4

Supply factors and utility or welfare

4.1. Utility function as the source of supply

Economists agree that an individual's supply behaviour concerning production factors and in particular labour derives from the maximization of his utility function under the restrictions prevailing. These restrictions are usually supposed to consist of the prices prevailing in the market compartments to which the individual has access and the individual's own characteristics. Agreement among economists stops when the question is asked whether the utility function can be specified by quantitative observations or *measurement*. The overwhelming majority of the profession denies this possibility. Cassel for this reason preferred not to use utility functions at all and to base economic science on supply (and demand) functions without an attempt to give them a deeper foundation. Here lies a danger of inconsistencies, however; and most economists prefer to maintain the concept of utility, adding, though, that only supply behaviour can be measured. Although the measurements so far presented in this study are not affected by the controversy on the measurability of utility and although the question of how income distribution can be influenced, can be dealt with without making use of such measurement, we will choose in favour of a quantitative specification of utility or welfare. One reason is the avoidance, already mentioned, of inconsistencies; another the gain of clarity obtained in the search for additional supply factors, one of the obvious lacunae in this field of research. A final reason for our choice is the necessity of a specification if we want to give content to the notions of an *optimal* and of an *equitable* income distribution – two different concepts, both of considerable practical importance. I want to add that my choice in favour of measurement of utility or welfare is shared, although with a different elaboration, by Van Praag and his collaborators [79, 80].

4.2. Specification of utility function

In order to avoid misunderstanding I want to state from the outset that I am not claiming that the treatment to be proposed and to be tried out in a first few steps is part of economic science only. It contains elements which may be named, respectively, (i) *methodological* or *philosophical* and (ii) *ethical* or *moralist*. What matters of course is not so much whether a contribution to the solution of a problem belongs to the realm of one or of another science or discipline; rather it is whether indeed a contribution is made.

To start with I submit [65] that three elements enter into a utility or welfare function, to be called variables, parameters and constants or coefficients. *Variables* are elements changing over time and among individuals. Some of them can be chosen by the individual (or the household) considered, others cannot. The most important choice variable in our context is the choice of a *job* or *occupation.* Examples of variables not under his control are changes in wage or salary scale or the general price level, in brief the socio-economic environment. This statement reflects my opinion that man-made monopolies are less important than is sometimes believed. The theory can be adapted to a different belief, however. Whereas in our context the choice of one's job is the most important one, for young generations the choice of *education* is the central one, not unrelated, of course, to an implied choice of occupation. Jobs can be represented by figures such as are supplied by *job evaluation;* or, alternatively, occupations can be represented by their occupational status [33].

Parameters are supposed to be constants for the individuals (or households) in the time span considered, but to vary among them. They characterize the individuals, households, or even groups of these under consideration. There are two groups of parameters: those characterizing the *productive capabilities* and those characterizing the *consumption needs*. Among the former, cognitive capabilities as shown by IQ stand out, but are by no means the only ones. Other features, corresponding for instance to the aspects used in job evaluation, are important too. We may think of the capacity to decide and act independently, as required for leading positions; of the willingness to co-operate with others; persistence, technical intuition, and so on and so forth. For lack of information we may sometimes replace a real parameter by a dummy. In our first step of applying the theory we will take as an important productive parameter *years of schooling* completed, which in actual fact is not exactly a parameter, but probably approximately so. Psychologists such as Duncan [23] and Jencks [33] have dug deeper already and eventually their work should be integrated into ours.

Of the second group of parameters, those reflecting needs, we think of *family size* as an important dummy; even though for certain problems again, it is essential to go back to its determinants. If in coming periods birth control is to play an important role, we must be aware of the desire to have at least a few children as a very important component of happiness. In the context of a poor population without an old-age pension children are a form of social insurance – and so on. Another group of need parameters is based on the individual's *health* and may therefore be represented by some health indicator. For larger groups of households these parameters may not in the average vary much and may therefore be disregarded. Inquiries based on individual households cannot neglect them.

Coming to the *coefficients*, these are the intensities by which variables and parameters affect utility or welfare; and here our methodological approach is that they are considered the same for all individuals/households considered. As soon as some student of welfare maintains that such an intensity is not equal for all, he is talking about a parameter, not a coefficient. The set of coefficients should be seen as characterizing the human species or some subspecies, in contradistinction to other animals. They may also be considered as an expression of some sort of '*equivalence*' of all human beings, dealt with in more depth in Chapter 8.

Mathematically the coefficients also depend on the functional shape given to the utility or welfare function and here we have, as is well known, some degrees of freedom. We have a choice between all functions which can be obtained from each other by monotonous transformations and we are free to choose some appropriate form as long as observations do not refute it. A simple example chosen to be used in this book is

$$\omega_i = \ln\{x_i - c_0 s_i - \tfrac{1}{2}c_2(s_i - v_i)^2\}, \tag{4.1}$$

where

ω_i is household i's utility,
x_i income after tax in thousands of guilders per annum,
s_i represents the job held by the head of the household, expressed in required level of schooling (in units of three years),
v_i actually attained level of schooling in the same units,
c_0 and c_2 coefficients.

This choice is extremely simple, as a first step to illustrate our procedure. It contains two variables, namely x_i and s_i, and one parameter v_i. It may be interpreted as a function of *consumable income corrected for two inconveniences*, namely the effort required* by the job $c_0 s_i$ (or, if c_0 were found to be negative, the additional satisfaction, on top of that of income, experienced in the job), and the inconvenience (or *tension*) experienced by a possible divergency between required and actual schooling. This tension is supposed to be equal for positive and negative divergencies of the same size. Since generally the tension will not be large, the quadratic function

* There is empirical evidence that working hours differ little for fully-employed persons in a large spectrum of jobs; hence no attempt was made to introduce this variable.

is a sufficient approximation in the neighbourhood of the minimum tension. An alternative function will be discussed in Section 4.6.

No other term in v_i has been included. There are two ways of interpreting this further simplification. One is that the only way in which a higher capability v_i affects welfare is that the level of reference for the *tension* is shifted, but that the only thing that matters for the tension is the divergence in years between required and actual schooling. The other interpretation is that we *cannot know* an additional (say, linear) term in v_i, since by definition nobody can shift from one value of it to another. If we supplement this ignorance with the assumption that a higher level of capability affects welfare positively, leaving out the term in v_i means that we underestimate welfare more the higher v_i is and that our estimates of welfare differences are underestimations. As we shall see, this strengthens our main conclusion about the inequity of the present income distribution (Chapters 7 and 8).

4.3. Two methods of measuring welfare

So far in the Netherlands two methods of estimating the welfare of various individual households or groups of households have been tried out. These two methods have been used alternatively for many other purposes as well. The first is to *interview* the (heads of) households themselves. This method has been used by Van Praag and collaborators [79, 80], who interviewed 3000 members of the Belgian and 3000 members of the Dutch consumer union. The second method consists of observing people's *behaviour* rather than their stated opinions. The advantage of the first method is that a set of figures is obtained, based on a uniform set of questions and also containing figures for individuals with unequal capabilities; a disadvantage is that experience in other cases has shown that opinions do not always correspond with later behaviour. Somebody earning Dfl. 15,000 at the moment of the interview in Van Praag's inquiries does not only state what level of satisfaction he feels with this actual income, but also states the level he would experience with an income of Dfl. 20,000. If he later actually earns the latter amount, will he maintain his statement? The material collected already suggests he will not ('preference drift' is observed). The advantage of the second method is that only behaviour is observed, but the disadvantage is that this excludes comparison between households whose heads show differing capabilities, as already observed. This isn't true if one accepts our argument to include no other term with v_i.

It goes without saying that Van Praag and I are working on an attempt to use his material in order to cross-check my findings. The results, which at the moment of writing, are not yet available in final form, will be published elsewhere. In this book the results of the second method only will be dealt with.

4.4. A test based on material for the Netherlands, 1962

The main test made of the utility function (4.1) was based on statistical material for the Netherlands in 1962.

(a) The essential characteristic of the material is its subdivision of the labour market into 21 compartments, later reduced to 19, each compartment being characterized by its s and v, where s is assumed to be the upper quartile of the corresponding v-distribution. The statistical material used is shown in Table 4.I and the way it has been estimated is described in the Annex. Since v is supposed to reflect the (only) parameter in the utility function and free competition is assumed to exist, a free choice of all individuals with a given v_i between the compartments showing that v is possible; and the coexistence of various values of s in these compartments enables us to test the relation, derived from (4.1),

$$\begin{aligned} x_i - c_0 s_i - c_1 v_i - \tfrac{1}{2}c_2(s_i - v_i)^2 \\ = x_j - c_0 s_j - c_1 v_i - \tfrac{1}{2}c_2(s_j - v_i)^2. \end{aligned} \tag{4.2}$$

Some terms in v_i cancel each other out. We can use as observation pairs any two pairs with equal v_i for all values of i.

Table 4.I
Estimated income distribution of the Netherlands, 1962, according to schooling (v), required schooling (s) and degree of independence (W, alternatively W').

s		v 2	3	4	5	6	W (W')	Description of main groups included
6	f			0.1	0.3	0.9		Free professions;
	l			42.3	42.3	42.3	2	Directors of limited
	l'			19.0	19.0	19.0	(6)	companies
	x			14.0	14.0	14.0		
5	f		1.5	1.1	0.5	0.7		Independent industrialists;
	l		19.5	19.5	19.5	19.5	1	Teaching staff of secon-
	l'		14.2	14.2	14.2	14.2	(5)	dary and third level
	x		11.3	11.3	11.3	11.3		
4	f	0.7	5.3	4.4	1.3	0.1		Civil servants in general
	l	10.2	10.2	10.2	10.2	10.2	0	service; Private employees
	l'	9.7	9.7	9.7	9.7	9.7	(2)	
	x	8.3	8.3	8.3	8.3	8.3		
3	f	0.1	4.0	1.0	0.1			Independents in trade
	l	15.0	16.2	17.4	18.6		1	and services
	l'	12.0	12.5	13.0	13.4		(3)	
	x	9.9	10.2	10.5	10.8			
2	f	10.7	6.6	0.1				Primary school teachers;
	l	8.9	10.8	12.7			0	Part of administrative
	l'	8.5	10.2	11.3			(2)	personnel; Police; Mining
	x	7.4	9.0	9.4				workers; Farmers
2	f	46.6	13.8					Workers, incl. Retail trade
	l	5.3	8.4				0	employees; Part of
	l'	4.9	8.1				(2)	administrative personnel
	x	4.4	7.1					

Symbol definition :
f : Frequency in per cent of number of taxpayers.
l : Income in thousands of guilders.
l' : Labour income contained in l.
x : Labour income after tax.

The correlations found are unsatisfactory, however, cf. Table 4.II, cases A–E. As a consequence, the theory in its initial form *had to be rejected.* The correlation coefficient between $x_i - x_j$ and $z_{ij} = (s_i - v_i)^2 - (s_j - v_i)^2$ indicated by x and z in Table 4.II, is low (0.14) and the regression coefficients, even after the introduction of other explanatory variables, are insignificant. A new variable which seemed to have high explanatory power will now be discussed.

(b) The description of the social groups under consideration suggested that an additional capability is relevant, namely the ability to make independent decisions. The groups considered were divided into three groups with decisions of increasing importance; the groups are represented by a parameter W. Persons having to make decisions of minor importance only and usually described as socially dependent were given the value $W = 0$; persons heading small and medium-sized enterprises, $W = 1$; and heads of limited companies or persons in liberal professions, $W = 2$. As can be seen from cases A–E in Table 4.II, satisfactory correlations were only obtained when W was included. This state of affairs constitutes a classic example of the necessity of introducing an additional element into a theory in order to fit the facts in a more satisfactory way. The basic idea may also be expressed as the capability of giving guidance to other people, a capability given a central place in income distribution by Tuck [73]. An alternative scale W' was chosen and is also shown in Table 4.I.

(c) The *interpretation* of the role of W poses some questions, however. Must we consider W as a parameter, as suggested, which implies that it indicates an innate personal capability; or must we consider it as a job variable? I opted for the first interpretation, which gives to W a role comparable to the role of v in the simplest model. This implies that we have to consider as separate groups not only those having a different v, but also those having a different set of v and W; Table 4.III indicates the arrangement of our statistical material corresponding to that interpretation. This arrangement reduces the number of observations to six pairs (each with another set v, W).

It appears that significant results can now be obtained (Table 4.II, cases F, G and H). They differ because in case F the

Table 4.II
Results of testing the two theories by equating utility between cells with same v (Theory I, cases A through E) or between cells with same v and W (cases F, G, H).

Case no.	N	Regression coefficients for x on							Simple corr. between x and				Const. term	Mult. corr. coeff.
		z	w	q	q'	q''	m	v	z	w	q	v	e	
A	16	0.16 (0.13)	1.58 (0.63)	−0.20 (0.20)	–	–	–	–	0.14	0.89	−0.83	–	+0.57	0.91
B	16	0.17 (0.14)	–	−0.10 (0.37)	–	–	–	–	0.14	0.89	−0.83	–	+0.49	0.88
C	16	0.16 (0.13)	1.61 (0.96)	−0.21 (0.35)	–	–	−0.01 (0.37)	–	0.14	0.89	−0.83	–	+0.57	0.91
D	12	0.10 (0.21)	–	–	−0.01 (0.21)	0.17 (0.08)	–	–	0.14	0.89	−0.83	–	+0.19 (0.45)	0.80
E	16	0.15 (0.13)	2.13 (0.29)	–	–	–	–	–	0.14	0.89	−0.83	–	+0.55 (0.27)	0.90
F[a]	6	0.44 (0.12)	–	–	–	–	–	1.10 (0.40)	0.63	–	–	−0.20	−3.6 (1.42)	0.91
G[b]	6	0.28 (0.28)	–	–	–	–	–	−0.17 (0.44)	0.79	–	–	−0.70	+1.47 (3.34)	0.79
H[b]	6	0.32 (0.10)	–	–	–	–	–	–	0.79	–	–	–	+0.9 (0.5)	0.79

[a] The lowest row of observations in Table 4.I has been excluded.
[b] The lowest two rows of observations in Table 4.I have been combined.

Symbol definition :
m : wv.
w : $W_1 - W_2$.
N : Number of observations of differences between two cells with same utility.
x : Differences in labour income after tax between two cells or $x_1 - x_2$.
z : $s_1^2 - s_2^2 - 2v(s_1 - s_2)$; s schooling required, v schooling available.
q : $W_1^2 - W_2^2 - v(W_1 - W_2)$, W degree of independence (as in Table 4.I).
q' : $W_1'^2 - W_2'^2 - v(W_1' - W_2')$, W' degree of independence (same, but 5 levels instead of 3).
q'' : $W_1'^2 - W_2'^2$.

lowest line of Table 4.I has not been used, since the category for $v = 2$ contains a group of holiday workers which somewhat depresses the income figure. Since, however, the figure covers also a large portion of ordinary workers, it was decided to include them, by combining them with the other group with $s = 2$. This yielded cases G and H. Since in case G, in addition, the influence of v became insignificant, case H was added. This provided us with some further insight into the sensitivity of e, representing the influence of s on x and hence on ω. While in case F (valid for the half of the population with high incomes) the influence of s was found to be negative, in case H (valid for the population as a whole) this influence was found to be positive. These findings are interesting since their meaning is that for higher incomes the attractiveness of a higher s job surpasses its disutility, while for lower incomes the opposite applies. Notwithstanding its lower multiple correlation coefficient we take case H as the most satisfactory result because (i) its coefficients take an intermediate place between those of cases F and G and (ii) because its coefficients are theoretically more attractive, while (iii) it will keep us, because of the sign of the constant term, on the safe side in our estimation of the optimal and the just income distribution. Written in the form of (4.2) the relationship runs:

$$x_1 - 0.45s_1 - \tfrac{1}{2}\cdot 0.64(s_1 - v_j)^2 = x_2 - 0.45s_2 - \tfrac{1}{2}\cdot 0.64(s_2 - v_j)^2, \qquad (4.3)$$

where

$s_1 = 2, 2, 3, 2, 3, 3;$
$s_2 = s_1 + 2;$
$v_j = 2, 3, 3, 4, 4, 5.$

This corresponds with

$$\omega = \ln\{x - 0.45s - 0.32(s - v)^2\} \qquad (4.4)$$

as the resulting utility or welfare function.

(d) Whether W is a parameter or a variable, in both cases we must face the problem of what the corresponding variable

Table 4.III
Distribution of labour income after tax over values of s, v and W (degree of independence).

	v	2		3		4			5			6		
s	W	0	1	0	1	0	1	2	0	1	2	0	1	2
6								14.0			14.0			14.0
5					11.3		11.3			11.3			11.3	
4		8.3		8.3		8.3			8.3			8.3		
3			9.9		10.2		10.5			10.8				
2		7.4		9.0		9.4								
2		4.4		7.1										

or parameter must be to define the 'tension' concept introduced in Section 4(b). For lack of observations the only alternative I investigated assumes that in the case where W is considered a job variable rather than a personal parameter, the corresponding parameter is again v. This requires the introduction of an expression $(W - \beta v)^2$ instead of W as an explanatory variable, where β is a scale coefficient to be estimated from regression analysis. The results of this attempt will be found in Table 4.II, cases A–E, from which it becomes clear that:

(i) correlations were substantially improved ($R \gg r_{xz}$) by the introduction of $W_1 - W_2$ or w, as shown even in $r_{xw} = 0.89$, and

(ii) attempts to introduce $(W - \beta v)^2$ were unsuccessful; they did not improve the multiple correlation and the expression $(W - \beta v)^2$ obtained insignificant regression coefficients; from this I conclude that W is a personal parameter rather than a job variable to be approached by schooling.

(e) As long as in fact parameters v or W are innate, incomes will depend on them as a consequence of a relative *scarcity* of these human properties. Comparisons of the type of (4.2) using different values of v (or W) instead of s and the same values of s on both sides will not reflect the psychological impact on ω of v (or W), but its relative scarcity, that is, the impact of the difference between demand for and supply of the capability considered. Since the frontier between scarcity and power is difficult to draw, one may even interpret the additional scarcity income received by those endowed with a high value of v or W as an income derived from 'power' and see it as an element of 'exploitation'.*

4.5. A test based on material for Illinois

The basic tool of the simplest test shown in the previous section is a two-entry table showing incomes of individuals arranged according to education (horizontal) and occupation (vertical), both measured in comparable units. In the primitive stage of the testing shown vertical differences are the object of free choice, whereas horizontal differences are scarcity differences due to the impossibility of shifting horizontally. For lack of more information and by way of experiment, census data for Illinois were used with the heroic assumption that the main non-farm occupational groups (h) placed in the order of their average incomes, require the education, in terms of school years completed, indicated in the left-hand column of Table 4.IV.

The sum of the totals of the occupational groups given by Dougherty [22] falls short by 441 (thousand persons) of the sum of the totals of the educational groups. In order to arrive at an identical grand total these 441 without occupation were distributed proportionally over the occupation classes.

Thus, the column and row indicated 'total' in Table 4.IV were given. In order to fill in the cells of the table the method of the northwest-corner rule was applied, well known from the 'Hitchcock problem' [30, 43]. The method may be interpreted as an 'optimum' use of the available manpower in the various occupation classes. Here optimum means the *cheapest way of having the jobs done*, in the sense of using people with the lowest education available first and only adding people with higher education as long as not sufficient cheaper manpower

* Interesting enough, a case of 'exploitation' by *human* capital.

Table 4.IV
Estimated employment according education h' and occupation h; Illinois, 1959, in thousands of male equivalents.

h	h'								Total	Denomination
	0	3	6	8	10	12	14	18		
0	54	155	.	.	.	.	.	.	209	Labourers
3	.	12	287	.	.	.	.	.	299	Service workers
6	.	.	127	638	.	.	.	.	765	Operatives
8	.	.	.	90	455	.	.	.	545	Clerical workers
10	.	.	.	.	198	439	.	.	637	Craftsmen
12	.	.	.	.	.	233	.	.	233	Sales workers
14	.	.	.	.	.	88	286	.	374	Professionals, technicians
18	.	.	.	.	.	.	15	304	319	Administrators, proprietors
Total	54	167	414	728	653	760	301	304	3381	

was available. As could be expected, only a number of cells around the main diagonal are non-empty, but enough of them to make vertical comparisons possible, that is comparisons of groups with the same education but in different occupations.

The next step is to estimate the incomes for the non-empty cells. As income we take the average of the column and of the row average income. Thus, for cell (1; 1) we take the average of 2.8 and 3.9, rounded to 3.4; and so on. This produces Table 4.V. In the 'average' column and row the cell figures have been averaged, using as weights the employment figures of Table 4.IV. In brackets the true averages have been added and it will be seen that the deviations between our 'estimated' and the observed average incomes are not large.

The units chosen for h' are, as already said, years of schooling; in order to attain a crude degree of commensurability between h and h' the eight occupation classes have been numbered in the same way, suggesting that these figures

Table 4.V
Average earnings for the cells of Table 4.IV, estimated as the averages of the column and the row average average observed and placed in parentheses; Illinois, 1959, in thousands of US $.

h	h'								Average
	0	3	6	8	10	12	14	18	
0	3.4	3.6							3.5 (3.9)
3		3.7	4.1						4.1 (4.0)
6			4.6	5.1					5.0 (5.1)
8				5.1	5.7				5.6 (5.2)
10					6.4	6.9			6.7 (6.6)
12						7.5			7.5 (7.8)
14						8.2	9.0		8.7 (9.3)
18							9.7	11.1	11.0 (10.8)
Average	3.4	3.6	4.2	5.1	5.9	7.2	9.0	11.1	
	(2.8)	(3.3)	(4.2)	(5.0)	(6.2)	(7.1)	(8.6)	(11.3)	

represent the education required for each of the occupation classes.

Since income after tax appears in our welfare function, Table 4.VI has been calculated from Table 4.V by the deduction of personal income taxes as prevailing in 1959.

As announced, the estimation of coefficients c_0 and c_2 will be based on the comparison of vertical variations in Table 4.VI, since these refer to groups with the same parameter h'. We exclude from the eight cases involved the lower element in column $h' = 12$, since the $h = 12$ and $h = 14$ professions not only differ in h value, but also in the degree of independence needed for the two groups of professions; this leaves us with seven cases. We followed a similar procedure for the Netherlands in Section 4.4.

For each pair of observations we assume that their welfare figures are equal since free mobility between the two is possible; hence,

$$x_1 - c_0 h_1 - \tfrac{1}{2}c_2(h_1 - h')^2 = x_2 - c_0 h_2 - \tfrac{1}{2}c_2(h_2 - h')^2, \qquad (4.5)$$

Table 4.VI
Incomes after personal income tax for the cells shown in Tables 4.IV and 4.V.

h	h'							
	0	3	6	8	10	12	14	18
0	3.1	3.3						
3		3.4	3.7					
6			4.2	4.6				
8				4.6	5.1			
10					5.7	6.2		
12						6.7		
14						7.2	7.8	
18							8.4	9.5

where suffix 1 refers to the lower and suffix 2 to the higher h in the pair.

Equation (4.5) reduces to

$$x_1 - x_2 = c_0(h_1 - h_2) + \tfrac{1}{2}c_2\{h_1^2 - h_2^2 - 2h'(h_1 - h_2)\}. \tag{4.6}$$

Replacing the left-hand side by x, the expression in parentheses on the right-hand side by h and the expression in brackets by z, (4.6) reduces to

$$x = c_0 h + \tfrac{1}{2}c_2 z. \tag{4.7}$$

For six pairs of observations contained in Table 4.VI this equation can be tested, assuming that no constant term appears on the right-hand side. Writing $\bar{x}$ for the average and x' for $x - \bar{x}$, and similarly for h and z, the assumption can be written

$$\bar{x} - \gamma_0\bar{h} - \tfrac{1}{2}\gamma_2\bar{z} = 0, \tag{4.8}$$

where γ_0 and γ_2 are the estimates of c_0 and c_2. Assumption (4.8) links γ_0 and γ_2 and reduces our regression to a single regression because

$$\gamma_2 = -0.33 + 2.29\gamma_0. \tag{4.9}$$

The dependent variable is $X = x' + 0.164z'$ and the independent one is $H = h' + 1.14z'$, for the first regression. For it we find

$$X = 0.144H \qquad (R = 0.98), \tag{4.10}$$

and for the second regression,

$$X = 0.149H.$$

Transformed back into the original variables we find, for the two regressions,

$$x = 0.144h + 0.00z,$$

respectively,

$$x = 0.149h + 0.007z, \tag{4.11}$$

meaning that in Frisch's [26] bunch-map-analysis way $c_0 = 0.146 \pm 0.003$ and $c_2 = 0.07 \pm 0.07$.

The utility function for Illinois thus becomes

$$\omega = \ln\{x - 0.146h - 0.03(h - h')^2\}. \tag{4.12}$$

For a comparison with the Dutch utility function (4.4) we have to transform the latter, where x is expressed in guilders (the buying power of which was roughly that of \$ 0.5) and schooling s or v in units of 3 years. Expressed in \$, h and h', the Dutch utility function then becomes

$$\omega_{NL} = \ln 2 + \ln\{x - 0.075h - 0.018(h - h')^2\}. \tag{4.13}$$

This means that the disutility of work requiring one more year of education is half as large for the average Dutchman as for the man from Illinois and the disutility of the 'tension' one-third as large.

4.6. Test results for some other American states

In a future publication some alternative tests of the measurement of utility functions for seven American states, including Illinois, will be discussed in detail. Some provisional results may be given here, in order to widen the background of our first attempts, given in this chapter. As a consequence of the results obtained with the material for six additional states, the method described in Section 4.5 has been supplemented with a second method, to be called the *linear*, as distinct from the *quadratic tension method* shown in that section. The words quadratic and linear refer to the 'tension term', specified as $\frac{1}{2}c_2(h - h')^2$ in Section 4.5 and alternatively as $\frac{1}{2}c'_2|h - h'|$, where the tension between education required, h, and actual education completed, h', is measured by the absolute value of their difference $z' = |h - h'|$ instead of by the squared difference. One reason why this second method was applied was that the results show a high degree of sensitivity of the main coefficient c_0 to this choice. Another reason was that the quadratic tension method sometimes yielded c_2 values with the wrong sign. This was also the case with the linear tension method, however. A further complication of the linear method is the identity of Δh [cf. equation (4.7)] and $\Delta z'$, preventing the identification of c_0 and c'_2. This complication was removed by assuming the value of $\frac{1}{2}c'_2$ to be equal, for all states, to the value found for Wisconsin, which is the smaller of the two values with the correct sign obtained. Since this value is small and two of the states yielded $c'_2 < 0$, a third theory, the '*simple linear theory*', was also applied, where $c'_2 = 0$. In all cases the restriction (4.8) had to be obeyed.

The main findings have been summarized in Table 4.VII. Some of these findings are:

(i) Correlations between X and H appearing in equation (4.10) are very satisfactory for three more states, but much less so for New York and Michigan. For New York the heterogeneity of the labour force may be an explanation. Because of these lower correlation coefficients R it is of some importance to distinguish between diagonal and first regression coefficients; theoretically the latter are more relevant, since we can take it for granted that our equations omit some significant additional independent variables.

(ii) The values obtained for c_0 applying the linear or no

Table 4.VII

Values of $c_0, \frac{1}{2}c_2$ and R found for seven American states using three methods: (i) quadratic tension, (ii) linear tension, (iii) no tension ('simply theory').

State	Value of c_0 using						Value of $\frac{1}{2}c_2$ using	
	(i)			(ii)		(iii)	(i)	(ii)
	Diag. regr.	1st regr.	R	c_0	R			
Cal	0.283	0.244	0.86	0.120	0.74	0.113	0.052	*
Ill	0.146	0.144	0.98	0.148	0.39	0.144	0.006	0.019
NY	0.350	0.250	0.71	0.124	0.43	0.116	0.083	*
Mich	0.227	0.171	0.75	0.111	0.71	0.103	0.038	*
SoCa	0.092	0.091	1.00	0.096	0.58	0.091	−0.004	−0.015
Tex	0.108	0.107	0.98	0.116	0.40	0.113	−0.005	−0.026
Wis	0.108	0.106	0.98	0.098	0.98	0.094	0.008	0.0077

Note: Asterisk means cases where $\Delta h = \Delta z'$ for all observations and c_2 is indeterminate.

tension version are considerably more uniform and closer to the values found for the Netherlands in equation (4.13).

(iii) The correlation coefficients found for the linear tension method are much lower. A closer look at the figures shows that some of the higher R for the quadratic theory are due, however, to the common term in z' of X and H, which becomes preponderant for some states. Anyway the R's are correlation coefficients between income differences and differences in h and much lower than the correlation between incomes and h.

4.7. Appendix on estimation of figures of Table 4.I

These figures have been estimated mainly with the aid of (i) the Netherlands Census of Population of 1960 which gives information on the school achievements of a number of two-digit social groups; (ii) tax statistics for 1964 giving incomes before and after tax for a smaller number (21) of socio-economic groups, and (iii) income distribution statistics for 1966 where labour income and income from other sources are shown. The exact names of the three sources are: (i) 13e Algemene Volkstelling, Deel 8 (13th General Census, Part 8) p. 62, 72 ff.; (ii) Nota over de Inkomensverdeling, Bijlage 15, Miljoenennota 1970 (Memorandum on Income Distribution, Annex 15, Budget Proposals 1970) p. 15; (iii) Inkomensverdeling 1966 en Vermogensverdeling 1967 (Income Distribution 1966 and Distribution of Wealth 1967) pp. 80–83.

Sources (i) and (iii) are publications of the Netherlands Central Bureau of Statistics; source (ii) is a parliamentary document. All can be obtained from the Staatsuitgeverij (Government Stationery Office), The Hague.

Since source (ii) refers to a smaller number of social groups than source (i), the latter's groups have been combined so as to cover the larger groups of (ii); even these have not been kept separate, however, but have been combined if they showed almost equal incomes and almost equal social status (especially comparable degrees of independence). As already observed in the main text, v is the actual amount of schooling completed,

Table 4.VIII
Comparison of US median and Dutch upper quartile schooling.

US groups		NL groups		US median (18–64 years old)	NL upper quartile
No.	Brief description	No.	Brief description		
A	Professionals, technicians, etc.	0	Professionals, specialists	16.3	15
B/I	Farmers and farm workers	4	Agrarian occupations, fishing	8.3	9
C	Managers, officers, proprietors	1	Managers, high officers	12.5	12
D	Clerical and kindred workers	2	Administrative occupations	12.4	9
E	Sales workers	3	Commercial occupations	12.5	9
F/G	Craftsmen, foremen, operators	7/8	Handicraft and industry	10.2	9
H	Service workers	6/9	Services	10.0	9
J	Labourers excl. farmers, miners	5	Mining, peat digging	8.7	6

as measured in periods of 3 years, and s the corresponding upper quartile. For approximately ten one-digit social groups, s has been compared with the median of school years (in units of 1 year) completed by the American active population between 18 and 64 years (US Department of Commerce, Current Population Reports, Series P 20, No. 207, p. 27; and US Census of Population 1960, Educational Attainment, Final Report PC (2) – 5B, p. 136). There was no systematic deviation, apart from a few years average difference. The comparison is crude anyway because the Dutch figures are expressed in multiples of three years only; the number of years of schooling is not given in the publications, but only the frequency distribution over five main school types. The reader should again be reminded of the illustrative character of the entire econometric material offered in this chapter. The figures for the large social groups are given in Table 4.VIII.