LECTURE 3

HOW DOES SCARCITY SHOW UP IN OUR RELATIONS?

3.1. The essence of a quantitative relationship

Since these lectures deal with quantitative research on income distribution, the central unit in which this type of knowledge is expressed can be seen to be an equation in the mathematical sense, or, in plain language, a *quantitative relationship*. In order to evaluate what has been accomplished by the authors to be discussed, I propose to start this part of my inquiry by trying to state what is essential to each of these units appearing in a quantitative model, supposedly tested statistically, and concentrating on relations reflecting *behaviour.* The last phrase implies that I am not interested in definition equations or balance identities.

In principle a relation presents a *causal connection,* with the aid of which we are able to calculate (or estimate) the value of a dependent or endogenous variable from the values of a number of independent or exogenous variables (endogenous or exogenous only for the relation considered). Among the exogenous variables there may be a random residual term reflecting some unspecified exogenous variables.

3.2. Examples of misleading relationships

In a number of cases a relation shown and even tested in a statistically satisfactory way (that is, showing a high correlation coefficient and low standard deviations of the regression coefficients) may be *misleading.* This means that under certain circumstances it will lead to erroneous results. This *unreliability* is a criterion to be derived not from statistical criteria but from criteria derived from the theory of the subject matter covered.

The clearest example is an *incomplete* relation, that is, one where a relevant exogenous variable is missing. Thus, in a demand equation a demand factor, which we know to be relevant from other sources of knowledge, may have to be present. Or, among the independent variables determining income, a relevant noncognitive ability may have been overlooked.

One reason why a relevant factor may have been left out is that in the statistical observation it did not vary. Observations about American incomes are all influenced by the American demand structure for different types of labour, and *that structure is the same for all observations*. We cannot find its impact and our relation will not necessarily apply to other countries. This applies to the relationship established by Mincer (1974), as I have set out elsewhere (Tinbergen, 1975, p. 54).

An important category of misleading equations are those in which *unidentified dummies* appear. A dummy is not by itself unidentifiable. We may measure some exogenous variables, such as the amount of schooling, in primitive ways, for instance by the number of years of schooling, or even by the figures 1, 2 and 3 for the level of schooling completed. This is a crude way of measuring, but generally not harmful. There is a more refined way of measuring if weights are given to each successive year or level, expressing their income formation power. A beautiful example of how this is done has been given by De Wolff and Van Slijpe (1972), using Husén's material. Similarly, the degree of independence of a job can be expressed by a dummy (cf. Tinbergen, 1975, p. 65). Dummies will become misleading, however, if the author does not *specify* which variable they represent. This happens if they are introduced for geographical areas, or for religious groups or for type of school; that is, if the underlying characteristics which really matter are not indicated. Which qualities, for instance, of various religious groups are responsible for differences in income? Is it that they are more motivated, or better 'mixers', or stronger? Is it how they will change over time or how they can be influenced? This is what matters for incomes policies, the ultimate goal of our research.

3.3. Statistical reasons for nonreliability

Alongside the theoretical criteria for nonreliability there are some well-known *statistical* criteria. Both have to be taken into account, to the extent possible. I already reminded you of the well-known question of identification of the relation to be tested and will discuss this aspect in some more depth a bit later (cf. sections 3.4–3.7).

First I want to list some well-known criteria, such as one detecting what Frisch called *multicollinearity*. Talking as a nonstatistician let me explain in the simplest terms what lies at the heart of this phenomenon. If two or more of the exogenous variables are moving – from one observation to the other – in an exactly parallel way, their impacts on the endogenous variable simply cannot be disentangled. The only way out here is additional information on those impacts, or, as the case may be, reasonable guesses or assumptions. A clear example within our subject is the inquiry made by Taubman (1974) into the education, jobs, and incomes of *identical twins,* which permitted him to rule out the purely genetical factor. His thesis is that other scholars have *overestimated* the impact of schooling. The implication is that some of the relevant causes other than education are so highly correlated with schooling that their influence has remained hidden. The point illustrates the rather primitive stage of our research, both with regard to theory and with regard to data available.

A similar phenomenon must be at stake when we find – which we do – that relations tested on *single individuals* show considerably lower correlations than relations tested on *average group values* for the independent variables. Groups may be inhabitants of local communities (for instance, Burns and Frech, 1970), provinces (Tinbergen, 1975, p. 44–45) or states, or age and schooling groups (Mincer, 1974), and so on. The fact that group figures show so much higher correlations must imply that relevant variables not explicitly mentioned in these relations are highly correlated with the exogenous variables included, or that they hardly vary between the groups, meaning that for the individuals within the groups they practically compensate each other. Thus the influence of age probably is hidden in relations based on local averages, since average ages vary too little among localities.

The other statistical criteria helping us to detect unreliability are the classical measures of standard deviations of regression coefficients, measures for randomness of residuals, and absence of serial correlation between residuals.

3.4. Differences in problem setting

In order to avoid formulating misleading relationships, we have to be perfectly clear about *what problem* we want to set out and answer by the relation formulated and tested. This may be the place to briefly summarize some economics of demand and supply analysis of markets; in our case markets of capital and mainly labour. A *demand* relation expresses the quantity traded X in terms of the price P and one or more demand factors D. Demand factors are exogenous variables affecting the attitudes of buyers. On the labour market they will refer to the organizers of production, who exert the demand for labour. In simple models the volume of production may act as a demand factor. In more complicated models the prices of other, competing factors of production, and hence of other types of labour, may constitute further demand factors.

A supply relation is a relation between the quantity traded X and price P, together with one or more supply factors S. Supply factors are exogenous variables, such as their parameters (abilities, needs, etc.), affecting the attitude of suppliers and hence employees.

Taking the demand and supply side of the market together, X and P are the endogenous variables and D and S the exogenous ones. The former can be expressed in terms of the latter which provides us with two equations, called by statisticians the reduced form of the system. Economists sometimes speak of the price equation $P = P(D, S)$ and the turnover equation $X = X(D, S)$ where price or quantity traded are expressed in terms of demand and supply factors, the numbers of which may vary. Each of the four relations so far discussed solves a different problem and we should be aware of which of the problems we are dealing with whenever we offer research results, in particular of econometric testing. As a consequence of specific assumptions some exogenous variable may be absent in one of the equations. The most common example in our subject, the labour market, is complete inelasticity of supply, meaning that all individuals in a certain quality class are offering themselves for employment regardless of the price. In that case the supply equation reduces to one in which the price does not occur. Another example may be the case where in the material studied, the demand factors, do not vary among observations and D disappears from both the demand equation and the price and turnover equations.

The simplest situation, in which only one demand and one supply factor are used, may be shown explicitly in the following formulae. Here we use lower-case letters to indicate logarithms, which may just as well be used and for which some concrete computations have been made (Tinbergen, 1975a, 1975b). Writing the supply and demand equations with the additional procedure to express all variables as deviations from their averages,

$$x = \alpha_1 s + \alpha_2 p \,, \qquad \text{(supply)} \qquad (1)$$

$$x = \beta_1 d - \beta_2 p \,, \qquad \text{(demand)} \qquad (2)$$

and writing the reduced form equations as

$$p = -\pi_1 s + \pi_2 d \,, \qquad \text{(price)} \qquad (3)$$

$$x = \xi_1 s + \xi_2 d \,, \qquad \text{(turnover)} \qquad (4)$$

we can easily verify that

$$\pi_1 = \alpha_1/(\alpha_2 + \beta_2) \,, \qquad \pi_2 = \beta_1/(\alpha_2 + \beta_2) \,, \qquad (5)$$

$$\xi_1 = \alpha_1 \beta_2/(\alpha_2 + \beta_2) \,, \quad \xi_2 = \alpha_2 \beta_1/(\alpha_2 + \beta_2) \,, \qquad (6)$$

Two particular cases are mentioned. The first is inelastic supply:

$$x = s \,, \qquad (7)$$

if by s we simply mean the numbers available in the various quality groups; we can substitute the values $\alpha_1 = 1$ and $\alpha_2 = 0$ in eqs. (5) and (6), obtaining

$$\pi_1 = 1/\beta_2 , \qquad \pi_2 = \beta_1/\beta_2 . \tag{8}$$

The second case is constant demand structure:

$$x = -\beta_2 p , \tag{9}$$

leading to

$$\pi_1 = \alpha_1/(\alpha_2 + \beta_2) , \qquad \pi_2 = 0 . \tag{10}$$

We did not add the turnover equations since we will deal mainly with prices (incomes). They could be easily added.

There is a further category of relationships to mention when we are interested in *distribution* problems. The endogenous or dependent variable will then be some measure of income distribution. Taking the standard deviation as this measure, we obtain

$$\sigma_p^2 = \pi_1^2\sigma_s^2 - 2\pi_1\pi_2 r_{sd}\sigma_s\sigma_d + \pi_2^2\sigma_d^2 , \tag{11}$$

where σ_s are standard deviations and r_{sd} is the correlation coefficient between s and d. If we have more than one supply or demand factor, the relationship will become more complicated.

Taking some ratio between two incomes (i.e., prices) as a distribution measure, the result may contain, apart from distribution measures of the supply and demand factors, their average levels, since the definitions of p, s and d are deviations from their averages and these averages need not always cancel out.

The main point to be made about the present question is that one has to be quite careful in identifying a relation formulated and tested. Several authors have attempted to establish relations pretending to explain *income as a function of a number of exogenous variables.* Since p appears in eqs. (1), (2) and (3) we have to clearly state whether we claim our relation to be a supply, a demand or a price equation in the terminology here used. There are a number of important results, for instance those of Mincer (1974), Chiswick (1974), Taubman (1974) and Thomas (1973), where this question arises. In all of them only supply factors, namely various human qualities, appear as explanatory variables. This eliminates the possibility of their being a demand equation; but are they a supply or a price equation? Possibly they are both, but with a special assumption added, namely that α_1 and α_2 are very large; otherwise there should have been a term in x. Alternatively, the assumption is that π_2 is negligible, meaning that the term $\beta_1 d$ in the demand equation has been neglected. This seems to me to be the more reasonable assumption, probably adhered to by all the authors quoted (except perhaps Thomas). The smallness of the term $\beta_1 d$ cannot, however, have been caused by a low value of elasticity β_1, but rather by the constancy of d, because all observations refer to one country. This, then, at the same time characterizes the *limits of validity of the relations discussed.* They are unable to answer the question of what will happen if demand structure in the U.S. changes. My extremely simple exercise to estimate (3) (Tinbergen, 1975b) is in that respect more general and covers a larger variety of situations. Thomas' relation has the disadvantage, already mentioned, that it uses dummies for the six Yugoslav republics, without identifying them as either demand or supply factors.

3.5. Differences in modelling

Apart from differences in problem setting, different ways of modelling are possible. One well-known choice the model builder has is whether to let a variable take a few *discrete* values, for instance, education required and actual education at three levels only, or to let that variable be a continuous variable. The same applies to other qualities. We have already discussed the degree of independence. In the model for the Netherlands, I used this approach to inquire on the impact of schooling required and actual schooling on income, and, next, to determine the optimal income distribution

(Tinbergen, 1975, ch. 7). This meant that five types of labour were introduced, indicated by two indices each, the first referring to the schooling level required and the second to the actual level: (11), (21), (22), (32) and (33). Other possible combinations were omitted on the assumption of scarcity of qualified labour in comparison to less qualified labour. It would be uneconomic to use, for instance, people with a third-level education for jobs requiring second-level education as long as the former are scarce. And around 1962, the year to which the model applies, the number of people in such positions was indeed small. With this approach the maximization of utility allegedly sought by each group took the mathematically simple form of stating that incomes for people with education level 2, for instance, in jobs 3 and 2 should be such as to give equal satisfaction.

In another model which I used for the formulation of the demand–supply theory (Tinbergen, 1956), I chose in favour of a *continuously* changing variable for each of the two qualities considered most important. In the article they were indicated by the symbols t_1 and t_2 for the actual qualities and by s_1 and s_2 for the required qualities. Welfare ω was assumed to depend on income l and the four other variables (where the t's are parameters in our present terminology) according to the formula

$$\omega = \omega_3 \log l - \tfrac{1}{2}\omega_1 (s_1 - t_1)^2 - \tfrac{1}{2}\omega_2 (s_2 - t_2)^2, \quad (12)$$

where the 'tension', discussed earlier, was used. With this style of modelling the maximization of welfare by the appropriate choice of a job takes the form

$$\frac{\partial \omega}{\partial s_i} \equiv \omega_3 \frac{\partial \lambda}{\partial s_i} - \omega_i(s_i - t_i) = 0 \,, \qquad i = 1, 2 \,, \qquad (13)$$

where $\log l = \lambda(s_1, s_2, ...)$ is the wage scale. This leads to an elegant and very simple mathematical setup. The wage scale was specified as a quadratic function in the s's:

$$\lambda \equiv \lambda_{00} + \lambda_{10}s_1 + \lambda_{01}s_2 + \tfrac{1}{2}\lambda_{20}s_1^2 + \lambda_{11}s_1s_2 + \tfrac{1}{2}\lambda_{02}s_2^2 \,, \qquad (14)$$

and one of the results requires our attention. The coefficient λ_{10} of s_1 turned out, for demand and supply to balance, to be

$$\lambda_{10} = \frac{\omega_1}{\omega_3}\left(\frac{\bar{s}_1}{\sigma_1} - \frac{\bar{t}_1}{\tau_1}\right) , \qquad (15)$$

where barred symbols stand for averages over all individuals and σ_1 and τ_1 for the standard deviations of s_1 and t_1, respectively. Expression (15) clearly is a measure of the scarcity of quality 1: it contains the difference between the average values of s_1 and t_1, measured in standard deviations. It is too bad, of course, that the tension theory is not so relevant (as seems to follow from my research so far).

Whereas my primary intention is to illustrate the difference in modelling between the two approaches summarized, the example brings us automatically to the second subject, the main subject of this lecture, namely to show more explicitly how the element of scarcity shows up in a number of formulae explaining incomes. In eq. (15), on the one hand, scarcity is determining the *coefficient* in the wage scale to be given to quality number one. In eq. (3), on the other hand, scarcity is expressed as a weighted difference $\pi_2 d - \pi_1 s$ between demand factors and supply factors, which are variables, and hence scarcity also appears as a *variable*. A simple example may be given as an illustration. As already mentioned, s sometimes can be interpreted as the (log of the) number of people available with the quality considered. In addition, d can be understood to be the (log of the) numbers *wanted* by the organizers of production at the zero value of p, which is the average log of prices. In two recent attempts to give precise content to these concepts (Tinbergen, 1975a, 1975b), I used two methods to estimate these numbers d for various qualities of labour. For Mexico I estimated the numbers needed if the organizers of production had in mind a further development of the country towards the industrial structure of Japan – a situation which they may reach in 10 to 30 years (applying the growth rates of Japan and of Mexico, respectively). For the U.S. I estimated the numbers needed if the organizers of production had in mind the professional structure 10 or 20 years ahead, derived from the changes over the years 1949–1969. In both cases I got remarkably good fits for eq. (3), with simple demand elasticities (equal for all types of labour because of the imposed logarithmic definition of p, d and s) not far from my favoured value of unity.

3.6. Analytical vs. normative problems

A very important classification of problems to be dealt with is the one into analytical or explanatory and normative or optimizing problems. One way of formulating the difference is to state that in the first category we consider the means or instruments of policy to be given and the results the unknowns of the problem, whereas in the last category the policy objectives or aims are considered given and the means or instruments the unknowns.

In our problem area, analytical problems are to find the ultimate determinants of incomes and income differences. Most of what was dealt with in what precedes was indeed of an analytical character. Normative problems have to start from an aim formulated by some policymaker or by the author himself and then ways and means found to attain such aims. We gave some attention to them in section 2.1, where concepts such as equity and optimal income distribution were discussed. We propose to take up questions of incomes policies in the widest sense in Lecture 4.

Generally speaking, the relations composing a model for an economy need not be different in order to enable the investigator to tackle either an analytical or a normative problem. It is rather a question of what the unknowns of the problem are. An interesting recent study using a very detailed model for both purposes, and including the income distribution problem, is one concerning the Republic of Korea (Cohen, 1975). A number of studies about income distribution and employment have been tackled by the International Labour Office as part of the World Employment Programme (1975).

3.7. Short-term and long-term problems

Economics traditionally has emphasized the time component in its problems, notwithstanding the Keynesian dictum that in the long run we are all dead. What matters, however, is also the life or death of our offspring. Econometric models, which came into being in the latter 1930s, are in many respects very similar to the psychologist's 'path analysis', already in use two decades earlier, with the clear difference that the *time structure* has been given much more care in the former. Various types of time periods enter into general economics, such as the age or lifetime of human beings, of trees and other capital goods and of pigs and other consumer goods, or the time taken by a production process, or, finally, delays or lags due to psychological reaction periods.

Also our subject shows various relevant time dimensions. We already discussed the number of years of experience and the number of years of schooling, the former with an upper limit of around sixty years, the latter with a usual upper limit, for formal schooling, of twenty years, and many shorter relevant lags. In his study of the market of college-trained technicians, Freeman (1971) makes use of the lag in supply flow due to the normal course length of four years.

In a less precise way the difference between short-term and long-term problems was already discussed, when we took up the subject of substitution elasticities between various types of labour. In my Dutch model the difference was due to the distinction between individual decisions which take for granted productivity differences between various types of labour, and the collective impact of all individual decisions on this productivity, this being the result of macro-shifts in manpower from one compartment of the labour market to another. The time factor in income distribution problems becomes quite evident if we consider the consequences of an increase in the supply of qualified manpower of one kind or another. Considerable increases require time because most schooling takes place in certain age groups and several vintages may therefore be involved in order to attain a change as assumed or desired. This necessitated our study of long-term changes extending over half a century or a century, for instance the race between technological development and education. It is said of some changes in attitude or performance that they take generations. Our data hardly enable us to verify these statements.

3.8. Sample size and dispersion

The way in which relations formulated and tested express the impact of scarcity on incomes is also influenced, finally, by the size and diversity or dispersion of our statistical material, or sample. I am not thinking so much now of the number of observations, which plays its own well-known role in, for instance, the reliability of regression coefficients, measured one way or the other. I am rather thinking of whether the sample or its individual observations refer to one country or many countries, to municipalities, districts, provinces, or states within one or within a number of countries, and the validity of the findings connected with these aspects. For a demand–supply theory to be applied, I think, for instance, that figures about individual municipalities are not appropriate, since commuters have one place where they live and another where they meet demand for them.

I already touched upon the limited validity of an equation referring to an area in which some relevant variables do not vary (cf. sections 3.2 and 3.4). This drawback can be overcome by considering observations for the same country at *different points of time*; or by taking data for various geographical units and adding a demand factor for these units. This is how I tried to use the material Dr. Chiswick was kind enough to make available to me. The condition must be fulfilled, then, that not many people commute between states.