IMPLICIT SOCIAL PREFERENCES IN THE NORWEGIAN SYSTEM OF INDIRECT TAXATION

Vidar CHRISTIANSEN*
Institute of Economics, University of Oslo, Oslo, Norway

Eilev S. JANSEN
Institute of Economics, University of Oslo, and Central Bureau of Statistics, Oslo, Norway

Received August 1977, revised version received April 1978

The revelation of implicit social preferences is a fresh field of econometrics. In this paper the theoretical setting is a model of optimal indirect taxation. A parametric preference function is specified, which makes it possible to separate and quantify three different effects. First, it provides a condensed quantitative measure of the degree of income inequality aversion. Second, a set of parameters evaluate external social costs induced by the consumption of certain commodities. Finally, the function allows estimation of implicit equivalent income scales. The authors consider the results as a source of information about an important part of Norwegian tax policy.

1. Introduction

A marginal income unit for a four-person household with an income of 25,000 (N.kr.) is worth three times as much in welfare terms as a marginal income unit given to a three-person household with an income of 60,000 (N.kr.). This indicates the sort of conclusions we may draw from the present analysis in which we make an empirical assessment of the implicit social preferences in the Norwegian system of indirect taxation. Under certain assumptions we find that the social welfare function integrates to be close to a log-linear function of income. A simultaneous evaluation of external social costs due to the consumption of certain commodities reveals plausible results. Finally, we have made an attempt to estimate implicit equivalent

*A more complete version of the current paper can be found in Christiansen and Jansen (1977) [see footnote 10 for a correction]. An earlier version was presented at the Marstrand Meeting of Nordic economists in June 1977. We are indebted to Erik Biørn and Erik Garaas for their kind assistance in providing us with necessary data for this project. We are also grateful to Leif Johansen, who made helpful comments on our draft manuscript, and Gunnar Bramness, whose interest in our work has been very stimulating. We would also like to thank the editor and two anonymous referees for valuable suggestions.
income scales which leads to a surprisingly high consumer unit number assigned to the second adult in a two-person household.

But the basic ideas behind our work go beyond this. We believe that such results may be used as a starting point for integrating distribution concerns into cost–benefit analysis. Furthermore, comparisons with subsequent analyses in related spheres of economic policy may reveal inconsistencies which imply that the decisions may be improved. Implications of this kind are interesting aspects of the present work.

Analyses of this type are very scarcely represented in the literature. Some works have been published on implicit macroeconomic preferences. For a representative example see Nijkamp and Sommermeyer (1977). Analyses concerned with microeconomic subjects like income distribution topics are even more rare. Weisbrod (1968) provides an example concerned with income distribution weights, but his method is quite different from the one adopted in this work. Some attempts have also been made to derive implicit preferences from income tax schedules. For references and critical discussion see Christiansen (1977) and Stern (1977). But as far as we know, no analysis has been reported which is close in scope and method to the one presented in this paper.

The organization of this study follows the principle of going from more general considerations on social preferences implicit in economic policy (section 2) to the more specific theme of this paper. In section 3 we consider income distribution preferences implicit in public policy. The general model of optimal indirect taxation is presented in section 4. Some practical problems which arise when we apply the model to Norwegian tax policy are discussed in section 5, leaving a brief survey of the data for the appendix. The econometric approach is presented in section 6, while the empirical findings are interpreted in section 7. A brief evaluation of our efforts is given in section 8.

2. Social preferences implicit in economic policy

The implicit preferences in economic policy are defined as the preferences which make the actual policy optimal given these preferences. This definition does not claim that implicit preferences have to be identical to the true preferences of a decision-maker. They may, however, coincide, and this is of course a particularly interesting case. To infer implicit preferences theoretical optimality conditions for economic policy are combined with observations of the actual policy and other relevant data.

There are several reasons why it may be interesting to explicate implicit social preferences.

By revealing implicit preferences in terms of a number of strategic parameters (like e.g. an implicit measure of inequality aversion) we may get
more information about how strongly the policy actually promotes certain objectives than we are able to infer from direct observation of the policy. This may be a useful contribution to policy assessment, and may give rise to more explicit discussion about the preferences underlying economic-political decisions.

A possible outcome is that the implicit preferences which emerge from the analysis are inconsistent with the aims and opinions expressed in political statements and different from what the politicians are willing to accept as their true preferences. Such a result may indicate that the actual policy is in fact non-optimal and thus provide an incentive for revising the policy.

If the implicit preferences estimated on the basis of some kind of economic policy are taken to represent the true preferences of the decision-makers, these preferences may be used as input in other analyses to find the optimal policy in other fields. The underlying idea is that certain preference considerations are more duly allowed for in some kinds of decisions than in others. For instance the distribution preferences implicit in tax and transfer policy may be used to evaluate the distributive effects of public projects, which up to now have received little attention.

Since there are several categories of public decisions, we can in principle assign one set of implicit values of the preference parameters to each of these categories. By comparing implicit preference parameters from different categories of decisions we may detect inconsistencies which imply lack of efficiency of the policy as a whole. It may well be that revelation of such inconsistencies could contribute significantly to improvements of public policy.

Let us now turn to some of the more fundamental problems in implicit preference analysis. Since political preferences are moulded through a process where several considerations are simultaneously effective in a very entangled fashion, it may be rather difficult to interpret implicit preferences. At least conceptually we may assume that the political body in power has basic preferences which can be thought of as the guidelines which would be followed if the decision-makers obtained dictatorial power. But since the government in a democratic society is likely to be influenced by political and tactical considerations, it seems very difficult, not to say impossible, to identify the basic preferences. It seems more appropriate to define a concept of modified preferences as the preferences with respect to economic conditions which prevail after political considerations have been influential. We shall assume that the political decision-making can be simulated as if it is determined by modified preferences, while political considerations are influential only indirectly by affecting the preferences. Political statements may well be most appropriately interpreted as expressions of modified preferences since politicians must take account of political reactions before they formulate their statements.
It is often argued that the actual decisions will converge to a policy which has majority backing. This need not be a problem in our approach: It may simply be taken to mean that the modified preferences behind the actual policy have to be such as are accepted by a majority. More serious problems arise if the uses of certain policy instruments as such are subject to political constraints. It may be that pressure groups do not explicitly reject certain political objectives (like e.g. stronger income redistribution) as such, but oppose the use of specific policy measures which are believed to promote objectives they disapprove of. This may impose political feasibility constraints on the policy design itself, which are not easily allowed for. In the present study, however, we do not see any strong reason to worry about this problem.

3. The income distribution preferences implicit in public policy

By income distribution preferences are meant preferences among different income distributions or different income redistributions per se. The economic feasibility or the costs of income redistribution are not taken into account when defining the preferences as such.

A convenient way to express income distribution preferences is by means of a schedule of distributive weights given to marginal changes of the incomes of individuals at different income levels. In order to adopt this approach we shall assume that the social welfare weights attached to marginal income changes at different income levels may be expressed as a function of income:

\[ w(r), \]

where \( r \) denotes level of initial income. It will be termed the marginal welfare function.

To make its role in evaluating the income distribution more clear, suppose that there are \( I \) individuals earning incomes \( r^1, \ldots, r^I \). \( w^i \) is the distributive weight or the marginal welfare of income assigned to individual \( i \), \( w^i/w^j \) is the distributive weight given to individual \( i \) relative to the one given to individual \( j \). If \( w^i/w^j = 1 \), the income distribution between individual \( i \) and individual \( j \) may be said to be equitable since nothing is gained by a marginal income transfer from one individual to the other. If \( w^i/w^j > 1 \), there is a distributive inequality which implies that an extra unit of income for individual \( i \) is judged to be more valuable from a welfare point of view than an extra income unit for individual \( j \).

The marginal welfare function is usually taken to be derived from an additively separable welfare function where individual incomes enter in a symmetric way:
\[ W = \sum_{i=1}^{J} \Omega(r^i), \]  
\[ w(r) = \frac{\partial \Omega(r)}{\partial r}. \]  

It is important to notice exactly what the essential content of the additivity assumption is. Since any increasing monotonic transformation of (2) is also an admissible welfare function, the point is that the relative distributive weight \( w_i/w_j \) depends exclusively on \( r^i \) and \( r^j \) for any \( i, j \). For conditions for additive separability see Gorman (1968). The additivity assumption is often criticized [see e.g. Sen (1973)]. Let us briefly consider one case related to the usual criticism. Suppose there are three persons, A, B and C. A’s income is greater than B’s income. Imagine that C’s income increases from that of B to that of A. It is often claimed that distributive weights should then allow for the fact that B’s relative position is clearly made worse. This may be said to be done even by an additive function in the sense that the distributive weight given to B relative to the average weight given to other people increases. We think that this expresses the main consideration, and find it less essential that the distributive weight given to B relative to that given to A alone remains unchanged. In general our judgment is that for most purposes the essential features of income distribution preferences can be captured by an additive welfare function.

Since distributive weights appear in optimality conditions for income distribution policy, (1) is a convenient analytical tool in revealed income distribution preference analysis. Our approach is to reveal the implicit income distribution preferences by estimating the marginal welfare function implicit in actual distribution policy.

To make estimation possible the marginal welfare function must be given a specific parametric form, which will imply a certain kind of income distribution preferences. This calls for careful consideration of the a priori information and insight which may be relevant to take account of in the choice of functional form. In revealed preference analysis the most appropriate sources of information seems to be political discussions, political programmes and other relevant political announcements and publications. Such sources may provide a rough picture of the preferences in question which enables the analyst to confine his interest to a certain class of preferences which excludes a large number of functional forms.

Johansen (1974) contains a list of different approaches to establish preference functions. Among the approaches which are suggested are imaginary interviews, inference from planning documents and revealed preference methods. Imaginary interviews and inference from planning documents imply making use of information gathered by attending political deliberations and
reading relevant documents. The approach we have suggested above indicates how information of this kind may be used to establish the necessary basis for a numerical revealed preference analysis. This leads to a synthesis of the three methods cited from Johansen.

It seems to be a common political judgment that the same relative income changes for all individuals imply distributive neutrality. For instance in analyses of distributive effects of price changes, policy changes etc. it is usually concluded that the economic inequality is reduced (increased) if lower incomes increase relatively more (less) than higher incomes while proportional changes in all incomes are considered as neutral. The same notion of distributive neutrality is put forward in Johansen (1965, pp. 321-22) in connection with distributive effects of taxation.

Using the analytical concepts we have introduced, we shall define a vector of income changes as being neutral from a distribution point of view if it leaves all the relative distributive weights \(w'/w\) unchanged. The marginal welfare functions which reflect this apparently widely accepted definition must be of the form

\[ w(r) = D r^d, \]

where \(D\) and \(d\) are constant parameters. \(D\) is a positive but otherwise arbitrary constant. The value of \(D\) implies a conventional choice of 'units of welfare'. If a marginal transfer of income from a person with higher to a person with lower income is considered to enhance welfare, \(d\) must be negative.

We have decided to let the marginal welfare function specified in (4) represent income distribution preferences in our empirical studies.

Given (4) we find by simple integration that \(\Omega(r)\) of the additive welfare function (2) is given by

\[ \Omega(r) = \begin{cases} D^* r^{d^*} & \text{if } d \neq -1, \\ D \ln r & \text{if } d = -1, \end{cases} \]

where \(D^* = D/(d+1)\) and \(d^* = d + 1\). A possible additive constant term is here omitted.

4. Optimal tax design: The theoretical model

The indirect taxation, including subsidies as negative taxes, has several purposes. The fiscal purpose is obviously an essential one. Income re-

---

1Among other things this type of function has been applied by Atkinson in defining a possible measure of economic inequality [see Atkinson (1970)].
distribution is recognized as an important purpose. Moreover, it is a purpose to change the consumption of certain commodities generating external effects.

In this context we are primarily concerned with income distribution preferences. We should, however, get a distorted picture of the redistributive purpose of indirect taxation if we did not allow for the evaluations of external effects. The fiscal aspect is allowed for by the requirement that the government proceeds from indirect taxation amount to a given total, \( T \). Thus we shall confine our interest to the optimal choice of indirect tax rates subject to this constraint.

Let us assume that there are \( H \) households, \( i = 1, \ldots, H \), who are the consuming units. Suppose there are \( m \) goods, \( j = 1, \ldots, m \), which are subject to indirect taxation. The quantities of these goods are denoted by \( x^j_i \), which is the amount of good \( j \) consumed by household \( i \). \( X^j = \sum x^j_i \) is the total amount of good \( j \). All other goods are amalgamated into one mixed good, the quantity of which is denoted by \( z \). Thus \( z' \) is the amount of the mixed good consumed by household \( i \). The price of the mixed good is set equal to 1. The prices of the \( x \)-goods excluding excises and subsidies are denoted by \( q_1, \ldots, q_m \). The tax per unit of \( X_j \) is \( s_j, j = 1, \ldots, m \). We assume that there is complete shifting of taxes into consumer prices so that \( q_1, \ldots, q_m \) are taken to be independent of \( s_1, \ldots, s_m \). This assumption is discussed in section 5 below.

Each household is assumed to be endowed with a utility function \( U^i(z^i, x^i_1, \ldots, x^i_m) \) of the standard type. The preferences of the government are assumed to be given by a social welfare function containing \( U^1, \ldots, U^H \) as arguments. To take account of possible external effects we simply introduce the total amounts of the \( x \)-goods in the welfare function in addition to the individual utility levels. This is more convenient than including the externalities in all the utility functions, but the interpretation does not need to be different. One possible interpretation is that the complete individual utility function is additively separable into \( U^i(\cdot) \) as given above plus \( \sum_j U^i_j(X^j_i) \) which indicates how household \( i \) is affected by externalities generated by the total consumption of each good. The individual contributions to these totals may be ignored by each individual. An alternative or additional interpretation is that there are paternalistic considerations associated with certain commodities. Thus we arrive at the social welfare function

\[
W(U^1, \ldots, U^i, \ldots, U^H, X_1, \ldots, X_m),
\]

where \( U^i = U^i(z^i, x^i_1, \ldots, x^i_m) \), \( i = 1, \ldots, H \). We assume that \( (6) \) is of the Pareto type, i.e. it is an increasing function of \( U^1, \ldots, U^H \).

\[
w_j = \frac{\partial W}{\partial X_j} = \begin{cases} > 0 & \text{if } X_j \text{ generates social benefits.} \\ < 0 & \text{if } X_j \text{ generates social costs.} \\ = 0 & \text{otherwise.} \end{cases}
\]
The consumption expenditure of each household, \( i \), is assumed to be exogenously given equal to \( r^i \). The behaviour of each household, \( i \), is assumed to be as if it maximizes the utility function \( U'(z^i, x^i_1, \ldots, x^i_m) \) subject to the budget constraint \( r^i = z^i + \sum_j (q_j + s_j)x^i_j \), where \( (q_j + s_j) \), \( j = 1, \ldots, m \), are the consumer prices which are taken to be exogenously given. This leads to ordinary household and aggregate demand functions for each commodity which inserted into (6) give us the indirect welfare function. For given household incomes and producer prices this is a function of commodity taxes alone.

The social optimization implies maximizing the social welfare function with respect to \( s_1, \ldots, s_m \) subject to the fiscal requirement \( \sum_j s_j X_j = T \). We form the Lagrangian:

\[
L = W(U^1, \ldots, U^m, X_1, \ldots, X_m) - \kappa \left( \sum_j s_j X_j - T \right).
\]

\( \kappa \) is the Lagrange multiplier associated with the fiscal constraint. We introduce the simplified notations \( p_k = q_k + s_k \), \( \forall k \), and \( w^i = \partial W/\partial U^i \cdot \partial U^i/\partial r^i \), \( \forall i \). \( w^i \) is the marginal welfare of income assigned to household \( i \). The first order conditions for the social optimization problem can then be written as:

\[
-\sum_i w^i x^i_k + \sum_j w_j \left( \frac{\partial X^i_j}{\partial p_k} \right) - \kappa T_k = 0, \quad k = 1, \ldots, m, \tag{9}
\]

where \( T_k = \sum_j s_j (\partial X^i_j/\partial p_k) + X_k \).

Let us now turn to the economic interpretation of (9). Each equation of (9) requires that three kinds of marginal effects offset each other at optimum. The first term in (9) is the welfare loss incurred by the consumers because of the extra burden which is imposed by increasing \( s_k \). The second term measures the net external benefits due to the consumption adjustments which are induced by changing \( s_k \). The last term expresses the welfare effect acquired by lowering other taxes to offset the revenue effect of increasing \( s_k \).

5. Application of the model to Norwegian tax policy

Personal taxation in Norway is based partly on direct taxes on income and wealth and partly on indirect taxes. Indirect taxation consists of a general value added tax and excises on certain consumption categories. Two elements in the system of personal taxation are especially important for income redistribution. Those are the progressive income tax and excises

\footnote{For similar optimality rules in the vast literature on optimal taxation, see e.g. Atkinson and Stiglitz (1976) or Sandmo (1976).}
combined with subsidies, which may be regarded as negative excises. In the present study our interest is in the latter element. The theoretical model presented in the preceding section will be applied to derive implicit social preferences embodied in the actual design of excises and subsidies in Norway. Throughout this paper indirect taxes or commodity taxes are used as synonyms of excises on consumption goods. The analysis will be based on data from 1975. Our approach is to observe the situation in which all tax rates including excise rates are fixed and the economy is assumed to be adjusted to these taxes.

The remark may be made that even though there are other taxes, like income tax combined with children allowances, the optimality conditions for excises and subsidies are still valid as partial optimality conditions for these tax instruments. If simultaneous optimization of all tax instruments takes place, the suboptimization of indirect taxes is a necessary part of the total solution. The crucial point is that the existence of other kinds of tax policy does not rule out the use of indirect taxation. Atkinson and Stiglitz (1976) have shown that under certain conditions, including that the utility function be weakly separable between labour and all consumption goods, no commodity taxation needs to be employed if an optimal income tax is possible. This is a very interesting theoretical result, but there are other reasons for believing that non-zero indirect taxation may still be part of an optimal tax system in Norway and other countries. Sole reliance on income tax may give too strong incentives to tax evasion. Furthermore, it may be desirable for political reasons to use a variety of tax bases, and there may be political consideration behind the revenue requirement associated with each tax base.

In our model possible effects on work effort do not appear. This can be accounted for either by assuming that such effects are negligible, or by assuming that the welfare effect of changes in work effort due to a marginal tax increase is the same no matter by means of which excise rate the extra tax is being raised. The latter assumption implies that the welfare effect of (dis)incentives to work effort is given when \( T \) is fixed, and is consequently not affected by the structure of indirect taxation which is being optimized within the model. This assumption seems to be more relevant the more incentive effects are dominated by income effects.

Effects on work incentives of changing excise rates are usually assumed to be small, and rarely used as arguments in tax debates. Institutional obstacles to free adjustment of work effort may be one reason, at least in the short run. But in fact there is very little knowledge about the effects of excises on work effort, and theory does not provide much guidance for guesswork.\(^3\) To

\(^3\)Suppose for example that a typical leisure good becomes more expensive. The relative price of recreation will then increase, and the demand will normally decrease, tending to reduce the demand for leisure as well. On the other hand it may be economical to substitute leisure time for some of the purchased good in the "production" of recreation. So the demand for leisure may change in either direction because of substitution effects.
assume that effects on work incentives do not affect the optimal structure of indirect taxation is therefore rather arbitrary. Since, however, no alternative seems more convincing, there is hardly any reason to adopt a more complicated assumption.

Up to now we have treated $X_1, \ldots, X_m$ as quantities of homogeneous goods. In practice it is neither manageable nor desirable to specify all the basic goods which are subject to indirect taxation. Therefore we have aggregated the basic goods into 15 consumption categories. The categories are: (1) flour and grain for food, (2) bread, cake, etc., (3) meat and eggs, (4) canned food, (5) milk, cream etc., (6) cheese, (7) butter, (8) margarine, etc., (9) chocolate, (10) non-alcoholic beverages, (11) beer, (12) wine and liquor, (13) tobacco, (14) petrol and oil, and (15) cosmetic articles. A sixteenth category, fish, is given a special treatment. The fish subsidies in Norway seem to be more determined by fishery policy and administrative considerations than by general income distribution considerations. Therefore we have decided a priori to treat fish subsidies as exogenously given.

Choosing certain representative commodities within each group, a Laspeyres price index may be constructed for each group. If the indirect taxes on all goods in a certain group were removed the corresponding price index would change. We let the relative change in the price index define the excise rate on the index good as a share of consumer price [see Christiansen and Jansen (1977, pp. 22-26)]. Indices with this property are calculated by the Central Bureau of Statistics of Norway [see CBS (1976, p. 143)]. The weights of the indices are based on the Survey of Consumer Expenditure 1973 [CBS (1975)]. In our analysis all indices are normalized to unity in 1975.

The definition of the excise rates above reflects our assumption of complete shifting of commodity taxes into consumer prices. This is a standard assumption in all Norwegian analyses of distribution effects of indirect taxes [see e.g. Biørn (1975)]. It is well known that this assumption is compatible with a completely elastic supply of the goods in question. This is the case for typical imports like good categories (1) and (12)-(15). The industries which deliver good categories (2) and (8)-(11) are mainly sheltered industries and we may assert that the long run marginal cost curves in these industries are horizontal. See Johansen (1965, pp. 331-337) for a general examination of these arguments.

The same arguments are not applicable to the agricultural products (3)-(7).

---

4It follows that no optimality condition is introduced for fish subsidies. However, fish subsidies appear in the total tax revenue constraint (8) and will therefore appear in the optimality conditions for the other categories [being included in the $\Sigma p_i(\partial X_j/\partial p_i)$-terms of (9)].

5Norway imports 80 per cent of her flour and grain for food [good category (1)].

6I.e. their products are marketed at home under conditions that leave them free of foreign competition [see Aukrust (1977, p. 110)].
They are all heavily subsidized (see table 6), and the domestic markets are almost completely sheltered from foreign competition. However, the error we make by assuming complete shifting will be quite small if the demand for these goods is sufficiently inelastic. Since rather inelastic demand is a typical characteristic of these goods, we therefore consider it as admissible to stick to the complete shifting assumption. This complies with the view that the farmers might have been ensured approximately the same income level by higher market prices as an alternative to subsidies—an alternative which is politically rejected because of its distribution effects among the consumers.

We may notice that the complete shifting assumption is of no significance for the actual consumption data we shall apply in (9) since they are based on actual consumer prices. What we use in our application is that there is complete shifting at the margin. This implies that marginal tax changes do not generate income transfers between farmers and consumers which might have welfare effects.

As far as social costs are concerned, political information has led us to believe that external social costs may be attributed to four of the consumption categories appearing in the analysis. These categories are beer, wine and liquor, tobacco, and petrol and oil. No other consumption categories will be held responsible for social costs.

In order to compare income levels for households of different sizes we define the equivalent one person income, \( r_e(n,r) \), as the income which is considered to make a single person just as well off as a household with \( n \) members and income \( r \). The usual method is to characterize the size of each household by its total number of consumer units, \( N(n) \). This is used to compute the equivalent income as the actual income per consumer unit, i.e. \( r_e(n,r) = r/N(n) \). We apply the term equivalent income scales to \( N(n) \).

The equivalent income scales we have used are found in Bojer (1977), where such scales are computed under the assumption that equal budget percentages for foodstuff imply equivalent incomes. Bojer’s estimated equivalent scales for the year 1973 are given in table 2 of section 7.2.

A parametric specification of the equivalent income function \( r_e(n,r) \) which yields Bojer’s results for an appropriate set of parameters, is

\[
 r_e(n, r) = r[1 - a(n - 1)^c],
\]

where \( a \) and \( c \) are parameters. The inverse of the expression in square brackets is interpreted as the equivalent income scales \( N(n) \). For \( a = 0.42 \) and

\[ Bojer’s estimates of equivalent income scales are based on the Survey of Consumer Expenditure in Norway in 1967 and 1973 [CBS (1969) and CBS (1975)]. Notably, the budget percentage of food in different households is subject to particular attention when indirect taxation, including food subsidies, is concerned. \]
The types of data required in the following, are recognized by observing the optimality rules for excises set forth in the theoretical model [(9) of section 4]. In the appendix we survey the actual data we have used.

6. On the method of estimation

We have used a nonlinear least squares method to derive estimators for the parameters in the social preference function. Each optimality condition gives us a functional relationship in these parameters, (9), and the form of this function is the same for all goods.

In order to make (9) operational from an econometric point of view, we have inserted the parametrizations previously introduced. This is done by applying the marginal welfare function (4) to equivalent income $r$, given by eq. (10), making $w$ in (9) a parametric function of $r$ and $n$ when all prices are given. Then (9) takes on the following form:

$$-\sum_r \sum_n (r - ar(n-1))(\frac{d}{\partial n})^r X_k^n + \sum_j w_j (\frac{\partial X_j}{\partial p_k}) = 0,$$

$$k = 1, \ldots, m.$$  \hspace{1cm} (11)

where $X_k^n$ is total consumption of good $k$ by households with income $r$ and size $n$ (see the appendix).

The total net proceeds from indirect taxation on each good are of course widely different. In order to avoid problems with differences in scale, we have divided through (11) by $\kappa \cdot T_k$. Random disturbance terms, $u_k(k = 1, \ldots, m)$ are introduced in the model to allow for random, non-systematic errors in the attempts to adjust the taxes to satisfy the optimality conditions. The $u_k$'s may of course also include errors of measurement in the data.

$$f_k(a, c, d, \kappa, w_1, w_2, w_1, w_1) = -\sum_r \sum_n (r - ar(n-1))(\frac{d}{\partial n})^r X_k^n/\kappa T_k$$

$$+ \sum_j w_j (\frac{\partial X_j}{\partial p_k})/\kappa T_k - 1$$

$$= u_k.$$  \hspace{1cm} (12)

We have no prior information about the random terms in (12). There is in particular no adding-up constraint in the sense that the sum of the error terms should equal zero. Bearing in mind that we have eliminated differences in scale in (12) it is at least an admissible choice to assume that the
V. Christiansen and E.S. Jansen, Social preferences 229

The variance–covariance matrix is diagonal with the same variance for all error terms. That is, we assume: $E(u_k) = 0$, $\text{var}(u_k) = \sigma^2$ and $\text{covar}(u_k, u_h) = 0$, $k = 1, \ldots, m$, and $h = 1, \ldots, m$ ($h \neq k$).

We consider each optimality condition as one realization of this function in the unknown parameters. Thus, (12) provides $m$ sample points for the subsequent estimation.

Our procedure has been to minimize the total sum of squares (SS) with respect to the parameters for the sample at hand.

$$SS = \sum_k [f_k(a, c, d, \kappa, w_{11}, w_{12}, w_{13}, w_{14})]^2.$$  \hspace{1cm} (13)

Following Wolberg (1967, p. 60) one can obtain unbiased estimates for the variances of the (nonlinear) least-squares estimators, $\hat{x}_h$, as

$$\hat{\sigma}_x^2 \approx \left( \frac{SS}{m - p} \right) H_{hh}^{-1},$$  \hspace{1cm} (14)

where $m$ is the number of equations in (12), $p$ is the number of unknown parameters and $H_{hh}^{-1}$ is the $h$th diagonal element of the inverse of the Hessian of the total sum of squares function. The formula (14) is more accurate the closer the estimating equations are to linearity. As is clear from (12) these are nonlinear in the parameters, but they are linear in $X^T_c \beta$ and $(\partial X_j/\partial \beta_k)/\beta$. Approximate confidence intervals [of level $100 \cdot (1 - \beta)$] for the parameters are given by

$$\hat{x}_h - \hat{\sigma}_x \cdot t(\beta/2, m - p) < x_h < \hat{x}_h + \hat{\sigma}_x \cdot t(\beta/2, m - p),$$  \hspace{1cm} (15)

where $\hat{\sigma}_x$ denotes estimated values and $t(\beta/2, m - p)$ is the $100 \cdot \beta/2$ percentage point of the $t$-distribution of $(m - p)$ degrees of freedom.

We have based the estimation on a quasi-Newton method to minimize $SS$ in (13) as a function of $p$ independent unknown variables (i.e. the parameters). The method uses difference approximations to both the derivatives

---

8This is a crucial assumption. Heteroscedastic error terms would mean that the model is unidentified with the data at hand. Our numerical results have, however, an alternative interpretation which applies even if the stochastic model is invalidated for this or some other reason. Consider the case where the optimality conditions (11) are fulfilled exactly. In principle we could then find the $p$ unknown parameters by solving any subset of $p$ equations chosen from the $m$ equations in (11) ($p < m$). For the reasons we have given above there will be some deviations from exact fulfillment of the optimality conditions. This means that we will get ($^2$) different solutions depending on which equations we include in the subset. Our numerical findings from minimizing the sum of squares function (13) may be interpreted as an average of these ($^2$) solutions. By this procedure we utilize the information which is contained in all the $m$ equations in (11).
and the Hessian of the function. Moreover, the computer program provides an approximation to $H_{hh}^{-1}$ needed in formula (14).

From a numerical point of view this procedure raises two major problems. Firstly, the optimization routine may find a local minimum, and it is impossible to eliminate the possible existence of a better minimum than the one we find. The best way to cope with this problem is to test whether different initial values of the parameters give the same minimum or not.

Secondly, to obtain convergence, it is necessary that we provide starting points which are reasonably close to the minimum values (or at least of the right order of magnitude). In order to overcome this difficulty we have chosen a stepwise estimating procedure. As we shall see each step may be interpreted as implying a specific hypothesis on political behavior. It is in this spirit we have used the term 'variant' for each step.

**Variant I:** We keep all parameters, except $d$ and $K$, fixed at constant values. The equivalent income parameters, $a$ and $c$, are fixed at $a = 0.42$ and $c = 0.33$, which imply equivalent income scales which coincide with those found in Bojer (1977), see section 5. Furthermore, we assume that no social costs are taken into consideration, i.e. $w_j = 0$, $\forall j$.

**Variant II:** As Variant I, except that we allow $w_j$ ($j = 11, \ldots, 14$) to vary.

**Variant III:** All eight parameters are free to vary during the minimization.

In this way we are gradually increasing the dimension of the minimization problem. With only two unknown parameters it is possible to make a grid-search for good starting points, (i.e. low function-values) over a wide range of values for the parameters. Combined with some a priori considerations of what may be plausible parameter-values, the minimizing parameter-values from each variant tell us where to concentrate the grid-search for reasonable initial values in the next variant. A concise record of this procedure is given in Christiansen and Jansen (1977, pp. 59–65).

**7. Empirical results**

In section 2 the implicit social preferences are defined as the preferences which make the actual policy optimal given these preferences. Basically it is

---

9 The actual minimizing algorithm used is the subroutine E04CDF of the NAG-library, Oxford University Computing Laboratory. A program description is given in the NAG Library Manual (1975), and Gill, Murray and Pitfield (1972). The computations were carried out at the Computer Center Blindern-Kjeller on a CDC 6600 computer.

10 The minor deviations between the results reported in this article and the corresponding figures in Christiansen and Jansen (1977) are due to a data error (in the excise rate of chocolate), which was discovered after the issuing of the research report. A complete rerun of the estimation was made, replicating the procedure described in the reference.
on this background we shall interpret the empirical content of the results from the estimation described in the preceding chapter. In Variant I we simulate the social optimization as if the only concern is with the distributional effects of the indirect taxation, while Variant II implies that allowance is also made for social costs induced by the consumption of some categories of goods. In Variant III the analysis is extended to the case where even the equivalent income parameters \( a \) and \( c \) are revealed through the optimality conditions for the system of indirect taxation.

The main results from the estimation are reported in table 1, summarizing the point estimates and the corresponding estimated standard deviations for the parameters in each of the three variants. The scaling of \( \kappa \) and the \( w_j \)'s in table 1 is rather arbitrary, but for the purpose of discussing the statistical properties of the estimates this is of minor importance.

In Variants I and II all parameters are significant at a 5 per cent level, except the social cost parameters for beer and wine & liquor \((w_{11} \text{ and } w_{12})\) of Variant II. In Variant III only the parameter \( d \) is significant at the same level. The differences observed for the estimated standard errors of the same parameters between Variant II and Variant III seem difficult to explain. However, the estimated standard errors may well be biased due to the fact that they are estimated on the basis of approximations at two different stages of the estimation.

We shall now turn to a more detailed discussion of the empirical findings for each category of parameters.

7.1. The elasticity of the marginal welfare of income

Our specification of the marginal welfare function \( w(r_e) = r_e^d \) implies that the quantitative properties of the income distribution preferences are embodied in the parameter \( d \). This makes \( d \) a particularly interesting parameter. As explained earlier \( d \) is the elasticity of the marginal welfare of income. In the special case where the welfare function is interpreted as a utilitarian additive welfare function \( d \) may be interpreted as a constant elasticity of the marginal utility of income.

When \( d \) is negative, income redistribution from households with higher incomes to those with lower incomes is considered as desirable in itself. The absolute value of \( d \) may be taken as a measure of the degree of income inequality aversion: The greater the absolute value of \( d \), the stronger the preferences in favour of equal distribution of income. As we have described above the Variants I to III imply that we reveal the implicit value of \( d \) in the Norwegian tax policy under three alternative sets of assumptions. In Variant I the estimated \( d \) is \(-1.71\), while Variant II reveals the value \(-0.87\).

We may conclude that the implicit inequality aversion is substantially lower when social cost considerations are supposed to matter than in the
Table 1
Summary of empirical results
(estimated standard errors are written in parentheses).

<table>
<thead>
<tr>
<th>Variant number*</th>
<th>Equivalent income parameters</th>
<th>Income elasticity of the marginal welfare of income</th>
<th>Shadow price of total tax revenue</th>
<th>Social cost parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha$</td>
<td>$\beta$</td>
<td>$\gamma$</td>
<td>$\kappa$</td>
</tr>
<tr>
<td>I</td>
<td>0.42</td>
<td>0.33</td>
<td>-1.706</td>
<td>-0.695 $\cdot 10^{-7}$</td>
</tr>
<tr>
<td>II</td>
<td>0.42</td>
<td>0.33</td>
<td>-0.867</td>
<td>-0.202 $\cdot 10^{-3}$</td>
</tr>
<tr>
<td>III</td>
<td>0.504</td>
<td>0.212</td>
<td>-0.879</td>
<td>-0.246 $\cdot 10^{-3}$</td>
</tr>
</tbody>
</table>

*The total sum of squares (SS) for each of these variants is:

- Variant I: $SS = 0.2911$
- Variant II: $SS = 0.0416$
- Variant III: $SS = 0.0373$
case where such considerations are excluded a priori. Since there is rather strong evidence in the political debate supporting the hypothesis that social cost considerations are relevant, we believe the assumptions of Variant II to be more plausible than those of Variant I.

Variant III differs from Variant II in that the equivalent income parameters are supposed to be unknown and are subject to estimation. Actually the corresponding estimate of \(d\) is \(-0.88\) which differs very little from the estimate in Variant II. The stability of the estimate of \(d\) when going from Variant II to Variant III adds to the evidence that a \(d\)-value near \(-0.9\) may be taken as a concentrated quantitative expression of the implicit income distribution evaluation in the Norwegian system of indirect taxation.

A \(d\)-value equal to \(-1.0\) implies that the individual welfare function integrates to yield the logarithm of income, according to (5) in section 3. It follows that the distributive weight in this case becomes the inverse of the equivalent income. With \(d = -0.9\) the corresponding social welfare function is quite close to a log-linear function. To illustrate the implications of this finding, a number of corresponding isowelfare curves (or social indifference curves) are depicted in figure 1 for the case of a two-person community. The incomes of the two persons are denoted by \(r_1\) and \(r_2\). The slope of the indifference curves has the usual interpretation as the marginal rate of substitution between income given to person 1 and income given to person 2. The isowelfare curves are characterized by constant elasticity of substitution equal to \(1/d = -1.1\).

In order to judge whether our estimated \(d\)-values should be considered as representing a strong or a moderate inequality aversion, it may be of interest to make a few comparisons. There are no precedent empirical estimates of the same kind in the literature, but a marginal welfare function of the type we are applying has been used in some studies of income taxation and for assessing income distribution effects in cost-benefit analyses. A study of British income taxation reported in Stern (1977) suggests that a \(d\)-value about \(-2\) seems to give tax rates not too dissimilar from those ruling in the U.K. Examples from CBA are given by McGuire and Garn (1969) and Nwaneri in his analysis of the equity aspects of the well-known cost-benefit analysis of the Third London Airport [see Nwaneri (1970)]. McGuire and Garn use a parameter corresponding to \(d\) which assumes a value of \(-2.5\), while Nwaneri investigates the implications of the two alternative values \(-2\) and \(-2.5\). Our findings indicate that a lower absolute value of \(d\) would be appropriate in similar cost-benefit assessments in Norway. On this background it also seems adequate to characterize an absolute \(d\)-value below 1.0 as expressing a moderate inequality aversion.

7.2. The equivalent income parameters

The equivalent income parameters, \(a\) and \(c\), are fixed in Variant I and II.
In Variant III implicit values of the equivalent income parameters are estimated. The estimates are $\hat{a} = 0.59$ and $\hat{c} = 0.21$, which deviate considerably from the parameter values $\bar{a} = 0.42$ and $\bar{c} = 0.33$ derived by fitting our function to Bojer's equivalent income scales (see section 5).

To recognize the implications of these discrepancies we need to clarify the meaning of $a$ and $c$. We see from (10) that the equivalent income scale for two adults, $n = 2$, is $N(2) = (1 - a)^{-1}$. If $a$ is less than 0.5, $N(2)$ is less than 2 which implies that a change of household size from one to two persons is accompanied by economies of scale. The deviation of $a$ from 0.5 may be taken as a measure of these economies of scale. Furthermore we see that for $n \geq 2$ the equivalent income loss, $a(n-1)^{\gamma}r$, increases with household size.
when \( c > 0 \), but at a rate which is more moderate the lower the value of \( c \).\(^{11}\) Thus \( c \) will reflect the relative needs of children and possible economies of scale when \( n \) becomes greater than two.

It follows that \( \tilde{a} = 0.42 \) clearly implies economies of scale of going from a one person to a two person household, whereas the implicit \( \tilde{a} = 0.59 \) implies considerable diseconomies of scale. The implications of different parameter values are exposed by contrasting the corresponding equivalent income scales \( N(n|\tilde{a}, \tilde{c}) \) and \( N(n|\hat{a}, \hat{c}) \) as done in table 2.

| Household size \( n \)         | Bojer (1977) \( N(n|\tilde{a}, \tilde{c})^* \) | \( N(n|\hat{a}, \hat{c})^* \) |
|-------------------------------|-----------------------------------------------|---------------------------------|
| 1 (single adult)              | 1                                             | 1                               |
| 2 (two adults)                | 1.73                                          | 1.72                            |
| 3 (two adults, one child)     | 2.14                                          | 2.12                            |
| 4 (two adults, two children)  | 2.55                                          | 2.52                            |
| 5 (two adults, three children)| 2.96                                          | 2.97                            |
| 6 (two adults, four children) | 3.37                                          | 3.50                            |

\( N(n|\tilde{a}, \tilde{c})^* = [1 - a(n - 1)]^{-1} \), \( \tilde{a} = 0.42, \tilde{c} = 0.33, \hat{a} = 0.59, \hat{c} = 0.21 \). Confer formula (10).

The most conspicuous feature of the implicit equivalent income scales is the high value for \( n = 2 \). The value of 2.44 in fact implies that, according to the implicit judgment being revealed, a two-person household needs 2.44 times the income of a single adult to be in an equivalent economic situation. This result is quite contrary to the established opinion on this matter.

For \( n \geq 3 \) households with \( n \) members need not be far from \( n \) times the income of a single person to be socially equalized according to the implicit parameters. The high values of these equivalent income scales are due to the high value of \( \tilde{a} \). \( \hat{c} = 0.21 \) is much lower than \( \tilde{c} = 0.33 \). This implies that in comparison with a second adult children contribute relatively less to reduce the household’s equivalent income according to the estimated implicit parameters than according to Bojer’s figures. It may, however, be interesting to notice that the costs of having children in a household measured by absolute loss of equivalent income is of the same order of magnitude in the two approaches. The cost of one child is about 11 per cent of actual income.

\(^{11}\) \( c \) is the elasticity of the equivalent income loss with respect to \((n - 1)\). When \( 0 < c < 1 \), the marginal effect of \( n \) on the equivalent income loss is positive, but decreasing in \( n \), which can be interpreted as economies of scale with respect to the number of children.
in Bojer's case and about 9 per cent in the implicit preference approach. For a second child the corresponding figures are 8 and 6 per cent.

Altogether, the most striking implication of our findings is the low relative weight given to one-person households as the income of other households are heavily deflated by equivalent income scales to yield considerably higher distributive weights. Compared to the alternatives in table 2 the implicit equivalent income scales are clearly disfavourable to single persons. Thus the value judgment which maximizes the optimality of the actual indirect taxation may be said to give little weight to the costs and benefits of single person households.

Having made these statements, it should be noted that the estimated standard errors of \( \hat{a} \) and \( \hat{c} \) indicate a high degree of uncertainty for these parameter-estimates. Indeed, if we make use of the approximate confidence interval formula (15), we find that both \( \hat{a} \) and \( \hat{c} \) lie well inside these intervals, which are \( a \in (-0.11, 1.29) \) and \( c \in (-0.49, 0.91) \).

7.3. The distributive weights

The estimates in Variants I–III of the distributive weights given to households of different size and with varying income, are presented in table 3. We may repeat that the distributive weight assigned to a household is the welfare weight or the social weight given to a marginal income unit for this household. In all three cases the marginal welfare of income for a married couple with no children and disposable income equal to 25,000 N.kr. is taken as the welfare unit, so that the distributive weight assigned to a household of this type is set equal to one. We may call this household the standard household. Our choice of standard of measurement implies that the distributive weight assigned to some household expresses the welfare weight given to a marginal income unit for this household relative to the weight given to a marginal income unit accruing to the standard household.

To give a few examples of how the table may be read, we may draw attention to the mid-section of table 3. We see that in this variant a marginal income unit given to a one-person household enjoying an income of 10,000 is equivalent in terms of welfare to 1.3802 additional income units to the standard household. And we see that according to the implicit welfare judgment a marginal income unit for a three-person household with an income of 25,000 is worth more than twice as much in welfare terms as a marginal income unit given to a three-person household with an income of 60,000. Moreover, we observe that the dispersion of the distributive weights is much greater in Variant I than in the Variants II and III. This is of course a direct implication of the different \( d \)-values in the two cases. The differences in distributive weights are greater in Variant III than in Variant II. This applies especially to the vertical variations and less to the horizontal
Table 3

Distributive weights for households of type \((r, n)\).*

<table>
<thead>
<tr>
<th>(n)</th>
<th>10,000</th>
<th>15,000</th>
<th>25,000</th>
<th>40,000</th>
<th>60,000</th>
<th>100,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.8849</td>
<td>0.9438</td>
<td>0.3948</td>
<td>0.1771</td>
<td>0.0887</td>
<td>0.0371</td>
</tr>
<tr>
<td>2</td>
<td>4.7741</td>
<td>2.3904</td>
<td>1.0000</td>
<td>0.4485</td>
<td>0.2246</td>
<td>0.0939</td>
</tr>
<tr>
<td>3</td>
<td>6.7837</td>
<td>3.3967</td>
<td>1.4209</td>
<td>0.6373</td>
<td>0.3191</td>
<td>0.1335</td>
</tr>
<tr>
<td>4</td>
<td>9.1360</td>
<td>4.5745</td>
<td>1.9137</td>
<td>0.8583</td>
<td>0.4298</td>
<td>0.1798</td>
</tr>
<tr>
<td>5</td>
<td>12.0937</td>
<td>6.0555</td>
<td>2.5332</td>
<td>1.1362</td>
<td>0.5689</td>
<td>0.2380</td>
</tr>
<tr>
<td>6</td>
<td>15.9823</td>
<td>8.0025</td>
<td>3.3477</td>
<td>1.5015</td>
<td>0.7518</td>
<td>0.3145</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(n)</th>
<th>1.3802</th>
<th>0.9710</th>
<th>0.6235</th>
<th>0.4148</th>
<th>0.2918</th>
<th>0.1874</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.2136</td>
<td>1.5574</td>
<td>1.0000</td>
<td>0.6653</td>
<td>0.4680</td>
<td>0.3005</td>
</tr>
<tr>
<td>2</td>
<td>2.6464</td>
<td>1.8619</td>
<td>1.1955</td>
<td>0.7953</td>
<td>0.5595</td>
<td>0.3593</td>
</tr>
<tr>
<td>3</td>
<td>3.0787</td>
<td>2.1660</td>
<td>1.3908</td>
<td>0.9253</td>
<td>0.6510</td>
<td>0.4180</td>
</tr>
<tr>
<td>4</td>
<td>3.5505</td>
<td>2.4979</td>
<td>1.6040</td>
<td>1.0670</td>
<td>0.7507</td>
<td>0.4820</td>
</tr>
<tr>
<td>5</td>
<td>4.0911</td>
<td>2.8783</td>
<td>1.8482</td>
<td>1.2295</td>
<td>0.8650</td>
<td>0.5554</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(n)</th>
<th>1.0138</th>
<th>0.7098</th>
<th>0.4530</th>
<th>0.2997</th>
<th>0.2098</th>
<th>0.1339</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.2379</td>
<td>1.5669</td>
<td>1.0000</td>
<td>0.6615</td>
<td>0.4632</td>
<td>0.2956</td>
</tr>
<tr>
<td>2</td>
<td>2.8192</td>
<td>1.9739</td>
<td>1.2598</td>
<td>0.8334</td>
<td>0.5835</td>
<td>0.3724</td>
</tr>
<tr>
<td>3</td>
<td>3.4199</td>
<td>2.3945</td>
<td>1.5282</td>
<td>1.0110</td>
<td>0.7078</td>
<td>0.4518</td>
</tr>
<tr>
<td>4</td>
<td>4.1053</td>
<td>2.8744</td>
<td>1.8345</td>
<td>1.2136</td>
<td>0.8497</td>
<td>0.5423</td>
</tr>
<tr>
<td>5</td>
<td>4.9359</td>
<td>3.4559</td>
<td>2.2056</td>
<td>1.4591</td>
<td>1.0216</td>
<td>0.6520</td>
</tr>
</tbody>
</table>

*The weights are calculated by the formula \(w = (r - \tilde{a}(n - 1)^2) / 25,000 - \tilde{a} \)

\(25,000 (2-1)^2\), i.e. all the weights are measured relative to the distributive weight given to a household with \(n = 2\) and \(r = 25,000\).

7.4. Estimates of the implicit valuation of social costs

We assumed at the outset that the consumption of four categories of goods, namely petrol, tobacco, beer and other alcoholic beverages, may generate social costs which are reflected in the excises on these commodities. The crucial point in the estimation of these social costs is that we have separated the marginal cost elements from the income distribution elements in the respective taxes.

The social costs are measured in terms of welfare. The units by which welfare is measured may be chosen quite freely by selecting an appropriate numeraire. One choice of measurement scale was implied by formula (11), the results of which are given in table 1. But the results may of course be variations in the tables, and it is mainly due to the differences in the equivalent income parameters discussed above.
transformed to a new scale. In the present context $-\tilde{\kappa}$ seems to be a convenient choice of welfare unit. The social costs are measured in public budget crowns, or to be more precise: a social cost of one unit is equivalent to the gross welfare loss inflicted on the taxpayers by collecting an extra tax crown. This implies that if the marginal consumption of some commodity has negative social effects, the social cost is measured by the greatest government expenditure financed by taxes, which it would be worth while granting to offset the social consequences.

Table 4

<table>
<thead>
<tr>
<th>Good category</th>
<th>Variant II</th>
<th>Variant III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Petrol and oil</td>
<td>0.30</td>
<td>0.31</td>
</tr>
<tr>
<td>Tobacco</td>
<td>0.42</td>
<td>0.42</td>
</tr>
<tr>
<td>Beer</td>
<td>0.09</td>
<td>0.06</td>
</tr>
<tr>
<td>Other alcoholic beverages</td>
<td>-0.01</td>
<td>-0.04</td>
</tr>
</tbody>
</table>

*For commodity $j$ $IMC(j)$ is given by the formula $IMC(j) = \tilde{w}_j/\tilde{\kappa}$.

Table 4 states the estimates of the implicit social cost due to the consumption of the four commodities in question, measured in terms of the chosen welfare units. The estimates from Variants II and III are only slightly different. We shall comment on Variant II only, bearing in mind that the main comments apply to Variant III as well.

The social cost associated with the consumption of petrol is directly related to the marginal cost of motor traffic, which is a major topic in applied transport economics. Marginal cost pricing would imply that the marginal costs of motor traffic which are borne by the government, e.g. the costs of road maintenance, are reflected in the petrol tax. The marginal cost element in the petrol tax is estimated at 0.30 per unit of expenditure on petrol and oil consumption in 1975. To find an estimate of the corresponding marginal social cost per vehicle-kilometer, which is the marginal cost concept most frequently applied, we need an estimate of the outlay on petrol and oil per vehicle-kilometer. This is estimated at 0.22 N.kr. The valuation in terms of budget crowns of the marginal social cost per vehicle-kilometer implied by the tax system is then estimated at $0.30 \times 0.22 = 0.07$.

---

12 This figure follows from the usual assumption that a passenger car uses 0.1 liter petrol per kilometer, which is used in traffic analyses made at the Institute of Transport Economics, Oslo, Norway. The average price of petrol in 1975 was 2.08 N.kr. per liter [source: The Norwegian Petroleum Institute, Oslo]. A rough estimate of the expenditure on oil per kilometer is obtained by taking 5 per cent of the petrol expenditure per kilometer (0.208 N.kr.).
This result may be compared to the estimates of actual marginal social costs per vehicle-kilometer worked out at the Institute of Transport Economics. An analysis reported in NOU 42 (1972, p. 43) based on 1972-prices yielded an estimate of 0.11 per vehicle-kilometer. On this background the implicit valuation in the tax system of the marginal cost of motor traffic as estimated in our analysis may be said to be rather low.

The implicit valuation of the social costs of smoking is higher per unit of expenditure than that of petrol and oil consumption. We may conclude that implicitly substantial social costs are attributed to the use of tobacco.

The social cost implicitly associated with beer drinking is rather low and is in fact not significantly different from zero.

In the case of other alcoholic beverages (wine and liquor) we find that implicitly the marginal cost is given a negative value. The reservation should be made at once that the estimate is not significantly different from zero. Nevertheless, from a priori expectations it is a surprising result. The relatively high tax rate on alcoholic beverages in Norway is commonly conceived of as motivated by the desire to reduce the amount of alcohol consumption because of its harmful social effects. At first glance our estimate of the implicit evaluation of the consumption of alcohol seems to be totally incompatible with this view. This may, however, be a too hasty conclusion.

In the first place, even if consumption of alcoholic beverages may cause harmful social effects which, together with redistribution purposes, motivates a special alcohol taxation, it is conceivable that the marginal social cost may be judged to be zero after taxes have been imposed, which allow for the redistributive aspect. However, this explanation does not seem to be quite satisfactory. It seems to be a widely held political opinion that a further reduction of alcohol consumption would reduce social costs.

Our result may then be subject to alternative interpretations. The first alternative is that we have revealed an inconsistency between what the politicians say to be their judgments and the preferences which actually govern their decisions. The second alternative is that the politicians fail to see the optimal policy implications of their preferences so that the tax policy is revealed to be non-optimal.

A third alternative may offer a very relevant explanation. Up to now we have neglected the existence of home production of alcoholic beverages. A further increase in the rate of alcohol taxation may well discourage purchases of alcoholic beverages while at the same time encourage the legal home production of wine and the illegal home production of liquor. Even smuggling may be a matter in this context. Therefore a marginal decrease in the official purchases of alcoholic beverages may be offset or even exceeded by an increase in the consumption of alcoholic beverages available from other sources. This explanation may well reconcile our finding with the official views on alcohol policy allowing for the facts of real life behaviour.
7.5. **The shadow price of the tax on each commodity**

A marginal tax increase imposes a welfare loss on the taxpayers. The evaluation of this loss depends on the distribution of the additional tax payments among households of the different types. The distributive features vary among the taxes on the different goods. When distributive weights are estimated, we can calculate the welfare losses generated by alternative tax changes. It is particularly interesting to consider tax changes which have the same effect on the total tax revenue. We define the shadow price of a commodity tax as the welfare loss due to a marginal increase in the tax which adds one unit to the total tax revenue. We let \( \kappa_k \) denote the shadow price of the excise on good \( k \), and \( \hat{\kappa}_k \) denote the estimate. The welfare scale is fixed by using \( \hat{\kappa} \) as the welfare unit. Then \( \hat{\kappa}_k \) is defined by

\[
\hat{\kappa}_k = \left( -\frac{\partial W}{\partial s_k} \right) / T_k(-\hat{\kappa}), \quad \forall k.
\] (16)

Observe that \( \hat{\kappa}_k \) is defined as a positive entity. If the optimization were perfect, \( \hat{\kappa}_k \) would be equal to 1 for all \( k \). But because of the random element in the determination of the excise rates there will be certain discrepancies. When one shadow price is lower than some other shadow price, it is possible to achieve a welfare gain by lowering the commodity tax with the higher shadow price and just offset the loss of tax revenue by enhancing the commodity tax with the lower shadow price. A table of estimated shadow prices provides information about which commodity taxes have low shadow prices and should be raised, and which commodity taxes have high shadow prices and should be lowered to yield a welfare gain. Our shadow price estimates for Variants I and II are presented in table 5. The estimates for Variant III are very close to those of Variant II.

We find that the estimated shadow prices of Variant II (and III) are much closer to unity than those of Variant I. This is a consequence of the fact that we have introduced social costs which give a more realistic model specification. For the good categories (12)–(14) the discrepancies of Variant I are captured by the social cost parameters.

8. **Concluding remarks**

This project has been motivated by the hope that it may be possible to bridge the gap between formal optimality theory and real life decision-making in the sphere of economic policy. Our efforts to pursue this task have brought us well outside the well-blazed paths of quantitative analysis. Reviewing what we have done, our lack of satisfactory data and all the assumptions being made to be able to complete this analysis, we have found it both comforting and stimulating to read what Leif Johansen has written in
Table 5  
The shadow price of the tax on each commodity group. Ranked from lowest to highest value.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Shadow price ( \tilde{r}_k )</th>
<th>Commodity group</th>
<th>Shadow price ( \tilde{r}_k )</th>
<th>Commodity group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.77</td>
<td>3 meat</td>
<td>0.86</td>
<td>3 meat</td>
</tr>
<tr>
<td>2</td>
<td>0.77</td>
<td>12 wine and liquor</td>
<td>0.94</td>
<td>4 canned meat and fish</td>
</tr>
<tr>
<td>3</td>
<td>0.86</td>
<td>10 non-alcoholic beverages</td>
<td>0.96</td>
<td>6 cheese</td>
</tr>
<tr>
<td>4</td>
<td>0.89</td>
<td>15 cosmetic articles</td>
<td>0.97</td>
<td>8 margarine etc.</td>
</tr>
<tr>
<td>5</td>
<td>0.89</td>
<td>4 canned meat and fish</td>
<td>0.97</td>
<td>1 flour for food</td>
</tr>
<tr>
<td>6</td>
<td>0.94</td>
<td>11 beer</td>
<td>1.00</td>
<td>14 petrol and oil</td>
</tr>
<tr>
<td>7</td>
<td>0.96</td>
<td>6 cheese</td>
<td>1.00</td>
<td>11 beer</td>
</tr>
<tr>
<td>8</td>
<td>0.98</td>
<td>8 margarine etc.</td>
<td>1.00</td>
<td>12 wine and liquor</td>
</tr>
<tr>
<td>9</td>
<td>1.00</td>
<td>1 flour for food</td>
<td>1.00</td>
<td>13 tobacco</td>
</tr>
<tr>
<td>10</td>
<td>1.01</td>
<td>9 chocolate</td>
<td>1.00</td>
<td>10 non-alcoholic beverages</td>
</tr>
<tr>
<td>11</td>
<td>1.05</td>
<td>2 bread etc.</td>
<td>1.03</td>
<td>2 bread etc.</td>
</tr>
<tr>
<td>12</td>
<td>1.08</td>
<td>5 milk etc.</td>
<td>1.03</td>
<td>5 milk etc.</td>
</tr>
<tr>
<td>13</td>
<td>1.10</td>
<td>13 tobacco</td>
<td>1.05</td>
<td>9 chocolate</td>
</tr>
<tr>
<td>14</td>
<td>1.10</td>
<td>7 butter</td>
<td>1.06</td>
<td>7 butter</td>
</tr>
<tr>
<td>15</td>
<td>1.32</td>
<td>14 petrol and oil</td>
<td>1.08</td>
<td>15 cosmetic articles</td>
</tr>
</tbody>
</table>

the Introduction to his book about the MSG-model [Johansen (1960)]. It says:

It has become a habit among economists who are quantifying their models to say something like the following: ‘The principal aim of the present study is methodological. The statistical data are very scanty and unreliable. The quantitative analysis should therefore be considered merely as an illustration of the method’. Having given this statement, the author would hardly be blamed for using the statistical data uncritically. However, in many cases statements of this sort are hardly intended to be accepted literally, and if they were, such an author would, in my opinion, lay himself open to a criticism no less serious than that of drawing conclusions too uncritically. In short, he might very well be guilty of not seeing and using all information which is contained in the data and in the quantitative results derived by his analysis. Taking into account on the one hand the scarcity of quantitative information about economic relationships, and on the other the need for such information, any lack of efficiency in the use of the sources or an understatement of the significance of the results may be no less blameworthy than too hasty conclusions.

In this spirit we would like to announce that we consider this work not only as an exercise in revealed preference methods, but as a source of information about an important part of Norwegian tax policy. Our estimate
of the parameter $d$ provides a condensed quantitative measure of the actual endeavour to promote income redistribution through the tax rates under survey. Even if the estimate may be very uncertain, the values sometimes suggested for $d$ in numerical examples of distributive weights are so arbitrary and widely dispersed that even an indication of the order of magnitude may prove valuable. At least we feel ourselves that we have gained an insight into the nature of this part of income distribution policy which is not acquired simply by observing excise rates and consumer data.

By introducing social cost parameters we have been able to separate the social cost effects from the income distribution effects. This is a new approach to the social marginal cost problem which may reveal interesting information about the properties of actual tax rates, which is relevant when alternative tax programs are considered.

Our results bring in their wake a large number of challenging questions. For instance, would a similar analysis of a related sphere yield results which are consistent with our findings? What would the optimal use of other income distribution instruments be like if based on the implicit preferences derived in the present analysis? How would public project selection be affected by adopting our distributive weights in cost–benefit analyses? We hope that these interesting questions will stimulate further research on economic policy along these lines.

Appendix: The data

The excise rates ($s_k$) are computed in the Central Bureau of Statistics in accordance with the description in section 5 [see CBS (1976, pp. 34–42)]. The excise rates as a share of consumer price for goods included in the analysis are given in table 6. The excise rate for fish is $-0.1144$.

The simultaneous distribution of households by size and income. We have decided to let households of six different compositions represent all households in Norway. We denote these households as the representative households. The six categories are the households with a single adult and those consisting of a married couple with no children, one child, two children, three children or four children below 16 years of age. The few households actually having more than four children are included in the last category.

The main reason for leaving out other types of households is that they are too difficult to handle. There is established no basis for comparing the economic situations of these strongly inhomogeneous households as it is for the representative households. Furthermore, there exists strong evidence for asserting that income distribution policy in Norway is dominantly influenced by how the representative households are affected [Confer Aukrust and Borgenvik (1969) and Djørn (1975)]. Thus we assume that the decision-
Table 6
The excise rates as a share of consumer price for the good categories included in
the analysis.

<table>
<thead>
<tr>
<th>Category number</th>
<th>Category</th>
<th>Excise rate as a share of consumer price</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Flour and grain for food</td>
<td>-0.8192</td>
</tr>
<tr>
<td>2</td>
<td>Bread, cake etc.</td>
<td>-0.1104</td>
</tr>
<tr>
<td>3</td>
<td>Meat and eggs</td>
<td>-0.2740</td>
</tr>
<tr>
<td>4</td>
<td>Canned food</td>
<td>-0.2024</td>
</tr>
<tr>
<td>5</td>
<td>Milk, cream etc.</td>
<td>-0.7590</td>
</tr>
<tr>
<td>6</td>
<td>Cheese</td>
<td>-0.4302</td>
</tr>
<tr>
<td>7</td>
<td>Butter</td>
<td>-0.2652</td>
</tr>
<tr>
<td>8</td>
<td>Margarine etc.</td>
<td>-0.2713</td>
</tr>
<tr>
<td>9</td>
<td>Chocolate</td>
<td>0.1760</td>
</tr>
<tr>
<td>10</td>
<td>Non-alcoholic beverages</td>
<td>0.1305</td>
</tr>
<tr>
<td>11</td>
<td>Beer</td>
<td>0.2955</td>
</tr>
<tr>
<td>12</td>
<td>Wine and liquor</td>
<td>0.7579</td>
</tr>
<tr>
<td>13</td>
<td>Tobacco</td>
<td>0.5414</td>
</tr>
<tr>
<td>14</td>
<td>Petrol and oil</td>
<td>0.4495</td>
</tr>
<tr>
<td>15</td>
<td>Cosmetic articles</td>
<td>0.2170</td>
</tr>
</tbody>
</table>

making takes place as if the whole population were made up of the representative households.

To obtain an operational concept of disposable income, $r$, we have decided to use total consumption expenditure as a proxy for disposable income.

The distribution of households by size and income which we have arrived at is based on the following assumptions:

(i) The distribution of households by composition was the same in 1975 as in 1973 and conforms to that of the Survey of Consumer Expenditure in 1973 [CBS (1975)].

(ii) The average total nominal consumption expenditure of each type of household increased by the same percentage from 1973 to 1975 (22 per cent).

The distribution of households by total consumption expenditure in the Survey of Consumer Expenditure is given by household numbers in expenditure intervals. We have worked out our distribution assuming that all households in an expenditure interval have the same expenditure which is set equal to the average of the household expenditures actually reported as belonging to the interval in question.

The total consumption of each good category ($X_k$) is collected from the ordinary national accounts of the Central Bureau of Statistics.

The consumption of various goods by each type of household, ($X_{it}'$). A recent study [CBS (1976)] by Mr. Erik Biørn and Mr. Erik Garaas of the Central Bureau of Statistics furnished us with a starting-point for estimating these entities. They provided the budget shares in 1975 for the relevant 15 good categories by each type of household ($x_{it}'$). These data are constructed
by a rather elaborate procedure from the Survey of Consumer Expenditure 1973. For a detailed record of this, see CBS (1976, p. 143).

We have proceeded to calculate the total consumption in 1975 of the different goods within each type of household. Let \( y_{r,n}^k \) denote such consumption figures corresponding to the number of representative households in the Survey of Consumer Expenditure 1973 by size and income, \( A(r,n) \). This leads up to

\[
y_{r,n}^k = a_{r,n}^k \cdot r \cdot A(r,n).
\]

(A1)

An estimate of \( X_{r,n}^k \), i.e. total consumption of good \( k \) by each type of household, should now be readily at hand but for two complications: We do not know the total number of households in Norway, and even if we knew this, the estimate of \( X_{r,n}^k \) at which we arrive would almost certainly not add up to equal the national accounts data for total consumption, \( X_k \), when summed over \( r \) and \( n \).

We have chosen to estimate \( X_{r,n}^k \) by

\[
X_{r,n}^k = y_{r,n}^k \cdot \frac{X_k}{\sum_r \sum_n y_{r,n}^k}.
\]

(A2)

This means in effect that we have maintained the relative consumption profiles for each good by household size and income as derived from the Survey of Consumer Expenditure 1973, but we have adjusted the levels of all \( X_{r,n}^k \) to be consistent with the national accounts data.

The Cournot derivations of the various goods are computed by means of the price elasticities used in MODIS\(^{13}\) and the national accounts figures for total consumption of the various goods. The price elasticities are essentially estimated on the basis of Frisch's complete scheme [Frisch (1959)].

\(^{13}\)The official annual planning model in Norway. For a description see for instance Bjerkholt and Longva (1975).

References


Christiansen, V., 1977, The theoretical basis for deriving distributive weights to be used in cost-benefit analysis, Memorandum from the Institute of Economics, University of Oslo, 25 April.

Christiansen, V. and E.S. Jansen, 1977, Implicit social preferences in the Norwegian system of indirect taxation, Memorandum from the Institute of Economics, University of Oslo, 13 August.

Frisch, R., 1959, A complete scheme for computing all direct and cross demand elasticities in a model with many sectors, Econometrica 27, 177-196.


