Pitfalls in the Statistical Construction of Demand and Supply Curves

by

Ragnar Frisch
Professor of Economics in the University of Norway

1933

HANS BUSKE / VERLAG / LEIPZIG

36

VORBEMERKUNG DES HERAUSGEBERS

Seit Henry L. Moore, vor nunmehr bald zwei Jahrzehnten — ungeachtet aller Einwände von Marshall und Edgeworth — mutig die statistische Analyse von Nachfragekurven in Angriff genommen hat, ist ein neues Sondergebiet mathematisch-ökonomischer Forschung entstanden. Daß die Geltung des Gesetzes von Nachfrage und Angebot an die Voraussetzung geknüpft ist, daß die „übrigen Bedingungen" gleich bleiben, wußte man sehr wohl und übersah auch keineswegs, daß diese Bedingungen in der Realität nie erfüllt sein können. Feinsinnig ausgesponnene mathematisch-statistische Verfahren sollten über diese Schwierigkeiten hinweghelfen. Allein, wiewohl das Problem theoretisch klar gestellt war, hat die mathematische Analyse gar zu oft zu Scheinlösungen geführt. Professor Ragnar Frisch sucht zu zeigen, welche Fallgruben des Forschers harren und wie der Gefahr einer Fehlanalyse begegnet werden kann.

Die kritischen Auseinandersetzungen über Voraussetzungen und Grenzen der einzelnen mathematisch-statistischen Verfahren sind fast ausschließlich in der amerikanischen Literatur erfolgt. In Deutschland haben diese die Fragen der praktischen Marktforschung eng berührenden Diskussionen noch kaum einen Widerhall gefunden. Wir sind daher genötigt, die Schrift von Professor Ragnar Frisch in englischer Sprache herauszugeben.

Wir haben von Anfang an unsere besondere Aufmerksamkeit der Analyse von Nachfragekurven gewidmet. Als Heft 2 unserer Reihe ist die Untersuchung von Hans Staehle: „Die Analyse von Nachfragekurven in ihrer Bedeutung für die Konjunkturforschung" erschienen, die erste Arbeit über den Gegenstand in deutscher Sprache überhaupt. In Heft 10 hat Henry Schultz, einer der bedeutendsten Schüler von Henry L. Moore, unter dem Titel „Der Sinn der statistischen Nachfragekurven" bereits eine umfassende Darstellung der gesamten Problematik zu geben versucht, während M. Ezekiel in Heft 9 gezeigt hat, wie die Analyse von Nachfragekurven in den Dienst der „Preisvoraussage bei landwirtschaftlichen Erzeugnissen" gestellt werden kann. Wir hoffen in nächster Zeit eine weitere Schrift über die Analyse von Nachfragekurven herausbringen zu können, die auch dem mathematisch weniger bewanderten Leser die Möglichkeit geben wird, sich mit dem Aufgabenkreis dieses besonders verheißungsvollen Zweiges ökonomischer Forschung vertraut zu machen.

Frankfurt a. M., im Januar 1933

E. Altschul

CONTENTS

1. INTRODUCTION. FICTITIOUS DETERMINATENESS CREATED BY RANDOM ERRORS.

On the road to statistical demand and supply curves there are many pitfalls. Some of them lie in the difficulty of knowing exactly how reliable the data are and under what conditions they were collected. Others are of a mathematical sort and are connected with the variability type of the data. The object of the present paper is to discuss some points regarding the latter aspect of the problem.

Some of the methods that have been proposed for the statistical construction of demand and supply curves involve a mathematical apparatus of a very dangerous sort. I shall especially consider a certain form of pitfall that occurs in this field in various forms of disguise, and which we may characterize by the catchword „fictitious determinateness created by random errors". The nature of this pitfall may be illustrated by the following simplified example. Suppose there exists in (x, y) coordinates a straight line whose angular coefficient c we want to determine by observing points on the line. Let (x_1, y_1) and (x_2, y_2) be two sets of observations. The angular coefficient is then equal to

(1.1)
$$c = \frac{y_1 - y_2}{x_1 - x_2}$$

This furnishes a determination of c, provided the two observations do not coincide, that is to say, provided we do not have $x_1 = x_2$ and $y_1 = y_2$. In this latter case the right member of (1. 1) is of the form $\frac{0}{0}$ and furnishes no determination of c. This will be the situation if the observations are absolutely correct, not affected by errors of observation.

Now suppose that there are present errors of observation. Suppose that each actual observation is of the form (X, Y) where

$$X = x + \xi$$
$$Y = y + \eta$$

x and y being the systematic variables and ξ and η being errors of observation. The actually observed slope

$$C = \frac{Y_1 - Y_2}{X_1 - X_2} = \frac{(y_1 - y_2) + (\eta_1 - \eta_2)}{(x_1 - x_2) + (\xi_1 - \xi_2)}$$

may or may not be an approximation to the correct slope. If we are in a situation where $x_1 = x_2$ and $y_1 = y_2$, then C is of the form: a random error divided by another random error. In other words the slope computed is entirely meaningless although to the appearance this slope is a determinate magnitude. And the situation will be similar if the systematic displacements $x_1 - x_2$ and $y_1 - y_2$ are small as compared with the error differences $\xi_1 - \xi_2$ and $\eta_1 - \eta_2$. We may for brevity express this fact by saying that the computed slope is meaningless whenever we are n e a r l y in a situation where $x_1 = x_2$ and $y_1 = y_2$.

In the above example the situation is so simple that no real danger exists of fooling oneself. But when we come to more complicated procedures a real danger of this sort arises. This is in particular the case when we approach a statistical material with the object of determining numerically the constants of certain theoretical laws that we have worked out a priori. In fact, in such a case it will always be possible to deduce an infinity of relations which the material must satisfy if it shall really be a material that has emerged under the influence of our postulated theoretical laws. There may for instance be sundry relations connecting the moments of the material, or connecting certain discrete values in the material etc. In general these relations will contain also the parameters we are attempting to determine. In such cases a nearly irresistible temptation arises to select a number of these sundry relations, e q u a l i n n u m b e r t o t h e u n k n o w n p a r a m e t e r s, and consider these relations as a statistical determination of the parameters. If it should be found that the theory, as first conceived, was not general enough, the temptation arises to generalize the theoretical concept at liberty, introducing new parameters and simply increasing the number of equations used accordingly.

Such a procedure will be nonsense unless it can be definitely shown that the systematic variations exhibited in the material are such that they would produce a determinate solution even though no random errors were present. But such an investigation of the nature of the systematic as distinguished from the accidental variations in the material is a very subtle matter. Sometimes it may for instance happen that the very assumptions back of the theoretical scheme adopted, are incompatible with the existence of a statistical material where the systematic variations lead to a determinate solution. To these questions there is frequently paid too little attention. And the appearance of the numerical results themselves as obtained by this or that arbitrarily selected system of equations will as a rule contain no warning signal by which to distinguish between significant determinateness and indeterminateness. In practice the random element will indeed always be present and make the result appear as if it should be quite determinate and significant. I believe that much work that has been done in multiple correlation both on prices and otherwise is meaningless for this reason. The usual computation of standard errors of the parameters involved will in general not be a safeguard against the kind of indeterminateness I have in mind. This question is related to the theory of cluster types discussed in my paper „Correlation and Scatter in Statistical Variables".[1])

In the present paper I shall not follow up this multiple correlation aspect of the fictitious determinateness created by random errors. Instead I shall take up the two-variable problems of determining at the same time a neo-classical demand curve and supply curve from the same price-quantity material. The present analysis does not wind up with any definite demand and supply curve method applicable under a variety of circumstances. But it gives a discussion of various possible cases and shows how various assumptions about the underlying situation leads to certain values of the demand and supply elasticities. (See for

1) Nordic Statistical Journal 1929. See also a joint paper by Professor Bruce D. Mudgett and me in the Journal of the American Statistical Association, December 1931.

instance formulae (3. 13), (3. 14), (3. 18) and (3. 19)). And above all it attempts to point out the pitfalls that must be avoided.

To make the discussion concrete a considerable part of it will be devoted to a criticism of a particular method which has been propagated in this field and which, in my opinion, is a characteristic example of one of these ingenious methods that are fundamentally unsound for the above mentioned reasons. I mean Dr. Leontief's method of constructing demand and supply curves [2]). In many cases the coefficients obtained by this method are, I believe, entirely meaningless, their magnitude being determined essentially by the random disturbances in the material. And in those cases where they have a sense, they do not as a rule express demand and supply elasticities, but simply express the historical trend connection between price and quantity.

The analytical tool I am using is more general than the one used by Leontief. But at the same time it is more elementary. It contains for instance no reference to the least square minimizing procedures used by Leontief. These minimizing procedures are in fact quite superfluous. They veil the true nature of the problem rather than shed light on it. All Leontief's results can be obtained simply by a few elementary considerations regarding the moments of a sum.

2. STATEMENT OF ASSUMPTIONS AND DEFINITION OF THE SYMBOLS USED.

In works on the statistical determination of demand and supply curves it is frequently assumed that the demand curve has an elasticity which is constant all along the curve and also constant over time. In other words, the demand curve is assumed to be such that it appears as a straight line when plotted on a double logaritmic scale. And the only change which takes place in this curve in the course of time, is that the curve is shifted up or down, its slope remaining unchanged. A similar assumption is often made for the supply curve. In themselves these assumptions are perhaps

2) Weltwirtschaftliches Archiv, July 1929.

questionable, but that is an aspect of the problem which I shall not discuss here. I simply grant the assumption of constant elasticities.

These assumptions being adopted, we can write the demand and supply functions:

(2.1) $$x = u + \alpha p \qquad \text{(Demand)}$$
(2.2) $$x = v + \beta p \qquad \text{(Supply)}$$

Here x stands for the log of the quantity demanded and supplied respectively, p stands for the log of the price, u and v are the shifts („Niveauverschiebungen" in Leontief's terminology), and α and β are the demand and supply elasticities respectively. The magnitudes x and p are directly observed as time series. The shifts u and v we also conceive of as time series, but they are not directly observable. The problem is to determine the magnitudes of α and β from the observation of x and p.

In order to characterize the four time series x, p, u and v we introduce in the usual way their arithmetic means $\bar{x}$, $\bar{p}$, $\bar{u}$, $\bar{v}$, taken over the interval of time considered. Further we introduce the moments

$$\begin{aligned} m_{xx} &= \Sigma (x - \bar{x})^2 & m_{uu} &= \Sigma (u - \bar{u})^2 \\ m_{pp} &= \Sigma (p - \bar{p})^2 & m_{vv} &= \Sigma (v - \bar{v})^2 \\ m_{xp} &= \Sigma (x - \bar{x}) \cdot (p - \bar{p}) & m_{uv} &= \Sigma (u - \bar{u}) \cdot (v - \bar{v}) \end{aligned}$$

where the summation Σ is extended to all observations in the interval of time considered. If we consider different materials, say the materials Nos. 1. 2. . . . we use the notation $m_{xx \cdot 1}$, $m_{xx \cdot 2}$, $m_{pp \cdot 1}$ etc. The magnitudes

$$r = \frac{m_{xp}}{\sqrt{m_{xx}\, m_{pp}}} \text{ and } \rho = \frac{m_{uv}}{\sqrt{m_{uu}\, m_{vv}}}$$

are the coefficients of correlation between x and p and between u and v respectively. The magnitudes

$$l = \sqrt{\frac{m_{xx}}{m_{pp}}} \text{ and } \lambda = \sqrt{\frac{m_{uu}}{m_{vv}}}$$

may be called the relative „violences" in the set (x, p) and (u, v) respectively. The number l expresses the intensity (the amplitude) of the fluctuations in x as compared with the intensity of the fluctuations in p. And λ has a similar meaning for the set (u, v). For convenience we shall also introduce the ratios

$$H = rl = \frac{m_{xp}}{m_{pp}} \qquad K = l^2 = \frac{m_{xx}}{m_{pp}}$$

$$h = \rho\lambda = \frac{m_{uv}}{m_{vv}} \qquad k = \lambda^2 = \frac{m_{uu}}{m_{vv}}$$

3. THE FUNDAMENTAL EQUATIONS.

From (2. 1) and (2. 2) we deduce

$$(x - \bar{x}) - \alpha(p - \bar{p}) = (u - \bar{u}) \tag{3.1}$$

$$(x - \bar{x}) - \beta(p - \bar{p}) = (v - \bar{v}) \tag{3.2}$$

First we multiply (3. 1) by (3. 2) and perform a summation over time, then we square (3. 2) and (3. 1) and perform a summation over time. This gives

$$m_{xx} - (\alpha + \beta)\, m_{xp} + \alpha\beta m_{pp} = m_{uv} \tag{3.3}$$

$$m_{xx} - 2\,\beta\, m_{xp} + \beta^2\, m_{pp} = m_{vv} \tag{3.4}$$

$$m_{xx} - 2\,\alpha\, m_{xp} + \alpha^2\, m_{pp} = m_{uu} \tag{3.5}$$

Taking the ratio between (3. 3) and (3. 4) and also the ratio between (3. 4) and (3. 5), we get

$$\begin{aligned} &(\alpha\beta - h\beta^2) - (\alpha + \beta - 2h\beta)H + (1 - h)K = 0 \\ &(\alpha^2 - k\beta^2) - 2\,(\alpha - k\beta)H + (1 - k)K = 0 \end{aligned} \tag{3.6}$$

The equations (3. 6) we shall call the fundamental equations. They form a system of two equation in the six parameters

$$(\alpha, \beta), (h, k) \text{ and } (H, K).$$

The equation obtained by taking the ratio between (3. 3) and (3. 5) would not be a new independent equation in these prameters.
The solutions of the two equations (3. 6) with respect to the various parameters are

$$h = \frac{\alpha\beta - (\alpha + \beta)H + K}{\beta^2 - 2\beta H + K} \tag{3.7}$$

$$k = \frac{\alpha^2 - 2\,\alpha H + K}{\beta^2 - 2\,\beta H + K} \tag{3.8}$$

$$\rho = \frac{\alpha\beta - (\alpha + \beta)\,H + K}{+\sqrt{(\alpha^2 - 2\,\alpha H + K)\cdot(\beta^2 - 2\,\beta H + K)}} \tag{3.9}$$

(3.10) $$H = \frac{\alpha - (\alpha + \beta) h + \beta k}{1 - 2h + k}$$

(3.11) $$K = \frac{\alpha^2 - 2\alpha\beta h + \beta^2 k}{1 - 2h + k}$$

(3.12) $$r = \frac{\alpha - (\alpha + \beta) h + \beta k}{+\sqrt{(1 - 2h + k) \cdot (\alpha^2 - 2\alpha\beta h + \beta^2 k)}}$$

(3.13) $$\alpha = l\left(r \pm (\lambda - \rho)\sqrt{\frac{1 - r^2}{1 - \rho^2}}\right)$$

(3.14) $$\beta = l\left(r \mp \left(\frac{1}{\lambda} - \rho\right)\sqrt{\frac{1 - r^2}{1 - \rho^2}}\right)$$

In the equations (3. 13) and (3. 14) we may attribute to the square root either the upper or the lower sign. This gives two sets (α, β), which both satisfy the fundamental equation. Even if we knew, not only the observed parameters r and l, but also ϱ and λ, the solution in (α, β) would therefore not be absolutely unique. But, apart from the selection of the sign, (3. 13) and (3. 14) furnish a solution in α and β whenever we are in a situation where we have reason for making a definite assumption about the correlation and about the relative violence in the shifts (u, v).

If the shifts are uncorrelated, that is $\rho = 0$, and λ finite, the equations (3. 10) to (3. 14) reduce to

(3.15) $$H = \frac{\alpha + \beta k}{1 + k}$$

(3.16) $$K = \frac{\alpha^2 + \beta^2 k}{1 + k}$$

(3.17) $$r = \frac{\alpha + \beta k}{+\sqrt{(1 + k) \cdot (\alpha^2 + \beta^2 k)}}$$

(3.18) $$\alpha = l\,(r \pm \lambda \sqrt{1 - r^2})$$

(3.19) $$\beta = l\,(r \mp \frac{1}{\lambda}\sqrt{1 - r^2})$$

4. THE COURNOT EFFECT AND THE TREND EFFECT. CLASSIFICATION OF THE VARIOUS POSSIBLE CASES ACCORDING TO THE NATURE OF THE SHIFT DISTRIBUTION.

From the definition equations (2. 1) and (2. 2) follows:

$$x = \frac{u\beta - v\alpha}{\beta - \alpha}$$

$$p = \frac{u - v}{\beta - \alpha}$$

Therefore, to any given scatter diagram in (u, v) corresponds a uniquely determined scatter diagram in (x, p) and vice versa (provided $\alpha \neq \beta$). And the nature of this correspondence is just determined by the two constants α and β. I propose to discuss what the nature of the (x, p) distribution will be under various assumptions regarding the nature of the (u, v) distribution.

The nature of the (u, v) distribution may best be described by distinguishing between the following four extreme cases: 1) Demand curve stability, i. e. $\lambda = 0$. 2) Supply curve stability, i. e. $\lambda = \infty$. 3) Bilateral and uncorrelated shifts, i. e. $\lambda \neq 0$ and finite, $\varrho = 0$. 4) Bilateral and highly correlated shifts, i. e. $\lambda \neq 0$ and finite, $\varrho = \pm 1$. Between these extreme cases there are, of course, intermediate cases, but for simplicity I shall here only discuss the „pure" types. The results of this discussion will in part exhibit well known facts, and in part be novel. Since we here need a complete classification, all cases must be discussed.

In the case of demand curve stability we have $h = 0$ and $k = 0$, hence by (3. 6)

(4.1) $$\alpha\beta - (\alpha + \beta)H + K = 0$$

(4.2) $$\alpha^2 - 2\alpha H + K = 0$$

From the last of these equations we obtain

(4.3) $$\alpha = H \pm \sqrt{H^2 - K} = l\,(r \pm \sqrt{-(1 - r^2)})$$

This shows that if a real solution in α shall exist, r^2 must now be equal to unity. Hence: In the case of demand curve stability, the (x, p) diagram must be perfectly organized, i. e. $r^2 = 1$. This is nothing but an algebraic statement of the fact which has been pointed out so clearly in Elmer Workings fundamental paper[1]) of 1927. Furthermore, since $r^2 = 1$, we get from (4. 3)

$$\alpha = \varepsilon l \tag{4.4}$$

where $\varepsilon = \text{sgn. } r = +1$ or -1 designates the sign of r. That is to say the regression in (x, p) coordinates has just the slope α. The elasticity β is now indeterminate, when the only available data are the (x, p) observations. From (4. 1) we see indeed that

$$\beta = \frac{\alpha H - K}{\alpha - H} = \frac{l^2 - r l \alpha}{r l - \alpha}$$

so that in the present case where $r = \varepsilon$, $\alpha = \varepsilon l$.

$$\beta = \frac{l^2 - l^2}{l - l} = \text{indeterminate.}$$

In the case where we have exactly $\lambda = o$, β will be exactly of the form $\frac{0}{0}$. And if we have nearly $\lambda = o$, then β will be of the form: random error divided by a random error. In both cases the value of β is of course meaningless.

It should be noticed that the above discussion holds good regardless of whether the case $\lambda = o$ is reached by a limiting process such that ϱ^2 tends towards zero or by a process such that ϱ^2 tends towards any number between 0 or 1, or even by a process such that ϱ^2 does not converge at all. We shall refer to the case of demand curve stability by saying that the (x, p) diagram now exhibits a *Cournot effect* on the demand side.

Also in the case of supply curve stability, i. e. $\lambda = \infty$, the organization of the (x, p) diagram must be perfect, i. e. $r^2 = 1$. It is now β that can be determined, namely by the formula

$$\beta = \varepsilon l$$

And α now becomes indeterminate. In this case we shall say that the (x, p) diagram exkibits a Cournot-effect on the supply side.

To sum up we can say: In both the stability cases the organization in the observation diagram (x, p) must be perfect. And the slope of the (x, p) regression line indicates in the demand stability

1) Quarterly Journal of Economics, February 1927. This paper gives an exceedingly interesting exposo of several points which have later come into the foreground of the discussion. It studies for instance the possibility of determing both the demand and the supply curve from the samo material.

case α (with β indeterminate), and in the supply stability case β (with α indeterminate).

In the case of bilateral, but uncorrelated shifts the appearance of the (x, p) scatter diagram will depend essentially on the magnitude of the relative violence in the shifts, i. e. on the parameter λ. Instead of λ itself it will here be more convenient to consider its square $k = \lambda^2$. How will r depend on k? From (3. 17) we see that if $\alpha = \beta$, r^2 will be equal to 1 for any magnitude of k. This is a trivial case without interest. If $\alpha \neq \beta$, r is a function of k. Let it be r(k). For $k = o$, $r = \text{sgn.}\ \alpha$. Assuming the demand elasticity to be negative we consequently have $r(o) = -1$. The derivative of r with respect to k, α and β being constants, is equal to

$$\text{(4.5)} \qquad \frac{dr(k)}{dk} = \frac{(\alpha - \beta)^2 \cdot ((-\alpha) + \beta k)}{2\,(\alpha^2 + (\alpha^2 + \beta^2)k + \beta^2 k^2)^{3/2}}$$

(3. 17) in connection with (4. 5) show that if the demand elasticity is negative and the supply elasticity is positive, r is increasing monotonically from $r(o) = -1$ to $r(\infty) = +1$ as k increases from o to ∞. This gives a description of the may in which the closeness of the organization in the (x, p) scatter diagram depends on the relative violence in the fluctuations of the shifts u and v, when these shifts are uncorrelated: There is high organization in the (x, p) diagram when either the demand or the supply curve is relatively stable. Otherwise there is a lack of organization (or at least this organization is far from rectilinear). From (3. 17) is seen that there is complete lack of correlation when the relative violence is equal to

$$\lambda = \sqrt{\frac{(-\alpha)}{\beta}}$$

For a lower λ there is a negative correlation. For a higher λ a positive correlation. The high correlation which we obtain in the present case for λ very small or λ very high is obviously nothing else than a Cournot effect of the kind discussed in the two stability cases.

Not only the closeness of the organization, but also the slope of the regression line in the (x, p) diagram will in the present case depend on k. The absolute value $l = \sqrt{K}$ of the slope of the diagonal mean regression[1]) in the (x, p) diagram is now simply equal to **the weighted square mean** of the elasticities α and β, the weights being the inverted squares of the standard deviations of u and v. We get indeed from (3. 16)

$$(4.6) \qquad l = \sqrt{\frac{\frac{\alpha^2}{\sigma_u^2} + \frac{\beta^2}{\sigma_v^2}}{\frac{1}{\sigma_u^2} + \frac{1}{\sigma_v^2}}} \quad (\text{when } \rho = 0)$$

Even if the shifts are uncorrelated, the absolute value of the slope of the (x, p) regression does therefore not tell us anything about the elasticities α and β. Trying to evaluate the elasticities α and β from the slope of the observed (x, p) regression and the observed (x, p) correlation, without making any **assumption** about the relative violence λ, would be the same as trying to evaluate the two parts of a cake by only knowing the size of the whole cake. But, if we have some reason for making an assumption about the relative violence in the shifts, that is about λ, then the elasticities α and β may be determined, namely by (3. 18) and (3. 19). This is one example showing the fundamental rôle played in the present problem by the relative violence between the fluctuations in u and v. We shall later see other examples of the same thing.

We now come to the case where neither the demand nor the supply curve is stable and where there exists a high correlation between the shifts. This correlation may be due to a **cyclical** connection between the shifts, or due to a **trend** connection between the shifts over the interval of time considered, or due to some other cause. In the present study I shall most of the time consider a high shift correlation as an expression for a trend relation between u and v. The case of a cyclical connection u and v I hope to be able to take up in a later study.

From (3. 12) we deduce when $\varrho = \pm 1$

$$(4.7) \qquad r = \frac{(1 \mp \lambda) \cdot (\alpha \mp \beta\lambda)}{\sqrt{(1 \mp \lambda)^2 \cdot (\alpha \mp \beta\lambda)^2}}$$

1) By the diagonal mean regression I understand the regression determined from the diagonal elements in the adjoint correlation matrix. (See formula (4. 23) in my paper „Correlation and Scatter ...“ in Nordic Statistical Journal 1929.) In two variables the diagonal mean regression is nothing else than the regression used by Lehfeldt (Economic Journal 1914).

Thus, unless $\lambda = \pm 1$ or $\lambda = \pm \frac{\alpha}{\beta}$ (upper sign when $\varrho = +1$, lover sign when $\varrho = -1$), we must have

$$r = \operatorname{sgn}(1 \mp \lambda) \operatorname{sgn}(\alpha \mp \beta\lambda) = \varepsilon$$

That is to say, apart from the specified exceptional cases, the assumption of a high correlation between u and v entails a high correlation between x and p. But this correlation between x and p must be interpreted in quite a different way from the high correlation between x and p occuring in the case of demand curve stability or supply curve stability. The correlation now considered is indeed not a Cournot effect but a **trend effect**. It is the trend in u and v, that produces a trend in x and p, and this fact is responsible for the (x, p) correlation. The slope of the (x, p) regression therefore now expresses neither the demand elasticity, nor the supply elasticity, but the historical trend relation between x and p. This historical trend relation would become nearly equal to the demand or the supply relation only if in addition to the high correlation between the shifts we should have the situation where one of the two curves shifts much more violently than the other, i. e. either $\lambda = o$ or $\lambda = \infty$. This is easily seen from the explicit expression for the slope of the (x, p) regression. This expression now becomes

$$\varepsilon l = \frac{\alpha \mp \beta\lambda}{1 \mp \lambda} \tag{4.8}$$

Here, instead of considering separately the two formulae abtained from (4. 8) by selecting successively the upper and lower sign, corresponding respectively to $\varrho = +1$ and $\varrho = -1$, we may simply consider the function

$$\varepsilon l = \frac{\alpha - \beta\lambda}{1 - \lambda} \tag{4.9}$$

but let λ vary from $-\infty$ to $+\infty$ instead of from o to ∞. A negative λ would then correspond to $\varrho = -1$, and a positive λ to $\varrho = +1$. The course of the function (4. 9) is as follows: As λ increases from $-\infty$ the function decreases monotonically from the (pos.) asymptotic level β, passing zero for $\lambda = \frac{\alpha}{\beta}$, being equal to the (negative) magnitude α for $\lambda = o$ and going down to $-\infty$ for $\lambda = 1$. In this point of singularity the other branch comes down from $+\infty$ and decreases monotonically towards the asymptotic (positive) level β as λ increases from $\lambda = 1$ to $\lambda = \infty$. Thus, also in the case of highly correlated shifts is it true that only when λ is near to o or ∞ will the observed slope in the (x, p) diagram have any significance as an expression for the demand or supply elasticity. As an example we may consider the situation where the demand shift is always a definite fraction, but a very small fraction of the supply shift. Here there would be a very high (u, v) correlation, but the observed (x, p) diagram would nevertheless exhibit a pronounced Cournot effect on the demand side.

As a special case of bilateral shifts we may consider equilateral shifts. This is the case where the fluctuations in u are just as violent as the fluctuations in v, in other words $\lambda = 1$. The case of equilateral shifts has a significance that is independent of units of measurement. The magnitudes u and v are indeed commensurable because they are both measured in the same units as x. In the case of highly and positively correlated equilateral shifts the slope of the (x, p) regression is infinite, as is seen from formula (4. 9), assuming $\alpha \neq \beta$.

5. CLASSIFICATION OF THE VARIOUS CASES ACCORDING TO THE NATURE OF THE OBSERVED DISTRIBUTION (x, p).

We may also classify the various cases according to the appearance of the observed diagram (x, p). Certain aspects of this analysis are already contained in the previous Section, but in order to get a complete picture of the situation it will be well to look at it also from the point of view of the (x, p) distribution. We shall distinguish betwen the following two principal cases: 1) the correlation in (x, p) is very high, 2) the correlation in (x, p) is very low.

If the observed correlation is perfect, i. e. $r = \varepsilon$ where ε is either $+1$ or -1, then by (3. 8) and 3. 9)

$$\lambda = \left| \frac{\alpha - \varepsilon l}{\beta - \varepsilon l} \right| \quad \rho = \frac{(\alpha - \varepsilon l)(\beta - \varepsilon l)}{+\sqrt{(\alpha - \varepsilon l)^2 (\beta - \varepsilon l)^2}}$$

This shows that if $\alpha = \varepsilon l$ but $\beta \neq \varepsilon l$, then $\lambda = 0$. If $\beta = \varepsilon l$ but $\alpha \neq \varepsilon l$, then $\lambda = \infty$. And if $\alpha \neq \varepsilon l$, $\beta \neq \varepsilon l$, then $|\varrho| = 1$. The trival case $\alpha = \varepsilon l$, $\beta = \varepsilon l$, which entails $\alpha = \beta$, is without interest in the present analysis. Thus, if the observed correlation between x and p is very high, there are three hypotheses to be considered:

1) The high (x, p) correlation is due to demand stability, in which case the observed regression slope is the demand elasticity.

2) The high (x, p) correlation is due to supply stability, in which case the observed regression slope is the supply elasticity.

3) The high (x, p) correlation is a trend effect in which case the correlation between the shifts must have been very high.

Which one of these three hypotheses is the correct one can, of course, not be decided only from the knowledge of the (x, p) distribution. But it is at least interesting to notice that if the observed (x, p) correlation is high, then the above three hypotheses are the only ones admissible. The case of bilateral and uncorrelated shifts is for instance excluded.

If we take account not only of the (x, p) distribution as represented by the swarm of observation points in (x, p) coordinates, but also take account of the shapes of the two time curves x and p, we can frequently get some basis for further conclusions. For instance, if the (x, p) correlation is high and if both the x and the p series exhibit a pronounced trend over the interval of time considered, it is very probable that the observed high (x, p) correlation is a trend effect, i. e. due to a high shift correlation, and is not a Cournot effect.

If the observed correlation between x and p is very low, it is simpler to indicate those hypotheses that are not admissible. If $r = o$ we have by (3. 8) and (3. 9)

$$\lambda = \sqrt{\frac{\alpha^2 + K}{\beta^2 + K}} \qquad \rho = \frac{\alpha\beta + K}{+\sqrt{(\alpha^2 + K)(\beta^2 + K)}}$$

The formula for λ shows that if both α and β are finite and different from zero, and if the observed K is finite, then λ must be finite and different from zero. And the formula for ϱ shows that, apart from the trivial case where $\alpha = \beta$ and the case where $K = o$ or ∞, ϱ^2 must be different from unity. Possibly we may have $\varrho = o$, and if this is the case, $l = \sqrt{(-\alpha)\beta}$. This shows that if there is very little correlation between x and p, then we know that:

1) We cannot have demand stability.

2) We cannot have supply stability.

3) We cannot have the trend situation.
4) But we may have the case of bilateral and uncorrelated shifts.

6. LEONTIEF'S PROBLEM AS IT APPEARS WHEN STRIPPED OF IRRELEVANT COMPLICATIONS.

Let us again start from the definitions (2. 1) and (2. 2). From these we deduce immediately (3. 1) and (3. 2). Multiplying these latter equations together and performing a summation over time we get (3. 3). Hence if u and v are uncorrelated, that is $m_{uv} = 0$

$$\text{(6.1)} \qquad \alpha\beta - (\alpha + \beta)H + K = 0.$$

This equation is all that is needed in order to deduce Leontief's results. (6. 1) shows in particular that if one of the two elasticities α and β is given, then the other is determined. It is even a simple rational function of the first, as is seen from (6. 1). And this holds good provided only that the shifts u and v are u n c o r r e l a t e d i. e. $\varrho = 0$, which is, of course, a much less severe restriction, than the assumption of their i n d e p e n d e n c e involved in the minimizing procedure which Leontief uses to show that one of the elasticities can be determined from the other in the case where $\varrho = 0$.

If we have two statistical materials Nos. 1 and 2, both having the same α and β, and u and v being uncorrelated in both materials, then

$$\text{(6.2)} \qquad \begin{aligned} \alpha\beta - (\alpha + \beta)H_1 + K_1 = 0 \\ \alpha\beta - (\alpha + \beta)H_2 + K_2 = 0 \end{aligned}$$

where H_1, H_2 and K_1, K_2 are the magnitudes H and K as determined in the materials Nos. 1 and 2. (6. 2) holds, of course, good regardless of whether the two materials are partly overlapping or not. The system (6. 2) is a linear system by which the two magnitudes $(\alpha + \beta)$ and $\alpha\beta$ may be determined. And furthermore, the two magnitudes $(\alpha + \beta)$ and $\alpha\beta$ are obviously the coefficients in the second degree polynominal whose roots are α and β. Now, solving the linear system (6. 2) with respect to the two magnitudes $(\alpha + \beta)$ and $\alpha\beta$, and using these two magnitudes as coefficients in a second degree polynominal, we find that α and β are the two roots of the equation in η

$$(6.3) \qquad \begin{vmatrix} 1 & H_1 & K_1 \\ 1 & H_2 & K_2 \\ 1 & \eta & \eta^2 \end{vmatrix} = 0$$

In other words the two „elasticities“ η_1 and η_2 in Leontiefs theory (determined by the formulae on pp. 30* and 31* of his paper) can be nothing else than the two roots of the equation (6.3). This follows from the above argument, and we can also check it by comparing with Leontief's formulae. Indeed, Leontief's magnitudes X and Z are in my notation

$$X_1 = Y_1K_1 \quad X_2 = Y_2K_2$$
$$Z_1 = Y_1H_1 \quad Z_2 = Y_2H_2$$

and his magnitude Y stands for the square moment of p. In other words $Y_1 = m_{pp.1}$ and $Y_2 = m_{pp.2}$

Further Leontief's magnitudes a, b, c are

$$a = X_1Y_1 - Z_1^2 = Y_1^2(K_1 - H_1^2)$$
$$b = X_1Y_2 + X_2Y_1 - 2Z_1Z_2 = Y_1Y_2(K_1 + K_2 - 2H_1H_2)$$
$$c = X_2Y_2 - Z_2^2 = Y_2^2(K_2 - H_2^2)$$

Now, Leontief expresses the elasticity η by means of his parameter λ thus

$$(6.4) \qquad \eta = \frac{\lambda Z_1 - aZ_2}{\lambda Y_1 - aY_2}$$

where λ satisfies

$$(6.5) \qquad \lambda^2 - b\lambda + ac = 0$$

The two roots λ give the two values of η. In order to show that Leontief's two magnitudes η are the same as the two roots of my equation (6. 3) we therefore only have to show that if we express λ by η by means of (6. 4) and insert this in 6. 5), we are led to an equation which is equivalent with (6. 3). Carrying out the substitutions indicated we get

$$\eta^2[aY_2^2 - bY_1Y_2 + cY_1^2] - \eta[2aY_2Z_2 - b(Y_1Z_2 + Y_2Z_1) + 2cY_1Z_1] + [aZ_2^2 - bZ_1Z_2 + cZ_1^2] = 0$$

And inserting here the expressions for a, b, c, Y, Z, in terms of my symbols we get

$$-Y_1^2Y_2^2(H_1 - H_2)\ [(H_1 - H_2)\eta^2 - (K_1 - K_2)\eta + (K_1H_2 - K_2H_1)] = 0$$

which is the equation (6. 3).

From now on I shall leave Leontief's formulae and handle his coefficients by means of the equation (6. 3). This reduction of the problem is a great help. It brings the problem back to its natural and simple form. The equation (6. 3) enables us to see much clearer on what properties of the material the results of the computation depend. Equation (6. 3) we shall call the two-material equation. Its explicite expression can be written in the form

$$(6.6) \quad \eta = \frac{K_1 - K_2 \pm \sqrt{(K_1 - K_2)^2 + 4\,(H_1K_2 - H_2K_1)\,(H_1 - H_2)}}{2(H_1 - H_2)}$$

If $H_1 = H_2$ and $K_1 = K_2$, the roots are not determinate. Otherwise they are. The behaviour of the roots in the vicinity of a situation where $H_1 = H_2$ and $K_1 = K_2$ will be discussed in the next Section.

7. THE SIMILARITY CASE: ONE ROOT OF THE TWO-MATERIAL EQUATION EXPRESSING THE EMPIRICAL REGRESSION SLOPE IN (x, p).

If we approach a situation where the magnitudes H and K in the two-material equation (6. 3) are such that $H_1 = H_2$ and $K_1 = K_2$, the roots of the equation become nearly of the form $\frac{0}{0}$. Does there still attach a meaning to the roots and if so what is this meaning?

For brevity we shall say that we are in a s i m i l a r i t y situation if $H_1 = H_2$ and $K_1 = K_2$. If we are e x a c t l y in such a situation, the numerical computation according to the formula (6. 6) would not lead to any determination of the roots at all, unless in the course of the computation we had introduced either some direct mistake or some slight inaccuracy due to the neglecting of decimals. However, this is not the question in which we are here primarily interested. In practice we will never have a situation where exactly $H_1 = H_2$ and $K_1 = K_2$. From an economic and statistical point of view the pertinent question is what the significance of the roots are when we are n e a r l y in a similarity situation, that is, when H_1 is nearly equal to H_2 and K_1 nearly equal to K_2.

If we have exactly $\varrho_1 = \varrho_2 = o$ and the assumption about constant elasticities is exactly fulfilled while there exists some slight difference between H_1 and H_2 or between K_1 and K_2 (or both), the equation furnishes the correct roots α and β. But again, in practice this is never the situation. There may perhaps be some reason for assuming that ϱ_1 and ϱ_2 are magnitudes very close to the zero, but it would be absurd to assume that they are rigorously equal to zero. Even if there is no systematic variation tending to produce a positive ϱ^2, the erratic element in u and v will always create s o m e deviation of ϱ from zero. Therefore, if we are nearly in a similarity situation we must reckon with the possibility that $\varrho_1 - \varrho_2$, $H_1 - H_2$ and $K_1 - K_2$ are small quantities o f t h e s a m e o r d e r o f m a g n i t u d e. That is why we need a closer discussion of the roots of the equation in the vicinity of a similarity situation. In the present Section I shall discuss what features of the observed (x, p) distribution the two roots are expressions for when approximately $H_1 = H_2$ and $K_1 = K_2$. The results thus obtained will then in the next Section be interpreted further in the light of possible assumptions about the (u, v) distribution.

A situation where $H_1 = H_2$ and $K_1 = K_2$ is the same thing as a situation where the two violences l_1 and l_2 are equal and also the two correlations r_1 and r_2 are equal. This means that l_1 and l_2 become nearly equal to the ratio $l = \frac{\sigma_x}{\sigma_p}$ that holds good for the whole material considered as a unity, and r_1 and r_2 become equal to the correlation r that exists in the whole material. Indeed, if $l_1 = l_2$ we have

$$l_1^2 = l_2^2 = \frac{m_{xx.1}}{m_{pp.1}} = \frac{m_{xx.2}}{m_{pp.2}} = \frac{m_{xx.1} + m_{xx.2}}{m_{pp.1} + m_{pp.2}} = l^2$$

And if further $r_1 = r_2$, we have

$$r_1 = r_2 = \frac{m_{xp.1} + m_{xp.2}}{\sqrt{m_{xx.1} \cdot m_{pp.1}} + \sqrt{m_{xx.2} \cdot m_{pp.2}}}$$

The denominator in the last expression is equal to

$$l_1 \cdot m_{pp.1} + l_2 \cdot m_{pp.2} = l\,(m_{pp.1} + m_{pp.2})$$
$$= \sqrt{(m_{xx.1} + m_{xx.2})\,(m_{pp.1} + m_{pp.2})}$$

so that $r_1 = r_2 = r$. Thus, as we approach a similarity situation, all the parameters l_1, l_2, r_1 and r_2 tend towards well defined values.

On the other hand when we approach a similarity situation, the parameter

$$(7.1) \qquad s = \frac{\frac{r_1}{r} - \frac{r_2}{r}}{\frac{l_1}{l} - \frac{l_2}{l}} = \frac{r_1 - r_2}{l_1 - l_2} \cdot \frac{l}{r}$$

approaches a $\frac{0}{0}$ form. The magnitude of s expresses the nature of the approach to the similarity situation, s expresses how that little divergency is constituted that separates the situation from being exactly a similarity situation. More precisely: s expresses whether the actual material may be looked upon as one where the rapidity with which l_1 and l_2 tend towards each other is less than or larger than the rapidity with which r_1 and r_2 tend towards each other. The magnitude s expresses just the relation between these two rapidities. And this relation has an important influence on the two roots (6. 6).

If the actual material is such that l_1 and l_2 lie much closer together than r_1 and r_2, then we get an approximate expression for the roots by letting first l_2 tend towards l_1 and then r_2 towards r_1 in (6. 6). On the contrary, if the actual material is such that r_1 and r_2 lie closer together than l_1 and l_2, then the approximate expression for the roots is obtained by first letting r_2 tend towards r_1 and then l_2 towards l_1. These two limiting processes do not give the same result. Indeed, the first process leads to

$$(7.2) \qquad \eta = \pm l$$

And the second process leads to

$$(7.3) \qquad \eta = \frac{l}{r}(1 \pm \sqrt{1 - r^2})$$

If r^2 is close to unity, the two roots (7. 3) will be nearly equal, both of them becoming nearly equal to the slope of the diagonal mean regression.

In terms of the parameter s the case leading to (7. 2) is characterized by $s = \pm \infty$, and the case leading to (7. 3) is characterized by $s = 0$. More generally if we are in situation where s is neither very large nor very small, the approximate expression for the roots must be computed by taking account of the actual size of s. How do the roots depend on s? This is seen as follows: Dividing the first row in (6. 3) by l_1 and the second row by l_2 we get

$$\begin{vmatrix} \frac{1}{l_1} & r_1 & l_1 \\ \frac{1}{l_2} & r_2 & l_2 \\ 1 & \eta & \eta^2 \end{vmatrix} = 0$$

From the second row in this determinant we subtract the first row. Then we divide the second row by $l_2 - l_1$. This gives

$$\begin{vmatrix} \frac{1}{l_1} & r_1 & l_1 \\ -\frac{1}{l_1 l_2} & \frac{r_2 - r_1}{l_2 - l_1} & 1 \\ 1 & \eta & \eta^2 \end{vmatrix} = 0$$

If we now let $l_1 \rightarrow l_2 \rightarrow l$ and $r_1 \rightarrow r_2 \rightarrow r$, the equation takes on the form

$$\begin{vmatrix} \frac{1}{l} & r & l \\ -\frac{1}{l} & sr & l \\ 1 & \eta & \eta^2 \end{vmatrix} = 0. \tag{7.4}$$

And the explicit solution of this equation is

$$\eta = l \, \frac{1 \pm \sqrt{1 - r^2 + r^2 s^2}}{r(1 + s)} \tag{7.5}$$

The formulae (7. 2) and (7. 3) are, of course, the special cases $s = \pm \infty$ and $s = 0$ of (7. 5).

The formula (7. 5) shows that the roots of the two material equation in the similarity case are approximately equal to the product of the absolute value l of the slope of the diagonal mean regression in the total material and a factor, which is a function only of r and s, namely the two-branched function

$$\psi(r, s) = \frac{1 \pm \sqrt{1 - r^2 + r^2 s^2}}{r(1 + s)} \tag{7.6}$$

The whole problem of studying the roots in the similarity case is thus reduced to the problem of studying the function of two variables defined by (7. 6), where r is the correlation coefficient between x and p in the total material and s is the parameter defined by (7. 1).

Let us first consider the case where the (x, p) correlation is very high, i. e. where we have approximately $r = \varepsilon$, ε being either $+1$ or -1. Putting $r = \varepsilon$ in (7. 5) we get

$$\eta = \begin{cases} \varepsilon l \\ \varepsilon l \dfrac{1-s}{1+s} \end{cases} \tag{7.7}$$

Consequently: **If the total (x, p) correlation is very high and if we have approximately $H_1 = H_2$ and $K_1 = K_2$, one of the roots of the two material equation is virtually independent of s, and is simply equal to the slope εl of the diagonal mean regression in the total material.**

In order to study ψ as a function of s for values of r different from ε, we first notice that if r changes sign, the only effect is that also ψ changes sign. It will therefore be sufficient to consider the function ψ when r is between 0 and $+1$. Furthermore, if s changes sign, each new root is obtained by dividing the other old roots into unity. In other words we have

$$\frac{1+\sqrt{1-r^2+r^2s^2}}{r(1+s)} = \frac{r(1-s)}{1-\sqrt{1-r^2+r^2s^2}}$$

Multiplying this equation out, we get indeed the identity

$$1-(1-r^2+r^2s^2) = r^2(1-s^2)$$

This fact is of help in plotting the function ψ. Such a plot is given in Fig. 1 for $r = 1$, 0. 99, 0. 95, 0.9, 0. 8 and 0. 1. Fig. 1 shows very clearly the situation that arises when the correlation is not perfect but still fairly high.

For instance, if we follow the ψ curve for $r = 0.9$ we see that o n e of the two branches is nearly always very close to $\psi = +1$. To the left of the point $s = o$ it is the branch obtained by attributing the sign minus to the square root, that is close to unity. To the right of this point it is the other branch that is close to unity. To the left of $s = o$ the first branch is lying between 0. 9 and 1. And tho the right of $s = o$ the second branch is lying

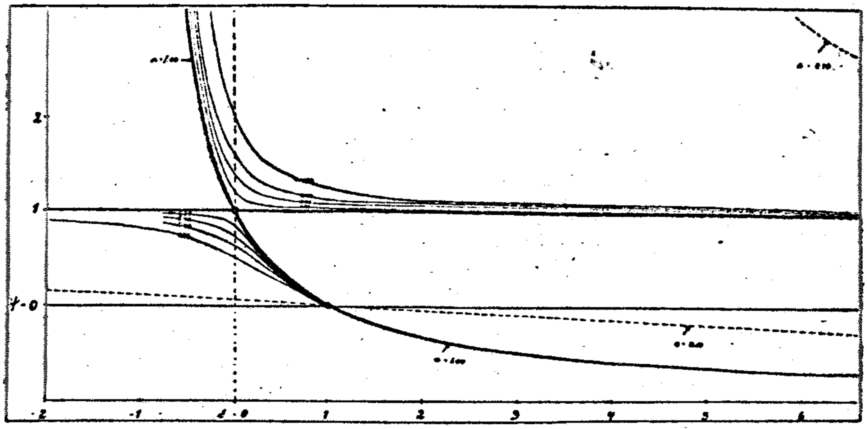

Fig. 1

between 1 and $\frac{1}{0.9}$ A similar remark applies tho the curve $r = 0.8$ only are the limits here 0.8 and 1 and 1 and $\frac{1}{0.8}$ respectively. Quite generally we see from (7.6) in conjunction with Fig. 1 that one (and, incidentally, only one) of the magnitudes ψ is always (that is for any s) lying between r and $\frac{1}{r}$. For $r = 0.99, 0.95, 0.9$ and 0.8 only that branch is plotted in Fig. 1 which is near to unity.

Since rl and $\frac{l}{r}$ are the slopes of the two elementary regressions, i.e. the regressions obtained by minimizing the sum squares of the deviation in the x and p directions respectively, we have the proposition: *In the similarity case there exists always* (i.e. whatever the magnitude of s) *one root of the two-material equation which gives a slope between the slopes of the two elementary regressions.* Thus if the (x, p) correlation is so high that the regression slope in (x, p) is clearly defined, one root of the equation has a meaning, namely this regression slope. But as the value of r^2 decreases towards zero the regression slope loses its meaning. That is, there is now a great difference between the slopes of the two elementary regressions and also between these slopes and the slope of the diagonal mean regression. And, naturally, the agreement between any of these slopes and the regression slope defined by one root of the two-material equation also becomes less exact.

In terms of the parameter s this latter point is characterized by the fact that there is now a considerable s-range, namely around $s = o$, where the root in question depends essentially on s. See for instance the dotted lines $r = 0.1$ in Fig. 1.

On the other hand we see from (7.6) and Fig. 1 that the o t h e r branch of ψ is essentially dependent on s, even if r^2 is close to unity.

Even if one root of the two-material equation has a meaning as the slope of the (x, p) regression, it need not have any meaning in terms of the elasticities α and β. As a matter of fact, so far I have not yet discussed the significance of the roots in terms of α and β at all. To this I now proceed.

8. A GENERAL DISCUSSION OF THE ROOTS OF THE TWO-MATERIAL EQUATION.

Utilizing the results of the last Sections we shall now study quite generally under what conditions the roots of the two-material equation, i. e. equation (6. 3), have a meaning and what this meaning is. In the classification of the various cases we shall use the nature of the (x, p) distribution as a primary principle of classification and the nature of the (u, v) distribution as a secondary principle. In other words the main classes will be determined by the parameters l_1, l_2, r_1 and r_2. And the subclasses will be determined by λ_1, λ_2, ϱ_1 and ϱ_2. In terms of these parameters the situation is as expressed in Table 1.

All the equal signs in Table 1 should be interpreted as „approximately equal to". In the case (1) of Table 1, we have a similarity situation, and the (x, p) correlation is low. Both roots will consequently depend essentially on the parameter s, as we have seen in Section 7. But s depends essentially on the small terms that express the deviation of $l_1 - l_2$ from zero and the deviation of $r_1 - r_2$ from zero. On what features of the (u, v) distribution will these small deviations depend? It will be sufficient to study $l_1 - l_2$. The analysis of $r_1 - r_2$ is similar. And since $l_1 - l_2 = \frac{K_1 - K_2}{l_1 + l_2}$, we may just as well study $K_1 - K_2$. By (3. 11) we have

$$(8.1) \quad (1 - 2h_1 + k_1)(1 - 2h_2 + k_2)(K_1 - K_2) = \alpha^2 A + \beta^2 B - 2\alpha\beta C$$

where

Tab. 1

Nature of the solution of the two-material equation.

			When the observed ratio between the intensity (amplitude) of the quantity fluctuations and the intensity of the price fluctuations is	
			Nearly equal in the two materials ($l_1 = l_2$)	Significantly different in the two materials ($l_1 \neq l_2$)
When the observed quantity price correlation is	Nearly equal in the two materials ($r_1 = r_2$)	Very low in both materials ($r_1 = r_2 = 0$)	(1) Both roots meaningless through indeterminateness. The shifts may or may not be correlated. Even if they are approximately uncorrelated the roots are meaningless.	(5) One root infinite the other zero. Both of them meaningless. The assumption about uncorrelated shifts not admissible.
		Very high in both materials ($r_1 = r_2 = \pm 1$)	(2) One root is meaningless through indeterminateness, the other root is determinate and is simply the slope of the (x, p) regression fitted to the total material. If there is a specific reason for assuming the high (x, p) correlation to be a Cournot effect on the demand side (i. e. $\lambda_1 = \lambda_2 = 0$), the slope of the (x, p) regression represents the demand elasticity in which case no conclusion can be drawn about the supply elasticy. If the high (x, p) correlation is a Cournot effect on the supply side (i. e. $\lambda_1 = \lambda_2 = \infty$), the regression represents the supply elasticity, in which case no conclusion can be drawn about the demand elasticity. If the high correlation is not a Cournot effect it must be a trend effect (i. e. $\rho_1 = \rho_2 = \pm 1$) in which case no conclusion can be drawn either about the demand elasticity or about the supply elasticity. The fact that the high (x, p) correlation is a Cournot effect does not exclude the existence of a high trend correlation between the shifts. But this trend correlation will not appreciably influence the (x, p) regression slope when we have a pronounced stability case.	(6) Both roots determinate. One gives the slope of the (x, p) regression in the first material, the other the slope of the (x, p) regression in the second material. The assumption about uncorrelated shifts is now admissible only in conjunction with the assumption that the observed high (x, p) correlation in one material is a Cournot effect on the demand side, and in the other material a Cournot effect on the supply side. In this extraordinary case the regression slopes give the elasticities. Otherwise the roots of the equation have no meaning in terms of elasticities, but will express the slope of the historical trend relation between x and p in the two materials.
		Neither very low nor very high	(3) Similar to the case (2), but the slope of the (x, p) regression is now less clearly defined.	(7) Similar to the case (6) but the slopes of the two (x, p) regressions are now less clearly defined.
	Significantly different in the two materials ($r_1 \neq r_2$)		(4) Both roots determinate. One gives the slope of the diagonal mean regression fitted to the total (x, p) material, and the other root gives minus this slope. The assumption about uncorrelated shifts is now admissible only in conjunction with the assumption that both elasticities are equal in absolute value, but of opposite sign. If these assumptions can be made, the elasticities are given by the (x, p) regression slope. Otherwise the slope of the (x, p) regression expresses neither the demand elasticity) nor the supply elasticity nor the trend effect, but must simply be looked upon as an empirical fit to the actually observed (x, p) scatter, this scatter being the composite effect of elasticities and shift characteristics. The only exception is the extraordinary case where both materials show high observed correlations, one positive the other negative (and regression slopes of equal absolute value), and this can be interpreted as due to a trendconnection that has been exactly reversed from one material to the other.	(8) As a rule both roots determinate The shifts may or may not be correlated. If they are uncorrelated the two roots give the demand and supply elasticities correctly. Otherwise the two roots are meaningless.

$$A = 2(h_1 - h_2) - (k_1 - k_2) \qquad B = k_1 - k_2 - 2(k_1h_2 - h_1k_2)$$
$$C = h_1 - h_2 - (k_1h_2 - k_2h_1)$$

If $H_1 = H_2$ and $\varrho_1 - \varrho_2 = o$, we have by (3. 15) $(k_1 - k_2)(\alpha - \beta) = o$, that is $k_1 = k_2$, since we may disregard the trivial case $\alpha = \beta$. In other words, if we are in case (1), and if we make the assumption that the shifts are nearly uncorrelated, then k_1 and k_2 must be approximately equal. But if k_1 and k_2 are approximately equal, we see from (8. 1) that the size of $(K_1 - K_2)$ depends essentially on the size of the small deviations from zero which ϱ_1 and ϱ_2 actually exhibit. This means that in case (1) the roots of the two material equation will depend essentially on whether the closeness with which ϱ_1 and ϱ_2 come to being equal to zero is more perfect or less perfect than the closeness with which l_1 equals l_2 and r_1 and r_2 equals o. It is clear that in any practical case it would be absurd to make any assumption about this relative closeness. The case (1) in Table 1 must therefore be characterized as a case where both roots are meaningless through indeterminateness.

In case (2) the roots of the equation are determined by (7. 7) or with a better approximation by (7. 5). From the discussion attached to these formulae we know that one of the roots has now a definite meaning which is virtually independent of s; it represents the slope of the total regression. What is the meaning of this total regression in terms of the elasticities α and β? This cannot be deduced only from the nature of the observed (x, p) distribution. It depends on the nature of the (u, v) distribution. In this respect we have just the three cases discussed in Section 5: The high (x, p) correlation may be either a Cournot effect on the demand side, or a Cournot effect on the supply side or a trend effect. If we want to interprete the first root of the equation in terms of the elasticities, we have therefore the three possibilities indicated under (2) in the table.

The fact that we have a Cournot effect in case (2) does not prevent the existence of a high (u, v) correlation. From the discussion connected with (4. 4) we have indeed seen that we get a Cournot effect no matter what the shift correlation is, provided only that we have either demand stability or supply stability. And in connection with (4. 9) we have particularly studied the Cournot effect when there exists a high (u, v) correlation.

The other root in the case (2) depends essentially on s. Obviously s always gives an expression for a certain feature in the actually observed material, namely the ratio $\frac{r_1-r_2}{l_1-l_2} \cdot \frac{l}{r}$. But this in itself is without interest in the present connection. We are here interested to know if s has any meaning in terms of the elasticities α and β. This it cannot have unless approximately $\varrho_1 = \varrho_2 = 0$, because this was the condition under which the two-material equation was derived. (We may disregard the case where the condition $\varrho_1 = \varrho_2 = 0$ is not fulfilled, but where the roots of the equation nevertheless *by coincidence* give the correct value of the elasticities). Therefore s and consequently the second root in the case (2) is meaningless unless we have approximately $\varrho_1 = \varrho_2 = 0$. And if we have approximately $\varrho_1 = \varrho_2 = 0$ we are again in the situation studied under case (1), namely the situation where the root depending on s is essentially influenced by the *size* of the small deviations from zero which ϱ_1 and ϱ_2 actually exhibit. The second root in the case (2) is therefore always meaningless.

From the discussion and the graph of the function (7.6) follows that the case (3) is similar to case (2) except for the fact that the (x, p) regression slope is now less clearly defined.

In case (4) the nature of the roots is best seen directly from (6.6). If r_1 and r_2 are significantly different while l_1 and l_2 are nearly equal, the two roots simly become $\pm\, l$. (Compare also (7.2)). That is to say one root is simply the slope of the diagonal mean regression, and the other root is minus this slope. What is now the meaning of the regression slope in terms of the elasticities? From (3.16) we see that if the shifts are uncorrelated in both materials, and if $l_1 = l_2$ we must have $(\alpha^2 - \beta^2)\,(k_1 - k_2) = 0$. We may exclude the case $\alpha = \beta$, which is a concretely inplausible case. The possibility $k_1 = k_2$ is also excluded because this by (3.17) would entail $r_1 = r_2$. Thus, in the case (4) the assumption $\varrho_1 = \varrho_2 = 0$ entails $\alpha = -\beta$. And if we have the situation $\varrho_1 = \varrho_2 = 0$ and $\alpha = -\beta$, then by (3.16) $\alpha^2 = \beta^2 = l^2$ so that the elasticities can now be determined by the diagonal regression slope.

On the other hand if the shifts are correlated in the case (4), the regression slope in the total material will still be determined by one of the roots, but this slope cannot have any significance in terms of the elasticities (except by coincidence). And it cannot as a rule represent a trend effect either. Indeed, the trend effect situation is the one where there is a very high correlation between the shifts, entailing a high observed (x, p) correlation. In the case (4) this can only happen if the observed correlation in one of the materials is near to + 1 and in the other material near to — 1. We assume indeed in the case (4) that the observed correlations are significantly different. Therefore, case (4) cannot represent a trend situation except in the extraordinary case where both materials show high observed correlations, one positive the other negative (and regression slopes of equal absolute value) and where there is a definite reason for interpreting this as due to a trend connection that has been exactly reversed from one material to the other.

Now, consider the case (5). If $l_1 \neq l_2$ but $r_1 = r_2 = r$, we get from (6. 6)

(8.2)
$$\eta = L \frac{1 \pm \sqrt{1 - \frac{l_1 l_2}{L^2} r^2}}{r}$$

where

$$L = \frac{l_1 + l_2}{2}$$

This shows that in the case (5), where approximately $r = o$

(8.3)
$$\eta = \begin{cases} o \\ (\text{sgn. } r). \infty \end{cases}$$

In this case the assumption about uncorrelated shifts is not admissible, because if $\varrho_1 = \varrho_2 = o$ and $r_1 = r_2 = o$ we must by (6. 1) have $l_1^2 = l_2^2$ ($= -\alpha\beta$). Therefore, in the case (5) the two roots (8. 3) have no meaning in terms of the elasticities. And since there is no organization in the (x, p) scatter they have no meaning in terms of the (x, p) regression slope either.

In case (6) the roots will be determinate. And their values are easily determined from (8. 2). If $r_1 = r_2 = \varepsilon$ (where ε is either + 1 or — 1), (8. 2) gives

(8.4)
$$\eta = \begin{cases} \varepsilon l_1 \\ \varepsilon l_2 \end{cases}$$

In other words: the first root is the regression slope in the first material and the second root the regression slope in the second material.

Do these two slopes have any meaning in terms of the elasticities α and β? If we are in the extraordinary situation where one of the materials exhibits a pronounced Cournot effect on the demand side and the other a pronounced Cournot effect on the supply side, obviously the two regression slopes considered would have a meaning in terms of α and β. But otherwise the roots will have no such meaning. In particular if the high correlation in the two materials is a trend effect, the difference in the (x, p) regression slope in the first and second material would simply express the fact that the historical trend connection between x and p has changed from the first to the second material.

Also in the more general case (7) we have a similar situation: If the correlations in the two materials are fairly high, so that the regression slope is fairly distinct in both materials, and if the shifts are really uncorrelated, then the demand elasticity must be something near the regression slope in one of the materials and the supply elasticity something near the regression slope in the other material. In fact if $\varrho_1 = \varrho_2 = o$ and $r_1 = r_2 = r$, we have by (6.1)

$$\alpha\beta - (\alpha + \beta)\, r l_1 + l_1^2 = o$$
$$\alpha\beta - (\alpha + \beta)\, r l_2 + l_2^2 = o$$

This shows that l_1 and l_2 must be the two roots of the equation

$$\alpha\beta - (\alpha + \beta)\, r l + l^2 = o$$

In other words, we must have

$$l = r\left[\frac{\alpha+\beta}{2} \pm \sqrt{\left(\frac{\alpha-\beta}{2}\right)^2 - \frac{1-r^2}{r^2}\alpha\beta}\right]$$

Developing the square root to the first approximation in $1-r^2$, we get

$$l_1 = \alpha\left[1 + \frac{\beta}{\beta-\alpha}\cdot\left(\frac{1}{r^2} - 1\right)\right] r$$

$$l_2 = \beta\left[1 + \frac{(-\alpha)}{\beta-\alpha}\cdot\left(\frac{1}{r^2} - 1\right)\right] r$$

which shows that, if the correlation is fairly high, the regression slope in one material must be near to α and the regression slope in the other material near to β.

Thus, the hypothesis that the shifts are uncorrelated is also in the more general case (7) admissible only in conjunction with the hypothesis that one of the materials exhibits a Cournot effect on the demand side and the other material a Cournot effect on the supply side.

In the case (8) there is a possibility that the roots shall give the correct magnitudes of the elasticities α and β. This possibility will be realized when and (apart from coincidence) only when the shifts have actually been uncorrelated.

9. THE MISTAKE IN LEONTIEF'S METHOD.

The implausible character of the results obtained by Leontief's method has already been pointed out by professor Henry Schultz [1]). Schultz has carried out a series of numerical computations comparing the Leontief coefficients with those obtained by the Moore-Schultz method. From this empirical comparison he concludes that there must be something wrong with Leontief's method. Schultz was quite right in his suspicion, although he did not arrive at a theoretical explanation of the trouble. Such an explanation we are now in a position to give. Utilizing the results of the preceding Sections it is indeed an easy matter to pin down exactly in what Leontief's mistake consists.

Let us for a moment disregard the stability cases, that is to say the cases where the (x, p) distribution exhibits either a pronounced Cournot effect on the demand side or a pronounced Cournot effect on the supply side. (See cases (2), (4) and (6) in Table 1.) These cases are without interest in the present connection. In these cases a simple regression fitted to the (x, p) scatter, or even a freehand curve drawn through the swarm of (x, p) observation points, would give an expression for that one of the two curves: demand or supply, which it is in the case at hand possible to determine from the data. If Leontief's method shall have any raison d'etre, it must be through its application to the non-stability cases.

In these non-stability cases there may occur many sorts of situations. Frequently we will have a situation where the shifts

1) See his essay „The Meaning of Statistical Demand and Supply Curves". Mimeographed February 1930 and later printed in the series „Veröffentlichungen der Frankfurter Gesellschaft für Konjunkturforschung", edited by Dr. Eugen Altschul.

u and v exhibit a pronounced trend over time. Occasionally there may, however, occur a situation where no such trends exist. It is obvious that the demand and supply elasticities as Leontief conceives them are n o t expressions for the historical trend connection between price-and quantity. The elasticities in Leontief's system are very definitely some sort of neo-classical demand and supply elasticities. Trends in u and v are therefore in his theory to be looked upon as a disturbing element that must be eliminated, before the true demand and supply elasticities can come to play. Leontief is also quite definitely of the opinion that his method is capable of eliminating such trends. He says for instance (p. 20) about Moore's and Schultz's trend eliminations: „Diese Trendausschaltung ist nur ein unvollkommener Ersatz des hier angewendeten theoretischen und statistischen Hilfsmittels der Niveauverschiebungen.“ Leontief does not seem to be aware of the fact that the existence of trends in u and v is just one of the most important cases where u and v become c o r r e l a t e d and where consequently the determination of the elasticities by his method becomes meaningless.

So far from eliminating the trends, one of the „elasticities“ determined by his method is, as we have seen, nothing else than just the slope of the trend relation. Indeed from (4. 7) follows that if a pronounced trend is present in u and v, then there must be present a high degree of trend-correlation in the (x, p) scatter. And from case (2) in the table of Section 8 follows that this trend effect in (x, p) is just what is expressed by that one of the two roots that is not meaningless.

On the other hand Leontief obviously thinks that his method will yield significant results even in the case where no trends are present in u and v. Let us see if this is correct.

In case (4) the roots have a meaning in terms of elasticities only in the exceptional case when the two elasticities are equal in magnitude but of opposite sign (and the shifts are uncorrelated). And if these assumptions can be made, we have just a case where an ordinary regression fitted to the (x, p) material would give the elasticities. In other words there would not be any use at all for Leontief's method. And in case (6) and (7) the roots will have a meaning in terms of elasticites only if we are in the exceptional situation where one of the materials shows a more or less pronounced Cournot effect on the demand side and the other a Cournot effect on the supply side. That is to say we are again in a situation where there is no use at all for Leontief's method, but where a straight-forward regression fitting would give the elasticities.

Thus the only situation in which there might be a meaning in using Leontief's method, is the case where all the following three conditions are fulfilled:

1. The relative (x, p) violence must be significantly different in the two materials on which the computations are built.

2. The (x, p) correlation must be significantly different in these two materials.

3. The shifts must be uncorrelated in both materials.

In other words the appearance of the (x, p) distribution must be the one corresponding to case (8) in Table 1. And in addition to this we must be in a situation where there is a definite reason for assuming the shifts to be uncorrelated.

Is there a great likelihood that we shall meet such a situation in practice? I think it is safe to say that it would be a veritable miracle if we should ever find a material satisfying all these conditions and having nevertheless the same demand and supply elasticities. It would even be a miracle, I think, if the two observable criteria 1) and 2) should be satisfied. In virtually all practical cases where it is plausible to assume that the elasticities have been constant I believe we shall have the situation where at least one, if not both of the conditions 1) and 2) are violated. If Leontief had discussed the conditions 1) and 2), I believe he would have found that they are not fulfilled in any of his data. But Leontief has not gone into any analysis of these conditions.

It is true that there are passages where he expresses the necessity of using two materials between which there exists some sort of difference, but he does not seem to understand what sort of difference this must be. In particular he does not seem to be aware of the fact that there is only one special kind of such difference that is compatible with his own assumptions. He seems to believe that any kind of difference will do. He says for instance (p. 28): „Wenn es sich um zwei Zeitreihen von Mengen und Preisen handelt, können z. B. aus den beiden Hälften des gesamten Zeitabschnittes zwei gesonderte Preis-Mengen-Gruppen zusammengestellt werden für die trotz verschiedener Punktverteilung eine gleiche Elastizität angenommen wird.... Erwünscht ist stets eine Zweiteilung, in der die beiden Teilsysteme möglichst verschieden sind, also möglichst wenig ähnliche Preis-Mengen-Kombinationen enthalten."

Professor Robert Schmidt in his paper in Weltwirtschaftliches Archiv 1930, has discussed the nature of the difference between the two materials a little more closely, but also he skips entirely those questions that are the pertinent ones in connection with the economic and statistical m e a n i n g of the coefficients computed. First of all Schmidt only takes account of the appearance of the (x, p) distribution, without making the slightest attempt to interpret this apparance in terms of the shift distribution. He does not even seem to recognize those cases where the appearance of the (x, p) distribution is definitely incompatible with Leontief's assumption about uncorrelated shifts. His analysis does therefore not give any answer to the question of whether the coefficients computed have any meaning a s e l a s t i c i t i e s. But even as an attempt at interpreting Leontief's coefficients in terms of the observed features of the (x, p) distribution Schmidt's analysis misses the point. It does for instance not tell us anything about that fact which is the most interesting in this connection, namely the fact that if the two materials are rather similar and if the (x, p) correlation is fairly high, one root simply gives the regression slope while the other is meaningless.

The reason for this negative result seems to be that Schmidt bases his whole analysis on a single coefficient of „Prägnanz" arrived at by a purely formal application of certain classical facts from the invariance theory of quadratic forms, without discussing the economic meaning implied in this mathematical process. In the light of our analysis in Section 7, it is obvious that a single coefficient of „Prägnanz" must be an utterly inadequate, or rather a misleading tool in the discussion of the significance of the roots of the two-material equation. We only have to think of the stability cases where o n e of the roots has a meaning while the other root is meaningless.

To sum up we may characterize Leontief's method thus: In the first place we have the simple cases of a Cournot effect where a freehand line or some simple regression fitted to the material would reveal that one or those elasticities that can possibly be determined from the data at hand. In these cases there is no use for Leontiefs method. In regard to the other cases, Leontief believes that his method is capable of eliminating the trend relation between price and quantity and give both the demand and the supply elasticity. But the fact is that if there is no trend relation, both Leontiefs coefficients are meaningless. And if there does exist a trend relation, one of his coefficients gives just this trend relation while the other coefficient is meaningless.

Needless to say, these critical remarks have not been motivated by any special wish to reduce the value of Leontief's work as compared with the value of other works in this field. Science progresses slowly by trial and error. Much experimental work in new directions — such as Leontief's — may be stimulating and valuable, although it is found on closer examination that the results obtained do not carry the meaning one had in mind at the outset.

And on the other hand there are many other investigations in this field that may be exposed to a similar sort of criticism. Leontief's work is here discussed only as a typical example in order to draw attention to a large and important group of pitfalls which, as I see it, have been nearly completely overlooked in the statistical works of recent years, particularly in the works that have been directed toward a numerical determination of the shape of various relationships of economic theory. In this field we need, I believe, a new type of significance analysis, which is not based on mechanical application of standard errors computed according to some more or less plausible statistical mathematical formulae, but is based on a thoroughgoing comparative study of the various possible *types of assumptions regarding the economic-theoretical set up*, and of the consequences which these assumptions entail for the interpretation of the observational data.

The discussion in the preceding sections is a humble attempt in this direction. I believe that in the future a considerable part of the efforts of the quantitative economists will be devoted to this type of significance analysis, and that this will open up new perspectives of the whole question of the development ot economics into a genuine science where the abstract-theoretical and the observational-statistical approaches have become effectively united.