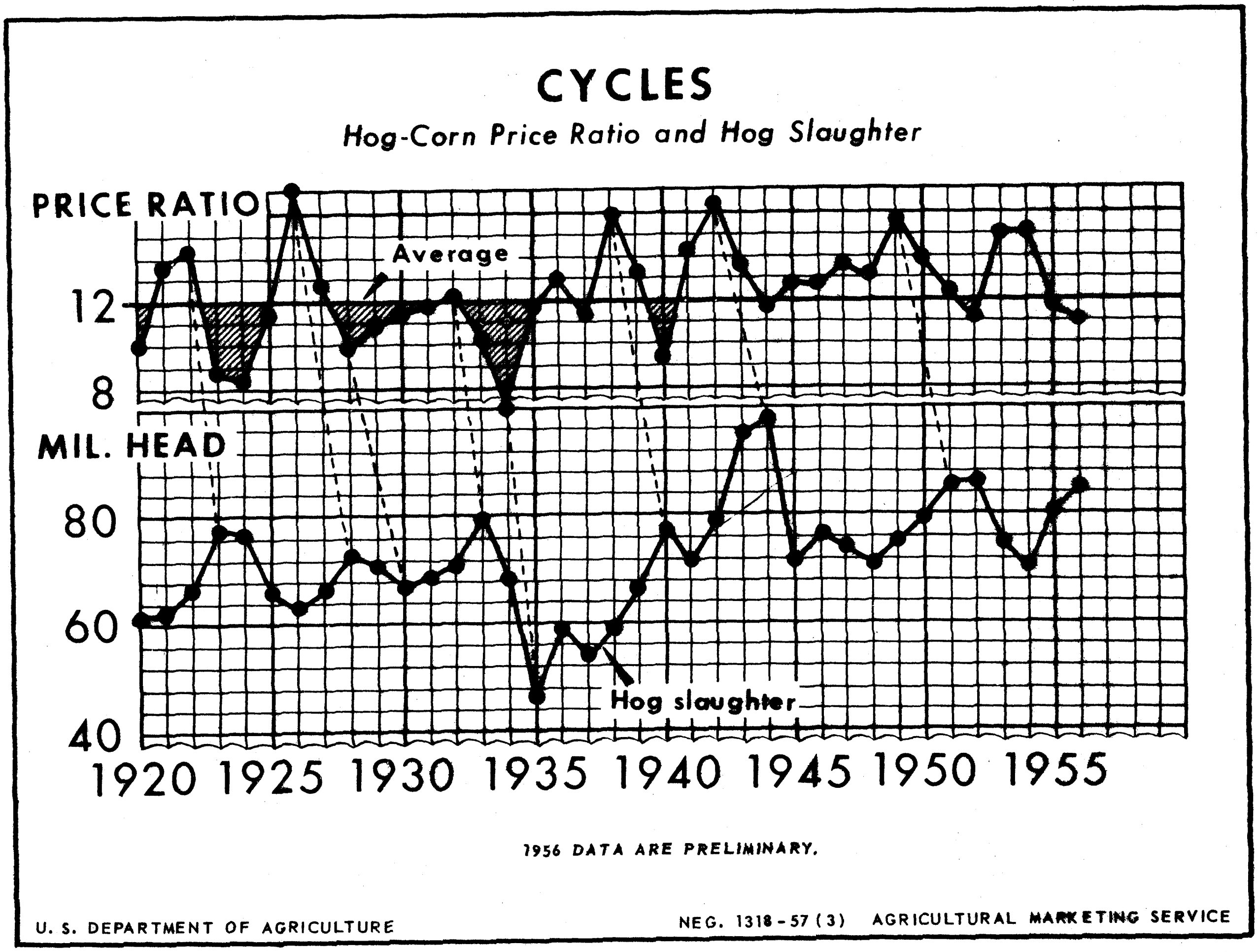


Figure 8

Hogs: Number slaughtered and hog-corn price ratio, 1920-56

Year	Slaughter of hogs	Hog-corn price ratio 1/	Year	Slaughter of hogs	Hog-corn price ratio 1/	Year	Slaughter of hogs	Hog-corn price ratio 1/
	Millions			Millions			Millions	
1920	61.5	9.8	1933 2/ ..	79.7	10.4	1946	76.1	12.6
1921	61.8	13.6	1934	68.8	7.0	1947	74.0	13.6
1922	66.2	14.4	1935	46.0	11.6	1948	70.9	13.0
1923	77.5	8.7	1936	58.7	13.0	1949	75.0	15.7
1924	76.8	8.2	1937	53.7	11.1	1950	79.3	13.7
1925	65.5	11.4	1938	58.9	16.0	1951	85.5	12.4
1926	62.6	17.0	1939	66.6	13.3	1952	86.6	11.0
1927	66.2	12.7	1940	77.6	9.2	1953	74.4	15.0
1928	72.9	9.9	1941	71.4	14.2	1954	71.5	15.0
1929	71.0	10.9	1942	78.5	16.5	1955	81.1	11.8
1930	67.3	11.4	1943	95.2	13.6	1956 3/ ..	85.5	11.1
1931	69.2	11.7	1944	98.1	11.6			
1932	71.4	12.3	1945	71.9	12.8			

1/ Number of bushels of corn required to buy 100 pounds of live hogs at local markets, based on average prices received by farmers for hogs and corn. Annual average is straight average of monthly ratios. 2/ Includes those slaughtered for Government account. 3/ Preliminary.
Agricultural Marketing Service.

SEASONAL VARIATION

Monthly Production of Pork and Prices Received by Farmers for Hogs, 1950-56

Many businesses are seasonal in nature. A department store has a lot of extra business just before Christmas and Easter. Hotels in Florida are full in mid-winter; while resorts in Maine do a big business in summer. The production and marketing of many agricultural products is strongly seasonal. This is simply a reflection of seasonal changes in the weather.

The farmer, the processor, and the distributor all are concerned with the seasonal movement of production, marketing, and prices. An understanding of these seasonal movements help determine the most profitable time to sell.

The accompanying chart is a typical example of a simple analysis of average seasonal variations in recent years. This particular diagram is concerned with production of pork and prices received by farmers for hogs. Each monthly figure was first expressed as a percentage of the 12-month moving average. Then the average percentage was computed for each month. For example, the production of hogs in January was 130 percent of the 12-month moving average centered on January. We simply plot the average percentage for production and prices on the chart, as shown. In general, the seasonal low point in prices is in the late fall and early winter months when production is at a seasonal high. Prices then usually rise and reach a seasonal high in mid-summer, soon after production has reached its seasonal low point.

This kind of analysis, of course, shows only what the average seasonal variation has been in past years. More detailed studies would be required to explain why the seasonal swings in production and prices vary from year to year.

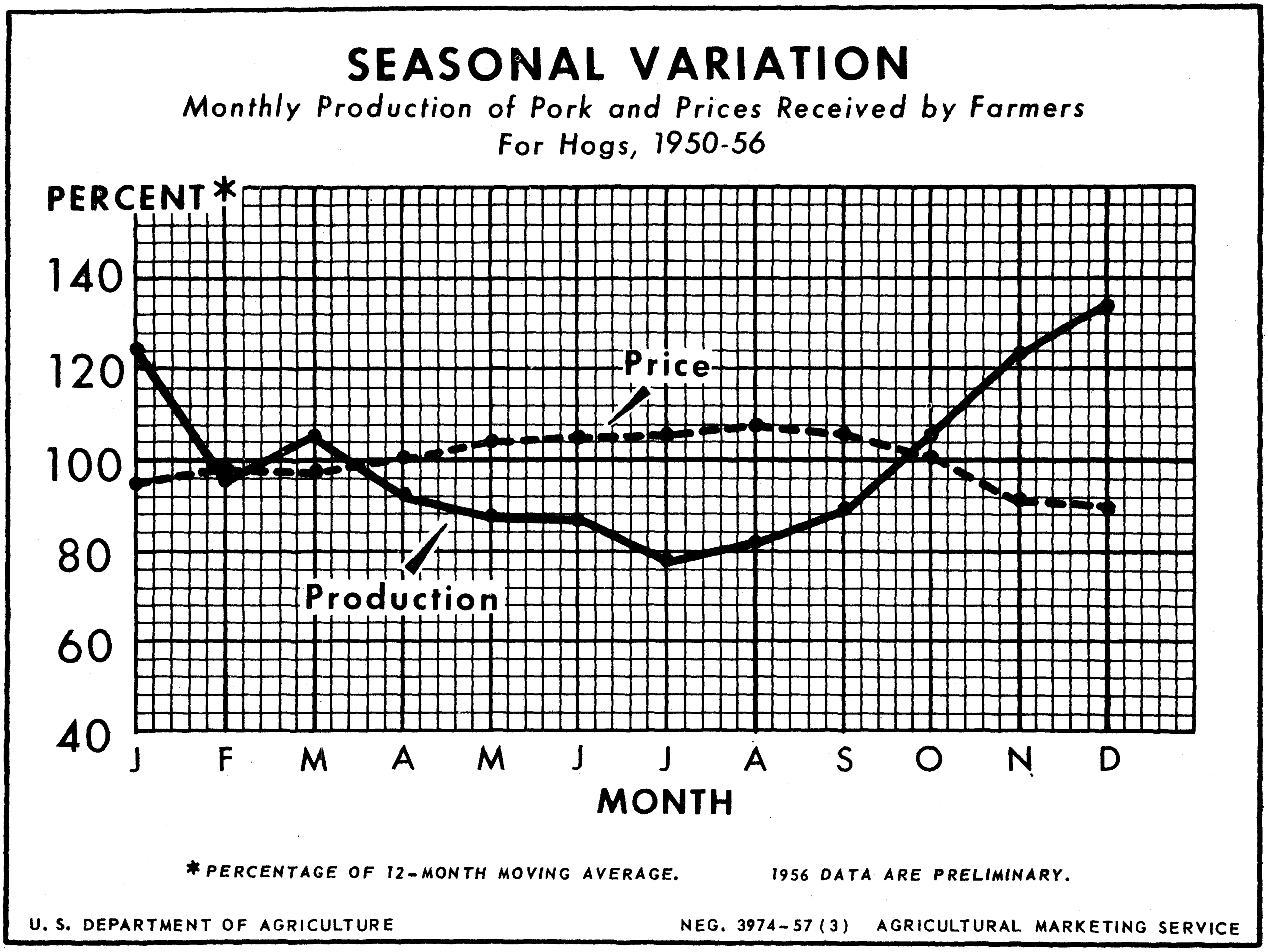


Figure 9

Production of pork and price received by farmers for hogs: Percentage of 12-month moving average, by months, average 1950-56 1/

Month	Production	Price
	Percent	Percent
January	124	95
February	95	97
March	105	97
April	92	100
May	87	104
June	87	105
July	78	106
August	81	108
September	88	106
October	106	101
November	123	91
December	134	90

1/ Production under Federal inspection.

Data revised and brought up to date from Breimyer, Harold F. and Kause, Charlotte A. Charting the Seasonal Market for Meat Animals. U. S. Dept. Agr. Agr. Handb. 83. 1955. pp. 4, 32.

Broiler Chick Placements and Marketings

In many cases the seasonal swings in production and prices do not remain fixed over a long period of time. Rather, they change gradually, reflecting changes in methods of producing and marketing.

A case in point is the accompanying diagram showing the seasonal patterns of broiler placements and marketings. This diagram was taken from a recent report by Martin Gerra, of the Agricultural Marketing Service. Important changes have been occurring in the broiler industry which have affected the timing of marketings. To study such changes in timing, it is often desirable to plot the data for each year separately. We show here the data for each year within two short periods--1941-43 and 1953-55--together with the average for each period. The picture for these two periods is strikingly different. Back in 1941-43 there was a sharp peak in broiler chick placements toward the end of the year, and placements were fairly stable from January through July.

By 1953-55 the seasonal swings in the broiler industry had become considerably different. The industry was geared to reach a peak in placements in early summer, so that the bulk of the marketings came in mid-summer when the demand for broilers is high. Placements then fell off rather regularly until they reached a low point in September or October.

An analysis of this kind can be of great practical value to broiler producers and distributors. The weekly reports on placements alone are good indications of probable marketings about 10 weeks later. This is not all. By studying seasonal diagrams, members of the industry can usually anticipate with some degree of accuracy the changes in placements that might be expected several weeks ahead. This, together with the lag of 10 weeks between placement and marketing, gives at least a general indication of probable changes in marketings over a period much longer than 10 weeks.

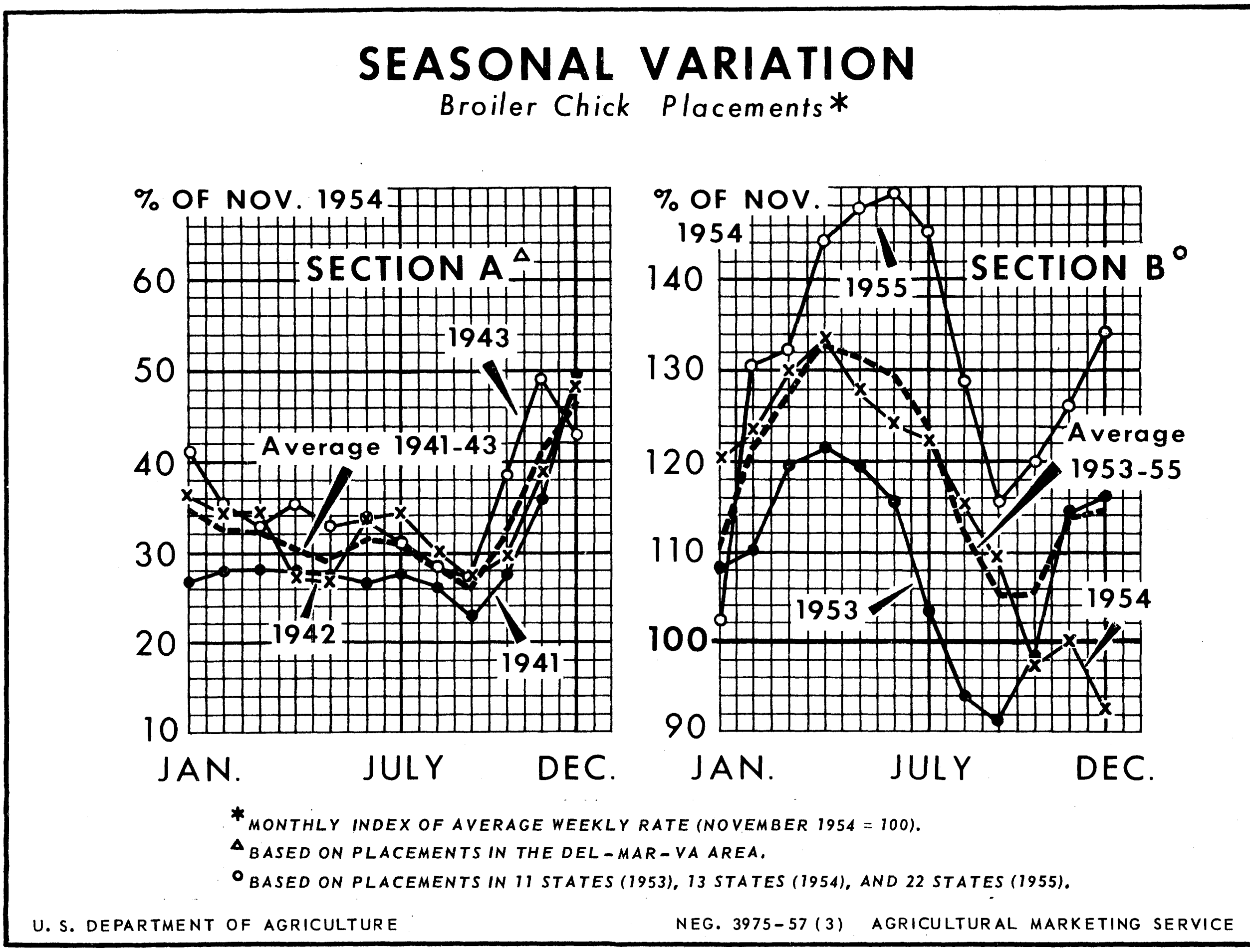


Figure 10

Broiler chick placements: Index numbers of average weekly rate, by months, 1941-43 and 1953-55

[November 1954=100]

Month	1941 1/	1942 1/	1943 1/	Average 1/	1953 2/	1954 2/	1955 2/	Average 2/
January	27.0	36.2	41.4	34.9	108.5	120.1	101.9	110.2
February	27.9	34.5	35.2	32.5	110.2	123.3	130.6	121.4
March	28.3	34.7	33.1	32.0	119.6	129.8	132.2	127.2
April	28.0	27.3	35.8	30.4	121.5	133.8	144.2	133.2
May	27.6	27.1	32.9	29.2	119.5	127.7	147.8	131.7
June	26.8	33.7	33.7	31.4	115.6	124.2	149.3	129.7
July	27.9	34.2	31.3	31.1	103.4	123.5	145.0	124.0
August	26.2	30.1	28.3	28.2	93.9	115.2	128.6	112.6
September ...	22.7	27.3	27.7	25.9	91.2	109.4	115.4	105.3
October	27.2	29.6	38.6	31.8	98.8	97.7	120.0	105.5
November	35.6	38.5	49.2	41.1	114.6	100.0	126.0	113.5
December	49.3	48.3	43.0	46.9	116.0	92.4	134.1	114.2

1/ Based on placements in the Del-Mar-Va area.
2/ Based on placements in 11 States (1953), 13 States (1954), and 22 States (1955).

Gerra, Martin J. Seasonal Changes in Broiler Chick Placements and Marketings. U. S. Agr. Mkt. Service, Poultry and Egg Situation, PES-183, May 1956. pp. 36-40.

Corn Yields Related to Nitrogen

A so-called "dot chart" is one of the handiest tools of economic analysis. The agricultural economist ordinarily must find the relation between two variables. Before putting numbers in a calculating machine, he should almost always draw a chart like the one on the opposite page.

This particular chart gives the results of experiments to determine the relation of corn yields to applications of nitrogen fertilizer. Each dot shows the yield obtained from some amount of fertilizer. For example, the dot to the left of the chart indicates that a yield of 64.6 bushels was obtained with no nitrogen at all. The second dot indicates that 90.4 bushels were obtained when 40 pounds of nitrogen were used. In this case, the several dots all lie fairly closely along a smooth curve which we have drawn freehand. This curve can be taken as an estimate of current yields to be expected from varying applications of nitrogen.

Note that in this case two alternative extensions of the curve are drawn at the right hand side of the diagram. The solid line is horizontal, indicating that the maximum yield is apparently obtained with an application of about 200 pounds of nitrogen to the acre and that additional applications neither increase nor decrease yields. The dotted curve suggests that yields begin to decrease with applications significantly above 200 pounds. So far as the observed statistics go, the dotted curve appears to fit the data slightly better than the solid line. This could, however, be a statistical accident. In deciding which extension to choose, the economist must consult with technical experts on soils and crops. Also, he should consider the results of other experiments.

It is possible to use a chart of this kind to determine the most profitable rate of fertilizer application. This matter is discussed in more detail on page 58. At present our sole purpose is to illustrate a graphic method of determining a simple regression.

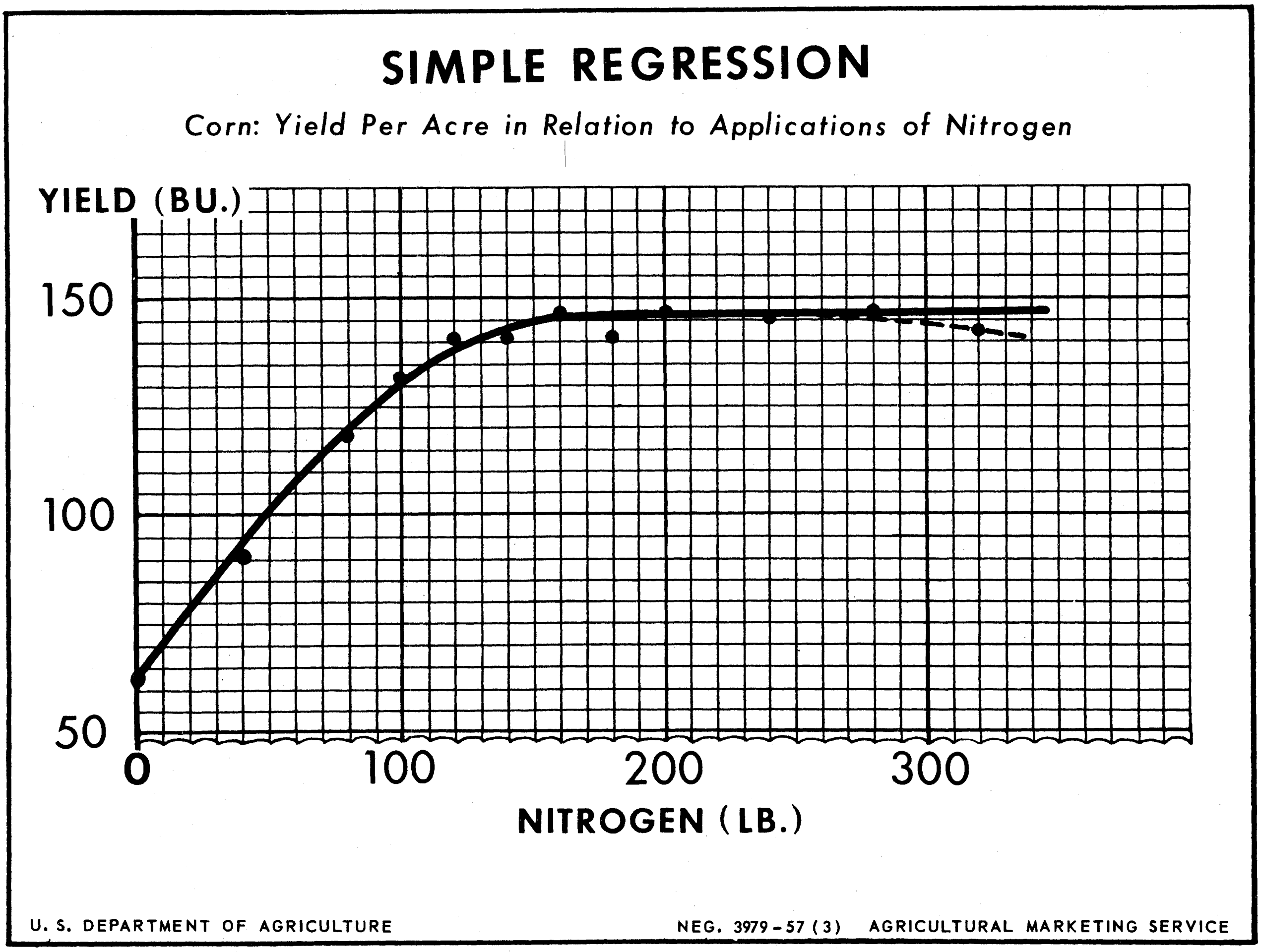


Figure 11

Corn: Yield per acre by specified quantity of nitrogen applied, Ontario, Oregon

Nitrogen applied	Yield of corn	Nitrogen applied	Yield of corn
Pounds	Bushels	Pounds	Bushels
0	64.6	160	146.8
40	90.4	180	141.2
80	118.2	200	147.1
100	132.4	240	145.8
120	140.7	280	147.4
140	141.0	320	143.8

Paschal, J. L., and French, B. L. A Method of Economic Analysis Applied to Nitrogen Fertilizer Rate Experiments on Irrigated Corn. U. S. Dept. Agr. Tech. Bull. 1141. 1956. p. 16.

Weekly Food Expenditures of Families, 1955

Economists have long been interested in the effect of family income on expenditures for food. Many specific studies have established two facts: (1) that the average high-income family spends more for food than does the average low-income family and (2) that the average high-income family spends a smaller proportion of its income for food than does the average low-income family.

We need rather precise estimates of the relation of family income to food expenditures in order to analyze some of the principal economic problems confronting agriculture. In my opinion, a great deal of confusion exists about such terms as "the income elasticity for food." We shall discuss elasticity in general on page 56. For the present, it is necessary only to note that this is a simple regression showing the average or expected weekly food expenditures associated with various levels of family income. It is a gross relation, not a net relation. It is a relation between expenditures and income, not between quantities and income.

Each dot on this diagram shows the average income of a group of families and the average weekly food expenditures of the same group. The dots all lie close to the smooth curve we have drawn. We have not plotted the last observation shown in the table, which normally would have gone some place on the extreme right of the chart. This represents food expenditures of families with incomes of over $10,000 a year. However, the report does not show the average income of these families. Therefore, it is not possible to plot this dot precisely. If we had plotted the dot at the lower limit of this class interval, it would have deviated widely from the curve shown.

This chart confirms one of the two facts given in the opening paragraph, namely that the average high-income family spends more for food than does the low-income family. A different chart would be needed to demonstrate that the average high-income family spends a smaller proportion of its income for food than does the average low-income family.

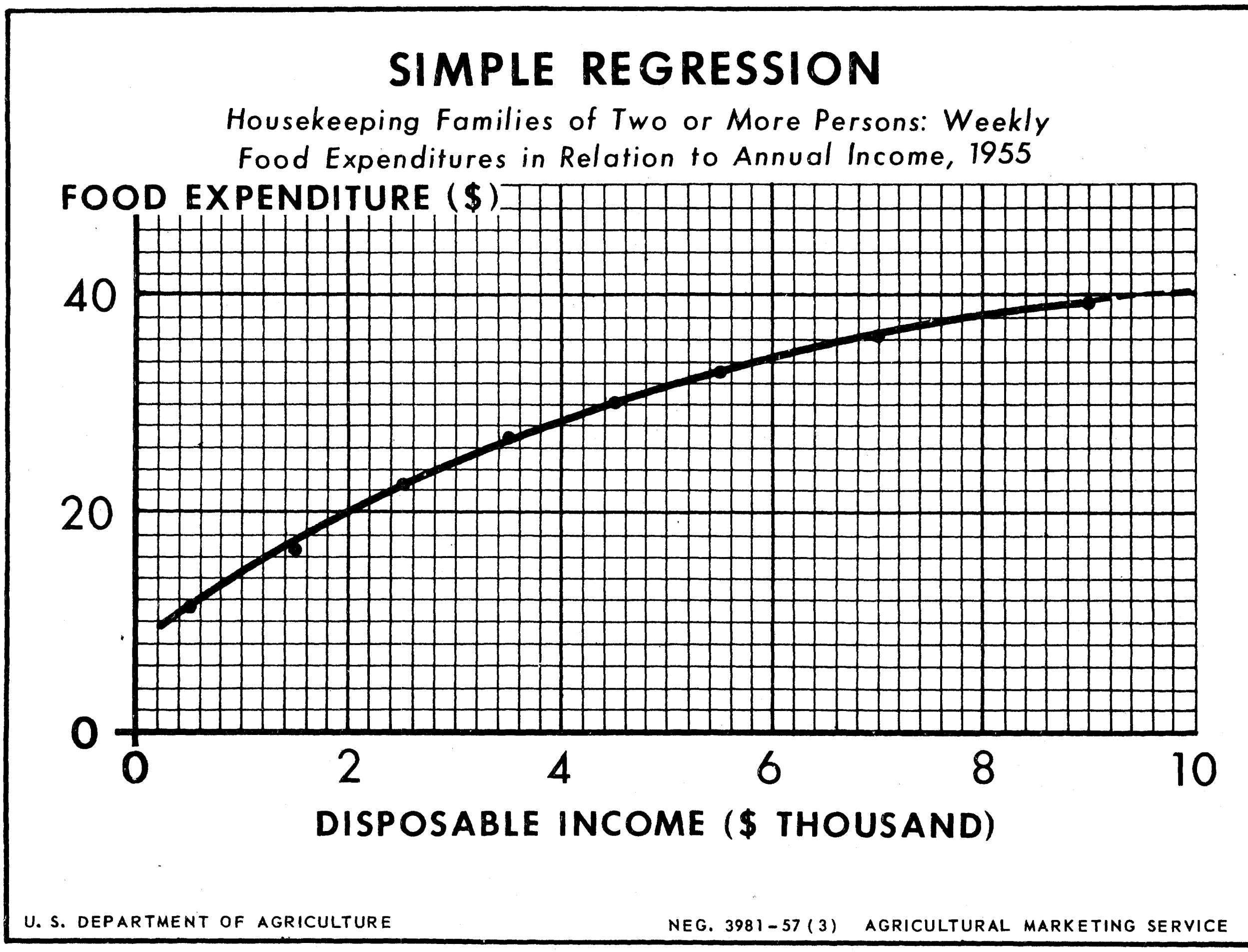


Figure 12

Housekeeping families of two or more persons: Weekly food expenditures by specified income groups, United States, ~~spring~~ 1955

Annual income after taxes	Food expenditures
Dollars	Dollars
Under 1,000	11.69
1,000 - 1,999	16.60
2,000 - 2,999	22.55
3,000 - 3,999	27.00
4,000 - 4,999	30.27
5,000 - 5,999	33.03
6,000 - 7,999	36.14
8,000 - 9,999	39.21
10,000 and over	52.44

Food Expenditures of Households in the United States. Household Food Consumption, 1955, Prel. Rpt. U. S. Dept. Agr. 1956. p. 4.

Cigarette Smoking Related to Age

This is a very simple dot chart based upon a recent survey of tobacco consumption. It shows how the average daily consumption of cigarettes is related to age of cigarette smokers.

The main reason for showing this particular chart is to emphasize that regression curves can take almost any shape. If we had followed the reprehensible practice of putting these data into the calculating machine and assuming that the relationship was linear, we would have come to the conclusion that there is practically no relation between age and smoking habits. Actually, there seems to be a decided relation, but a relation that is distinctly curvilinear. By far the highest rate of consumption is in the middle-aged groups of around 40 to 45 years. Younger and older smokers apparently consume fewer cigarettes. One might, of course, speculate on the reasons for this. It is not primarily a matter of income. These data have been tabulated separately by income groups and each income group exhibits a similar curve in relation to age. A possible reason is that young people get the habit rather slowly and they may turn to cigars and pipes as they get older. Anyway, whatever the reason, the evidence in the chart is rather clear.

Note the wavy lines representing the x-axis. This is a warning that the scale does not start with zero. In this case, if we started both scales with zero, the curve would not have shown up so well. But the statistician should always warn his readers when this is the case. Otherwise, the chart would give the over-exaggerated impression that the youngest and oldest age groups smoked practically no cigarettes at all. This is not true. The youngest-aged group smokes an average of 19.9 cigarettes a day, and the highest consumption of any group is only 22.9 cigarettes. A wavy line is not used on the y-axis, as this scale is not apt to be misleading. It was the opinion of the analyst that few youngsters below 15 years of age smoke a significant number of cigarettes per day.

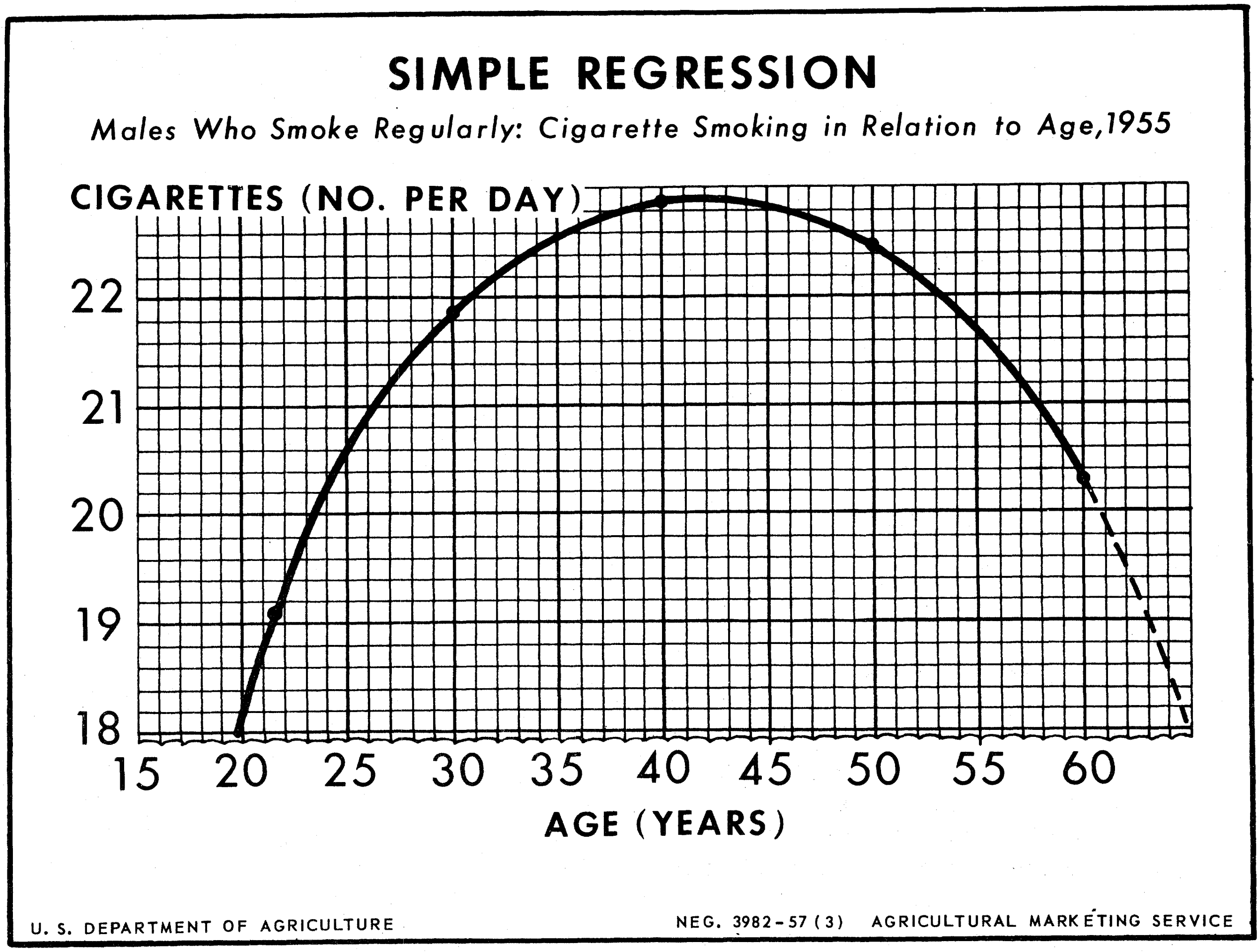


Figure 13

Cigarettes: Average daily consumption by males who smoke regularly, by age, 1955

Age	Consumption
Years	Number
18-24	19.1
25-34	21.9
35-44	22.9
45-54	22.5
55-64	20.3
65 and over	17.6

Smoking Survey. Bureau of the Census. 1955.

Onion Prices Related to Production

Section A of the chart in this example is a dot chart showing the relation between onion production and prices in the years 1939-56. You will note that the observations are scattered all around the diagram and that some of the highest prices occurred in the years of medium to large production. Also, some of the lowest prices occurred in years of low production.

This does not indicate a positively sloping demand curve. It indicates only that both prices and production increased during the period studied. To get a rough idea of the relation between production and prices, we have drawn a line from each observation to each succeeding observation. This is generally a good practice in dealing with time series. It quickly shows up any trend in the data and gives a rough idea at least of the slope of the curve.

In this particular case, section A suggests that we consider the relation of year-to-year changes in prices and in production. This relation is shown in section B. It appears that changes in production give a fairly good indication of expected changes in prices. The explanation is far from perfect. For example, if we had used the curve in section B to estimate expected changes in prices we would have been over 60 cents too low in 1952 and 80 cents too high in 1953.

A more accurate way of studying the relation between onion prices and production is discussed on page 40.

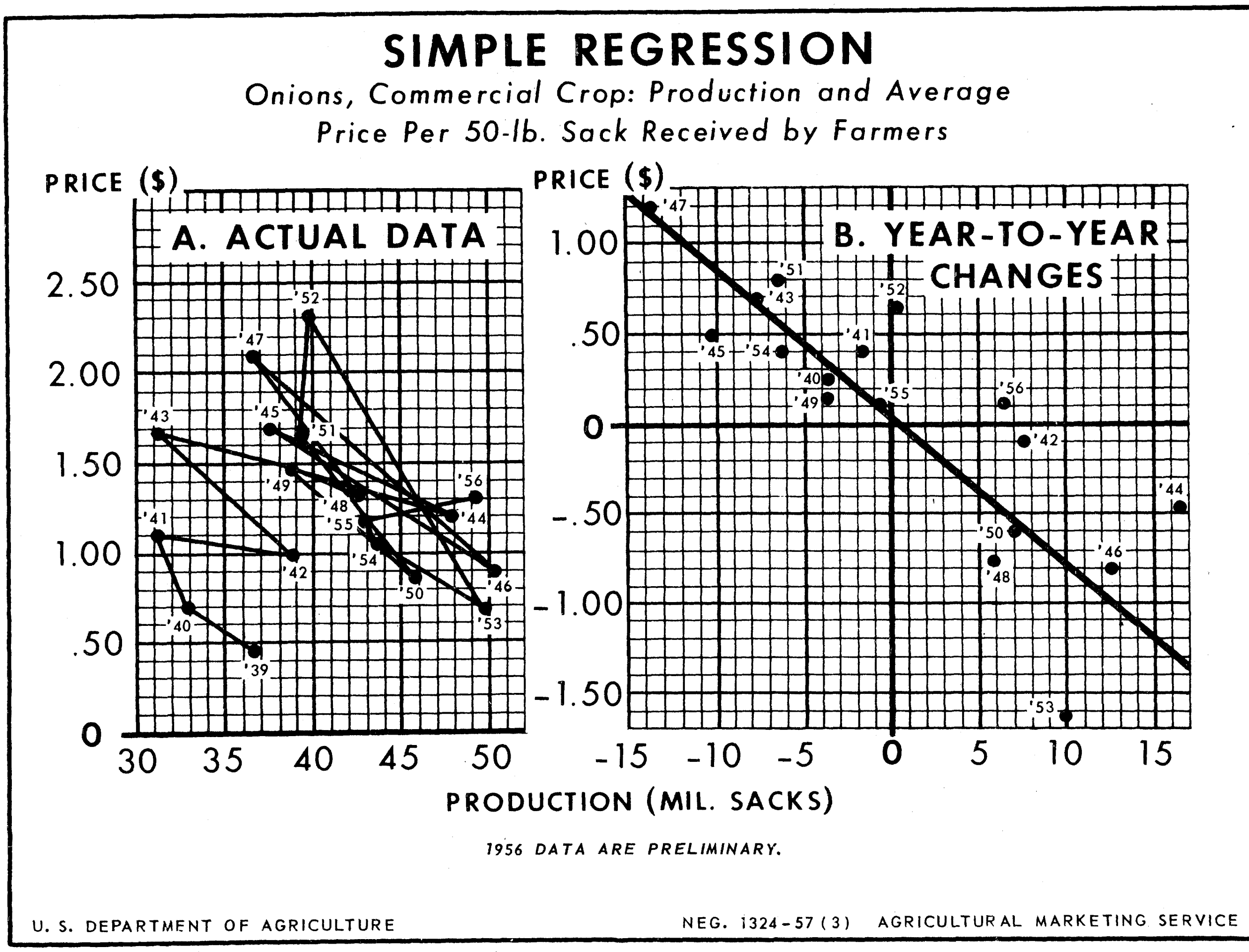


Figure 14

Onions, commercial crop: Production and average price per 50-pound sack received by farmers, 1939-56

Year	Production	Price	Change from preceding year in-- Production	Change from preceding year in-- Price	Year	Production	Price	Change from preceding year in-- Production	Change from preceding year in-- Price
	Million sacks	Dollars	Million sacks	Dollars		Million sacks	Dollars	Million sacks	Dollars
1939	36.6	0.45	---	---	1948 ...	42.5	1.32	5.8	-.76
1940	32.9	.70	-3.7	0.25	1949 ...	38.8	1.47	-3.7	.15
1941	31.2	1.10	-1.7	.40	1950 ...	45.8	.87	7.0	-.60
1942	38.9	.99	7.7	-.11	1951 ...	39.4	1.67	-6.4	.80
1943	31.3	1.68	-7.6	.69	1952 ...	-39.8	2.31	.4	.64
1944	47.9	1.20	16.6	-.48	1953 ...	49.8	.68	10.0	-1.63
1945	37.7	1.69	-10.2	.49	1954 ...	43.6	1.07	-6.2	.39
1946	50.4	.89	12.7	-.80	1955 ...	42.8	1.18	-.8	.11
1947	36.7	2.08	-13.7	1.19	1956 ...	49.4	1.30	6.6	.12

Agricultural Marketing Service.

Shifts in Demand for Beef and Pork

Simple regression is often useful in analyzing problems that are more complicated than those we have just considered. Actually, the demand for beef and the demand for pork are each affected by a number of different variables. Still, it is possible to discover certain basic relationships by simple 2-variable regressions.

The diagram facing this page is based upon an ingenious analysis by Shepherd, Purcell, and Manderscheid. 2/ They first allowed for the effects of population growth by using data on per capita consumption of beef and pork on the x-axis. They also allowed for changes in the general price level by deflating beef prices and pork prices. This was done by dividing retail prices by an index of per capita disposable personal income. While this is not the usual method of deflation, it seems to work well in this case.

In the original report all data were plotted, including those for the war years. As the authors point out, the relationship between consumption and prices in the war years was abnormal because of price controls and meat rationing. In order to simplify the diagram, I have not shown the data for the war years. However, the figures are shown in the table.

This chart suggests that we have several different regressions in each section instead of just one. It seems quite clear, for example, that the demand for beef was higher in the post-World War II years 1947-56 than in the prewar years 1925-41. Thus, it seems desirable to draw two lines through the data rather than the usual single line.

Now, looking at the pork data in the right hand part of the diagram, we find that the demand was apparently highest in the period 1925-31. It dropped somewhat in the period 1932-41 and remained about the same in 1947-52. But in 1953-56 it seems to have been substantially lower. Shepherd et al showed two regression lines for pork, but their data ran only through 1952. On our up-to-date chart, it seemed appropriate to me to draw a third line.

The Iowa publication discusses in detail the reasons for a rising demand for beef and a falling demand for pork. Among the principal reasons are the increasing urbanization of the country and a more even distribution of incomes.

This diagram could serve as a connecting link between simple regression and multiple regression. A study of the two parts of this diagram indicates, for example, that beef prices are affected not only by the supply of beef but also by the supply of pork. Also, pork prices are affected by the supply of beef as well as by the supply of pork. You can see this, for example, if you look at the years 1934 to 1937 in the left hand part of the diagram. In these years the price of beef was higher than indicated by the regression line. Apparently this was because the droughts of 1934 and 1935 severely reduced the supplies of pork in this period. Thus, if we wanted to get a more complete explanation of changes in the prices of beef and pork, we would need to consider more than two variables. This would take us into multiple regression, which is the subject of the next several diagrams. In fact, we might need to consider more than one equation.

2/ Shepherd, Geoffrey S., Purcell, J. C., and Manderscheid, L. V. Economic Analysis of Trends in Beef Cattle and Hog Prices. Iowa Agr. Exp. Sta. Res. Bull. 405, 1954, p. 737.

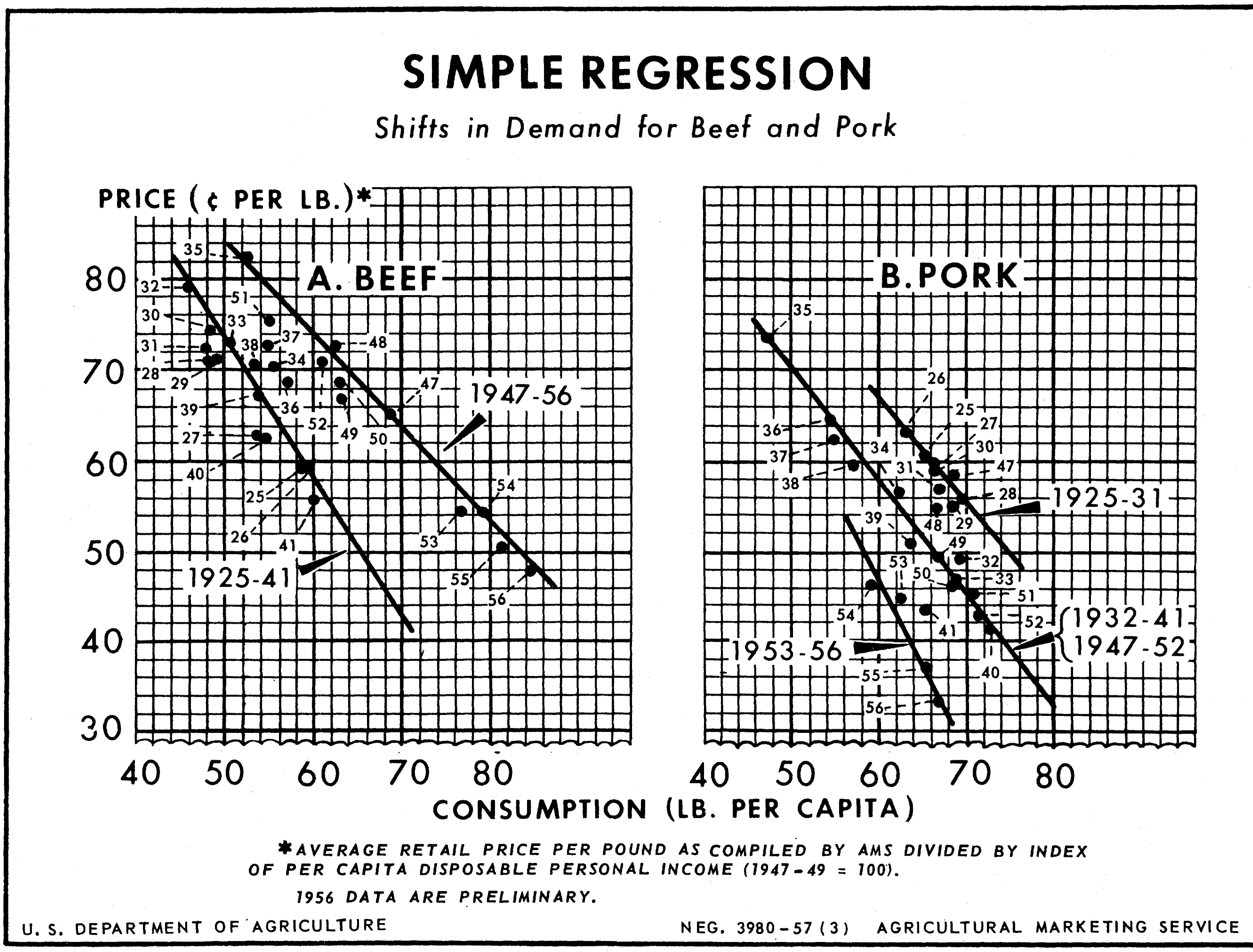


Figure 15

Beef and pork: Retail price per pound and consumption per capita, 1925-56

Year	Beef: Price of choice cuts 1/	Beef: Consumption	Pork, excluding lard: Price 1/	Pork, excluding lard: Consumption	Year	Beef: Price of choice cuts 1/	Beef: Consumption	Pork, excluding lard: Price 1/	Pork, excluding lard: Consumption
	Cents	Pounds	Cents	Pounds		Cents	Pounds	Cents	Pounds
1925	59.7	58.6	60.5	65.8	1941	56.0	60.0	43.9	67.4
1926	59.7	59.4	63.3	63.3	1942 3/	49.7	60.4	42.6	62.8
1927	63.0	53.7	59.9	66.8	1943 3/	45.9	52.5	39.2	77.9
1928	71.0	48.1	56.0	69.9	1944 3/	40.0	54.9	33.9	78.5
1929	71.1	49.0	55.0	68.7	1945 3/	38.6	58.6	33.4	65.7
1930	74.2	48.2	59.6	66.1	1946 3/	46.7	60.8	40.8	74.9
1931	72.1	47.9	57.0	67.4	1947	65.3	68.6	58.6	68.6
1932	79.0	46.0	49.5	69.7	1948	72.8	62.3	54.6	66.8
1933	73.1	50.8	47.3	2/68.7	1949	67.1	63.1	49.7	66.8
1934	70.2	2/55.2	56.6	2/62.2	1950	68.7	62.6	46.1	68.2
1935	82.2	2/52.2	73.9	47.7	1951	74.6	55.3	45.9	70.9
1936	68.4	2/57.3	64.4	54.4	1952	70.9	61.4	42.7	71.4
1937	73.0	54.4	62.2	55.0	1953	54.5	76.5	45.3	62.6
1938	70.2	53.6	59.9	57.4	1954	54.1	79.0	46.1	59.2
1939	67.8	53.9	51.0	63.9	1955	51.2	80.9	37.2	65.9
1940	63.4	54.2	41.5	72.4	1956	48.0	84.2	33.9	66.8

1/ Retail price as computed by Agricultural Marketing Service divided by index numbers of disposable personal income (1947-49=100). 2/ Consumption less use under Federal programs. 3/ War years omitted from diagram.
Agricultural Marketing Service.

Food Prices, Consumer Incomes, and Volume of Farm Marketings

Economists often work with time series; that is, with records of prices, production, and consumption over a period of time. When studying relations between time series, particularly if several variables are involved, it is a good practice to plot each series before drawing dot charts such as the ones we have just discussed.

Suppose, for example, that we were trying to discover the factors which affect retail food prices. Two of the factors that would doubtless come to mind are consumer incomes and the volume of marketings for food. Before rushing to the calculating machine or even drawing a dot chart, it would be a good idea to plot each series as we have done in this diagram and to study the changes which have occurred over a period of time.

In this case it is clear that there is high correlation between the food price index and per capita disposable income. In fact, the relationship is so pronounced that it tends to overshadow the effect of per capita food marketings. We might notice, too, that during the war years from 1941 to 1945 the relationships do not seem to be the same as in other years.

Comparisons of these three time series suggest that the correlation between the average price index for food and per capita disposable income would be reduced by deflating each series (for example, by dividing each of these by the consumer price index for all commodities). Such a computation would also reduce the magnitude of the gyrations to more nearly correspond to those for per capita marketings of food. The sharp rise in marketings of food during the war years and subsequent decline, which appears to have taken place independent of changes in the other series, suggests that the war years be omitted from the analysis. If a chart of this sort indicates pronounced trends in one or more variables, it suggests that the analysis might yield improved results if it were based on year-to-year changes in the variables.

In some cases, a comparison of time series will indicate a timelag between changes in one variable and changes in another. We saw previously that changes in slaughter of hogs occur several months after a change in the ratio of hog prices to those for corn.

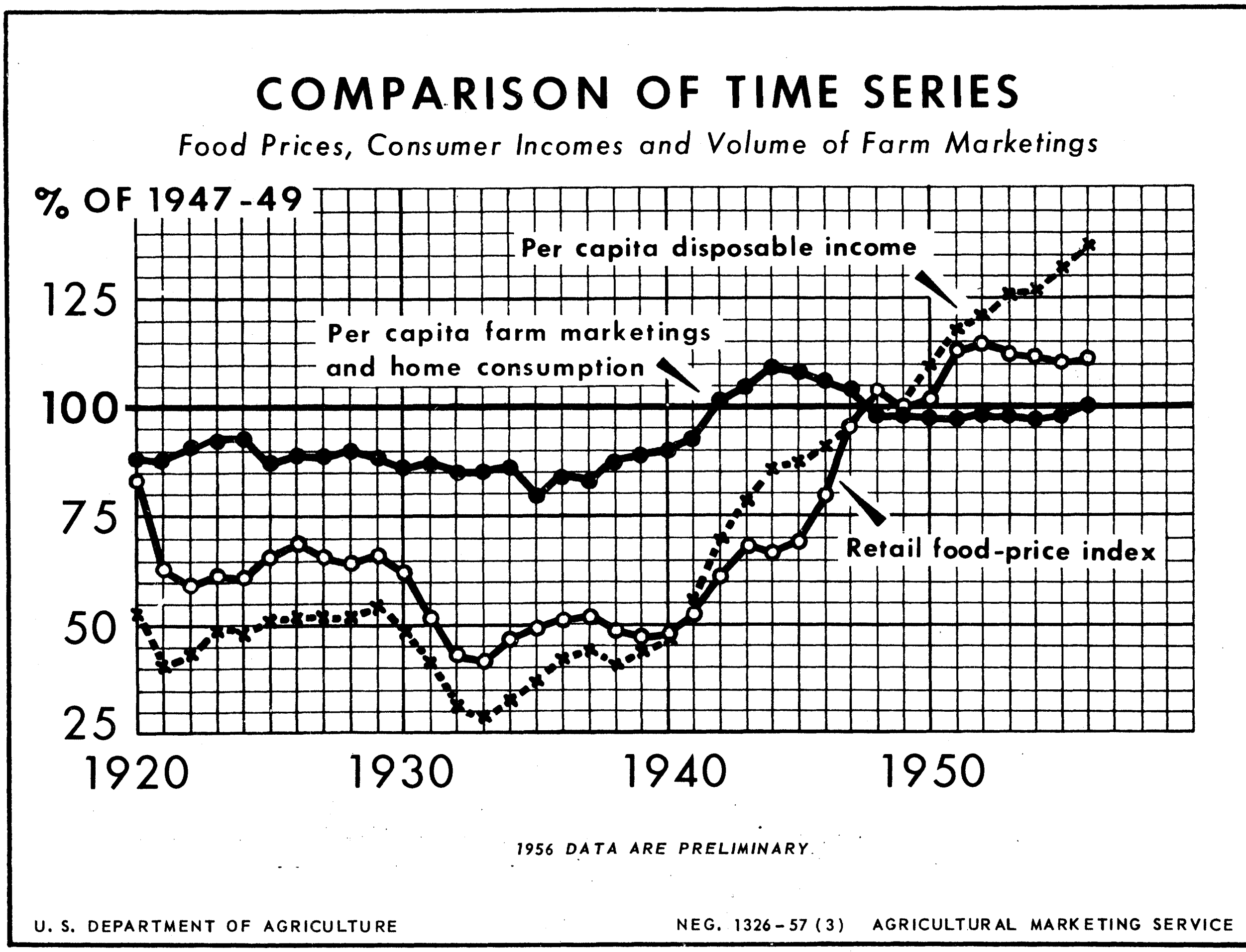


Figure 16

Price and marketing of food and disposable income: Index numbers, 1920-56

[1947-49=100]

Year	Food		Disposable income per capita
	Retail price	Farm marketings and home consumption per capita	
1920	83.6	88	52.8
1921	63.5	88	41.0
1922	59.4	91	43.7
1923	61.4	93	49.8
1924	60.8	94	49.3
1925	65.8	89	51.4
1926	68.0	90	52.6
1927	65.5	90	52.1
1928	64.8	90	52.7
1929	65.6	89	55.1
1930	62.4	87	48.8
1931	51.4	87	41.5
1932	42.8	86	31.4
1933	41.6	86	29.4
1934	46.4	87	33.2
1935	49.7	79	37.0
1936	50.1	84	41.8
1937	52.1	83	44.5
1938	48.4	87	40.8
1939	47.1	89	43.5
1940	47.8	91	46.5
1941	52.2	93	56.3
1942	61.3	101	70.4
1943	68.3	105	78.9
1944	67.4	109	85.6
1945	68.9	108	86.8
1946	79.0	106	91.0
1947	95.9	104	94.7
1948	104.1	98	103.3
1949	100.0	98	101.9
1950	101.2	96	109.8
1951	112.6	97	118.3
1952	114.6	97	122.1
1953	112.8	97	126.7
1954	112.6	97	126.6
1955	110.9	98	132.2
1956	111.7	100	137.6

Prices from Bureau of Labor Statistics, marketings of food from Agricultural Marketing Service, and income from Department of Commerce.

Per Capita Consumption Related to Deflated Per Capita Income and Food Prices

Multiple regression has been used to analyze a wide variety of economic problems. This is because most economic variables (such as prices and rates of consumption) are influenced by a number of different factors. Ordinarily the economist cannot conduct controlled experiments allowing only one of these factors to vary. Rather, he must try to unscramble market data in order to separate out the influence of each of several variables. Whenever the influences of the separate variables can be added together, the problem can be studied by multiple regression.

As in the case of simple 2-variable regression, the statistician can put the data for a multiple regression problem into the calculating machine and compute the answer by least squares. However, in my opinion there are many advantages to a graphic analysis of such problems. Such analyses were made popular by Louis H. Bean 3/ and have been used widely in the Department of Agriculture and in the State colleges.

A case in point is the relation of food consumption to income and food prices. If we can accurately measure these relationships, we have the basis for determining the so-called "income elasticities" and "price elasticities" for food consumption. James P. Cavin recently made a mathematical analysis of the data presented here. I shall illustrate how the data can be analyzed graphically.

In section A of this chart I have first plotted the data to show for each year indexes of deflated per capita income, together with the indexes of per capita food consumption. Each dot shows the pair of indexes for a particular year. (The years 1942 through 1947 were excluded from this analysis because food consumption was affected by such things as rationing and price control.) If we were to draw a line representing the simple relationship between food consumption and income, we would doubtless draw a curve which would be steepest at the left hand side of the chart, and which would become less steep as we move from left to right. However, we are not concerned with this simple relationship. We want a regression line which will be our best estimate of what the index of food consumption would have been if food prices had remained constant. An examination of the dots in section A and the price data in the table indicate a general tendency for food consumption to be reduced when food prices are high and to be increased when food prices are low. The regression line in section A is drawn with this in mind. For example, it is drawn considerably higher than the dots for the post-World War II period when real (or deflated) food prices were higher than in the prewar period.

Section B attempts to explain how food consumption was related to the level of real food prices. Specifically, each dot shows for some year the deflated food prices for that year, together with the deviation above or below a regression line in section A. For example, take the first year, 1922. The index of deflated food prices was 83.0, and the dot for 1922 in section A is 0.6 units above the regression line. This observation is plotted in section B with the coordinates 83.0 on the x-axis and +0.6 on the y-axis. Similarly, for each other year. After these dots are properly located in section B, we draw the regression line indicating the net effect of deflated food prices on food consumption. Then our job is done unless a further study of the data suggests a need for making some adjustment. In this case the two lines seem reasonably satisfactory.

The deviations from the regression line in section B indicate the amount of error that is made in estimating food consumption from the two independent variables--income and price. The largest errors are about two index points in 1926, 1935, 1937, and 1951. For most years our estimates are within one index point of the true figure. With the index of consumption varying between 87.8 and 104.0, this amount of error seems reasonably small.

3/ Bean, Louis H. A Simplified Method of Graphic Curvilinear Correlation. Jour. Amer. Statis. Assoc. 24:386-397, illus. 1929.

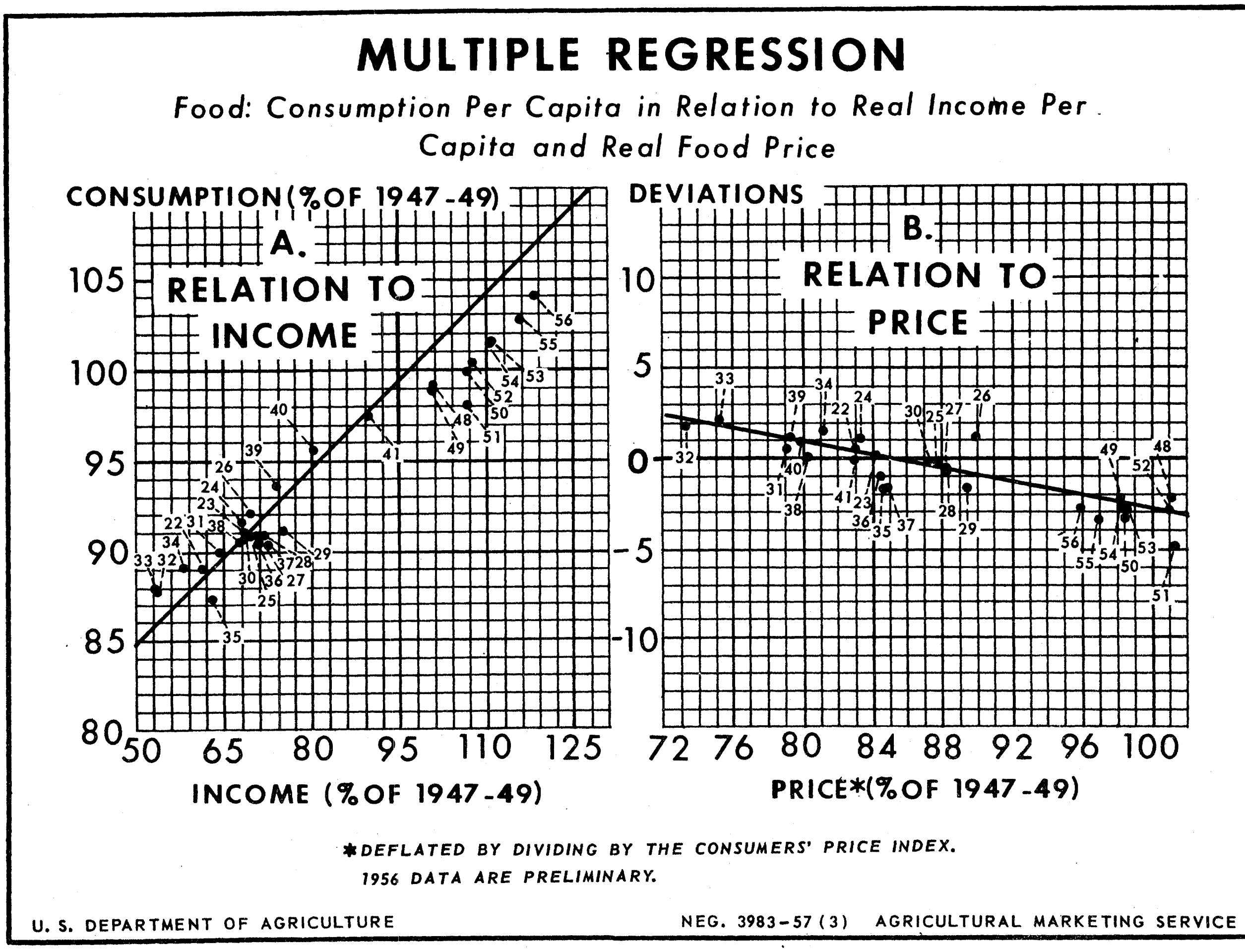


Figure 17

Index numbers: Consumption of food, disposable income, and retail food price, 1922-41 and 1948-56

[1947-49=100]

Year	Per capita: Consumption of food	Per capita: Disposable income 1/	Price of food 1/	Year	Per capita: Consumption of food	Per capita: Disposable income 1/	Price of food 1/
1922	89.0	61.0	83.0	1937	90.4	72.5	84.9
1923	90.9	68.3	84.2	1938	90.6	67.8	80.3
1924	91.5	67.4	83.2	1939	93.8	73.2	79.3
1925	90.9	68.5	87.7	1940	95.5	77.6	79.8
1926	92.1	69.6	89.9	1941	97.5	89.5	83.0
1927	90.9	70.2	88.3				
1928	90.9	71.9	88.4	1948	99.1	100.6	101.3
1929	91.1	75.2	89.5	1949	98.9	100.1	98.2
1930	90.7	68.3	87.4	1950	99.9	106.8	98.4
1931	90.0	64.0	79.1	1951	98.1	106.6	101.4
1932	87.8	53.9	73.3	1952	100.4	107.6	101.0
1933	88.0	53.2	75.2	1953	101.5	110.8	98.6
1934	89.1	58.0	81.1	1954	101.4	110.3	98.1
1935	87.3	63.2	84.7	1955	102.8	115.5	96.9
1936	90.5	70.5	84.5	1956	104.0	118.4	95.9

1/ Deflated by dividing by the Bureau of Labor Statistics Consumers' Price Index.

Agricultural Marketing Service.

Price of Corn Related to Price of Livestock and Supply of Feed Concentrates Per Animal Unit

The diagram illustrates an analysis of corn prices (X_0) related to two independent variables--prices of livestock and livestock products (X_1) and supplies of feed concentrates per animal unit (X_2). We know from theory and from general observation that high livestock prices tend to be associated with high prices of corn. We also know that large supplies of feed concentrates tend to be associated with low prices of corn. But we want to quantify these relationships--perhaps to forecast prices of corn.

Section A of this chart shows corn prices and prices of livestock and livestock products from 1936 through 1955. Before drawing the regression line, we try to take account of X_2. We draw several regressions for subsamples of data, commonly called "drift lines." Thus in 1948, 1949, and 1950, supplies of concentrates were from 1.05 to 1.07 tons. We connect these observations with a drift line. Similarly we connect the observations for 1940, 1941, and 1942, when supplies were 0.90 tons. After drawing all possible drift lines, we draw a net regression line the slope of which represents approximately an average of the slopes of the drift lines. In this case, a straight line happens to be satisfactory. In many cases, a curve would be indicated.

Section B shows how the residuals (departures from the first regression line) are related to X_2. These residuals are clustered closely around the regression line we have drawn. If a nearly perfect fit were not given by the dots around this line, the process of successive approximation would be used. Foote 4/ has shown that when we use this method graphically based on linear relationships, the slopes of the successive approximations tend to converge toward the value that would be obtained had we fitted a mathematical regression line by the method of least squares.

Graphic multiple regression requires a fair amount of imagination and some practice. But it often shows up important relationships that are not brought to light by grinding figures out of a computing machine.

4/ Foote, Richard J. The Mathematical Basis for the Bean Method of Graphic Multiple Correlation. Jour. Amer. Statis. Assoc. 48:778-788. 1953.

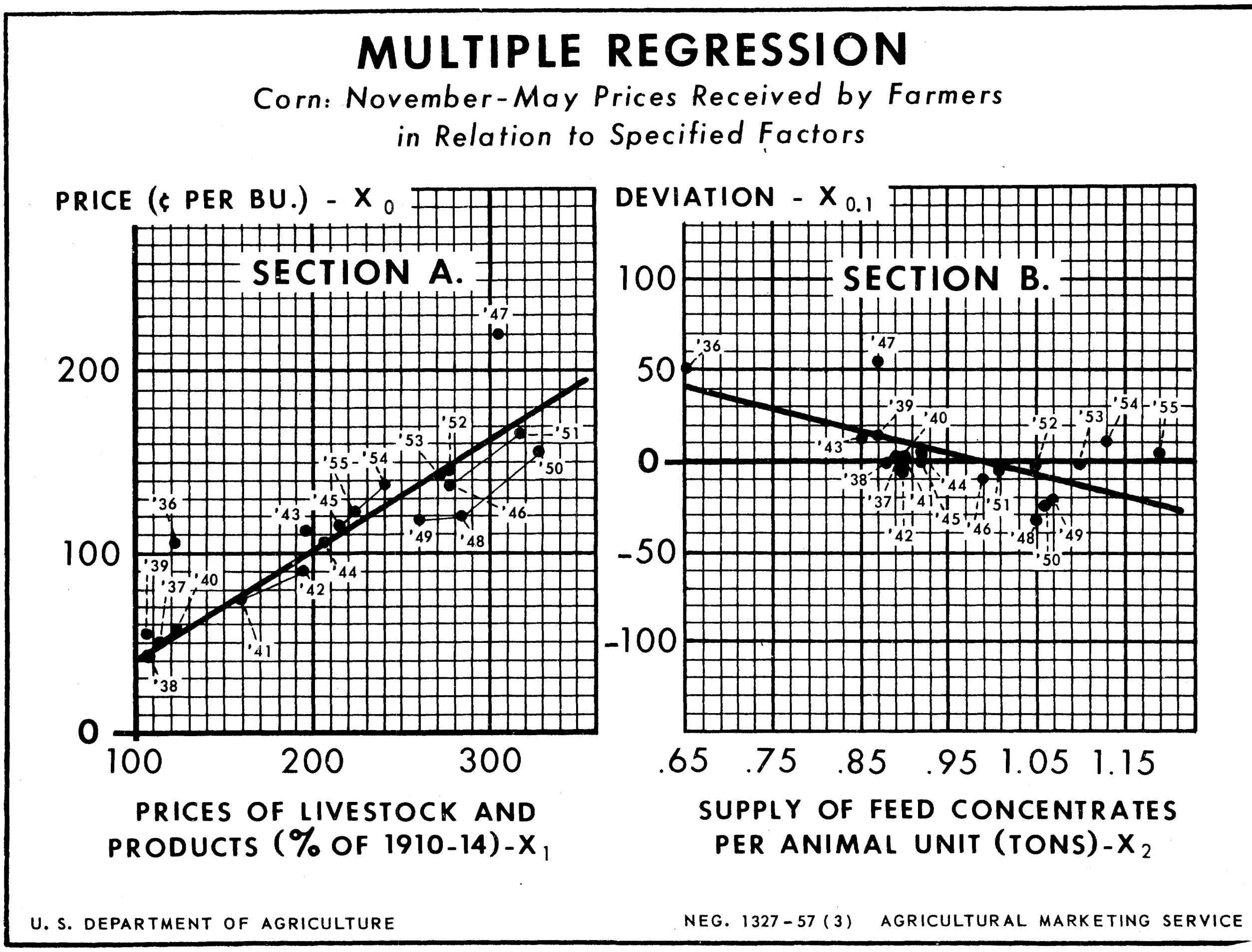


Figure 18

Corn: Price per bushel received by farmers and related variables, 1936-55

Period beginning	Price received by farmers (November-May): Corn	Price received by farmers (November-May): Livestock and products 1/	Supply of feed concentrates per animal unit 2/	Period beginning	Price received by farmers (November-May): Corn	Price received by farmers (November-May): Livestock and products 1/	Supply of feed concentrates per animal unit 2/
	Cents		Tons		Cents		Tons
1936	106	123	0.65	1946	138	278	0.99
1937	51	114	.89	1947	220	305	.87
1938	44	108	.88	1948	120	285	1.05
1939	55	107	.87	1949	118	260	1.07
1940	58	122	.90	1950	155	329	1.06
1941	74	159	.90	1951	166	318	1.01
1942	90	194	.90	1952	147	278	1.05
1943	112	196	.85	1953	142	271	1.10
1944	107	206	.92	1954	138	240	1.13
1945	115	215	.92	1955	121	224	1.19

1/ Index number, 1910-14=100. 2/ Year beginning October.

Computed from data in Foote, Richard J. Statistical Analyses Relating to the Feed-Livestock Economy. U. S. Dept. Agr. Tech. Bull. 1070. 1953. p. 6.

Yields of Corn in Illinois, 1934-55, Related to Reported Condition on September 1 and a Time Trend

The Agricultural Marketing Service estimates the probable production of many of the principal crops several months before they are harvested. Such advance estimates of probable production are based in part upon the judgment of farmers concerning "the condition of the crop as a percentage of normal." It is unnecessary here to explain in detail the concept of normal production. Statisticians have found that the farmers' reports as to current condition of crops is a fairly good indication of the yield that would occur with average growing conditions during the rest of the growing season.

The Division of Agricultural Estimates makes extensive use of dot charts in graphic analysis to interpret reported condition and to estimate probable yields. The accompanying chart, suggested by C. E. Burkhead of the Agricultural Estimates Division, Agricultural Marketing Service, illustrates how this can be done in the case of corn yields in Illinois. Section A of the chart is a scatter diagram relating the reported condition as of September 1 of each year from 1934 through 1956 to the final estimate of harvested yield. It is easy to see that there is some positive correlation between reported condition and final yield. When farmers report a condition of 90 to 95 percent of normal, the final yield tends to be high. When they report a low condition of, say, 40 or 50 or 60 percent of normal, the yield tends to be low. But this is not the whole story shown in the chart. When dealing with time series, it is always a good idea to label each dot to indicate the year, as we have done in this case. Notice that the dots for the early years are all in the lower top part of the scatter. In other words, a reported condition of 80 percent of normal today indicates a higher yield of corn than would have been suggested 20 years ago.

The solid, straight line drawn through this scatter is an estimate of the relation we might have expected between condition and harvested yield at about the middle of the period studied; that is, from around 1940 through 1950. To make a good estimate of corn yields today we need to consider not only the average relationship between reported condition and harvested yield for the whole period, but also a "net trend"; that is, the trend in yields after allowing for the average relationship shown in section A of the chart.

This is a problem in multiple regression. The corn yield is the dependent variable; that is, the variable we are trying to estimate. In this case there are two independent variables which are useful in estimating yield. The first of these is reported condition and the second is time. The effect of time is shown in section B. Here we have plotted for each year the deviation of the actual yield from the regression line shown in section A. Take the first year in the series--1934. The regression line indicates a yield of 30.5 bushels. The actual harvested yield was 21.5 bushels. So, there was a deviation (or "residual") of -9.0 bushels. Thus, in section B we indicate -9.0 for the year 1934. Similarly, for each of the other years in the series. When these dots are plotted, it is apparent that there was a definite net trend. It rose sharply from 1934 to about 1942. Then leveled off until about 1950. Since 1950 it has again risen sharply. Probably the sharp increase in the early years of the series was due mainly to the introduction of hybrid corn. The effect of this began to peter out in the 1940's. Since about 1950 there has been a new upward trend, probably due to increased use of fertilizer.

The use of an analysis of this kind can be illustrated by data for 1956. Farmers reported a condition 96 percent of normal. Section A indicates a yield of 58 bushels. Section B indicates that we should add 9.5 bushels to account for the trend; thus giving us an estimate of 67.5 bushels. Actually, the yield in 1956 turned out to be 68.0 bushels. In this case the two regression lines would have given us a good forecast of corn yields in Illinois, somewhat more accurate than we should expect in an average or typical year.

One of the difficulties with time series of this kind is that of extrapolating the net trend shown in section B. Each year we make a forecast we have to extrapolate beyond the range of observed data. We don't really know what the net trend in Illinois corn yields will be in the future and have to do some guessing.

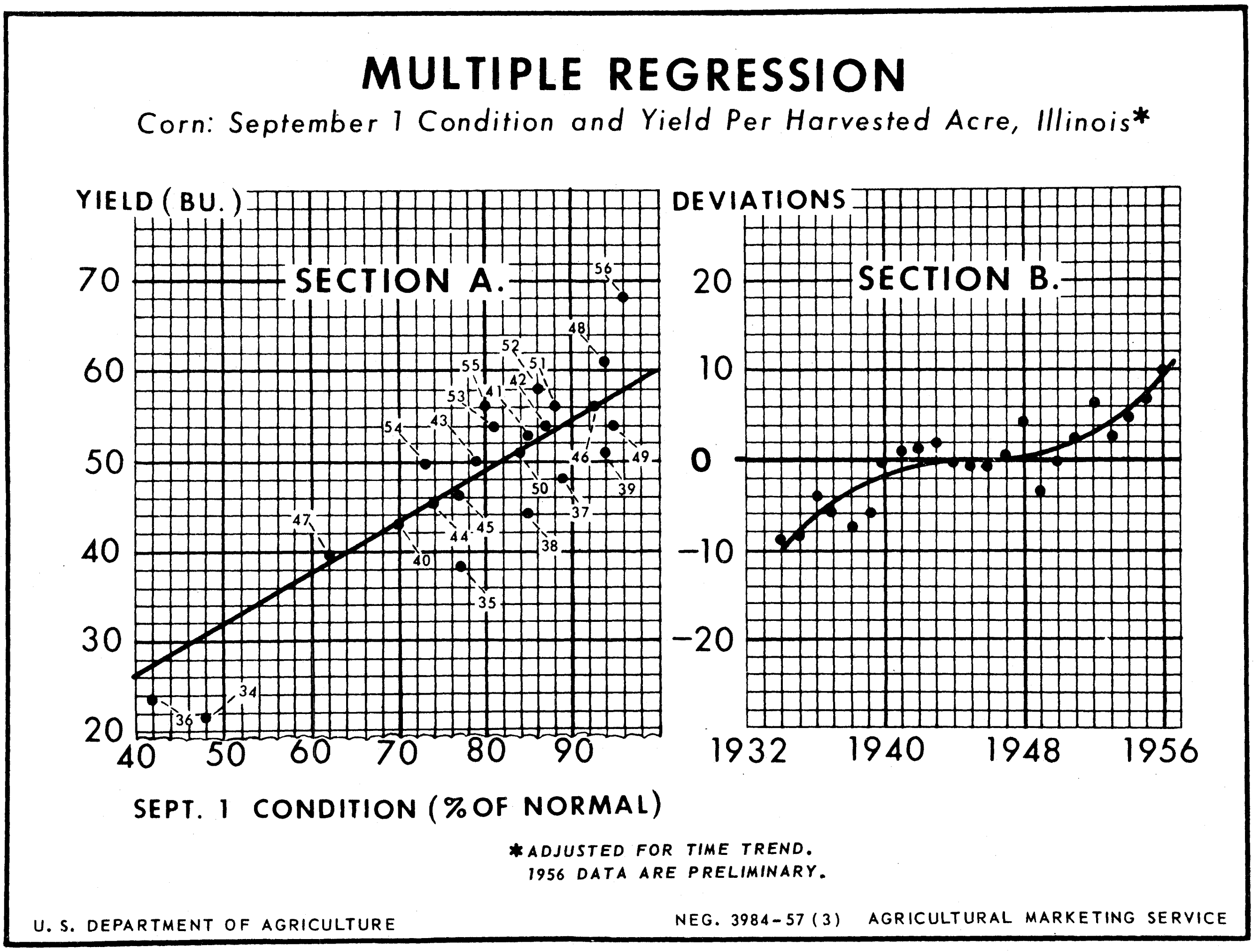


Figure 19

Corn: Condition September 1 and yield per harvested acre, Illinois, 1934-56

Year	Condition 1/	Yield	Year	Condition 1/	Yield
	Percent	Bushels		Percent	Bushels
1934	48	21.5	1946	93	56.0
1935	77	38.5	1947	62	39.5
1936	42	23.5	1948	94	61.0
1937	89	48.0	1949	95	54.0
1938	85	44.0	1950	84	51.0
1939	94	51.0	1951	88	56.0
1940	70	43.0	1952	86	58.0
1941	85	53.0	1953	81	54.0
1942	87	54.0	1954	73	49.5
1943	79	50.0	1955	80	56.0
1944	74	45.4	1956	96	68.0
1945	77	46.5			

1/ As a percentage of normal.

Data supplied by C. E. Burkhead, Agricultural Estimates Division, Agricultural Marketing Service.

Price of Late Onions Related to Production and Disposable Income

The data for this diagram, taken from Shuffett, 5/ are expressed as first differences (i.e. year-to-year changes) in logarithms. The rationale of this may be found in Shuffett's bulletin and need not concern us here. Graphic analysis will handle logarithms and first differences, as well as the unmanipulated data.

The main purpose of this diagram is to illustrate successive approximations to the true regression lines. We have already discussed the graphic determination of the net regression lines. So far, we have tacitly assumed that one approximation is enough. But in many cases the statistician should try two or more successive approximations.

The original data (here they are the first differences of logarithms) are plotted as in the regression charts we have already discussed. The black dots in section A show the joint scatter of production and price. The heavy line is our first approximation to the net regression of production on price. (Drift lines were drawn, but have been erased to keep from cluttering up the chart.) Deviations from this line were then plotted as heavy dots in section B. The solid line through these heavy dots is the first approximation of the net regression of disposable income on price.

So far, our analysis is the same as in several previous diagrams. We now proceed to make a second approximation. The deviations from the solid line in section B are now plotted as circles in section A. The dashed line, drawn through these circles, is our second approximation to the net regression of production on price. Then the deviations from this dashed line are plotted as circles in section B. A dashed line, drawn to fit these circles, is our second approximation to the net regression of disposable income on price.

This process can be continued to get as many approximations as needed. If done correctly, the successive approximations will converge to the true (least squares) regressions. Ordinarily two or three approximations are enough.

5/ Shuffett, D. Milton. The Demand and Price Structure for Selected Vegetables. U. S. Dept. Agr. Tech. Bull. 1105, pp. 38-43. 1954.

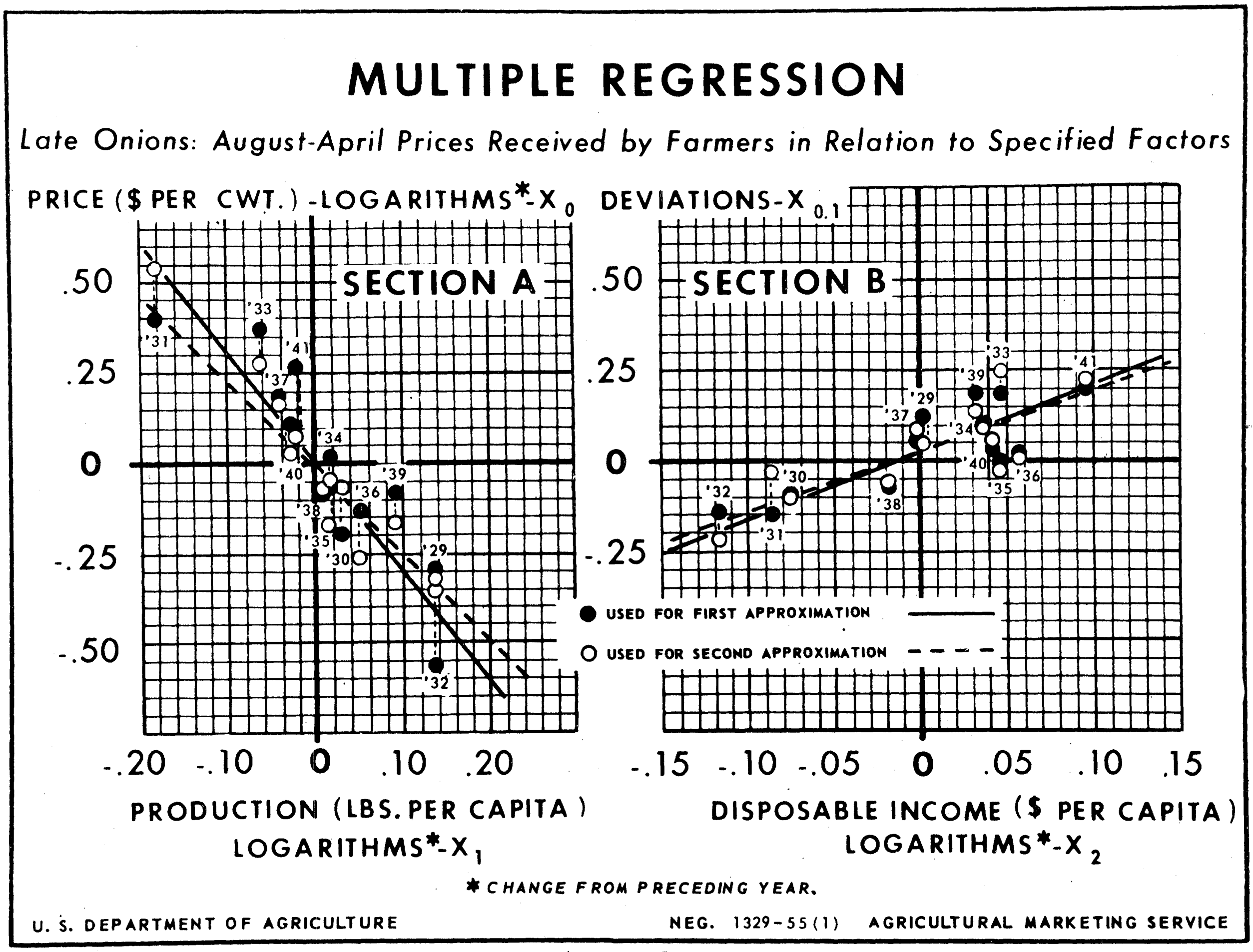


Figure 20

Late onions: Average price per 100 pounds received by farmers and related variables, August-April average, 1928-41

Period beginning	Actual: Price 1/	Actual, per capita: Production 1/ 2/	Actual, per capita: Income 3/	First difference of logarithms: Price	First difference of logarithms, per capita: Production	First difference of logarithms, per capita: Disposable income	Period beginning	Actual: Price 1/	Actual, per capita: Production 1/ 2/	Actual, per capita: Income 3/	First difference of logarithms: Price	First difference of logarithms, per capita: Production	First difference of logarithms, per capita: Disposable income
	Dol.	Lb.	Dol.					Dol.	Lb.	Dol.			
1928	2.54	6.65	658	---	---	---	1935	1.18	8.20	467	-0.062	0.017	0.046
1929	1.30	9.08	663	-0.291	0.135	0.003	1936	.86	9.23	534	-.137	.051	.058
1930	.82	9.75	557	-.200	.031	-.076	1937	1.30	8.36	532	.179	-.043	-.002
1931	2.02	6.41	456	.392	-.182	-.087	1938	1.06	8.59	509	-.089	.012	-.019
1932	.54	8.75	347	-.573	.135	-.119	1939	.88	10.57	546	-.081	.090	.030
1933	1.28	7.58	386	.375	-.062	.046	1940	1.12	9.93	601	.105	-.027	.042
1934	1.36	7.89	420	.026	.017	.037	1941	2.08	9.47	748	.269	-.021	.095

1/ Excludes quantities produced in market gardens for sale in nearby cities prior to 1939.
2/ Production divided by November 1 civilian population.
3/ Disposable income at annual rates divided by November 1 civilian population.

Shuffett, D. Milton. The Demand and Price Structure for Selected Vegetables. U. S. Dept. Agr. Tech. Bull. 1105. 1954. p. 43.

JOINT (3-DIMENSIONAL) REGRESSION

Yield of Corn In Relation To Applications of Nitrogen and Phosphoric Acid

Many problems of economic analysis can be handled by simple (2-variable) regression. We have considered several examples of problems that can be handled by this technique. Many other economic problems can be analyzed rather well by the use of multiple regression. However, the use of a multiple regression is limited to problems in which the effects of several variables can be added to one another, except where special transformations of the data are made, such as the use of logarithms. In mathematical terms multiple regression is limited to the analysis of problems that can be stated in the form

$$x_o = f_1(x_1) + f_2(x_2) + \ldots\ldots f_n(x_n) \quad (1)$$

Actually, many important problems in economic research cannot be handled satisfactorily by such an additive function. In many cases we must consider the more general relation

$$x_o = F\ (x_1,\ x_2,\ \ldots\ldots,\ x_n) \quad (2)$$

A case in point is the relationship of crop yields to various dosages of fertilizer. We considered on page 22 the relationship of corn yields to a single variable, the application of nitrogen. In this case the applications of potash and phosphoric acid were held constant. While this sort of analysis tells us something about response to nitrogen, researchers want to know the response to various combinations of nitrogen, potash, and phosphoric acid. Many experiments have been conducted in which all three of these have been varied. The 3-dimensional diagram facing the page shows how we can analyze the combined effects of two independent variables at a time. In this case we consider the combined effects of nitrogen and phosphoric acid ($P_2\ O_5$). The data are taken from a recent report of the Iowa Agricultural Experiment Station. 6/

In this case we have plotted the data with the application of potash (K_2O) held constant at 40 pounds to the acre. The applications of nitrogen (N) varied from 0 to 240 pounds. The application of phosphoric acid varied from 0 to 120 pounds. The location of each hatpin on the base of the diagram indicates a combination of nitrogen and phosphoric acid. The height of the hatpin above the base indicates the yield of corn. For example, one of the hatpins is located at the point corresponding to N = 0, P_2O_5 = 0, and the height of this hatpin corresponds to a yield of 32.00 bushels of corn to the acre. Similarly, each of the other hatpins shows the yield obtained by some combination of nitrogen and phosphoric acid. The relationship between these combinations of fertilizer applications and corn yield is quite apparent when one looks at a diagram of this kind. It is easier to see it in the original diagram than in the photograph. The highest corn yields were obtained by a combination of about 160 pounds of nitrogen and about 80 pounds of phosphoric acid. When the nitrogen application was increased to 240 pounds, yields were definitely reduced. Also, there is some indication of a reduction in yield when phosphoric acid is increased to 120 pounds. When no nitrogen is used, applications of phosphoric acid tended to decrease the yield. As increased amounts of nitrogen were used, applications of a considerable amount of phosphoric acid were beneficial.

This is a 3-dimensional dot chart. In principle, it is the same thing as the several 2-dimensional dot charts we have looked at. We want to visualize a graphic, 3-dimensional surface which describes the general nature of the relationship. Such a surface could be constructed either graphically or by fitting some proper form of mathematical function. Before choosing a mathematical function, however, the researcher would do well to sketch in a smooth regression surface, such as the one shown on the diagram. Such a sketch will show that any satisfactory mathematical function would permit an inverse relation between corn yields and phosphoric acid when nitrogen applications are low, and a positive relation between corn yields and phosphoric acid when nitrogen applications are high. None of the usual formulas used to describe the results of fertilizer application do this. Therefore, they will not fit these particular observations well. If other experiments should produce similar results, we ought to either look for another mathematical formula, or else be satisfied with the results we can get from graphic analysis.

6/ Brown, William G., Heady, Earl O., Pesek, John T., and Stritzel, Joseph A. Production Functions, Isoquants, Isoclines and Economic Optima in Corn Fertilization for Experiments with Two and Three Variable Nutrients. Iowa Agr. Exp. Sta. Res. Bull. 441, 1956.

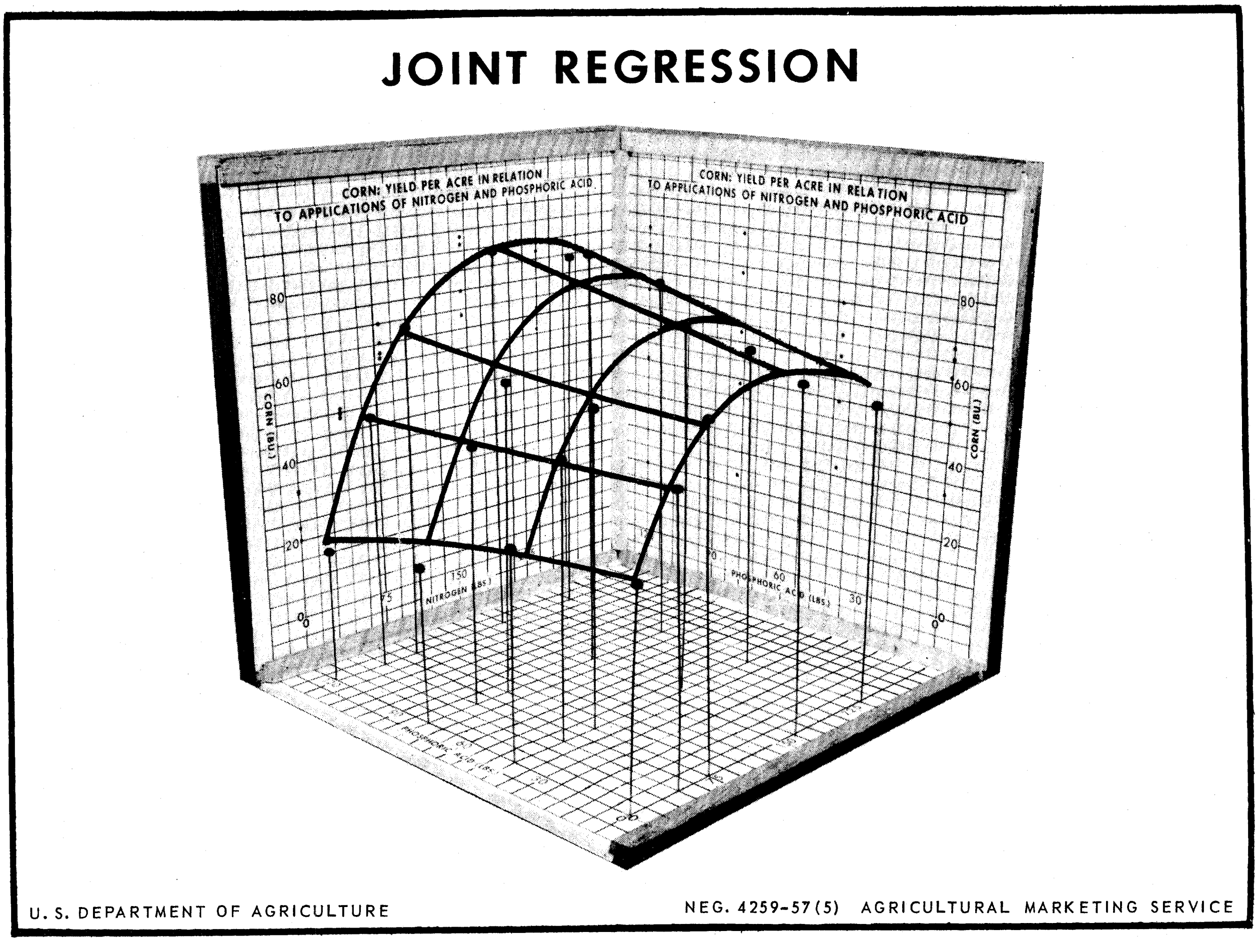


Figure 21

Corn: Yield per acre for given applications of phosphoric acid and nitrogen 1/

Phosphoric acid	Nitrogen in pounds				
	0	40	80	160	240
Pounds	Bushels	Bushels	Bushels	Bushels	Bushels
0	32.00	49.85	65.50	68.25	61.20
40	32.25	49.55	61.55	74.75	66.90
80	23.25	48.55	62.65	88.15	78.45
120	20.20	50.75	69.80	86.80	81.90

1/ With potash at 40 pounds per acre.

Brown, William G., Heady, Earl O., Pesek, John T., and Stritzel, Joseph A. Production Functions, Isoquants, Isoclines and Economic Optima in Corn Fertilization for Experiments With Two and Three Variable Nutrients. Iowa Agr. Expt. Sta. Research Bull. 441. 1956. p. 815.

USE OF ISOQUANTS TO STUDY JOINT REGRESSION

Yield of Corn in Relation to Applications of Nitrogen and Phosphoric Acid

Another technique for studying variation of a 3-dimensional surface is similar to that used in surveying and grading land. We can forget for the moment that the chart refers to yield of corn. Suppose that the vertical axis measures distances north and south, the horizontal axis measures east and west, and the numbers written by the dots on the diagram indicate the elevation of the land at various points and determined by surveyor's transit. Anyone used to maps would recognize that the land is gradually increasing in height as we move toward the right side of the diagram, becoming steeper as we move toward the upper right corner. You would also see that there are bumps and hollows. In simple regression we smooth in only one dimension. Here we are smoothing in two dimensions. We can describe the general lay of the land by a series of smooth contour lines.

Of course, we are not dealing here with land and contour maps. However, the general problem of joint regression is that of determining a series of isoquants. Whatever the three variables may be, an isoquant will show the combinations of two independent variables which correspond to a given value of the dependent variable. In the case illustrated by the diagram, it is clear that yields increase steadily with increases in nitrogen up to about 160 pounds per acre and then begin to decline. The effects of additional phosphoric acid are less for small applications of nitrogen than for large applications of nitrogen. Yields can be increased substantially by high level applications of both nitrogen and phosphoric acid, although the maximum combination, as indicated by this diagram, is reached by the use of perhaps 175 pounds of nitrogen and 75 pounds of phosphoric acid.

With a little practice anyone can draw isoquants graphically, as we have in this diagram, that give at least a general indication of the relationships involved. If the researcher wants to fit mathematical functions, the diagram should suggest the kind of function to use. Another technique which is sometimes used to study three-dimensional relationships is the "isometric projection." Those who are not familiar with isoquants may find such projections easier to visualize. But they are also harder to read accurately.

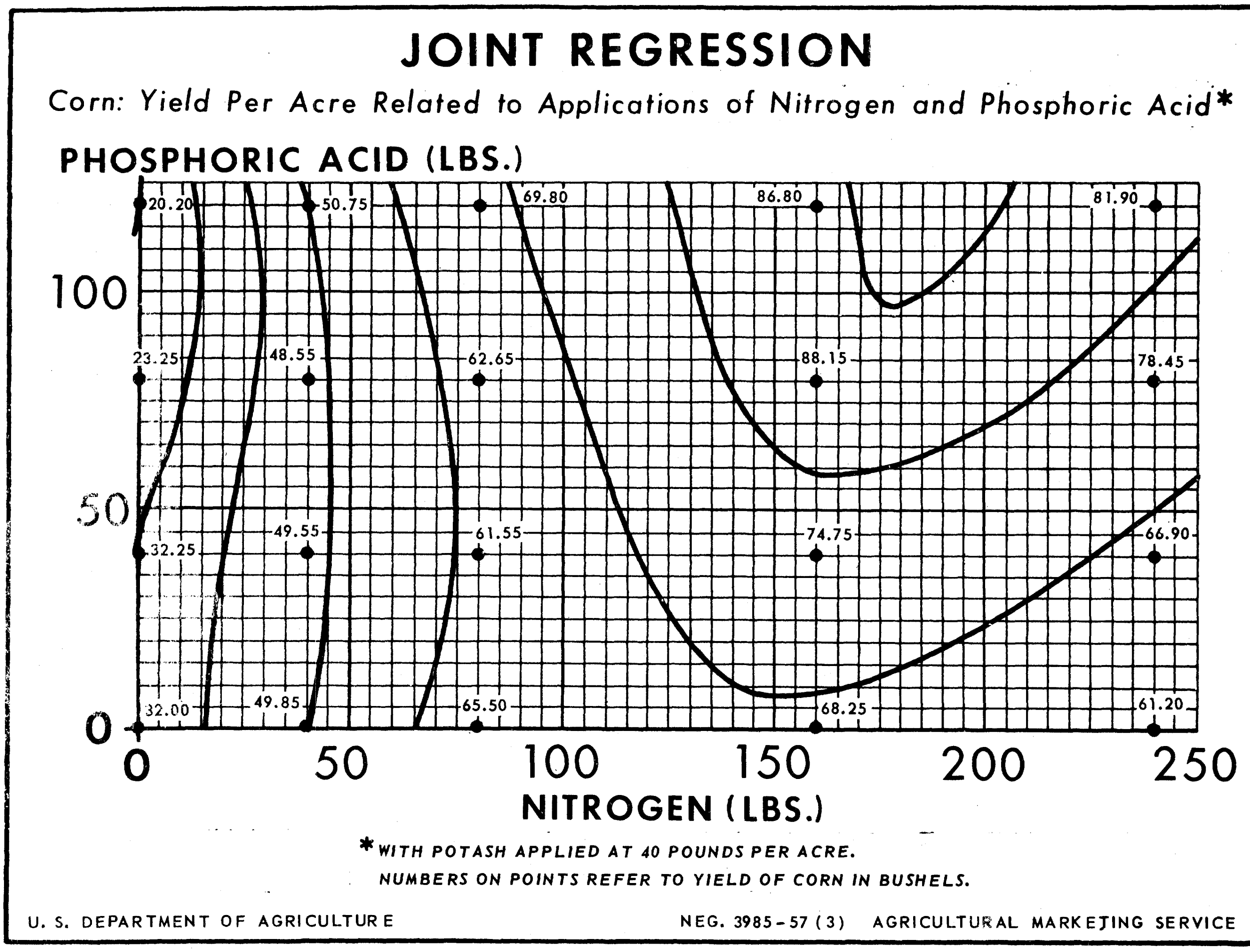


Figure 22

Corn: Yield per acre for given applications of phosphoric acid and nitrogen 1/

Phosphoric acid	Nitrogen in pounds				
	0	40	80	160	240
Pounds	Bushels	Bushels	Bushels	Bushels	Bushels
0	32.00	49.85	65.50	68.25	61.20
40	32.25	49.55	61.55	74.75	66.90
80	23.25	48.55	62.65	88.15	78.45
120	20.20	50.75	69.80	86.80	81.90

1/ With potash at 40 pounds per acre.

Brown, William G., Heady, Earl O., Pesek, John T., and Stritzel, Joseph A. Production Functions, Isoquants, Isoclines and Economic Optima in Corn Fertilization for Experiments With Two and Three Variable Nutrients. Iowa Agr. Expt. Sta. Research Bull. 441. 1956. p. 815.

Beef and Pork, 1947-56

The first edition of the handbook included a diagram labeled "Indifference Curves" based on data for beef and pork. Two periods were shown, 1921-31 and 1932-41. The present chart shows a family of curves for the post-World War II period 1947-56.

This is doubtless the most controversial diagram in the handbook. I tried to justify such a diagram in a recent paper. 7/ But not all economists accept my justification. Some of them think it is impossible to derive indifference curves from any analysis of market data.

In any case, the economists and statisticians that I know would agree that a diagram such as the one presented here gives a satisfactory explanation for changes in the ratio of beef prices to pork prices. Three observations are plotted for each year. For example, the x marked '56 indicates that in 1956 the per capita consumption of beef was 84.2 pounds and of pork was 66.8 pounds. The slope of the line drawn through that point indicates a price ratio of 1.42. In other words, in 1956 the consumer would have to give up 1.42 pounds of pork to get a pound of beef. After plotting for each year the consumption of beef and pork, together with the price ratio, we note that the ratio was not constant. It was high when pork consumption was high and beef consumption was low. It was low when pork consumption was low and beef consumption was high. This is as we would expect. The light curves drawn through this diagram are similar to the contour lines used in the diagram on page 45 to analyze a problem of joint regression. In drawing such lines we attempt to portray a smooth 3-dimensional surface, changing gradually and slowly. The slope of these curves is to be about the same as the slope of the neighboring heavy straight lines indicating observed price ratios. You will note that the observed price ratios for each year except 1947 are almost the same as those of the neighboring contour lines. The observed price ratio in 1947 was smaller than we would have estimated. This may be an indication that the postwar demand for meat did not become stabilized until after 1947.

Certainly this family of curves does not tell us anything about the total welfare or level of living of the typical American consumer. To analyze this we would have to study his entire expenditure pattern. However, in a sense at least, I believe the diagram does represent a "partial indifference surface" for beef and pork alone, indicating the amount of satisfaction obtained from these two commodities. This assumes, of course, that the satisfactions from beef and pork are independent of satisfactions obtained from other goods and services. Perhaps we have to take this assumption with a grain of salt. But don't forget that we make assumptions in most statistical analyses. For example, when deriving a demand curve from market data.

7/ Waugh, Frederick V. A Partial Indifference Surface for Beef and Pork. Jour. Farm Econ. 38:102-112, illus. 1956.

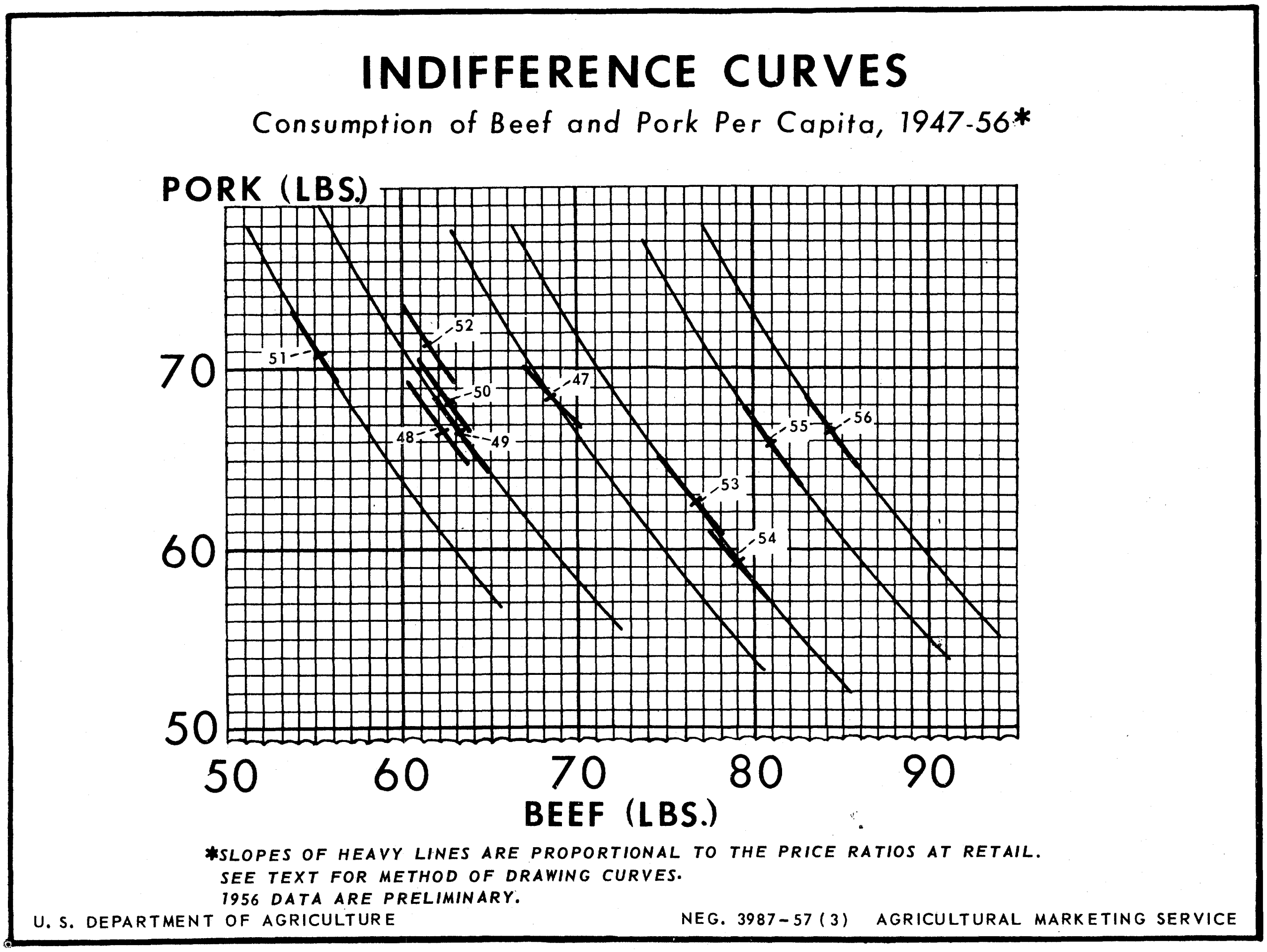


Figure 23

Beef and pork: Ratio of beef price to pork price at retail and per capita consumption, 1947-56

Year	Retail price per pound: Beef	Retail price per pound: Pork	Price ratio	Consumption: Beef	Consumption: Pork
	Cents	Cents		Pounds	Pounds
1947	61.8	55.5	1.11	68.6	68.6
1948	75.3	56.5	1.33	62.3	66.8
1949	68.4	50.6	1.35	63.1	66.8
1950	75.4	50.3	1.50	62.6	68.2
1951	88.2	54.3	1.62	55.3	70.9
1952	86.6	52.1	1.66	61.4	71.4
1953	69.1	57.4	1.20	76.5	62.6
1954	68.5	58.3	1.18	79.0	59.2
1955	67.7	49.2	1.38	80.9	65.9
1956	66.0	46.6	1.42	84.2	66.8

Agricultural Marketing Service.

Combination of Two Farm Enterprises

Programming is the planning of economic activities to maximize income or to minimize costs. In some cases it is reasonable to assume that the input-output relationships are approximately linear. For example, if we know how much seed, labor, and fertilizer is required to grow an acre of potatoes, we can assume that it will require about twice as much of each input factor to grow two acres of potatoes by the same process. In a similar manner, if we know the amount of protein, calcium, and other nutrients in a bushel of corn, there would be twice as much of each nutrient in two bushels of corn. These are linear relationships, and in cases of this kind we can estimate the optimum program by a technique known as linear programming.

The data on this chart show two possible farm enterprises in North Carolina and six input factors. 8/ To be feasible a combination of inputs must not require more than the available amount of any resource. In an analysis of this kind it is convenient first to compute for each enterprise the proportion of available resources needed to produce some arbitrary amount of net income. In this case we chose $10,000. For example, to get a net income of $10,000 from beef cattle would require 4.63 times as much spring land as the farmer has available. The left scale of the chart represents the proportions of available resources needed to get a net income of $10,000 from beef cattle. The right scale shows the proportion of available resources needed to produce $10,000 of net income from fall cabbage. If we had to choose one or the other of these enterprises, the choice should be fall cabbage, since the highest dot on the right scale is lower than the highest dot on the left scale. The limiting factor for fall cabbage is September-October labor. To get an income of $10,000 from fall cabbage would require 2.17 times as much September-October labor as the farmer has available. If he used all of his September-October labor on cabbage, his income would be $10,000 divided by 2.17, or $4,608. This is better than he could get from beef cattle alone.

However, this farmer could raise his income by combining beef cattle with fall cabbage. Each of the six lines drawn across the diagram show the proportion of some resource needed for various combinations of beef cattle and fall cabbage. The limiting factor for any combination is indicated by the top line at that point on the horizontal scale. A combination that is mostly beef cattle has as its limiting factor fall land. With combinations including 46 to 91 percent fall cabbage, the limiting factor is production capital. Finally, in combinations that are mostly fall cabbage and only a little beef cattle, the limiting factor is September-October labor. The minimax point (that is, the lowest of the maximum points for any combination) indicates that the most profitable combination of these two enterprises would use about 91 percent of (1) the available production capital and (2) the September-October labor to produce fall cabbage. The other 9 percent of these two limiting factors would be used for beef cattle. To get an income of $10,000 from these combinations would require almost twice as much of the two factors as are available. So the best the farmer could get with these two enterprises would be an income of a little over $5,000.

8/ See King, R. A., and Freund, R. J. A Procedure for Solving a Linear Programming Problem. N. C. Agr. Expt. Sta. Jour. Paper 503, 18 pp. 1953. (Processed.) This study lists 9 different inputs needed to carry on each of 6 different enterprises.

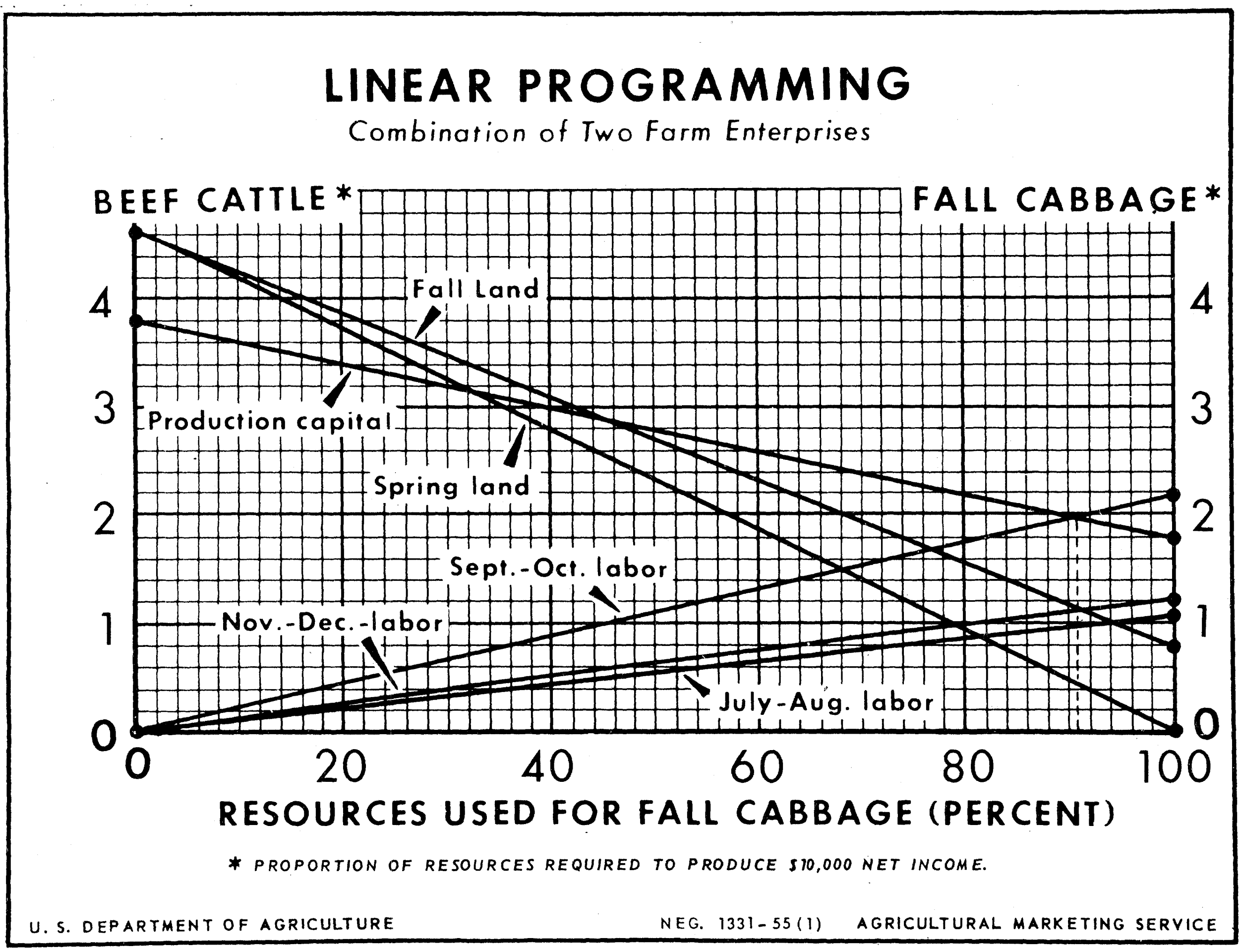


Figure 24

Beef cattle and fall cabbage: Proportion of available resources required to produce $10,000 net income for a farm, North Carolina

Resource	Proportion required	
	Beef cattle	Fall cabbage
Land:		
Spring	4.63	0
Fall	4.63	0.80
Production capital	3.78	1.80
Labor:		
July-August	0	1.08
September-October	0	2.17
November-December	0	1.22

King, R. A. and Freund, R. J. A Procedure for Solving a Linear Programming Problem. N. Ca. Agr. Expt. Sta. Jour. Paper 503. 1953. (Processed.) p. 13.

The Minimum-Cost Dairy Feed

Here is another diagram that is useful in linear programming. In this case, we want the least cost combination of feeds that will meet stated requirements. The prices of several feeds are given; also such requirements as total digestible nutrients and protein. We first compute the proportion of each requirement that could be supplied by $1 worth of corn, $1 worth of oats, and so on. The net result is shown on the table and plotted on the chart.

We then consider combinations of two feeds that will meet two requirements--those for total digestible nutrients and for protein. For $1 we could buy any combination lying along a straight line joining two dots. We have drawn such a line showing combinations of gluten and middlings. A balanced ration would lie on a line through the origin having a slope of 45 degrees. The point at which this line cuts the line connecting the points for gluten and middlings indicates a ration mostly of gluten with a small amount of middlings. It can be shown that this combination will meet the two requirements at less expense than either feed alone. This is true because (1) the line joining the two dots slopes downward to the right and (2) it crosses the 45-degree line. If these two conditions were not met, it would be less expensive to meet the two nutritive requirements from a single feed. Also, this combination is less expensive than any other combination of two feeds that would meet the two nutritive conditions. This is because no dot lies above the line (extended by dashes) joining the dots for gluten and middlings. If there were a dot above this line it would indicate that the cost would be reduced by substituting this feed for one of those in the combination. If the combination of gluten and middlings not only meets the requirements for total digestible nutrients and for protein, but also meets all other requirements, the combination we have found is the final answer--that is, it will meet all requirements at less expense than any other possible combination of feeds. This example is discussed in more detail in an article published in 1951. 9/

The one drawback to this kind of diagram is that we can consider only combinations of two feeds meeting two nutritive requirements. Yet, in many practical cases we want to study combinations of three or more feeds meeting three or more nutritive requirements. The next diagram will show how these principles can be extended to three dimensions--in other words, how we can find the least expensive combination of three feeds meeting three nutritive requirements.

9/ Waugh, Frederick V. The Minimum-Cost Dairy Feed. Jour. Farm Econ. 33:299-310, illus. 1951.

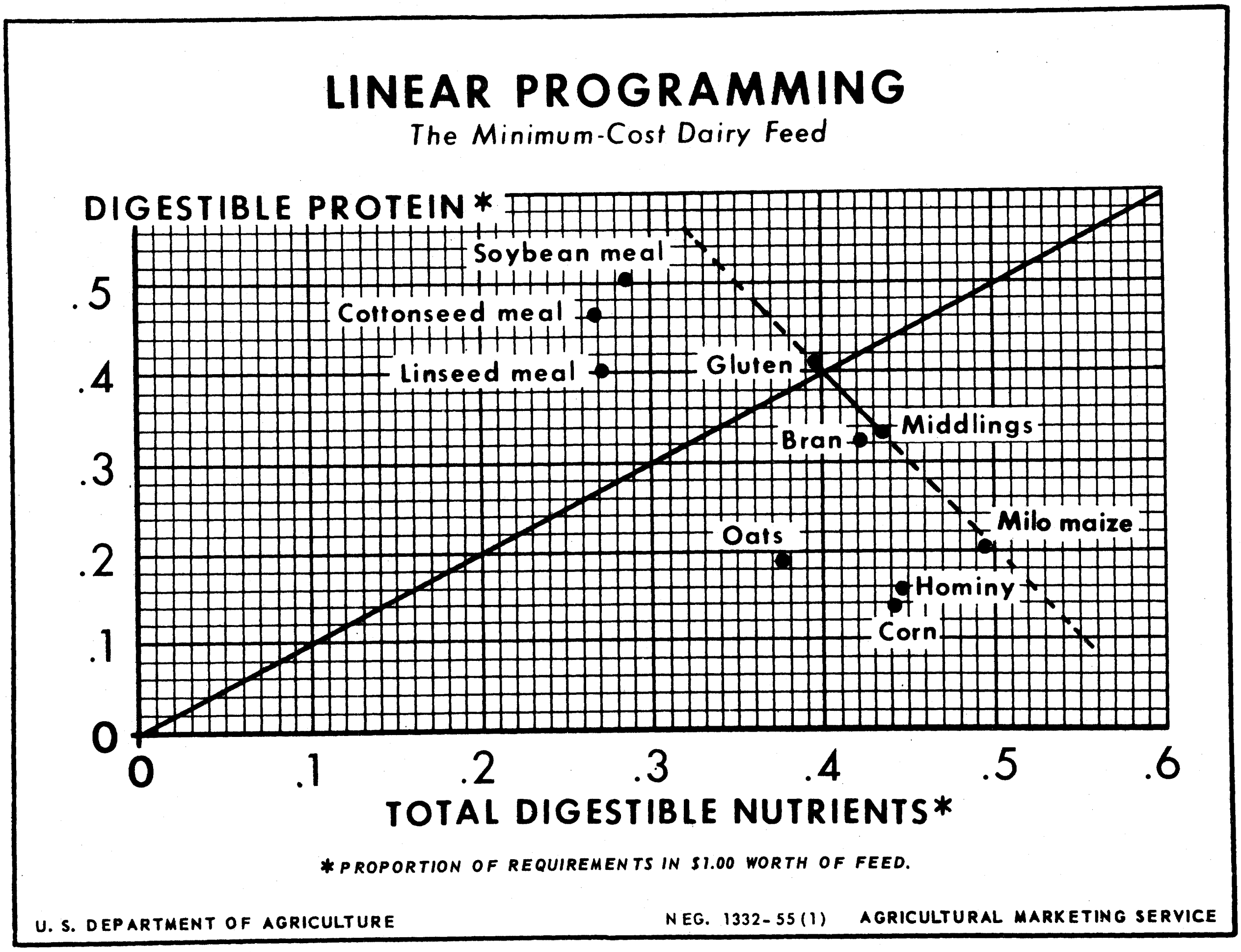


Figure 25

Dairy feed: Proportion of the requirements for protein and total digestible nutrients supplied by $1 worth of each feed

Feed	Proportion supplied: Digestible protein	Proportion supplied: Total digestible nutrients
Corn	0.136	0.441
Oats	.187	.375
Milo maize	.203	.495
Bran	.321	.423
Middlings	.332	.436
Linseed meal	.400	.272
Cottonseed meal	.464	.268
Soybean meal	.504	.286
Gluten	.412	.395
Hominy	.158	.448

Waugh, Frederick V. The Minimum Cost Dairy Feed. Journal Farm Economics. 33:299-307, illus. 1951.

LINEAR PROGRAMMING IN THREE DIMENSIONS

The Minimum-Cost Broiler Feed

The preceding diagram illustrated a method of finding the least-cost combination of two feeds meeting two nutritional requirements. In many practical cases, of course, we are concerned with combinations of more than two ingredients meeting more than two requirements.

In the case of broiler feeds, for example, poultry nutritionists consider about 20 different requirements which must be met by any satisfactory mixed feed. This makes a complete graphic analysis impossible because we cannot graphically portray 20 different dimensions. We can, however, work with at least three dimensions at once, as we have already seen in the chart on page 43. In that chart we were working with a 3-dimensional regression surface. Here our problem is quite different. But we are again working in three dimensions and can use the same kind of box to plot our data.

In this case each hatpin represents some feed ingredient. The hatpins are numbered at the base. For example, the hatpin numbered 1 represents the first feed shown in the table; that is, soybean meal. The head of each hatpin shows the percentage of the required amounts of protein, productive energy, and non-fiber that could be bought by one dollar's worth of some ingredient. For example, hatpin number 1 shows that one dollar's worth of soybean meal would buy 3.6 percent of the requirements of protein, 1.4 percent of the requirements of productive energy, and 1.63 percent of the requirements of non-fiber. Similarly, for the other hatpins. Each is identified by the same number as in the table.

Now, look at the wire which runs diagonally across the diagram. This wire corresponds to the 45 degree line in the chart on page 51. Any point on this wire would be a balanced ration; that is, it would have equal percentages of protein, productive energy, and non-fiber. The object of our programming is to get a feed which will have 100 percent of each of these requirements--and also 100 percent, or more, of the other requirements which are not considered here. To get 100 percent of all these requirements would take us far outside the scope of this diagram. However, we want to find a combination of feeds which will get us as far up on the wire as possible.

The optimum in this case is a combination of feeds 5, 10, and 13. These are meat and bone scrap, corn, and hominy. To show that this is the optimum combination, we have laid a plexiglass plane on top of the hatpins representing feeds 5, 10, and 13. (For $1 we could buy any combination of these three feeds lying on the plane represented by the plexiglass.) A balanced combination of these three feeds would be at the point where the plane is cut by the wire. At that point, we can get almost 2 percent of each requirement for $1. In other words we can get 100 percent of each requirement for a little more than $50. It is fairly easy to see graphically, even in a photograph, that this combination of feeds can get us higher up on the wire than could any other combination.

One further point should be noted. The hatpin near the top of the wire does not represent a feed ingredient. It simply represents the point at which an imaginary feed would have two percent protein, two percent productive energy, and two percent non-fiber. There is no such feed. It is plotted here only to locate the wire which is needed to find a balanced mixture.

The geometry of this problem could be discussed in more detail. For present purposes perhaps it is enough to indicate that for the solution to be an optimum the plane must slope downward in both directions. It does so in this case.

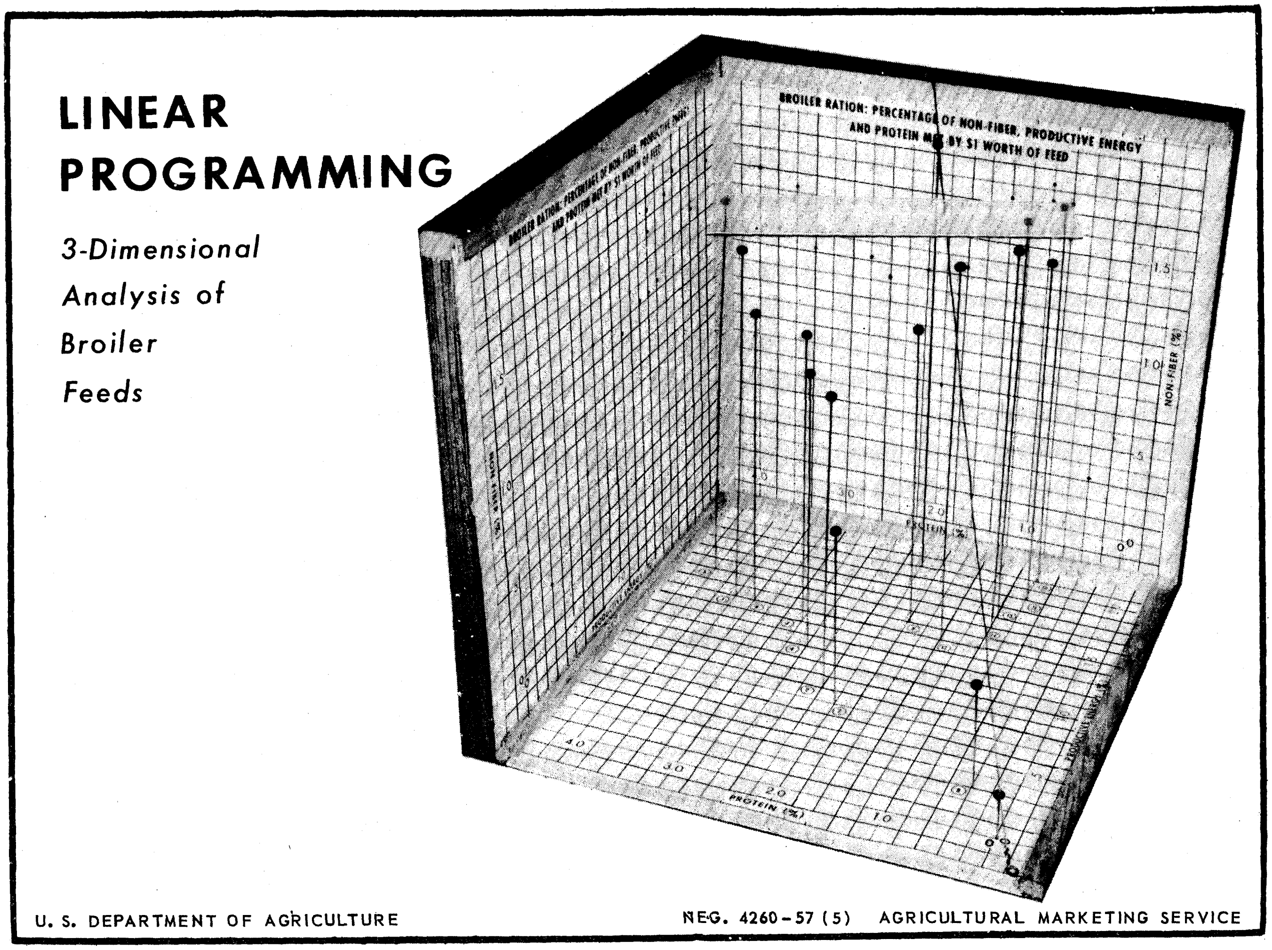


Figure 26

Broiler ration: Percentage of non-fiber, productive energy and protein met by $1 worth of feed, specified feeds

Number on chart	Feed	Percentage of required nutrients		
		Non-fiber	Productive energy	Protein
		Percent	Percent	Percent
	Meal:			
1	Soybean, 44 percent	1.63	1.4	3.6
2	Linseed, solvent	1.30	.8	2.2
3	Cottonseed, expeller	1.33	1.3	2.8
4	Gluten	1.24	1.1	2.6
5	Meat and bone scrap	1.70	1.6	4.0
6	Meat scrap	1.33	1.4	3.4
7	Fish meal, menhaden	.69	.7	2.0
8	Buttermilk, dried	.32	.3	.4
9	Corn distillers solubles, dried	1.35	1.5	1.8
10	Corn	1.75	2.1	.8
11	Milo	1.57	1.9	.8
12	Wheat, standard middlings	1.74	1.4	1.4
13	Hominy feed, yellow	1.83	1.7	1.0
14	Barley	1.75	1.6	1.0

AVERAGES

Gross Profit from Storage

In economic analysis we often want to compute the average of two or more points on a curve.

In this diagram the curve represents total returns to growers from sales of various amounts of eggs. In deriving these figures, allowance was made for the effect of disposable income on prices of eggs. The prices shown are those that might have been expected with income at its average level for the period 1940-48. A practical question is whether it would be profitable to store up the surplus in periods of large production and to sell it in periods of small production.

Suppose we produced 40 billion eggs in one period and 50 billion in another period. The returns for each period would be shown on the curve. The average for the two periods would be halfway between these two points. This average is indicated by the dot at the midpoint on the straight line joining the appropriate points on the curve. In this case it indicates a moderate gross profit from storage. That is, the gross income from selling 45 billion eggs in each period would be greater than the average income from selling 40 billion in the first period and 50 billion in the second. Costs of storage, handling, and any loss in quality would have to be deducted in order to determine whether net returns would be larger from storage.

It is easy to see that there will be a gross profit from storage if, and only if, the returns curve is concave downward. The degree of curvature is an important indication of the possible amount of gross profit.

Of course, this is only one of the many uses of averages. The economist-statistician often wants to compute average prices, average cost, average yield of a crop, and so on. When working with graphic diagrams, such averages can be computed graphically with little time or trouble. There is no need to read the numbers from the diagram, copy them on a piece of paper, add them, divide by two, and put the average back on the diagram. The simple arithmetic average of any two points on any curve can be located graphically by the graphic method explained here.

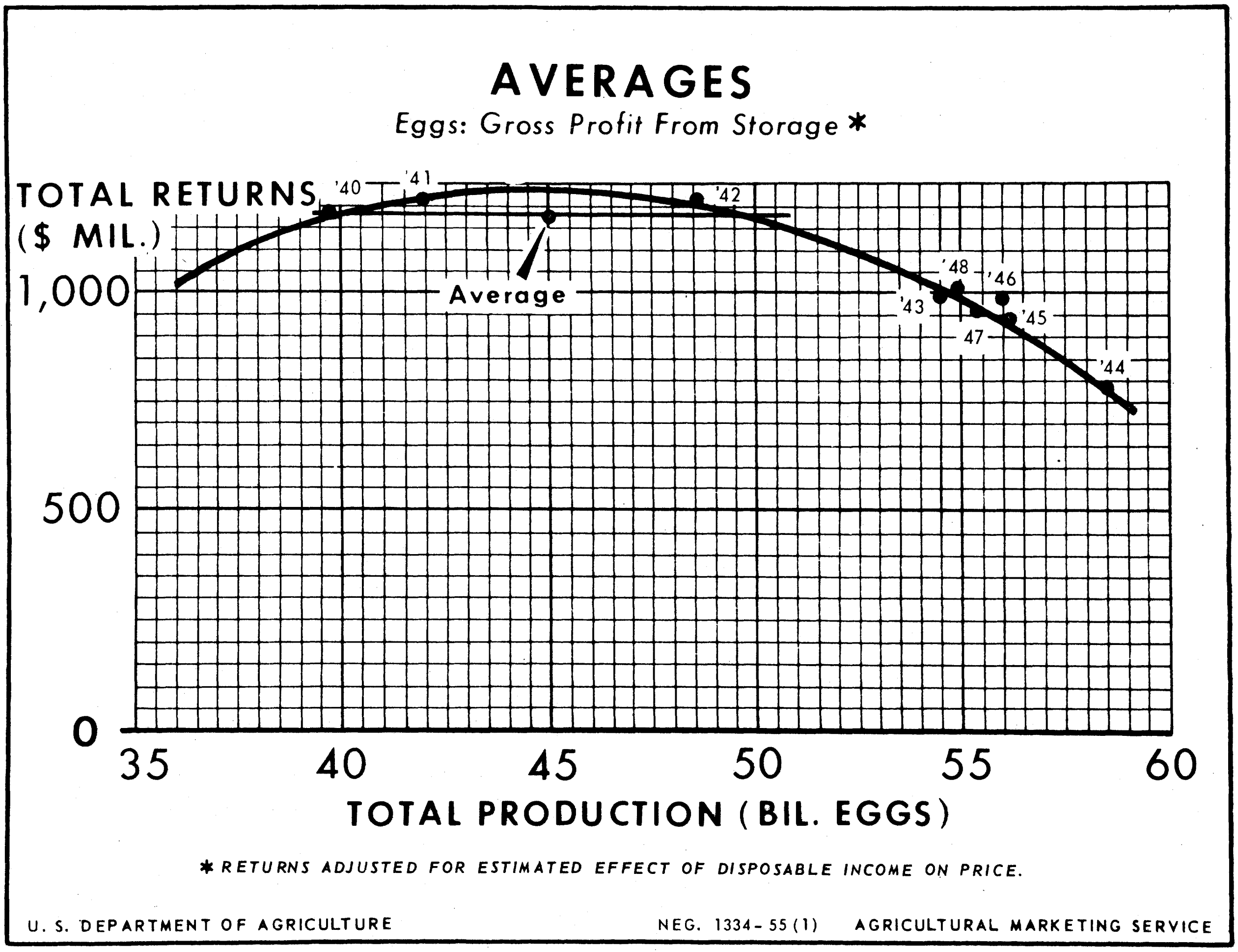


Figure 27

Eggs: Production, price per dozen received by farmers, and total returns, 1940-48

Year	Production	Price 1/	Total returns 1/
	Billions	Cents	Million dollars
1940	39.7	36	1,188
1941	41.9	35	1,225
1942	48.6	30	1,230
1943	54.5	22	990
1944	58.5	16	784
1945	56.2	20	940
1946	56.0	21	987
1947	55.4	21	966
1948	54.9	22	1,012

1/ Adjusted for estimated effect of disposable income on price.

Data derived from Figure 92 in Thomsen, Frederick L., and Foote, Richard J. Agricultural Prices. New York. 1952. p. 431.

ELASTICITY

Coefficient of Elasticity of Demand

Many economists have trouble with coefficients of elasticity. They are frequently concerned with the elasticity of demand--more precisely, with the elasticity of consumption with respect to price. The diagram shows how this can be measured graphically.

The curved line on the diagram represents an assumed demand curve for eggs. The scales for consumption and prices would not need to be shown. They are unimportant, because the coefficient of elasticity is invariant to changes in scale provided that the axes start at the origin. Suppose we want the coefficient of elasticity at the point (p=a, q=c). We draw the indicated straight line tangent to the demand curve at that point. The elasticity in question is -a/b. For this example, this equals -35.5 divided by 64.5 based on the scales shown. In terms of small squares on the grid, this equals -17.75 divided by 32.25. Either computation indicates an elasticity of -0.55.

This piece of graphics comes from Alfred Marshall. 10/ It derives from the definition of elasticity $\eta = \frac{dq}{dp} \cdot \frac{p}{q}$. Note that $\frac{dq}{dp} = -\frac{c+d}{a+b}$, and (by similar triangles) $\frac{c+d}{a+b} = \frac{c}{b}$. Also p=a, and q=c. So $\frac{dq}{dp} \cdot \frac{p}{q} = \frac{-c}{b} \cdot \frac{a}{c} = -a/b$.

Some economists have found the concept of elasticity so difficult that they have used "arc elasticity," or the "average elasticity of a curve." If the graphic approach to elasticity is used, there is little need for such concepts. The elasticity coefficient shown here is exact and easy to compute.

We should note that the concept of elasticity applies not only to demand curves--but to any curve. When we speak of the elasticity of demand we (usually) mean the elasticity of consumption with respect to price. But we might want the elasticity of cost of producing potatoes with respect to the amount of fertilizer used, for example. Whatever the curve, we can measure its elasticity at any point, using the same graphics as shown here.

10/ Marshall, Alfred. Principles of Economics. Ed. 8, pp. 102-103. New York. 1948. First published 1920.

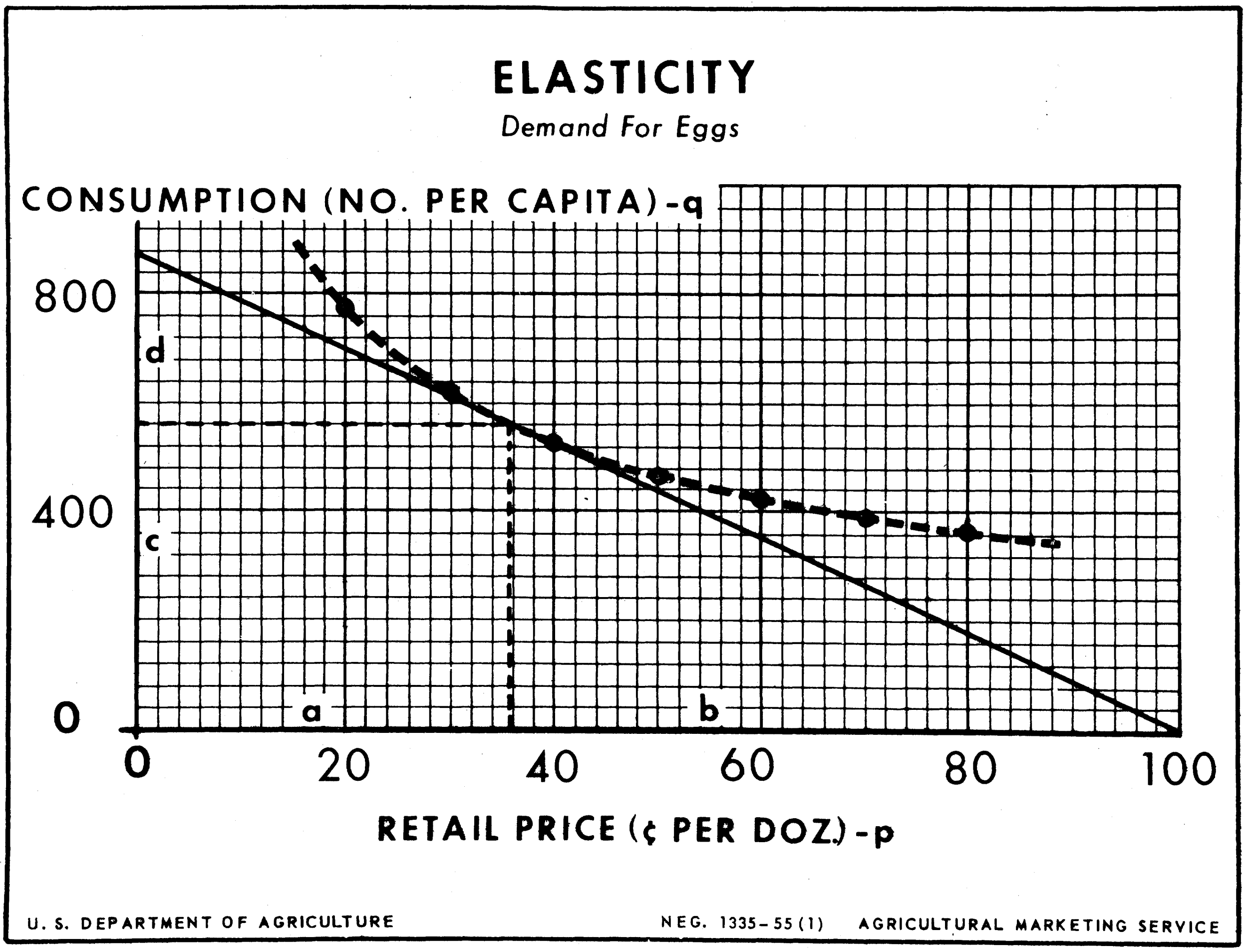


Figure 28

Eggs: Consumption per capita associated with given retail price per dozen

Consumption	Price
Number	Cents
777	20
621	30
530	40
469	50
424	60
390	70
362	80

Based on an assumed elasticity of demand coefficient of -0.55. See Foote, Richard J. and Fox, Karl A. Analytical Tools for Measuring Demand. U. S. Dept. Agr. Agr. Handbook 64. 1954. p. 40.

METHOD OF DETERMINING MOST PROFITABLE OUTPUT

Applications of Nitrogen on Corn

The chart on page 23 shows, based on experimental results, the relationship between observed yields of corn and the amount of nitrogen applied per acre. The discussion of that chart promised that we would discuss later a graphic analysis of the most profitable rate of fertilizer application. The chart facing this page is an attempt to carry out that promise. Several persons who read the first edition of this handbook suggested that this type of graphics be included. One of the men who suggested this was George G. Judge of Oklahoma Agricultural and Mechanical College. Professor Judge suggested the general type of chart shown on the facing page.

The dots and the heavy curve shown on this chart are the same as those shown on page 23, except for a change in the scale for corn yields. The straight line OA at the bottom of the chart shows the number of bushels of corn required to pay for various amounts of nitrogen when one bushel of corn will buy 10 pounds of nitrogen. This would be the case, for example, if a farmer could sell corn for $1.50 a bushel and could buy nitrogen for 15 cents a pound.

The most profitable rate of nitrogen application with the given price ratio is determined by drawing a tangent to the input-output curve parallel to line OA. To do this, we lay one edge of a transparent triangle along line OA, place a straightedge along one of the other sides of the triangle, and slip the triangle along the straightedge until the edge that was touching line OA now just touches the input-output curve. Then that edge of the triangle will be parallel to line OA and we draw the line indicated.

At this point the farmer would buy about 160 pounds of fertilizer to the acre and would expect to get a yield of about 146 bushels. If he used more fertilizer than this, it would cost him more than the value of the additional corn. If he used less fertilizer, the saving in fertilizer would be less than the value of the additional corn.

A similar analysis could be made, of course, with any assumed prices of corn and nitrogen. Changes in the price ratio would change the slope of line OA. Therefore, they would change the slope of the tangent and would move the point of contact between the tangent and the input-output curve. In this particular case, however, there would be little change in the most profitable rate of nitrogen application unless the price of nitrogen were greatly increased or the price of corn were greatly decreased. This is because the input-output function curves vary sharply in the neighborhood of the point that is most profitable with the price assumption shown on the chart. Assuming that our input-output function is approximately right, applications of much more than 160 pounds to the acre would obviously be unprofitable since the maximum corn yield is apparently attained with a fertilizer application of a little over 160 pounds of nitrogen.

This general type of chart is useful in analyzing a wide variety of economic problems. It is not limited to fertilizer applications, but applies just as well to such problems as the most profitable amount of concentrates to feed to dairy cows, or the most profitable amount of labor to be used in any operation whether on the farm or in marketing.

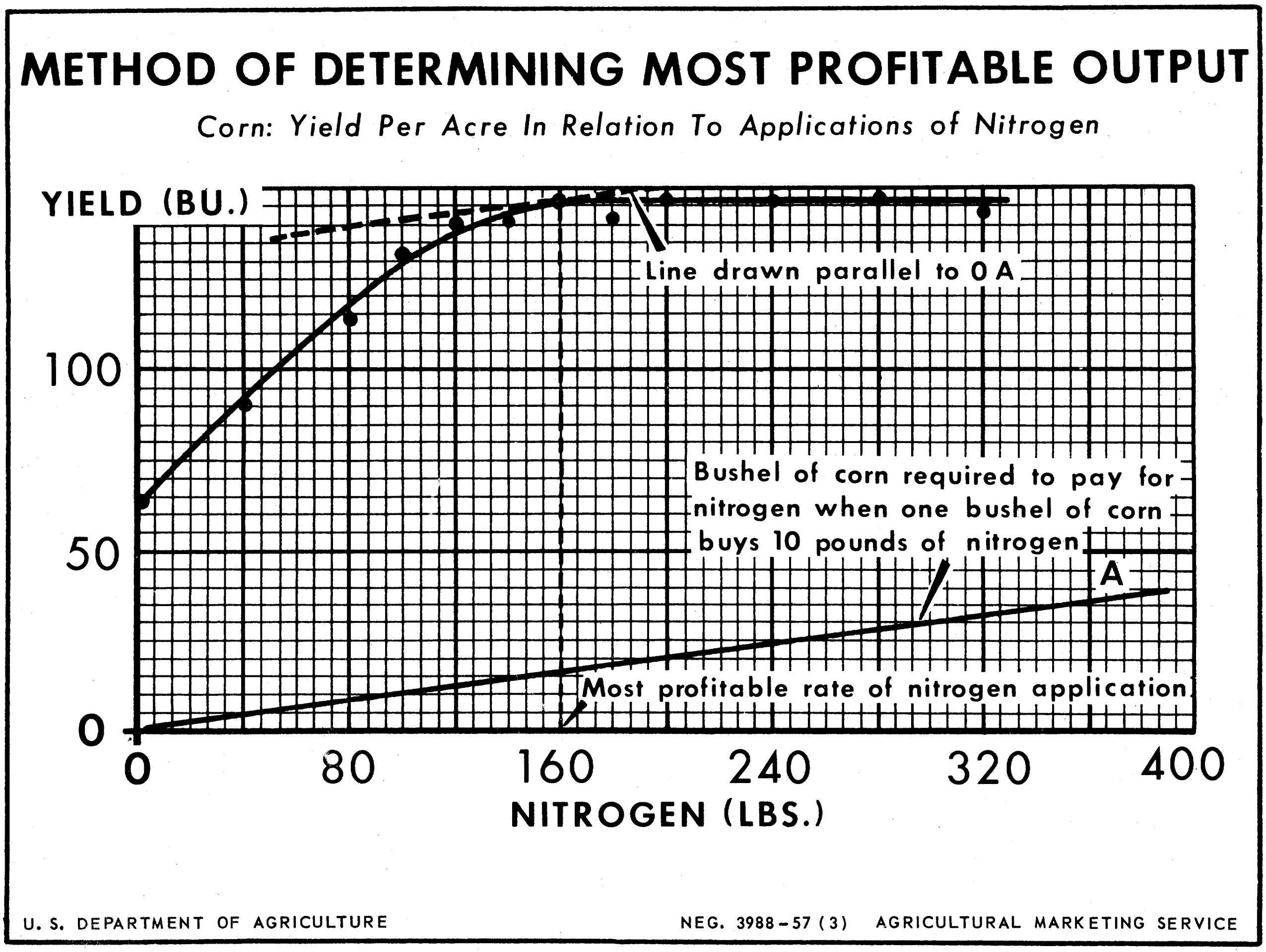


Figure 29

Corn: Yield per acre by specified quantity of nitrogen applied, Ontario, Oregon

Nitrogen applied	Yield of corn	Nitrogen applied	Yield of corn
Pounds	Bushels	Pounds	Bushels
0	64.6	160	146.8
40	90.4	180	141.2
80	118.2	200	147.1
100	132.4	240	145.8
120	140.7	280	147.4
140	141.0	320	143.8

Paschal, J. L., and French, B. L. A Method of Economic Analysis Applied to Nitrogen Fertilizer Rate Experiments on Irrigated Corn. U. S. Dept. Agr. Tech. Bull. 1141. 1956. p. 16.

DIFFERENTIATION

Some statisticians and economists find calculus a difficult subject. Differential calculus is relatively easy if you do it graphically. The differential at any point on a curve is simply the slope of a tangent drawn at that point. The tangent can be drawn easily with a transparent straightedge. The differential, $\frac{dy}{dx}$, is the slope of this tangent.

In figure 30, the slope of the straight line is 5.8 (that is, y increases 5.8 units for each increase of one unit of x). In figure 31, the slope is -0.002 (that is, y decreases 0.002 units for each increase of one unit of x). Thus, in figure 30, $\frac{dy}{dx} = 5.8$, and in figure 31, $\frac{dy}{dx} = -0.002$.

These differentials can be read most easily by drawing the dotted lines shown on the diagrams. These dotted lines are drawn parallel to the tangent and through the origin (the point x=0, y=0). To draw these parallel lines, place one side of a right triangle along the original curve, place a straightedge along another side of the triangle, and then slip the triangle along the straightedge. With a little practice it is very easy to draw parallel lines.

The slope of the tangent is the same as the slope of the dotted parallel line. It is measured by the height of the dotted line corresponding with one unit on the x-axis. In figure 30 it is 5.8. In figure 31 it would not be possible to read the height of the dotted line corresponding to one unit on the x-axis. So we read the height corresponding to 1,000 units. It is -2. So the slope is -2/1,000 or -0.002.

Graphic differentiation is quick and easy. It is important in any sort of marginal analysis.

We have not given data for these charts as the curves are purely hypothetical and are shown merely to illustrate the method.

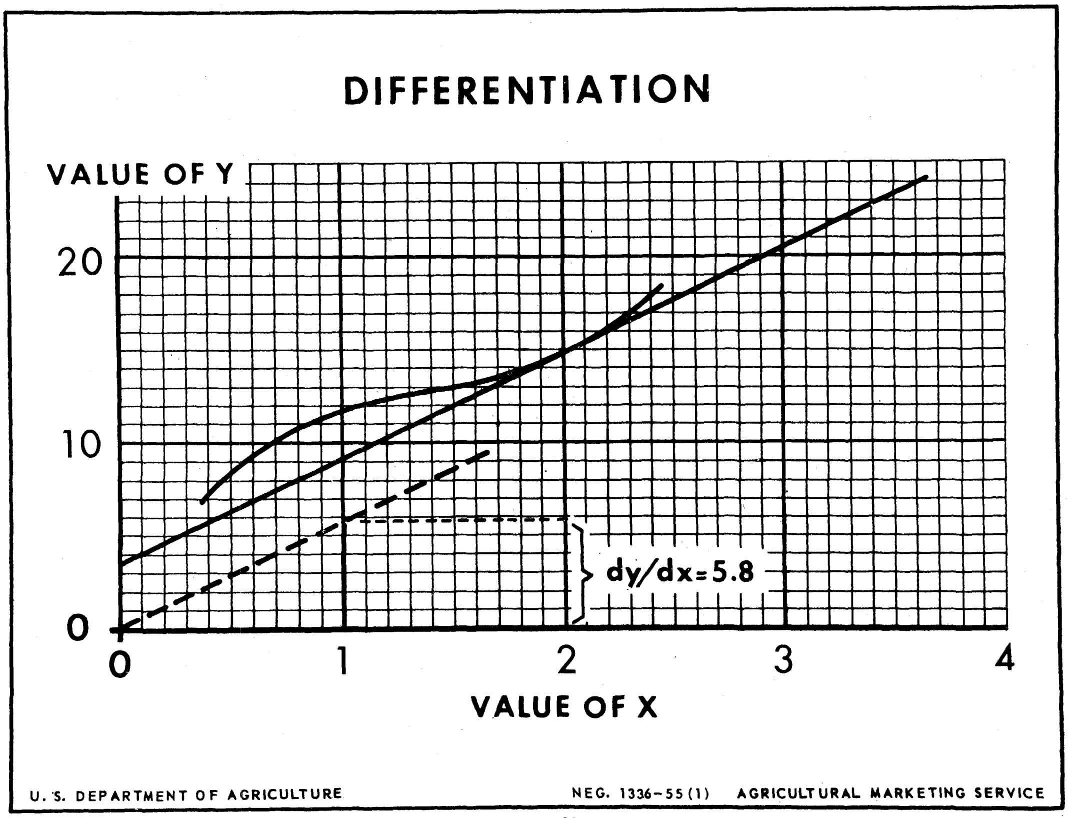


Figure 30

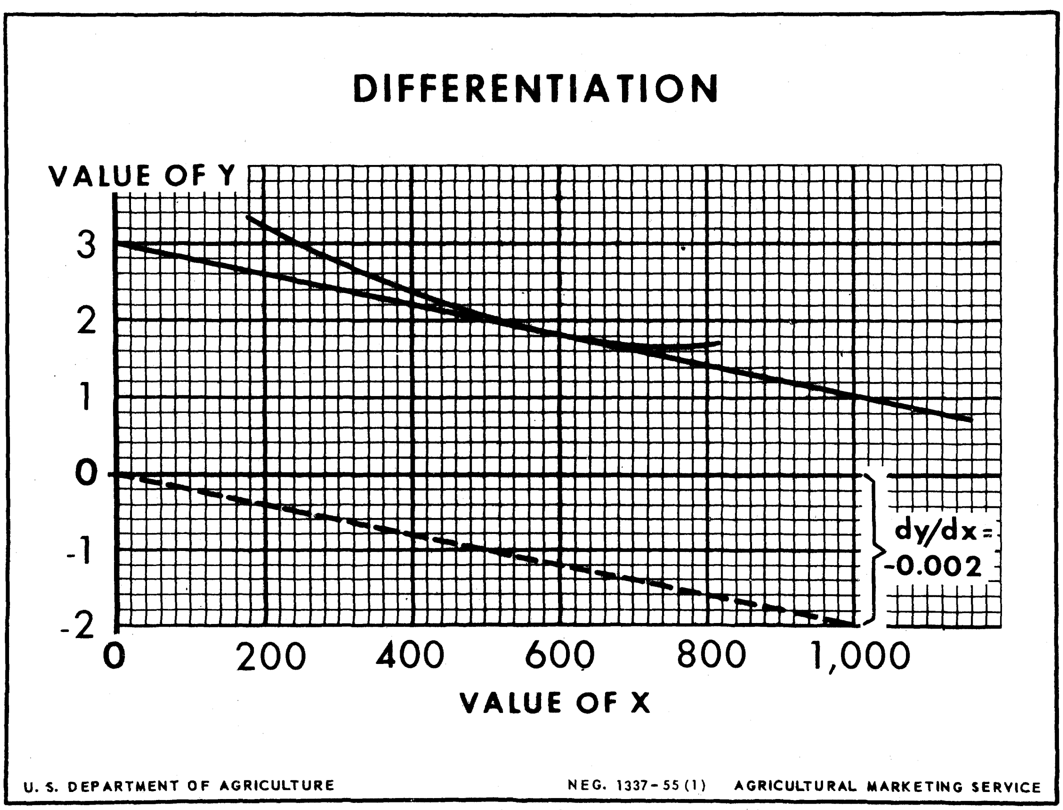


Figure 31

DERIVING A MARGINAL CURVE FROM AN AVERAGE CURVE

Marginal Returns

Often the economist has a demand curve showing estimates of average prices corresponding with a range of quantities sold. His problem may call for an analysis of marginal returns (or marginal expenditures of consumers). The easiest way to do this is to find graphically several points on the returns curve.

Robinson 11/ explained the geometry of this. Briefly, total returns are $R = pq$. We want $\frac{dR}{dq} = p + \frac{dp}{dq} q$. We can take any point on the demand curve, such as point A in our diagram (32 pounds, at an average price of 41 cents), and draw a tangent to the curve at that point. We then draw a line parallel to the tangent such that it cuts the price axis at the price indicated by the point on the demand curve (that is, at 41 cents). This parallel cuts a perpendicular dropped from A at point B, and the price equivalent of B measures the marginal returns corresponding to the quantity sold at point A on the demand curve. Here marginal returns are 10.5 cents when 32 pounds per capita are sold.

This is a simple process and can be done in five seconds. With a little practice you can quickly locate several points on the marginal returns curve, and then draw the whole curve.

11/ Robinson, Joan. The Economics of Imperfect Competition, p. 30. London. 1933.

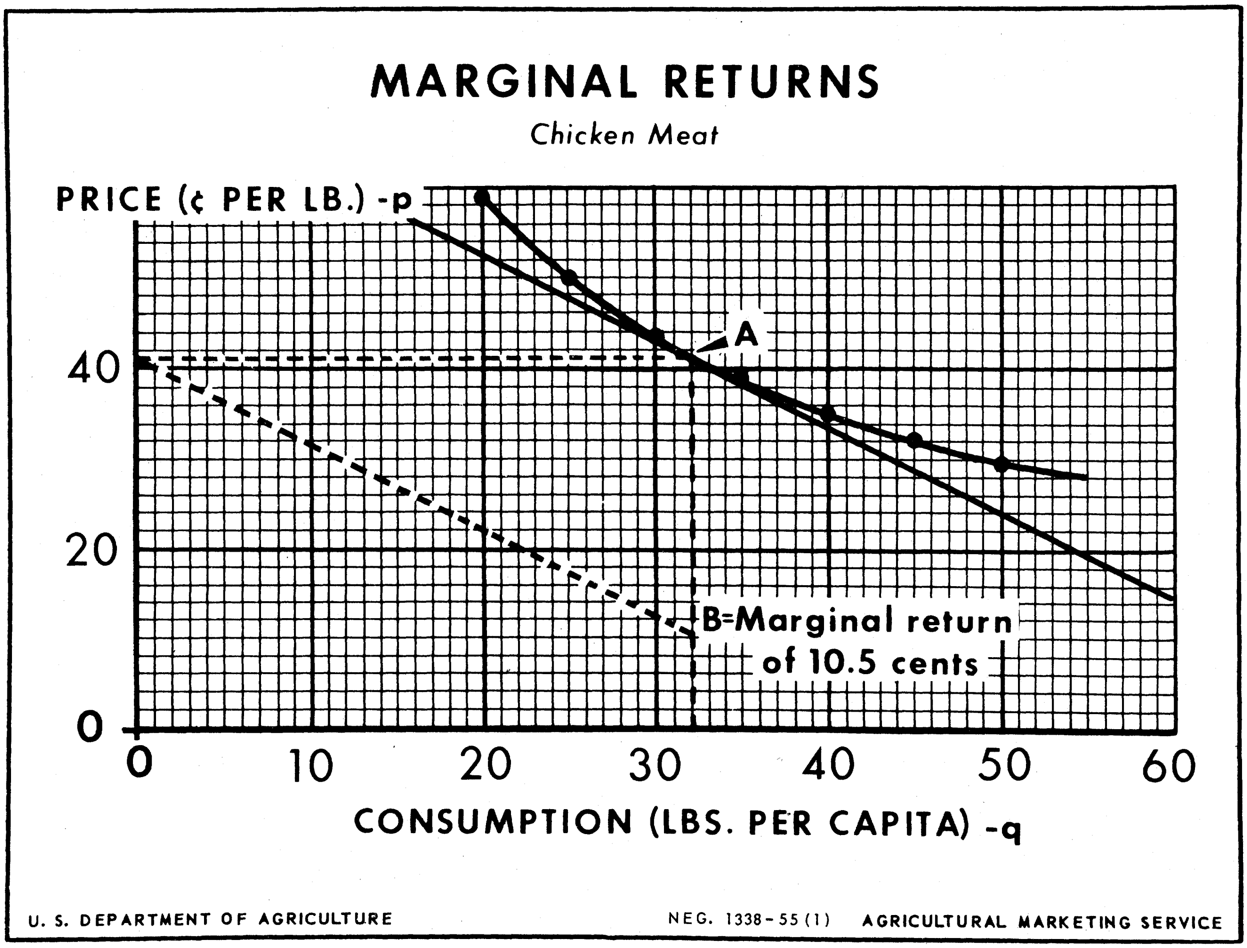


Figure 32

Chicken meat: Price per pound at retail associated with given levels of consumption per capita

Price	Consumption
Cents	Pounds
59.1	20
49.9	25
43.5	30
38.8	35
35.1	40
32.1	45
29.7	50

Regression coefficient based on the reciprocal of an assumed elasticity of demand coefficient of -1.33. See Foote, Richard J. and Fox, Karl A. Analytical Tools for Measuring Demand. U. S. Dept. Agr. Agr. Handbook 64. 1954. p. 40.

Marginal Costs

A marginal cost curve can be obtained from a curve of average costs by the same graphic procedure as that just explained for marginal returns. This process is illustrated in the diagram. In this instance, the marginal curve will be above the average curve. To find the marginal cost at point A in the diagram, we erect a perpendicular line at point A and draw a tangent to the average cost curve at this point. We also draw a horizontal line from point A to the cost axis and note the point at which this line cuts the axis. We then draw a line through this point that is parallel to the tangent. The cost at which this line cuts the perpendicular line is the marginal cost for the input represented by point A.

In the example used here, we show average costs of land and fertilizer per unit of output for given inputs of fertilizer applied to an acre of land. Point A applies to slightly more than $6 worth of fertilizer. For this amount, average costs per unit of output are about $0.237. Marginal costs, as indicated by B, are $0.292. As in the preceding example, several points on the marginal curve can be located as a basis for drawing the entire curve.

If xy is given as a fraction of x, as in these examples, we always can compute

$$\frac{d\,xy}{d\,x} = y + \frac{dy}{dx}\,x$$

by this process no matter what x and y represent.

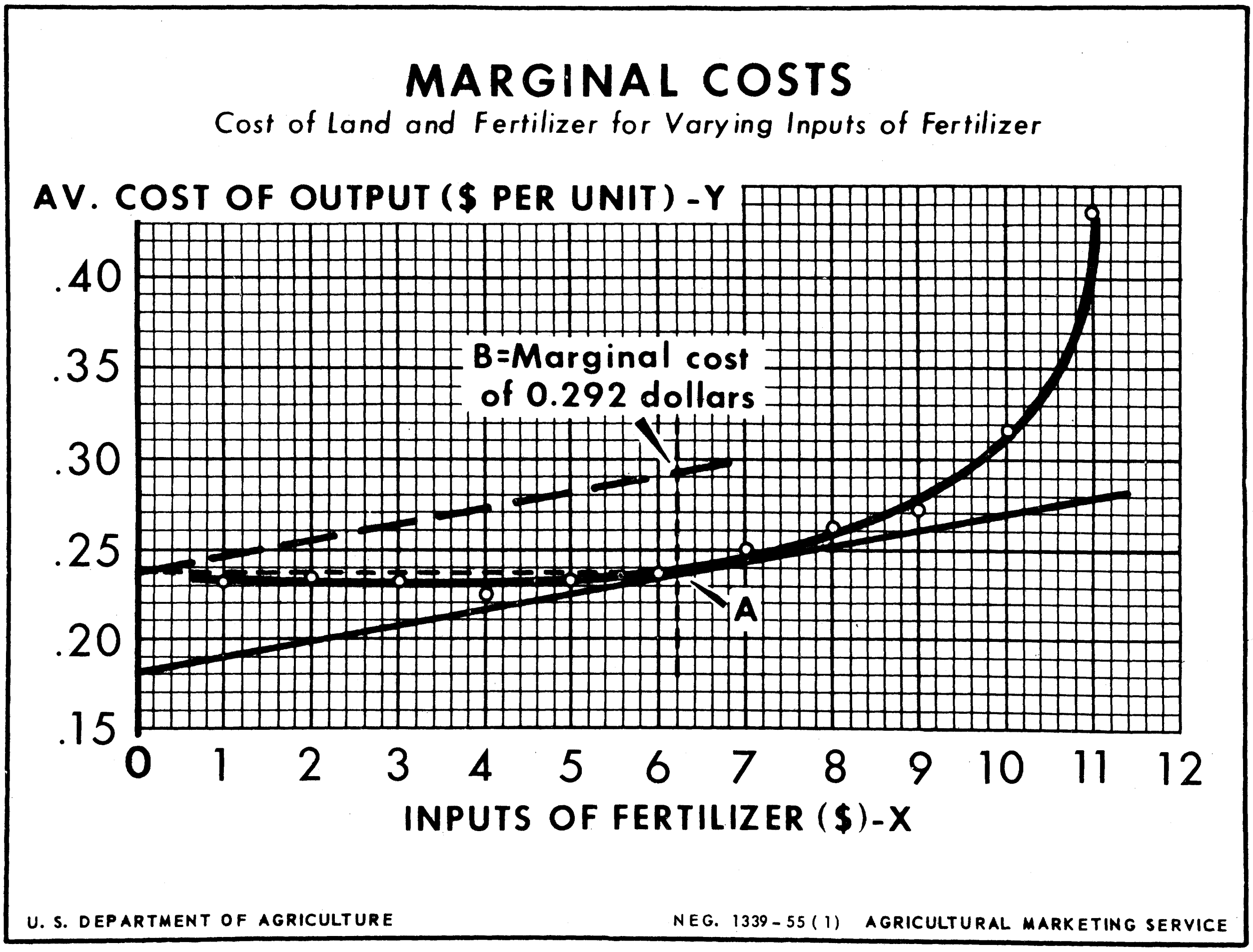


Figure 33

Total output of a given crop and cost per unit of output for given inputs

Cost of total input		Total output	Cost per unit of output for--		
Land	Fertilizer		Land	Fertilizer	Total
Dollars	Dollars		Dollars	Dollars	Dollars
10	1	47	0.213	0.0213	0.2343
10	2	51	.196	.0392	.2352
10	3	56	.178	.0536	.2316
10	4	62	.161	.0645	.2255
10	5	64	.156	.0781	.2341
10	6	67	.149	.0895	.2385
10	7	68	.147	.1030	.2500
10	8	69	.145	.1159	.2609
10	9	70	.143	.1287	.2717
10	10	64	.161	.1562	.3172
10	11	48	.208	.2294	.4374

Black, John D. Production Economics. New York. 1926. pp. 317-318.

ROOTS OF A POLYNOMIAL

$$x^3 - 1.2240\,x^2 + 0.3695\,x - 0.0183 = 0$$

Some of my good friends, including Professor Charles H. Merchant, of the University of Maine, think it is out of place to discuss the roots of a polynomial in a handbook dealing with graphic analysis. Roots of polynomials are used mainly in high-powered mathematical studies dealing with such things as canonical regression, component analysis, and cyclical variation. But graphics can help, even in these studies. So other friends have induced me to leave this piece in the handbook.

We have included this diagram to illustrate the use of graphics in connection with more elaborate mathematical techniques. The particular polynomial is taken from Tintner. 12/ Tintner was dealing with a problem of canonical regression. The largest root of the above equation indicates the squared correlation coefficient. We shall not bother to explain how the equation was obtained. We are concerned only with computing its roots--and especially its largest root.

The roots of a polynomial are values of x which satisfy the equation. There are many mathematical tricks for discovering such values of x. But the graphic method illustrated here is practical and easy.

We simply plot several values for x. Thus if x=0, the polynomial equals -0.0183; so we plot y= -0.0183 corresponding to x=0. If x=0.1, the polynomial equals 0.0074; so we plot y=0.0074 corresponding to x=0.1. We proceed to compute several points on the curve, $y = x^3 - 1.2240x^2 + 0.3695x - 0.0183$. When we have enough points, we draw a curve through them. Wherever this curve crosses the x-axis, it indicates a real root. In this case, the roots are approximately 0.06, 0.38, and 0.78. The canonical correlation is approximately equal to the square root of 0.78.

We could locate any of these roots more exactly by blowing up the part of the diagram near the root. Thus, we could draw a new diagram for the part of the curve between x=0.76 and x=0.80, plot the curve on a blown-up scale, and compute the largest root more accurately. This could be repeated until we obtained as many significant figures as wanted.

As a guide to the parts of the curve that must be plotted, we know that there must be as many roots as the degree of the curve. Here we have a third-degree polynomial, so we know that there must be three roots. Once we have located them, our job is finished. Sometimes we have multiple roots (that is, two or more roots at a single point) or imaginary roots. These also can be located by graphic means but these topics are beyond the scope of this handbook.

12/ Tintner, Gerhard. Econometrics. New York. 1952. Taken from equations (18) on p. 119, letting $x = \lambda^2$.

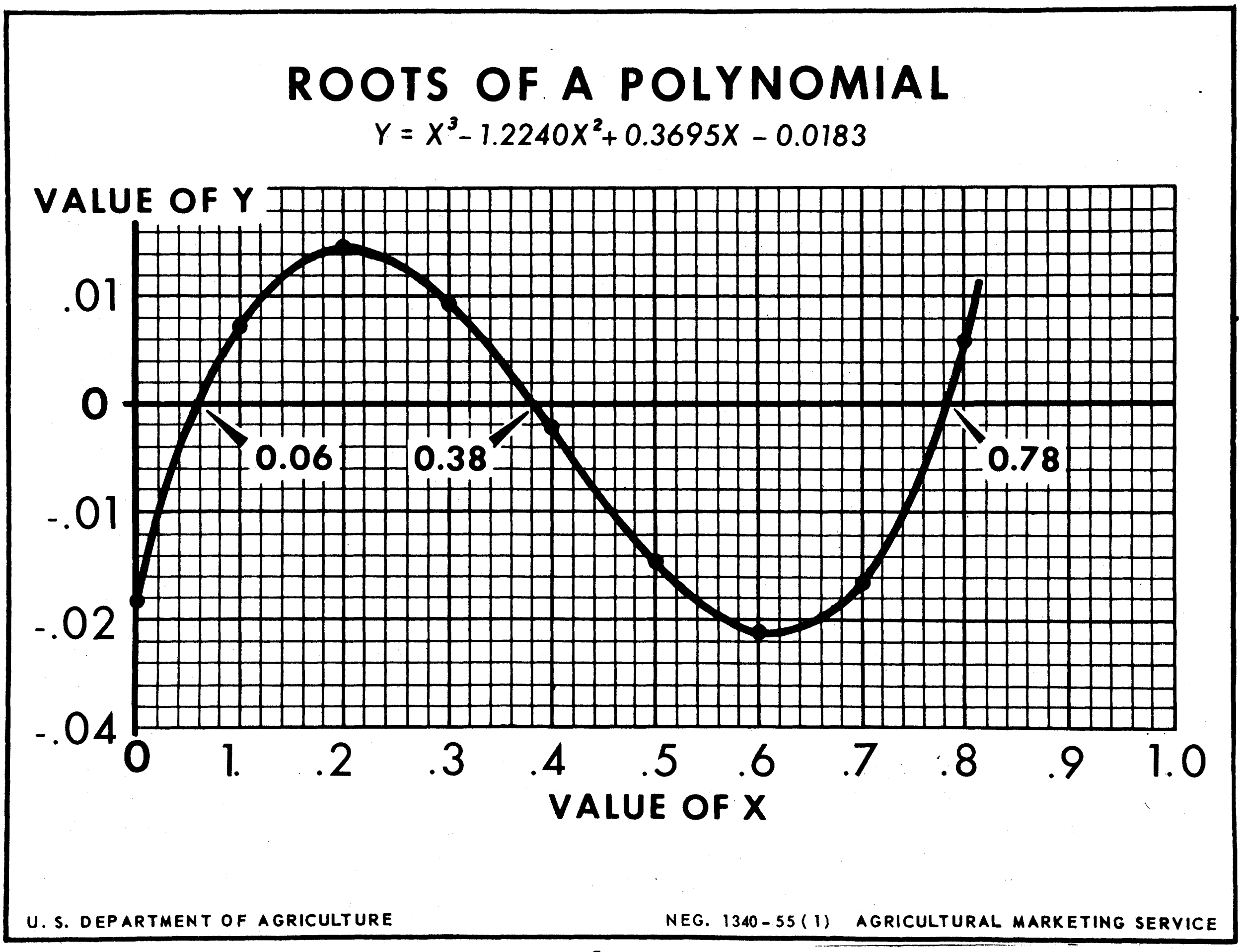


Figure 34

Values of a third-degree polynomial, y, at specified levels of x 1/

x	y
0	-0.0183
.1	.0074
.2	.0146
.3	.0094
.4	-.0023
.5	-.0146
.6	-.0212
.7	-.0164
.8	.0059

1/ $y = x^3 - 1.2240\ x^2 + 0.3695\ x - 0.0183$.

Data compiled using equations (18) as a basis and letting $x = \lambda^2$. Tintner, Gerhard. Econometrics. New York. 1952. p. 119.

SOLUTION OF SIMULTANEOUS EQUATIONS

Determining the Supply Below Which No Grain Should Be Stored in An Optimal Storage Program

We end this handbook with another use of graphics as an aid to mathematical computation. Statisticians often must solve two or more equations simultaneously. Various methods of solution are available, including the popular Gauss-Doolittle technique. But the equations can also be solved graphically. The diagram illustrates only the solution of pairs of equations. It is possible to solve any number of equations graphically by a process very similar to the Gauss-Doolittle method. But we shall not explain the procedure here. 13/

To solve any pair of equations, we substitute several successive values of x in each equation, compute the corresponding values of y, plot the value of y corresponding to each value of x, and draw a smooth curve through the observations. When these operations are performed for each equation, this gives us a pair of curves. Wherever the two curves cross one another, there is a solution of the two equations.

Any pair of linear equations will have one, and only one, real solution--except in the extreme case where the two lines are identical or parallel, where there are infinitely many or no solutions, respectively. Quadratic equations have up to four solutions to a pair of equations, depending on how they are situated one to another. For equations of any degree, solutions are real wherever the curves cross one another; otherwise they are imaginary.

Gustafson, in an unpublished manuscript 14/, outlines some methods for determining storage rules which are optimal in terms of certain economic criteria. These optimal rules can be obtained exactly by mathematical solutions that involve the use of calculus. However, a method is outlined by which approximate rules can be obtained by carrying out certain essentially arithmetic operations. One of the necessary computations requires the obtaining of a value for k, which represents the supply below which no grain should be stored. To obtain this value of k, we must find a solution for two curves which show the relation between k and another variable, L. One of these curves is obtained by a tabular method by which values of L are computed for given values of k. Results of this tabulation are shown by the curved line on the facing chart. The second relation is a linear one in which values of k are computed from the values of L obtained by the tabular method. The formula used is shown on the chart. The desired value for k is obtained from the intersection of these two curves. The chart shown here suggests a value for k of approximately 31.0 bushels per acre.

Gustafson suggests that instead of drawing the complete curves, the range within which the solution lies be determined by inspection so that only a pair of values for k and L respectively is involved. A greatly enlarged graph is then drawn, and the respective points connected by a pair of lines. When this method was used, a value for k of 31.01 bushels was given. In this instance a graphic solution was the only feasible method, since the mathematical formula for the curve plotted from tabular values was not known.

13/ The method is described in Maxfield, John E. and Waugh, Frederick V. A Graphic Solution to Simultaneous Linear Equations. Math. Tables and Other Aids to Computations, 5:246-248, illus. 1951.

14/ Gustafson, Robert L. Optimal Storage Rules for Grains, unpublished manuscript, 1957.

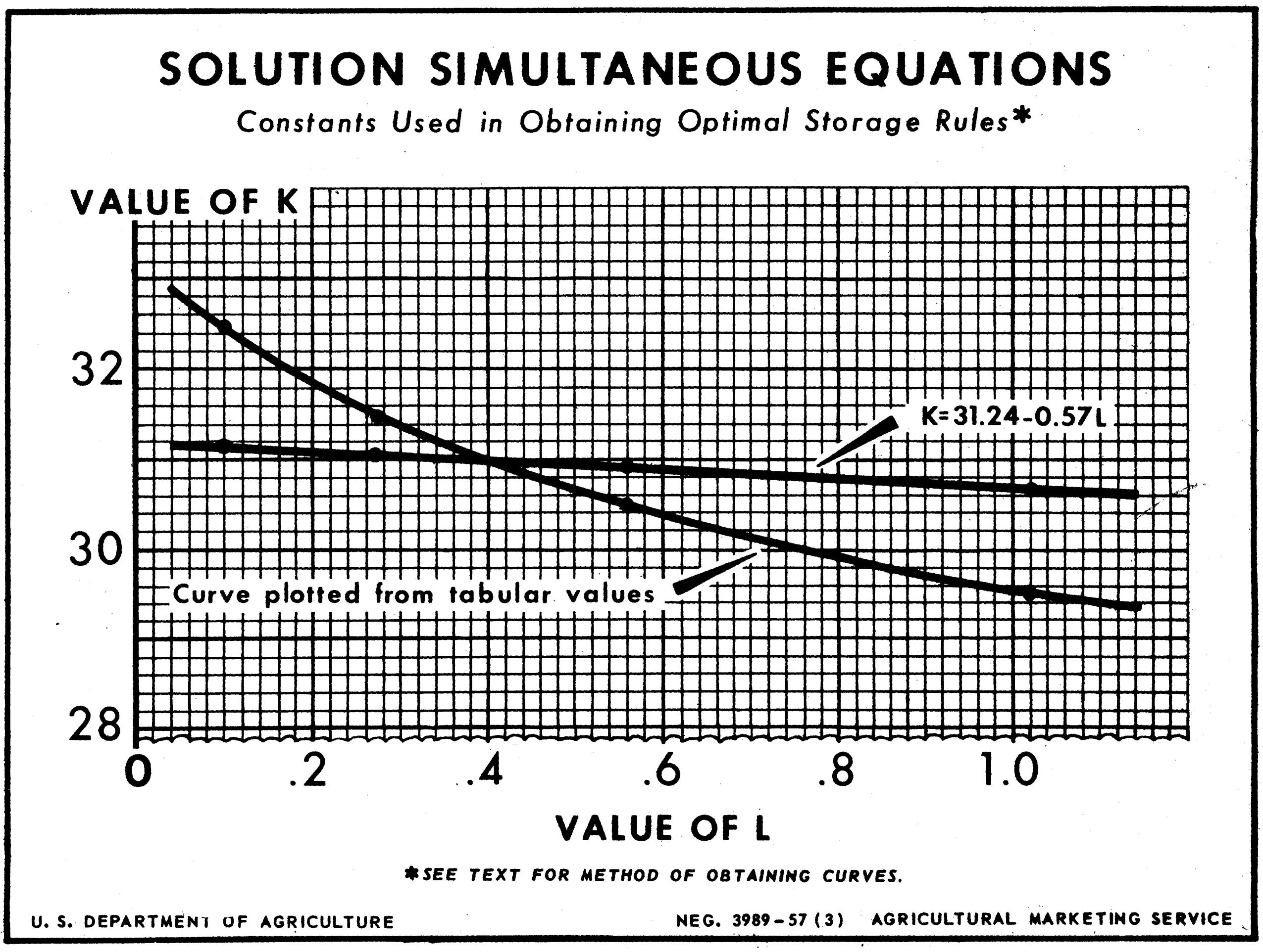


Figure 35

Values of k and L as computed in alternative ways

Tabular method		Value of k as computed from formula
Given value of k	Computed value of L	
29.5	1.02	30.66
30.5	.56	30.92
31.5	.27	31.09
32.5	.10	31.18

Gustafson, Robert L. Optimal Storage Rules for Feed Grains. Unpublished manuscript. 1957.

☆ U. S. GOVERNMENT PRINTING OFFICE : 1957 O - 430756