Discussion Papers No. 294, February 2001
Statistics Norway, Research Department

Knut R. Wangen and Erik Biørn

Individual Heterogeneity and Price Responses in Tobacco Consumption: A Two-Commodity Analysis of Unbalanced Panel Data

Abstract:
The paper presents a panel data analysis of tobacco demand. The purpose is threefold: (i) to measure income, own-price, and cross-price responses for two tobacco commodities: cigarettes and handrolling tobacco, (ii) to investigate sociodemographic effects, and (iii) to utilize the potential offered by panel data to investigate unobserved heterogeneity. The latter is crucial for commodities where consumers have different tastes and where users tend to become addicted. Several linear and non-linear specifications of a two-equation system are estimated, using an unbalanced panel data set of Norwegian households over the period 1975 - 1994 and a modified Maximum Likelihood procedure. Differences between the results based on the sub-sample of smokers and on the pooled sample of smokers and non-smokers are interpreted. For both commodities we find a high degree of unobserved heterogeneity and several robust and "heory consistent" patterns with respect to price responses and demographic effects.

Keywords: Tobacco. Panel data. Addiction. Cross-price response. Heterogeneity.

JEL classification: C33, C34, D12, I18

Acknowledgement: Earlier versions of the paper were presented at the Ninth International Conference on Panel Data, Geneva, 2000, and at seminars at the University of Oslo and Statistics Norway. We thank Jørgen Aasness, Jo Thori Lind, Kjersti-Gro Lindquist, Terje Skjerpen, and seminar participants for helpful comments. The research of Knut R. Wangen has been partly financed by the Norwegian Research Council (project no.111097/330).

Address: Knut R. Wangen, Statistics Norway, Research Department, P.O. Box 8131 Dep, N - 0033 Oslo, Norway. e-mail: knut.reidar.wangen@ssb.no

Erik Biørn, University of Oslo, Department of Economics, P.O. Box 1095 Blindern, N - 0317 Oslo, Norway. e-mail: erik.biorn@econ.uio.no

Discussion Papers comprise research papers intended for international journals or books. As a preprint a Discussion Paper can be longer and more elaborate than a standard journal article by including intermediate calculation and background material etc.

Abstracts with downloadable PDF files of
Discussion Papers are available on the Internet: http://www.ssb.no

For printed Discussion Papers contact:

Statistics Norway
Sales- and subscription service
N-2225 Kongsvinger

Telephone: +47 62 88 55 00
Telefax: +47 62 88 55 95
E-mail: Salg-abonnement@ssb.no

1 Introduction

Tobacco is commonly considered a commodity which is subject to addiction, like alcohol, coffein, and drugs. This addiction, reflecting, *inter alia*, past experience with tobacco smoking, may be strong. Its damage on the health status of many of its users is well documented by medical research, and policy interventions to curtail consumption are adopted in many countries. Such interventions include general warnings, restrictions on advertisement and sale, smoking-free areas, and excise taxes. In Norway, the excise taxes on tobacco are far higher than in most other countries[1], and for several years, advertisement of tobacco goods has been prohibited by law. Yet the average tobacco consumption is higher than in the other Scandinavian countries. Of considerable interest when discussing policy measures towards tobacco abuse may be estimates of the consumers' price and income responses for tobacco, not only direct price effects, but also cross-price effects for different tobacco commodities.

Psychological 'stocks of habits' – combined with genetic dispositions and attitudes towards health risks – are therefore potentially important factors when explaining observed tobacco consumption econometrically. These are additional factors to standard observable economic factors like income, prices, sociodemographic variables, etc. In a dynamic model of individual behaviour, addiction may be represented by a time-dependent variable incorporating the 'stock of habits' determined by each individual's past consumption [cf., *e.g.*, Lluch (1974) and Becker and Murphy (1988)]. Within a static model, habit effects can be considered as individual 'properties', represented as (components in) individual specific, *i.e.*, time invariant, latent variables. The latter framework may be the most convenient when data in the form of short panels from a large set of individuals are available. This is the case in the present paper.

There has been a growing interest in econometric analyses of tobacco consumption in recent years. The theoretically appealing aspects of habit formation in tobacco consumption and the need to evaluate potential policy measures have together spawned a large literature using several different approaches. Chaloupka and Warner (1999) offer an elaborate overview of issues in and contributions to the economic literature on tobacco. Most studies treat tobacco as one homogeneous commodity, although Chaloupka and Warner (1999, p. 16) mention a few studies of the substitution between cigarettes and smokeless tobacco and four studies on substitution between manufactured and handrolled cigarettes. The latter studies are of most interest for the present paper, as they do not give unambiguous evidence for price-induced substitution between manufactured and

[1] In December 1994, the average retail prices for 20 cigarettes and 50 grams of handrolling tobacco were USD 6.30 and USD 8.26, respectively. (The average exchange rate in 1994 was 7.05 NOK/USD.)

hand-rolled cigarettes. Our results indicate that such substitution does occur. This issue is also addressed in Wangen and Aasness (2001). Amongst the papers which disregard the composition of tobacco consumption, we will briefly mention three. Jones (1989) investigates the discrete-continuous choices of participation and consumption within a static 'double-hurdle model', using cross-sectional household expenditure data from the UK in 1984. This type of model provides useful information on the distribution of consumers' choices at a given point in time, although the addictive nature of nicotine suggests that tobacco consumption appears in a dynamic context. Chaloupka (1991) uses a model within the rational addiction tradition, with Becker and Murphy (1988) as a standard reference, and applies this dynamic model on micro data. In a recent study, Labeaga (1999) combines the discrete/continuous aspects of the double-hurdle model with the rational addiction framework, using unbalanced household panel data from Spain over the period 1977 – 1983.

The purpose of this paper is threefold. The first is to estimate, from unbalanced household panel data, within a *static two-equation model* framework, income and price responses (including cross-price responses) for two tobacco commodities, cigarettes and other smoking tobacco (mainly handrolling and pipe tobacco). The second purpose is to investigate the effects of sociodemographic variables like age, cohort, gender, and geographic location, on the tobacco consumption and its composition. These variables account for *observed heterogeneity.* Third, we want to explore *unobserved heterogeneity.* This can, to a large extent, be expected to be due to addiction and is commonly assumed to show larger variability for tobacco commodities than for most other consumption commodities. However, estimated unobserved heterogeneity can also represent the effect of valid, unobserved explanatory variables. None of these issues have, to our knowledge, been investigated previously in a framework with two tobacco commodities.

Our data set consists of a large set of unbalanced panel data (more than 26 000 observations from more than 18 000 households) from annual Norwegian household budget surveys for a twenty year period (1975 – 1994). The panel is rotating; some respondents are observed twice, at a one year interval, and some are observed only once. Because of the long sample period, the relative price variation is substantial along the time dimension. We are therefore able to obtain meaningful estimates of price responses. Our data set also makes it possible to estimate the covariance matrix of the latent household specific component of the two tobacco commodities along with the covariance matrix of the genuine disturbance vector, which will shed light on the addiction issue.

Why are genuine panel data essential for this kind of investigation? Unobserved individual effects, whether they are treated as random or fixed, cannot be identified from cross-section data alone. In the random effects situation, when only one observation of each respondent is available, such effects cannot be separated from the genuine disturbances, and hence the relative variation of the latent individual effects and the disturbances cannot be identified.

An interesting question, which has not been given much attention in the panel data literature, is whether and to what extent the variability of the estimated latent individual effect is sensitive to the list of observed regressors. One hypothesis may be that the estimated variance of the latent effect, in an absolute or relative sense, declines when additional explanatory variables are included in the model and the overall fit is improved. Another hypothesis may be that this variance is insensitive to the choice of regressors. Absence of correlation between the latent effect and the specified regressors may be a more critical assumption with some choices of regressors than with others. These issues are of particular interest in the case of tobacco demand, since, as stated above, we can conject that a large part of the variability of the individual effect for this commodity is due to variations in the 'stock of habits', genetic dispositions, and attitudes towards health risks. Characteristically, in a previous analysis of rotating panel data for the years 1975 – 1977, covering an exhaustive set of 28 consumption commodity groups and using a common functional form, Biørn and Jansen (1982, section 7.5) found that tobacco was the commodity for which individual heterogeneity represented the largest part of the estimated total disturbance (more than 70 per cent).

A main result of the present paper is that the variances of the latent individual effects tend to decrease when more variables (including square and interaction terms) are included in the model. Relative to the gross disturbance variances, the estimates of these variances are, however, high and fairly constant across model versions, about 60 – 70 per cent.

The qualitative pattern of the price elasticity estimates is robust over model variants, but the size of the estimates differ according to whether the data set includes all households or only smokers. An interpretation of these differences, as resulting from censoring, will be given. We find that our splitting of tobacco into two commodities has given value added as compared with treating it as one commodity. First, the income elasticity estimates differ substantially. Cigarettes tend to being a luxury, while we find signs that handrolling tobacco may be an inferior good. Second, the estimated price responses differ, but for both tobacco commodities, the estimated own price elasticities are negative and the cross-price elasticities are positive. Third, for both commodities we find a negative coefficient for the number of children, while for the number of persons in the other age groups all coefficients are positive. The coefficient values are generally higher for handrolling tobacco than for cigarettes, which indicates that substitution is important. Fourth, we find pronounced differences between the gender and the regional effects for the two commodities.

The rest of the paper is organized as follows. In Section 2, we dicuss some modeling problems, for instance the participation issue. Section 3 elaborates the econometric specification and the estimation procedure, which is a modified Maximum Likelihood (ML) procedure. The data set is described in Section 4. In Section 5, we present the empirical results, with focus on Engel and Cournot elasticities, age, cohort, and other demographic effects, as well as unobserved heterogeneity. Section 6 concludes.

2 Modeling problems

We use a static two-equation model framework with latent individual heterogeneity and with the consumption of cigarettes and smoking tobacco as endogenous variables. The unit of analysis is the household. We do not embed the two equations into a complete system of consumer demand, *e.g.*, derived from standard static consumer demand theory, with all adding-up and symmetry restrictions, etc. taken into account. Hence, the demand function for the 'third', remainder commodity is not part of our formalized model. The primary reason for this is that tobacco goods take a small part of the budget of an average household, in Norway only about 1.5 per cent (and zero for a substantial part of the households). This is one reason why our two equations should be interpreted as approximations to the tobacco demand equations in an underlying full demand model. Other reasons are given below. A complete system explaining along with tobacco foods, beverages, services, etc. would seem overdimensioned and overparametrized in relation to our basic focus. Neither is substitution between tobacco and other addictive commodities, for instance alcohols, in focus. On the other hand, price induced substitution between the two tobacco commodities is allowed for. Its magnitude may have important policy implications. The various sociodemographic background variables are intended to represent the effect of shifts in preferences.

The latent habit component of tobacco consumption, discussed in the introduction, may be specified econometrically as additive to the systematic part of the demand equation and included in individual specific (fixed or random) effects. Regression analysis using the fixed effects approach, unlike the standard random effects approach, is robust to potential correlation between the latent effects and the specified regressors [cf. Hsiao (1986, section 3.4)]. This may be important in our context, since, *e.g.*, income and demographic variables may be correlated with the latent heterogeneity.

Formally, the model is a two-equation regression system for unbalanced panel data with unobserved individual heterogeneity. To our knowledge, this is the first analysis using this kind of model for disaggregate tobacco commodities. A precise model description will be given in Section 3. Systems of regression equations and estimation methods for linear balanced panel data models with (random) error components are discussed in Avery (1977), Baltagi (1980), and Krishnakumar (1996). In Biørn (1999), this kind of model is adopted to data sets with unbalanced panels.

A problem that has to be addressed is the fact that many households in the data set (about 50 per cent) report *zero purchase* of both tobacco commodities, and some households purchase only one of them (see Table 2). From a modelling point of view, the situation is complicated since we, in general, do not know the reason for zero reporting. It may be due to (i) preference characteristics; the household consists of obstinate non-smokers only, (ii) there is at least one potential smoker in the household, but the actual price-income constellation motivates a corner solution, (iii) infrequency of purchases due to transaction costs [see Deaton and Irish (1984) and Keen (1986)], or (iv) misreporting.

Labeaga, Preston and Sanchis-Llopis (1998, p. 4 and table 1) report proportions of zeros in the Spanish Family Survey, where tobacco expenditure is recorded in weekly periods in up to eight successive quarters. Their results indicate that infrequency is not the major component of the proportions of zeros. In our data, the recording period is two weeks, which should reduce the importance of infrequency further. Since smoking to some extent is not 'socially acceptable', misreporting may be more important for tobacco than for other goods, either because of deliberate erroneous reporting or because of self-deception. However, studies have shown that measurement of smoking by self-report or by biochemical markers (blood, urine, hair etc.) gives approximately the same estimates of prevalence, cf., *e.g.*, USDHHS (1989, p. 265). Hence, the two most important reasons for zero expenditure seem to be (i) and (ii).

The data give evidence against assuming that all households maximize the same instantaneous utility function. In the data section, we argue that all households face the same set of prices on the two tobacco commodities. If the consumers were homogenous, the standard theory then predicts that, with the same vector of other covariates, they would choose the same consumption bundle (when neglecting noise). From this we can conclude that if the zero observations are mainly due to (i) and (ii), the preferences must vary in a non-trivial manner over the population of consumers. It is desirable to impose some structure on the preference variation, and at least two strategies have been suggested. Muellbauer (1988) assumes that the utility of current consumption is influenced by past consumption, and that rational consumers will account for this in their long-term consumption plan. Lluch (1974) suggests a model with the same basic ideas as Muellbauer, but formulates it differently. In Lluch's model, the instantaneous utility is a (time-invariant) function of the current consumption and a vector of consumption capital variables. Each consumption capital variable is a function of past consumption, and has properties similar to physical capital stock variables (*e.g.*, as used in production theory). Becker and Murphy (1988) has become a standard reference in the literature dealing with addictive goods. Their model is quite similar to Lluch's, but in addition they allow the instantaneous utility function to depend on stocks of personal and social capital. We, in contrast, represent preference variations by means of latent time invariant variables within a static model.

The high frequency of zero expenditures requires a clarified strategy for treating them. Basically there are three possibilities: (a) deleting them, (b) model them, or (c) treat them as any other observations. We have chosen both alternative (c) and a moderate version of alternative (a). Below we will outline a framework for interpreting the relationship between all three alternatives. The purpose of the study and the reasons for zero expenditures are fundamental when choosing the modelling strategy. Choosing alternative (c) can be justified by an argument given in Deaton (1990, p. 282): "The revenue effects of a tax change depend on how total demand is altered and not on whether changes take place at the extensive or intensive margins." If the object of interest is the average tax paid by different types of households, or one wants to quantify their typical behaviour for welfare analysis, alternative (c) provides the information needed and will usually be the most easily obtainable. Alternative (a) is also a practical solution, but depending on the reason for zero expenditures it may result in severely biased samples. In the present application, deleting all households with zero tobacco expenditures is very unattractive since this would leave us with less than 16 per cent of the total sample, cf. Table 2. Instead, we delete only households with zero expenditures on both types of tobacco, resulting in a sample without non-smokers. Alternative (b) is not pursued, as multi-equation discrete-continuous choice models for unbalanced panel data with unobserved heterogeneity are too difficult to implement econometrically at the present stage, not least because of the lack of available computer software. We will, however, outline a two-equation Tobit-type model with two thresholds[2] which will clarify some of the difficulties, but also, and more importantly, will make it possible to interpret differences in

[2]Generalized Tobit models are discussed in Heckman (1976), Amemiya (1984), Deaton and Irish (1984), Amemiya (1985, chapter 10), and Blundell and Meghir (1987), although not in a panel data context. As far as we can see, however, even disregarding the panel aspect, none of the models considered in these papers are similar to the two-equation model based on a truncated binormal distribution that we consider here.

estimation results from model strategies (a) and (c).

Assume, for simplicity, that the household's maximizing behaviour results in a *desired*, or *latent*, *expenditure* on cigarettes and handrolling tobacco,

$$
\begin{aligned}
y_C^* &= x\beta_C - \sigma_C u_C, \\
y_H^* &= x\beta_H - \sigma_H u_H,
\end{aligned}
\tag{1}
$$

respectively, where x is the (common) covariate vector, β_C and β_H are coefficient vectors, σ_C and σ_H are positive constants (to be interpreted as disturbance standard deviations) and (u_C, u_H) are standardized disturbances which are assumed to be independent of x and binormally distributed with zero means, unit variances, and coefficient of correlation ρ.[3] The latent expenditures may be positive, zero, or negative. The observed expenditure of a commodity is assumed to be equal to its latent expenditure when the latter is positive. When the latent expenditure is negative, the consumer does not want to use the commodity, and so the consumption is zero. The observed expenditures on the two commodities are then

$$
\begin{aligned}
y_C &= \max[y_C^*, 0], \\
y_H &= \max[y_H^*, 0].
\end{aligned}
\tag{2}
$$

This implies that y_C and y_H, conditionally on x, jointly follow a truncated binormal distribution.

Let $\phi(\cdot)$ and $\Phi(\cdot)$ denote the marginal density function and the cumulative distribution function, respectively, of the standardized univariate normal distribution, and let $\Psi(\cdot,\cdot)$ be the cumulative distribution function of the standardized binormal distribution with coefficient of correlation ρ. Let furthermore $\Phi_g = \Phi(x\beta_g/\sigma_g)$, $\phi_g = \phi(x\beta_g/\sigma_g)$ $(g = C, H)$, and $\Psi_{CH} = \Psi(x\beta_C/\sigma_C, x\beta_H/\sigma_H)$. Here, $\Phi_C = P(y_C > 0)$ and $\Phi_H = P(y_H > 0)$ are the marginal smoking probabilities for cigarettes and handrolling tobacco, respectively, and Ψ_{CH} in the probability of using both commodities. The probability of being a smoker, *i.e.*, of consuming *at least one* of the two commodities is then

$$
\begin{aligned}
\Phi_S &= P(y_C + y_H > 0) \\
&= P(y_C > 0, y_H = 0) + P(y_H > 0, y_C = 0) + P(y_C > 0, y_H > 0) \\
&= (\Phi_C - \Psi_{CH}) + (\Phi_H - \Psi_{CH}) + \Psi_{CH} = \Phi_C + \Phi_H - \Psi_{CH}.
\end{aligned}
$$

In Appendix we show that

$$
\begin{aligned}
\mathsf{E}(y_C|x) &= x\beta_C\Phi_C + \sigma_C\phi_C, \\
\mathsf{E}(y_H|x) &= x\beta_H\Phi_H + \sigma_H\phi_H,
\end{aligned}
\tag{3}
$$

[3] We here, for simplicity, neglect the panel property of the data.

and that

$$
\begin{aligned}
\mathsf{E}(y_C|x, y_C + y_H > 0) &= x\beta_C \frac{\Phi_C}{\Phi_S} + \sigma_C \frac{\phi_C}{\Phi_S}, \\
\mathsf{E}(y_H|x, y_C + y_H > 0) &= x\beta_H \frac{\Phi_H}{\Phi_S} + \sigma_H \frac{\phi_H}{\Phi_S}.
\end{aligned} \tag{4}
$$

In the next section, we present our framework for estimating approximations to $\mathsf{E}(y_C|x)$, $\mathsf{E}(y_H|x)$, $\mathsf{E}(y_C|x, y_C + y_H > 0)$, and $\mathsf{E}(y_H|x, y_C + y_H > 0)$. The estimation of the last two conditional expectations is based on a subsample of smokers only, while the estimation of the first two is based on the full sample. All approximations are specified as continuous in the exogenous variables, treating zeros and positive expenditures alike. Equations (3) and (4) suggest that the approximations should be interpreted as projections of an underlying discrete-continuous model [confer Olsen (1980), Goldberger (1981), and Greene (1981) for discussions of the single equation Tobit case, the latter two assuming full normality]. They could also be regarded as projections of more complex discrete-continuous models. The two-equation Tobit model above is quite restrictive since it assumes the same parametric structure for both smokers and non-smokers. Cragg (1971) suggests a more flexible model for the single equation case, allowing exogenous variables to have different effect at the intensive and the extensive margins. For tobacco this is reasonable. As an example, high income households may have smaller probability of being smokers than low income households, but if they do smoke they are likely to consume more. Cragg's model have been generalized into double-hurdle models, which consists of two parts, a Probit part and a Tobit part. Blundell and Meghir (1987) use a double-hurdle model to account for infrequency of purchase. They interpret the Tobit part as giving the actual consumption, allowing a corner solution, while the Probit part accounts for the difference between purchase and actual consumption in a two week expenditure survey. In contrast, Jones (1989) interprets both hurdles as a result of actual consumption decisions. The first decision, represented by the Probit part, is whether to be a smoker or not. The second desision is how much to smoke, given that the outcome of the first choice is to be a smoker. As this second decision is represented by a Tobit it allows for corner solutions. Jones also suggests that the start-decision and the stop-decision should be treated differently, leading to a trivariate model. The two interpretations of the double hurdles models suggest that an even more complicated model could be applied – combining Jones' double-hurdle with an additional hurdle to account for infrequency.

It is a potential problem that maximum likelihood estimators for dicrete-continuous models are sensitive to misspecification, even with respect to assumptions about the distribution of the error terms, see Godfrey (1988, Chapter 6). More robust estimators have been suggested, for instance the Least Absolute Deviation estimator of Powell (1984), but they are generally harder to compute.

As mentioned, the continuous projections of discrete-continuous models with latent heterogeneity are our main interest in this study. Since they are also quite easy to handle numerically they would in any case be valuable to have at hand before embarking on discrete-continuous modeles. In Wangen and Biørn (2001), we give a further discrete choice analysis of the smoking probabilities for the two commodities, within the framework of (binomial and multinomial) logit models.

3 Econometric framework and estimation procedure

This section elaborates the model specification and the estimation procedure for our unbalanced panel data set. Our method is, to a large extent, based on Biørn (1999), which gives a more detailed treatment.

All model versions we consider contain two equations, which are linear in the coefficients and can be written compactly as

$$y_{git} = x_{git}\beta_g + \alpha_{gi} + u_{git}, \qquad g = 1, 2;\ i \in S_p;\ p = 1, 2;\ t = 1, \ldots, p,$$

where g is the equation number ($g = 1$ represents cigarettes and $g = 2$ represents hand-rolling tobacco), i is the household number, p is the number of periods in which the households are observed, and t is the observation number. The N_1 households observed once have numbers in the index set S_1, and the N_2 households observed twice have their numbers in S_2. In total, there are $n = N_1 + 2N_2$ observations and $N = N_1 + N_2$ households. The endogenous variable in eq. g, y_{git}, is a scalar, x_{git} is a $(1 \times H_g)$ vector of exogenous variables (or transformations of such variables)[4], β_g is its $(H_g \times 1)$ vector of coefficients, α_{gi} is a random household specific effect which includes the latent habit component of commodity g, and u_{git} is a disturbance term. We consider these equations when we use data for all households, as linear approximations to (projections based on) (3), and when we use data for smoker households, as linear approximations to (projections based on) (4). We can write the two equations as

$$(5) \qquad y_{it} = x_{it}\beta + \alpha_i + u_{it} = x_{it}\beta + \varepsilon_{it}, \quad \varepsilon_{it} = \alpha_i + u_{it}, \quad i \in S_p;\ p = 1, 2;\ t = 1, \ldots, p,$$

where

$$y_{it} = \begin{pmatrix} y_{1it} \\ y_{2it} \end{pmatrix}, \ x_{it} = \begin{pmatrix} x_{1it} & 0 \\ 0 & x_{2it} \end{pmatrix}, \ \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix},$$
$$\alpha_i = \begin{pmatrix} \alpha_{1i} \\ \alpha_{2i} \end{pmatrix}, \ u_{it} = \begin{pmatrix} u_{1it} \\ u_{2it} \end{pmatrix}, \ \varepsilon_{it} = \begin{pmatrix} \varepsilon_{1it} \\ \varepsilon_{2it} \end{pmatrix}.$$

[4]Details will be given in Section 5.

We formally consider ε_{it} as a vector of 'gross disturbances' and assume that

$$
\begin{aligned}
&(6) \qquad \mathsf{E}(\alpha_i) = 0_{2,1}, \qquad \mathsf{E}(\alpha_i \alpha_j') = \delta_{ij}\Sigma_\alpha, \\
&(7) \qquad \mathsf{E}(u_{it}) = 0_{2,1}, \qquad \mathsf{E}(u_{it} u_{js}') = \delta_{ij}\delta_{ts}\Sigma_u, \\
&\qquad\qquad x_{it}, \alpha_i, u_{it} \text{ are uncorrelated,}
\end{aligned}
$$

where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$, and Σ_α and Σ_u are positive definite (but otherwise unrestricted) (2×2) matrices.

Let $\varepsilon_{i(p)}$ denote the stacked $(2p \times 1)$ vector of ε_{it}'s for the two equations and the p observations of individual i $(p = 1, 2)$. The composite covariance matrix $\mathsf{E}(\varepsilon_{i(p)}\varepsilon_{i(p)}') = \Omega_{\varepsilon(p)}$, of dimension $(2p \times 2p)$, for each balanced subpanel with p observations then has the form

$$
(8) \qquad \Omega_{\varepsilon(p)} = I_p \otimes \Sigma_u + E_p \otimes \Sigma_\alpha = B_p \otimes \Sigma_u + A_p \otimes (\Sigma_u + p\Sigma_\alpha),
$$

where E_p is the $(p \times p)$ matrix with all elements equal to one, and I_p is the identity matrix of order p, $A_p = E_p/p$, and $B_p = I_p - E_p/p$. The expression after the first equality sign follows from (6) and (7). This is a convenient way of rewriting the covariance matrix, since all columns of A_p are orthogonal to those of B_p, A_p and B_p add to the identity matrix, and both are symmetric and idempotent.

The *generalized least squares (GLS)* problem for estimating the joint coefficient vector β for known values of Σ_α and Σ_u is to minimize

$$
(9) \qquad Q = \sum_{p=1}^{2} \sum_{i \in S_p} [y_{i(p)} - X_{i(p)}\beta]' \Omega_{\varepsilon(p)}^{-1} [y_{i(p)} - X_{i(p)}\beta],
$$

where $y_{i(p)}$ and $X_{i(p)}$ are the stacked vector/matrix of y_{it}'s and x_{it}'s for the p observations of individual i, with respect to β for given Σ_α and Σ_u, subject to (8), using the fact that

$$
\Omega_{\varepsilon(p)}^{-1} = B_p \otimes \Sigma_u^{-1} + A_p \otimes (\Sigma_u + p\Sigma_\alpha)^{-1}.
$$

The GLS estimator of β is

$$
(10) \qquad \widehat{\beta}^{GLS} = \left[\sum_{p=1}^{2} \sum_{i \in S_p} X_{i(p)}' \Omega_{\varepsilon(p)}^{-1} X_{i(p)} \right]^{-1} \left[\sum_{p=1}^{2} \sum_{i \in S_p} X_{i(p)}' \Omega_{\varepsilon(p)}^{-1} y_{i(p)} \right].
$$

If Σ_α and Σ_u were unknown, but ε_{it} were known, unbiased estimators of these covariance matrices would be [see Biørn (1999, p. 4)]

$$
(11) \qquad \widehat{\Sigma}_u = \frac{W_{\varepsilon\varepsilon}}{n - N}, \qquad \widehat{\Sigma}_\alpha = \frac{B_{\varepsilon\varepsilon} - \frac{N-1}{n-N} W_{\varepsilon\varepsilon}}{n - \frac{\sum_{p=1}^{2} N_p p^2}{n}},
$$

where $W_{\varepsilon\varepsilon} = \sum_{p=1}^{2} \sum_{i \in S_p} \sum_{t=1}^{p} (\varepsilon_{it} - \bar{\varepsilon}_{i\cdot})(\varepsilon_{it} - \bar{\varepsilon}_{i\cdot})'$, $B_{\varepsilon\varepsilon} = \sum_{p=1}^{2} \sum_{i \in S_p} p (\bar{\varepsilon}_{i\cdot} - \bar{\varepsilon})(\bar{\varepsilon}_{i\cdot} - \bar{\varepsilon})'$, $\bar{\varepsilon} = \frac{1}{n} \sum_{p=1}^{2} \sum_{i \in S_p} \sum_{t=1}^{p} \varepsilon_{it}$, and $\bar{\varepsilon}_{i\cdot} = \frac{1}{p} \sum_{t=1}^{p} \varepsilon_{it}$ for $i \in S_p$, which are, respectively, the within variation, the between variation, the global mean, and the household specific means of the ε disturbances. Note that the estimator of the covariance matrix of the individual effects, Σ_α, utilizes disturbances from both households observed once and twice, while the covariance matrix of the genuine disturbances, Σ_u, is estimated from disturbances from those observed twice only.

The log-likelihood function of the endogenous variables, to be maximized in the full *Maximum Likelihood (ML) problem*, is:

$$L = -n \ln 2\pi - \frac{1}{2} \sum_{p=1}^{2} N_p \ln |\Omega_{\varepsilon(p)}| - \frac{1}{2} Q(\beta, \Sigma_u, \Sigma_\alpha). \tag{12}$$

Following Biørn (1999, section 4), we split the full ML problem into two conditional subproblems: (A) Maximization of L with respect to β for given Σ_α and Σ_u, and (B) Maximization of L with respect to Σ_α and Σ_u for given β. This motivates an iteration procedure as follows: In the first step we choose some initial values of Σ_α and Σ_u and solve subproblem A. The solution to subproblem A is then used as input in subproblem B; the solution to subproblem B obtained is next used as input in subproblem A, and so on. Oberhofer and Kmenta (1974) [see also Breusch (1987) and Baltagi and Li (1992)] give a set of assumptions which ensure that this kind of 'zig-zag' procedure generates at least one accumulation point, which will be a local maximum of L.

Splitting the maximization problem in this way, greatly simplifies the computation. Subproblem A is identical to GLS, so if the estimators of Σ_α and Σ_u used in the GLS iterations also were solutions to subproblem B, the GLS-iteration would generate the Maximum Likelihood (ML) estimators. However, this is not the case. Except for restrictive special cases, subproblem B does not even have a closed form solution. This can be seen from its first order conditions [cf. Biørn (1999, eq. (64))]:

$$\begin{aligned} \sum_{p=1}^{2} \left[N_p \Sigma_{(p)}^{-1} + N_p (p-1) \Sigma_u^{-1} \right] &= \sum_{p=1}^{2} \left[\Sigma_{(p)}^{-1} \widetilde{B}_{\varepsilon\varepsilon(p)} \Sigma_{(p)}^{-1} + \Sigma_u^{-1} \widetilde{W}_{\varepsilon\varepsilon(p)} \Sigma_u^{-1} \right], \\ \sum_{p=1}^{2} N_p p \Sigma_{(p)}^{-1} &= \sum_{p=1}^{2} p \Sigma_{(p)}^{-1} \widetilde{B}_{\varepsilon\varepsilon(p)} \Sigma_{(p)}^{-1}, \end{aligned}$$

where $\widetilde{W}_{\varepsilon\varepsilon(p)} = \sum_{i \in S_p} \sum_{t=1}^{p} (\varepsilon_{it} - \bar{\varepsilon}_{i\cdot})(\varepsilon_{it} - \bar{\varepsilon}_{i\cdot})'$, and $\widetilde{B}_{\varepsilon\varepsilon(p)} = p \sum_{i \in S_p} (\bar{\varepsilon}_{i\cdot} - \bar{\varepsilon})(\bar{\varepsilon}_{i\cdot} - \bar{\varepsilon})'$, and $\Sigma_{(p)} = \Sigma_u + p\Sigma_\alpha$. In the general case, numerical solution strategies should be considered for solving subproblems A and B iteratively.

To simplify the computations we have used the following modified iteration procedure:

1. Compute the OLS estimates of β and the residuals for each equation separately.

2. Estimate Σ_α and Σ_u by (11), letting the residuals replace the error terms.
3. Compute the GLS-estimator of β, using the Σ estimates from step 2.
4. Repeat steps 2 – 4 until convergence.

The numerical calculations for this iterative Feasible GLS (FGLS) procedure are performed by means of a computer program written in the Gauss software code by the authors.

4 Data

The data set is taken from the Norwegian Surveys of Consumer Expenditures, collected by Statistics Norway, for the years 1975 – 1994 and detailed official Consumer Price Indexes for the same period. The consumer survey data consist of a rotating panel in which roughly 30 per cent of the households participate in two subsequent years and the rest is observed once. The expenditure data are collected almost evenly throughout the year. Roughly 1/26 of the households participate between the 1st and the 14th of January, roughly 1/26 participate between the 15th and the 28th of January, and so on. Most of the expenditure data are reported in *two-week accounting periods*, and yearly expenditure is estimated simply by multiplying the two-week amount by 26. Expenditure on goods with a low purchase frequency rate (*e.g.*, certain durables) are reported in annual interviews.

Tables 1 – 3 contain summary information of the data set. Table 1 gives an overview of definitions, abbreviations, and some descriptive statistics for the variables.[5] Table 2 contains the user frequencies for one or both tobacco commodities. Table 3 reports the number of households observed once and twice in the data set, classified by year. It describes the rotation design of the data set, formally combining 19 balanced two-wave panels with 20 year specific cross-sections. For each year in the 20 year data period, on average about 900 households are observed once and about 200 households are observed twice, giving a total average of about 1300 reports from about 1100 households.

Total consumption expenditure excluding durables is the *income measure* used. The exclusion of durables is done mainly to reduce the number of extreme observations, but also for theoretical reasons. In the official definition of total consumption expenditure, purchases of durables are treated as any other commodity, and symmetrically, revenues from selling such commodities are counted as a negative expenditure. This, in fact, causes total consumption expenditure, including transactions in durables, to be negative for sev-

[5]Total consumption expenditure, age, and cohort have been rescaled to give a mean value of an order of magnitude equal to unity, in order to reduce round off errors in the numerical calculations. Confer Table 1.

eral households which have sold durables and to be extremely high for several households which have had large expenditures on such commodities during the observation period. In any case, our exclusion of durables should give a better proxy as an income measure.[6]

The *price indexes* are from the monthly official Consumer Price Index (CPI) and sub-indexes. The total CPI is used as deflator of the total consumption expenditure excluding durables, while for cigarettes and handrolling tobacco the corresponding detailed sub-indexes have been used. Following a simple set of rules, the monthly price indexes are converted to fit into the two-week periodization in the consumer survey.[7] The relative price between cigarettes and handrolling tobacco has been declining during the observation period, see Figure 1.

The CPI and its sub-indexes are reported only for the whole country, implying that all households are facing the same tobacco prices. However, this assumption may not be as strong as it seems; due to a recommended price policy there was very little, if any, intra-monthly dispersion of prices until early 1991. Probably, most of the variation after 1991 is caused by differences in vendors' mark-up. As far as we know, there is very little difference in prices between brands (within each group of the two tobacco goods) and no particular geographical variation. The neglect of inter-monthly variation in prices is appropriate for the period until 1991, but probably less accurate thereafter.[8]

The *expenditure on cigarettes and other smoking tobacco* is defined as the nominal expenditure divided by the detailed consumer price index of each item. This gives a measure of consumption that is proportional to physical consumption (measured in grams), each commodity having its specific factor of proportionality. Assuming that the total CPI does not differ substantially from the sub-index for durables, a similar deflating is made for the total expenditure on non-durables and the sub-indexes of cigarettes and handrolling tobacco. The average yearly consumption of the two tobacco commodities, measured in grams, is shown in Figure 2.

The *household size* is represented by the number of household members in four age intervals, $0-15$ years, $16-30$ years, $31-60$ years, and $61-99$ years. Four *characteristics of the head of household* are included. Age is measured in the observation year, cohort is

[6]We searched each of the 19 two-wave panels and the 20 cross-sections specified in Table 3 for univariate outliers. If an observation was more than twice the size of its closest neighbour when the observations in these subgroups were ordered by size, it was censored and set to twice the value of its closest neighbour. In all we censored six observations.

[7]For two-week periods which belong entirely to one calendar month, the respective months' indexes are applied directly. For periods overlapping two months the indexes are calculated as weighted arithmetic means of the two months' indexes, using the relative number of days in each month as weights.

[8]Since brand differences in quality are not reflected in prices, these two groups are quite homogeneous along the price dimension at each moment of time, and homogeneous in quality over the entire period.

(rescaled) year of birth, gender is one for females and zero otherwise, and activity is one if the head of household is economically inactive and zero otherwise. Finally, two sets of *geographical dummies* are included. The first set (west, mid, north, east) indicates in which trade region the household is located. The second set [rural, densely, city (the three largest cities)] indicates the population density in the residence municipality.[9]

For most of the variables, there is only small differences between the smoking and no-smoking households, with the two obvious exceptions of the consumption of cigarettes and handrolling tobacco (Table 1). We have not formally tested whether or not the two samples are drawn from the same population, but simply noted that demographic and geographic variables are quite close in the two samples.

5 Empirical results

Overview. Hierarchy of models and model nomenclature

We can divide the model's explanatory variables, contained in the vectors x_{1it} and x_{2it}, into four categories: (i) total expenditure and prices, (ii) household size variables, (iii) characteristics of the head of household (main income earner), and (iv) geographic dummy variables (see Table 1). Twenty model versions are considered, but for only a few we report coefficient estimates.

The version chosen as the *basic model* and the only one for which we report a full set of results (Tables 7 – 10), is a model in which all variables under (i) – (iv) are included and assumed to affect the consumption of both tobacco commodities *linearly*. This implies, for instance, that the Engel and the Cournot derivatives (in terms of deflated expenditure and prices) are constant. We use a nomenclature in which this specification is labeled *Model LLLL*, where the four characters refer to the groups of variables (i) – (iv), respectively, L symbolizing 'linear'. Quadratic terms and/or interaction terms in some of the variables, symbolized by Q in the model label, are included as additional regressors in some models. This makes it possible, to some extent, to examine the curvature of the demand functions and the sensitivity of their derivatives to changes in the background variables, and to test for linearity.[10] Throughout, the same functional form is assumed for both tobacco goods, *i.e.*, $x_{1it} = x_{2it}$. To keep the number of model versions tractable, we *a priori* disregard any kind of interaction between the four groups of variables, so that,

[9]In order to avoid the dummy trap, one category in each set is excluded in the linear regressions ("east" and "city" – which means Oslo). In the "quadratic" regressions (see Section 5), the category "east*city" is excluded.

[10]An additional argument for allowing for non-linearities is the interpretation of the equations we estimate as approximations. See Olsen (1980) for a discussion of a simpler case.

for instance, household size, age, and geographic region are not allowed to affect the Engel or Cournot derivatives. Neither do we include quadratic terms or interaction terms in the price variables and interaction terms involving the activity dummy. The model version which includes quadratic terms and interaction terms for each of the four groups of variables is thus labeled Model QQQQ, Model QLQL includes linear and square terms in total expenditure and linear, square, and interaction terms in age, cohort, and the gender dummy, and is linear otherwise, etc. Omission of a variable group is symbolized by O, so that, for instance, Model LLLO excludes all geographic dummies from an otherwise linear specification. The model versions (hypotheses) can be arranged in a *hypothesis tree*, such that, for instance, Models QLQL and LLLL are nested within QQQQ, Models LLLO, LLOL, LOLL are nested within LLLL.

With a few exceptions, all models are estimated by Maximum Likelihood, approximated as iterative FGLS, as described in Section 3, for two data sets, one including all observations and one including observations from smoking households only. A *smoking household* is defined as a household reporting positive expenditure on at least one of the two tobacco goods *in at least one of the years of observation* (two for the panel, one for the cross-section).

We 'structure' the discussion of our findings by successively focusing on different aspects of the results, starting with the goodness of fit of the various models. The differences between the results based on the full sample and on the sub-sample of smokers will be touched upon at some places, leaving a more systematic discussion and comparison to the next last section. In the final section, we compare selected ML/FGLS results with results based on other estimation methods.

Goodness of fit

The goodness of fit of the twenty models, expressed by their *log-likelihood values* (after omission of an irrelevant constant), is reported in Table 4. The number of unknown parameters in the likelihood function is given in column 1. We find that removing from the basic model, respectively, the household size variables, the age/cohort/gender variables, and the geographic dummies, all lead to a substantial drop in the log-likelihood function (compare Model LLLL with LOLL, LLOL and LLLO). In all cases, this drop is highly significant according to a likelihood ratio test,[11] which gives a clear evidence that all these groups of variables are significant in explaining tobacco consumption. Not unexpectedly, we also find that the income and price variables are highly significant; the log-likelihood value of Model LLLL exceeds that of Model OLLL by more than 450, even though the

[11]Strictly, Likelihood ratio tests based on the likelihood function value evaluated at the estimator point obtained by iterative FGLS (cf. the last part of Section 3) are only approximately valid.

latter only includes six fewer parameters.

Unobserved heterogeneity

We next consider the degree of heterogeneity in tobacco consumption as characterized by properties of the *distribution of the latent α vector*. This vector can be interpreted as including 'stock of habits' related to the two tobacco goods, as discussed in Section 1. An examination of the variation of its estimated covariance matrix, Σ_α, across model versions is interesting. The variances, $(\sigma_{\alpha 1\alpha 1}, \sigma_{\alpha 2\alpha 2})$, can be taken as indicators of the latent preference variation for cigarettes and handrolling tobacco, respectively, while the covariance, $\sigma_{\alpha 1\alpha 2}$, indicates the latent preference covariation between the two goods.

The estimated Σ_α for the twenty model versions are shown in Table 5. The variances based on observations from both smokers and non-smokers and the corresponding covariances are given in columns 1 – 3; similar estimates confined to smokers only are given in columns 6 – 8. Starting with Model LLLL and successively removing all regressors until we finally retain only the intercept term, we find that the estimates of $(\sigma_{\alpha 1\alpha 1}, \sigma_{\alpha 2\alpha 2})$ increase from (64.54, 42.67) to (72.77,47.16) when we use observations from all households and increase from (97.05,49.25) to (121.06, 55.08) when we include smokers only. The overall tendency is that these variances decrease when more variables (including square and interaction terms) are included in the model – in agreement with our expectations. In particular, the variances decrease when we include square terms in income, age, and cohort (compare the results in the first five rows of Table 5). This holds for both commodities, and the tendency is more pronounced when only smokers are considered than when also non-smokers occur in the data set.

Estimates of the ratio between $\sigma_{\alpha g\alpha g}$ and the gross disturbance variance, $\mathrm{var}(\varepsilon_{git}) = \sigma_{\alpha g\alpha g} + \sigma_{ugug}$ $(g = 1, 2)$, are given in Table 5, columns 4 and 5 (all households) and columns 9 and 10 (smokers). This ratio, ρ_g, can be interpreted either (i) as the coefficient of correlation between the two realizations of the gross disturbance ε_{git} from the households observed twice, or (ii) as a (dimensionless) measure of the *degree of latent habit*. The estimates of (ρ_1, ρ_2) increase from (0.6878, 0.7287) to (0.7149, 0.7487) when we successively go from the 'full' linear Model LLLL to a model with only an intercept term, using data from all households. Including smokers only, the corresponding ratios increase from (0.6421, 0.6313) to (0.6963, 0.6567). Thus, by and large, the ρ_g's are fairly constant across model versions, about 60 – 70 per cent, although they tend to decrease slightly with increasing size of the model. Maybe this is a characteristic of the habit structure of tobacco goods in Norwegian households. It should be remembered, though, that several non-economic variables which allegedly affect tobacco consumption, *e.g.*, measures related to ethnicity, religion, and education, are not included in our data set.

It remains an open question whether inclusion of these variables would have reduced the estimates of $\sigma_{\alpha g \alpha g}$ or ρ_g further.

The estimate of the 'preference covariance' $\sigma_{\alpha 1 \alpha 2}$ is positive when based on the data set for all households, but negative when only smokers are included. The different sign may be explained as follows. Inclusion of non-smokers makes the consumption predicted by the model higher than the actual consumption for non-smokers, and tends to make it lower for smokers. The whole sample contains a large proportion of non-smokers, *i.e.*, with zero consumption of both commodities in both periods. Non-smokers consume less than the (conditional) average, and do so systematically over time. This will give a tendency for the latent effect to be negative for both commodities, and thus lead to a positive correlation. The sample of smokers, one the other hand, contains a substantial share of zero observations of one of the commodities, since many smokers use only one kind of tobacco. If this pattern is systematic over time, smokers who only use cigarettes will tend to have a positive α_1 and a negative α_2, and *vice versa* for smokers who only use handrolling tobacco.

As our observation period is rather long, we are able to uncover possible trends or cyclical patterns in the estimated unobserved heterogeneity along with the overall 'structural' change in tobacco consumption illustrated in Figure 2. For this purpose, we have examined the residuals from the estimation of Model LLLL on the full data set separately for each of the 19 sub-panels. Results corresponding to those in Table 5, obtained by using (11) for each sub-panel, are reported in Table 6. Neither the estimates of the absolute variances $\sigma_{\alpha g \alpha g}$ nor their relative counterparts, ρ_g, paints a very clear picture. In the second half of the period the variance is somewhat higher than in the first half for cigarettes; for handrolling tobacco the variance shows a weakly negative trend (columns 1 and 2). This may suggest that the habit structure for tobacco is characterized by increasing latent heterogeneity in cigarette consumption and slightly decreasing heterogeneity in the consumption of handrolling tobacco during the 20 year period. Apart from two outliers for cigarettes, $\rho_1 = 0.34$ in the 1983 – 1984 panel and $\rho_1 = 0.41$ in the 1991 – 1992 panel (both of which reflect relatively high estimated genuine disturbance variances), the relative variances are fairly stable, about 0.60 – 0.80 for cigarettes and about 0.65 – 0.85 for handrolling tobacco. The corresponding 'overall' estimates in Table 5 are 0.69 and 0.73, respectively. We thus find no strong signs that time invariance of the covariance matrices Σ_α and Σ_u is invalid, although an improved goodness of fit could have been obtained by relaxing this assumption.

Engel and Cournot derivatives and elasticities

A full set of coefficient estimates of selected models are reported in Tables 7A–B (quadratic models) and 8A–B (linear models). We interpret the elasticities with respect to total expenditure, calculated from these estimates and evaluated at the overall sample mean of the regressors, given in the first section of Table 9 (rows 1 − 10), as estimates of the average Engel elasticity. Results from both samples suggest that at the mean income, cigarettes is a luxury good (Engel elasticity greater than one) and that handrolling tobacco is a necessity (Engel elasticity between zero and one) or a weakly inferior good (negative Engel elasticity). The average household then will, with increasing income, increase the consumption of cigarettes and keep the consumption of handrolling tobacco roughly constant. It is unlikely that individual and heterogeneous households adjust smoothly to a marginal increase in income, since only a fraction of the households uses both tobacco commodities. However, it seems reasonable that low income households are more inclined than high income households to choose the cheaper of the two substitutes, and that this gives rise to the estimated elasticities.

Examining the curvature of the Engel functions is also interesting. In Model QQQQ the coefficient of the squared total expenditure is significantly negative,[12] while that of its linear term is significantly positive. This applies to both commodities and both samples. For the sample of all households, the estimated functions have maxima at total expenditure 3.92 and 1.57 (corresponding to 392 000 and 157 000 1979-NOK) for cigarettes and handrolling tobacco, respectively. The corresponding maxima for the sample of smokers are 3.78 and 1.28. The maxima for handrolling tobacco are much closer to the sample mean of total expenditure (0.697) than the maxima for cigarettes. Concavity of the estimated functions is not surprising, considering possible saturation effects in total tobacco consumption. The closeness of the maximum point to the sample mean of handrolling tobacco strengthens our conclusion that it is an inferior good for the upper part of the income range. A strict interpretation of the concave quadratic function for cigarettes also implies that demand for cigarettes will decrease at high incomes. However, the number of observations in the income range where the function value decreases, is substantially lower for cigarettes than for handrolling tobacco.

The qualitative pattern of the price elasticities is robust across model variants and samples (compare the rows in the second and third section of Table 9), but the numerical values differ somewhat. For both commodities, the estimated own price elasticities are negative and quite large in absolute value, and the cross price elasticities are positive, suggesting that the commodities are substitutes, as predicted. In Model LLLL estimated on data for all households, the direct price elasticity for cigarettes is -1.700 and its cross price elasticity in the equation for handrolling tobacco is 0.788. Thus, the effect of a

[12] A 5 per cent significance level is used throughout this paper.

one per cent increase in the cigarette price add up to a decrease in total consumption of tobacco of 0.9 per cent. Similarily, a one per cent increase in the price of handrolling tobacco reduces the consumption of handrolling tobacco by 0.829 per cent and increases the consumption of cigarettes will increase by 0.825, adding up to a negligible decrease in total consumption.

Effect of household size variables

The household size is commonly considered an important determinant of household expenditure on most commodities. Our data do not permit us to model and analyze the intra-household decision process for the two tobacco commodities. Following Wangen and Aasness (2001), we interpret differences in the number of household members, *cet.par.* as differences in 'relative household income' – meaning that, for *a given total expenditure*, a household gets poorer if its size is increased by one person. We would expect an increased number of children (0–15 years) to have a negative effect on tobacco consumption, since the household is getting poorer and the newcomer is (presumably) a non-smoker. If a newcomer is an adult smoker, this will have a positive effect on the household tobacco consumption – it is an open question whether or not this effect is stronger than the effect of reduced income per person. On the other hand, we can expect substitution to affect the two tobacco goods oppositely – as poor households may be more inclined than rich ones to choose the cheaper commodity.

From the results for Model LLLL estimated on data for all households (Tables 7A and 8A), we find a negative coefficient estimate for the number of children for both tobacco commodities. For the number of persons in the other age groups, all coefficients are positive. The coefficient values are generally higher for handrolling tobacco than for cigarettes, indicating that substitution induced by changes in household size is important. Using the data set for the smoking households only (Tables 7B and 8B), we find higher coefficient estimates for handrolling tobacco than for cigarettes. A distinct feature of the latter sample is that all the four age group variables have positive coefficient estimates for the handrolling tobacco equation, whilst they are negative in the cigarette equation. This is quite reasonable since substitution effects should be expected to be more pronounced in the sub-sample of smokers than in the whole sample.

Effect of characteristics of the head of household

It is not straightforward to interpret the impact of characteristics of a particular household member on the consumption of the whole household – unless it is a one-person household. Economic inactivity of the head of household (main income earner) surely has a strong influence on the household income; the effects of gender, age, and cohort are less obvious. To some extent age is also related to economic inactivity. We do not have a structural theory for these variables, and therefore do not intend to give a complete interpretation of the estimated effects. The following interpretation is something between a description of the systematic differences between households, and a simplified structural 'analysis' under the assumption that the head is dominating the behaviour of the whole household.

In Model LLLL, the effect of economic inactivity is insignificant for cigarettes, but significantly positive for handrolling tobacco (Tables 7A–B and 8A–B). Hence, the households with inactive heads have a higher consumption of tobacco, and they tend to use the cheaper alternative. In the same model and for both samples, the coefficients of the gender dummy is significant in both equations, but have opposite signs.[13] This may reflect that it is less fashionable for women to smoke handrolling tobacco than cigarettes.

Age and cohort are interesting explanatory variables, as tobacco consumption may vary over the life-cycle and individuals born in the same year share a common history (including the impact of anti-smoking campaigns etc.). Their coefficients in Model LLLL show a quite similar pattern for the two samples – both coefficients are positive for cigarettes and negative for handrolling tobacco. Apart from the age coefficient for cigarettes in the sample of all households, all coefficient estimates are significant.

In the more general Models QLQL and QQQQ, there are no significant coefficients of age and cohort for cigarettes in neither of the samples. For handrolling tobacco many, but not all, of the estimates are significant. For Model QLQL, handrolling tobacco, we find that the estimated function is globally concave in age and cohort and declining over the sample range. At the outset we expected the curvature to be more pronounced than these results imply. However, we should keep in mind that the regressions include neither a trend variable – as the sum of the age and cohort variables equals current time – nor period dummies, and unmodeled trend effects may interfere. The quadratic functions in age and cohort add even more flexibility to the variation over time accounted for. Thus, although the results are not as easily interpretable as we could hope for, at least the structure we have modeled may serve as a correction of trend effects.

Since the relative prices only varies over time, and do so monotonically, unmodeled

[13] Consumption is measured in real expenditure, expressed 1979-NOK. For many purposes it is interesting to use weight units. Provided that the two commodities are (internally) homogeneous in prices, the weight measure is a proportional transformation of the deflated value of the expenditure, where the factor of proportionality for handrolling tobacco is roughly three times the factor for cigarettes. If the results were translated onto a weight scale, the effect of gender on total tobacco consumption would be opposite: Compared with males, females have a higher tobacco expenditure, but they buy a smaller physical amount.

trend effects might have a major impact on the estimated price coefficients, but this does not seem to be the case. In Tables 7A and 7B the absolute value of the own-price coefficient of handrolling tobacco in Model LLLL is higher than in Models QLQL and QQQQ. Otherwise, there is only negligible differences.

Effect of geographic dummy variables

The demographic dummies are significant, with opposite signs, in the two equations of Model LLLL (Table 7A), and the coefficient estimates based on the data set for smokers are roughly twice as large as those based on the whole sample. Compared with the base geographic region and the largest city, Oslo, households in all regions use less cigarettes and more handrolling tobacco. The more elaborate spesification of dummies in Model QQQQ have the same characteristics. In addition there is an indication that, with respect to smoking habits, the second and third largest cities (Bergen and Trondheim) are more similar to Oslo than to the other areas within their respective regions, *cet. par.*

Effect of sample and censoring

The results presented above do not invite making inference on the magnitude of the coefficient vectors β_C and β_H in eq. (1), which is a simplified representation of the equations which determine the latent expenditure of the two commodities. Comparing Tables 7A and 8A (based on data for all households, *i.e.*, the 'non-censored' data set) with Tables 7B and 8B (based on data for smoker households, *i.e.*, the 'censored' data set), we find substantial differences between the coefficient estimates. In general, the latter exceed, in absolute value, the former. This is not surprising, in view of their different interpretion; cf. the discussion leading up to eqs. (3) and (4) (when we, for simplicity, disregard the panel dimension of the model and data). We have that $\mathsf{E}(y_g|x) = \Phi_S \mathsf{E}(y_g|x, y_C + y_H > 0)$, $g = C, H$, for all x, where $\Phi_S \in (0, 1)$ is the overall smoking probability. If the smoking probability were a constant independent of the covariates, say p, all coefficients in Tables 7A and 8A should have been p times the corresponding coefficients in Tables 7B and 8B, and the Engel and Cournot elasticity estimates in Table 9, columns 1 and 3 (based on the complete data set) should have been equal to those in columns 5 and 7 (based on the censored data set). Obviously, the smoking probability is not independent of the covariates. Inspecting the coefficient estimates of the demographic variables for Model LLLL in Tables 8A and 8B, we note, for instance, that the estimates for adults (dem2, dem3, and dem4) for cigarettes and the estimates for children (dem1) for handrolling tobacco have opposite signs for the two samples. This contrasts with many applications of single equation Tobit models, where typically the OLS projection is biased towards zero as compared with the coefficients of the underlying equation in the latent variables.[14] Similar results have not yet been established for the bivariate Tobit, and thus we cannot, by this observation alone, conclude that more general discrete-continuous models should be applied. This topic will not be pursued here, but some aspects are discussed in Wangen and Biørn (2001), which contains a discrete choice analysis of the smoking probabilities for the two commodities, within the framework of (binomial and multinomial) logit models.

Other estimators

The estimates reported so far are based on the random effects specification of the model and obtained by the modified ML procedure described in Section 3. The final tables, Tables 10A–B contain coefficient estimates for Model LLLL based on four different estimators and the panel part of the data set only: The Ordinary Least Squares (OLS), the Feasible Generalized Least Squares (FGLS), the between household estimator (B), and the within household estimator (W). In a single equation context, both the OLS and the FGLS can be interpreted as matrix weighted averages of B and W, FGLS giving a relatively "smaller" weight to B and a relatively "larger" weight to W than OLS [see Hsiao (1986, section 3.3.2) and the examples related to Engel function estimation in Biørn (1994)].

From these results it is obvious that heterogeneity in tobacco consumption is important, not only as an issue on its own, but also for the conclusions which can be drawn about the coefficients of the demand functions. First, the estimates based on OLS, which disregards heterogeneity, differ markedly from the FGLS and the W estimates. For instance, the OLS estimate of the cross price effect of cigarettes has the (theoretically) wrong sign. Second, the FGLS and the W estimates also differ considerably, although all price coefficients of both these models have the right signs. The latter is the estimator we would obtain in a fixed effects specification (interpretation) of the heterogeneity, and this estimator, unlike the FGLS, is robust to correlation between the latent habit effects (including addiction) and the regressor vector in a random effects setting. On the other hand, it utilizes only the within variation in the data set – which, with only two observations of each household in the panel, means that it operates on the differences between the two observations and disregards any level information. This is a considerable disadvantage and gives the method a potentially low efficiency compared with the FGLS and the ML [confer Biørn (1994, table 8)]. Moreover, the estimates of coefficients of demographic variables and geographic dummies may be unreliable. These variables are close to being household specific and so a few households with non-zero differences may

[14] Identification problems and problems related to the choice of parameters of interest for models with censoring are discussed in Heckman (1990) and Leung and Yu (1996).

dominate the results. We therefore conclude that in the present context, despite its larger potential robustness towards error specification, the W estimator is no real competitor to FGLS and ML.

6 Concluding remarks

In this paper we have presented an empirical demand study of two closely substitutable tobacco commodities, cigarettes and handrolling tobacco, based on combined panel data and cross-section data from more than 18 000 Norwegian households. The panel aspect of the data set has been used to identify unobserved individual effects in tobacco consumption, part of which can be attributed to addiction. The model used is formally a system of two static regression equations with random individual heterogeneity, which is estimated by a modified, stepwise Maximum Likelihood procedure.

Different linear and non-linear model versions are estimated and we find a tendency that the variances of the latent individual effects decrease when more variables (including square and interaction terms) are included. The tendency is more pronounced when only smokers are considered than when also non-smokers are included in the data set. Relative to the gross disturbance variances, however, these variances are, however, fairly constant across models, and as large as about 60 – 70 per cent. This can be taken as a support to the addiction hypothesis even if the estimated heterogeneity probably also reflects the effect of explanatory variables related to, *inter alia*, ethnicity, religion, and education, which are not observed in the present data set. We have been unable to detect trends or cyclical patterns in the estimated degree of unobserved heterogeneity along with the overall 'structural' change in tobacco consumption over the data period, 1975 – 1994, although some year to year variation is found.

For both tobacco commodities, the estimated own price elasticities are negative and the cross-price elasticities are positive, as predicted by consumer theory. The qualitative conclusions that cigarettes is a luxury good and handrolling tobacco is a necessity or a weakly inferior good at the sample mean is robust over model variants and hold both when the data set includes all households and only smokers.

Finally, some attention has been given to differences between estimation results based on the full sample and the censored sample with only observed smokers included. A fuller discussion of estimation problems for two-commodity discrete-continuous choice tobacco demand models for panel data with heterogeneity is left for future research. A more limited discrete-choice analysis of tobacco consumption, using the same panel data set, is given in Wangen and Biørn (2001).

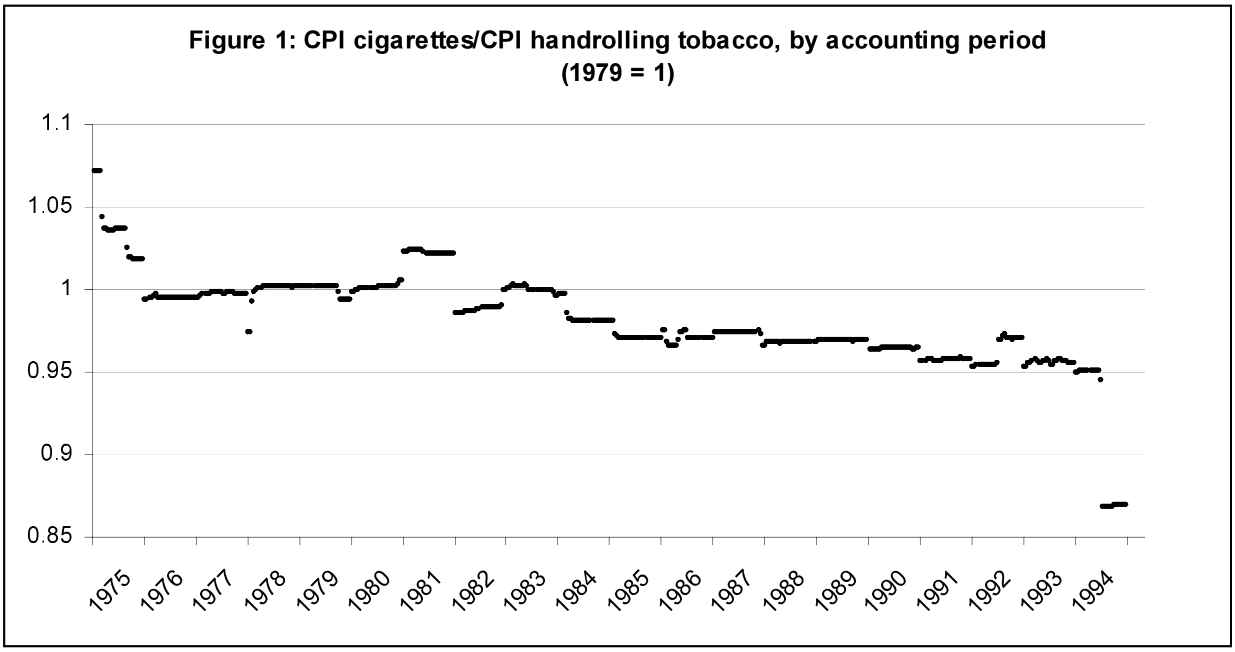
Figure 1: CPI cigarettes/CPI handrolling tobacco, by accounting period
(1979 = 1)
1.1
1.05
1
0.95
0.9
0.85
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994

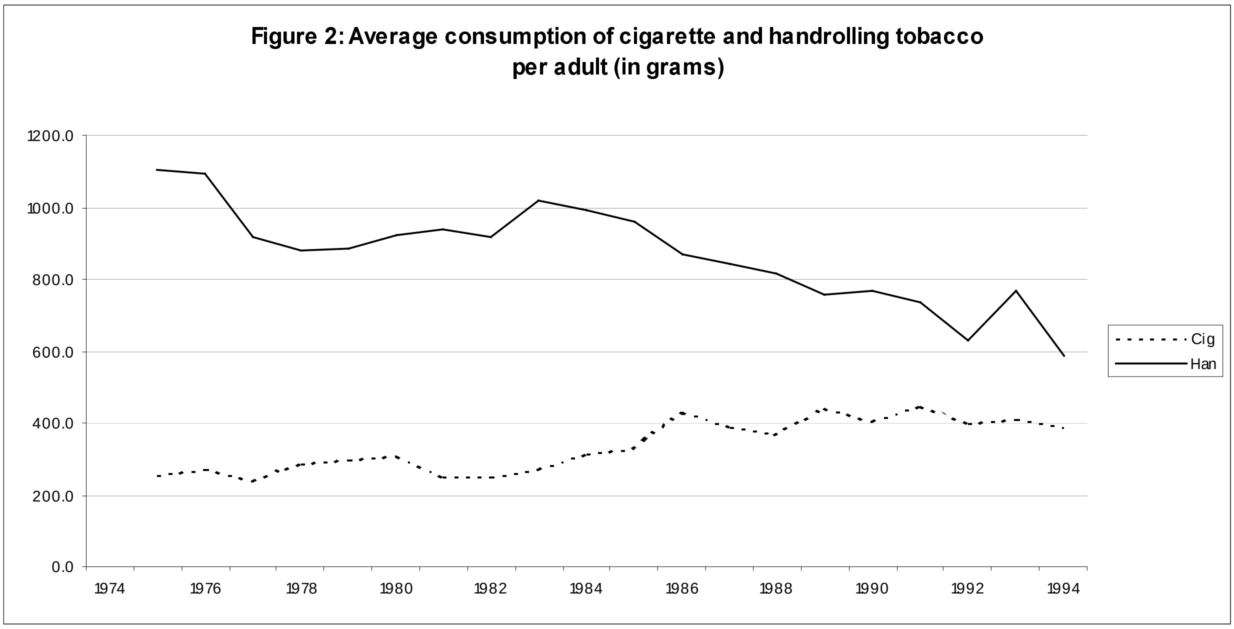
Figure 2: Average consumption of cigarette and handrolling tobacco
per adult (in grams)
1200.0
1000.0
800.0
600.0
400.0
200.0
0.0
1974
1976
1978
1980
1982
1984
1986
1988
1990
1992
1994
Cig
Han

Table 1: Descriptive statistics				
	All households		Smokers	
	Mean	St.dev.	Mean	St.dev.
Cig	3.646	10.089	7.109	13.185
Han	4.586	7.937	8.941	9.158
Texp	0.679	0.445	0.734	0.443
Pcig	1.207	0.177	1.201	0.175
Phan	1.235	0.215	1.228	0.212
dem1	0.734	1.034	0.808	1.037
dem2	0.664	0.840	0.798	0.870
dem3	1.067	0.880	1.187	0.857
dem4	0.426	0.720	0.308	0.643
Age	4.856	1.636	4.542	1.483
Coho	5.616	1.782	5.911	1.632
Gend	0.230	0.421	0.190	0.392
Inac	0.273	0.445	0.219	0.414
west	0.241	0.428	0.227	0.419
cent	0.142	0.349	0.147	0.354
nor	0.083	0.276	0.090	0.286
east	0.535	0.499	0.536	0.499
rur	0.244	0.429	0.228	0.419
dens	0.558	0.497	0.574	0.494
city	0.198	0.398	0.198	0.399

Explanation:

Endogenous variables
Cig = (Expenditure on cigarettes in nominal kr.)/(CPI for cigarettes[a])
Han = (Expenditure on handrolling tobacco in nominal kr.)/(CPI for handrolling tobacco)

Exogenous variables
Group 1
Texp = 0.001*((total expend.)-(expend. on durables))/(Total CPI)
Pcig = (CPI cigarettes)/(Total CPI)
Phan = (CPI handrolling tobacco)/(Total CPI)
Group 2
dem1 = Number of persons in age group [0,16)
dem2 = Number of persons in age group [16,31)
dem3 = Number of persons in age group [31,61)
dem4 = Number of persons in age group [61,99)
Group 3
Age = 0.1*(age of head of household)
Coho = 0.1*((year of birth, head of household)-1880)
Gend = 1 if head is female, 0 otherwise
Inac = 1 if head is economically inactive, 0 otherwise
Group 4
west = 1 if residence is in the western trade region, 0 otherwise
cent = 1 if residence is in the central trade region, 0 otherwise
nor = 1 if residence is in the northern trade region, 0 otherwise
east = 1 if residence is in the eastern trade region, 0 otherwise
rur = 1 if residental municipality is rural (with less than 50% of residents in densely populated area), 0 otherwise
dens = 1 if residental municipality is densely populated (with 50% or more of residents in densely populated area (exept Oslo, Bergen and Trondheim)), 0 otherwise
city = 1 if resident in a city (Oslo, Bergen or Trondheim), 0 otherwise

[a] All CPI=100 in July 1979.

Table 2: User status by year. Relative frequency, per cent				
Year	**None**	**Handroll only**	**Cigarettes only**	**Both**
1975	44.3	31.5	7.6	16.6
1976	43.7	32.0	8.7	15.6
1977	48.1	26.7	10.3	14.9
1978	51.5	24.0	9.4	15.1
1979	49.7	26.5	8.4	15.3
1980	51.2	24.9	9.3	14.6
1981	51.0	26.2	7.6	15.1
1982	51.1	27.8	8.6	12.5
1983	49.8	28.7	9.2	12.4
1984	50.9	26.4	8.7	14.0
1985	49.6	25.1	11.0	14.3
1986	49.4	22.4	13.5	14.7
1987	50.9	20.2	13.6	15.2
1988	52.5	20.8	12.7	14.0
1989	50.6	19.3	15.3	14.7
1990	53.1	20.6	13.0	13.3
1991	54.3	17.7	14.8	13.2
1992	52.2	17.5	15.3	14.9
1993	50.4	19.8	15.4	14.4
1994	55.6	16.2	15.1	13.2

Table 3: Number of observations in different cross-sections and panels			
Cross sections		**Two-year panels**	
Year	**obs.**	**Panel**	**obs.**
1975	904	-	-
1976	727	**1975-1976**	426
1977	551	**1976-1977**	466
1978	551	**1977-1978**	416
1979	1050	**1978-1979**	426
1980	730	**1979-1980**	406
1981	1120	**1980-1981**	404
1982	1007	**1981-1982**	420
1983	1029	**1982-1983**	488
1984	1084	**1983-1984**	416
1985	1118	**1984-1985**	436
1986	1103	**1985-1986**	438
1987	907	**1986-1987**	328
1988	1068	**1987-1988**	328
1989	801	**1988-1989**	356
1990	814	**1989-1990**	380
1991	873	**1990-1991**	388
1992	955	**1991-1992**	434
1993	899	**1992-1993**	432
1994	1142	**1993-1994**	386
Sum	18433	**Sum**	7774
Total	26207		

Table 4: Log-likelihood values			
Model	**No. of parameters**[a]	**All households**[b]	**Smokers**[c]
QQQQ	82	-424.90	-103.11
QLQL	52	-532.68	-184.39
LLQL	50	-600.48	-260.08
QLLL	42	-618.69	-233.29
LLLL	40	-687.91	-309.73
OLLL	34	-1152.87	-761.30
LOLL	32	-1188.08	-563.40
LLOL	32	-841.89	-375.85
LLLO	30	-958.42	-604.69
OOLL	26	-1757.14	-1089.23
OLOL	26	-1465.70	-953.47
OLLO	24	-1526.21	-1150.02
LOOL	24	-1494.02	-672.05
LOLO	22	-1530.36	-941.00
LLOO	22	-1118.03	-672.56
OOOL	18	-2349.99	-1343.56
OOLO	16	-2157.75	-1510.53
OLOO	16	-1845.41	-1328.00
LOOO	14	-1863.28	-1072.26
OOOO	8	-2776.95	-1762.46

[a] Total numbers of parameters, i.e. coefficients and elements in Σ_α and Σ_u
[b] Log-likelihood value + 212000
[c] Log-likelihood value + 117000

Table 5: Characteristics of disturbance covariance matrices										
	All households					Smokers				
Model	$\sigma_{\alpha1\alpha1}$	$\sigma_{\alpha2\alpha2}$	$\sigma_{\alpha1\alpha2}$	ρ_1 [a]	ρ_2 [b]	$\sigma_{\alpha1\alpha1}$	$\sigma_{\alpha2\alpha2}$	$\sigma_{\alpha1\alpha2}$	ρ_1 [a]	ρ_2 [b]
QQQQ	64.09	41.82	2.91	0.6872	0.7239	96.23	47.90	-17.63	0.6431	0.6241
QLQL	64.33	42.32	2.99	0.6883	0.7273	96.87	48.73	-17.83	0.6450	0.6292
LLQL	64.31	42.38	3.05	0.6867	0.7274	96.86	48.79	-17.75	0.6416	0.6293
QLLL	64.55	42.60	3.00	0.6893	0.7285	97.05	49.17	-18.05	0.6455	0.6311
LLLL	64.54	42.67	3.07	0.6878	0.7287	97.05	49.25	-17.96	0.6421	0.6313
OLLL	68.09	42.67	2.88	0.7011	0.7286	108.21	49.33	-19.25	0.6710	0.6318
LOLL	65.31	44.56	3.05	0.6910	0.7382	98.93	51.87	-19.59	0.6486	0.6447
LLOL	65.01	43.29	3.10	0.6899	0.7328	98.15	50.01	-18.65	0.6459	0.6362
LLLO	65.49	42.91	2.71	0.6892	0.7296	101.04	49.89	-19.49	0.6476	0.6340
OOLL	68.95	44.84	3.98	0.7033	0.7388	108.85	52.01	-18.55	0.6714	0.6442
OLOL	68.89	43.83	2.60	0.7041	0.7352	111.05	51.10	-21.34	0.6783	0.6415
OLLO	70.11	42.92	2.40	0.7067	0.7296	115.78	50.04	-21.37	0.6845	0.6352
LOOL	65.94	45.57	2.71	0.6938	0.7422	101.07	52.80	-21.05	0.6562	0.6487
LOLO	66.68	44.97	2.40	0.6944	0.7399	105.05	53.03	-22.21	0.6602	0.6497
LLOO	66.00	43.53	2.76	0.6913	0.7338	102.28	50.67	-20.22	0.6517	0.6390
OOOL	70.38	46.79	4.29	0.7080	0.7473	112.44	53.90	-20.90	0.6805	0.6518
OOLO	71.02	45.21	3.34	0.7092	0.7403	116.95	53.11	-21.31	0.6864	0.6488
OLOO	71.05	44.05	2.15	0.7101	0.7362	118.68	51.75	-23.40	0.6915	0.6445
LOOO	67.53	46.05	1.95	0.6981	0.7443	108.12	54.18	-24.12	0.6704	0.6546
OOOO	72.77	47.16	3.62	0.7149	0.7487	121.06	55.08	-23.93	0.6963	0.6567

Table 6: Characteristics of disturbance covariance matrices by panel. LLLL model for all households					
Panel	$\sigma_{\alpha1\alpha1}$	$\sigma_{\alpha2\alpha2}$	$\sigma_{\alpha1\alpha2}$	ρ_1 [a]	ρ_2 [b]
1975-1976	67.14	61.56	4.12	0.8252	0.7973
1976-1977	59.62	53.33	0.04	0.7820	0.7456
1977-1978	48.36	45.42	3.81	0.7882	0.7363
1978-1979	66.81	33.56	5.58	0.7289	0.6325
1979-1980	71.74	38.43	0.70	0.7155	0.7204
1980-1981	50.99	39.70	5.00	0.6282	0.6502
1981-1982	54.87	45.70	2.93	0.7576	0.7675
1982-1983	54.89	48.16	0.04	0.7299	0.7343
1983-1984	24.54	40.28	0.95	0.3383	0.6851
1984-1985	55.78	33.19	4.17	0.6765	0.5645
1985-1986	84.90	47.39	2.55	0.7902	0.7471
1986-1987	83.92	49.27	5.39	0.6933	0.8049
1987-1988	90.19	31.19	6.51	0.7985	0.5402
1988-1989	71.77	40.58	2.39	0.6950	0.7705
1989-1990	87.65	30.29	3.81	0.6961	0.6520
1990-1991	100.04	46.31	1.06	0.7968	0.8641
1991-1992	51.28	37.89	3.65	0.4073	0.8008
1992-1993	64.10	37.16	0.47	0.6005	0.7847
1993-1994	58.16	35.73	1.72	0.7058	0.7756

[a] $\rho_1=\sigma_{\alpha1\alpha1}/(\sigma_{\alpha1\alpha1}+\sigma_{u1u1})$

[b] $\rho_2=\sigma_{\alpha2\alpha2}/(\sigma_{\alpha2\alpha2}+\sigma_{u2u2})$

Table 7 A: Coefficient estimates and standard errors in selected nonlinear models. All households

	Cigarettes						Handrolling tobacco					
	QQQQ		QLQL		LLLL		QQQQ		QLQL		LLLL	
	Coef	St. err.	Coef	St. err.	Coef	St. err.	Coef	St. err.	Coef	St. err.	Coef	St. err.
Const	-18.9027	28.5141	-17.2278	28.4810	2.5752	2.0986	-30.6020	22.3588	-19.6624	22.3847	14.9728	1.6472
Texp	6.7604	0.3229	6.7531	0.3201	4.5085	0.1606	1.3322	0.2484	1.3506	0.2465	0.1401	0.1248
Pcig	-6.2200	2.4409	-6.2783	2.4417	-6.2321	2.4453	4.4513	1.8877	4.1524	1.8907	3.6194	1.8938
Phan	3.0193	2.0510	2.9939	2.0518	3.0235	2.0081	-3.2168	1.5875	-3.0937	1.5902	-3.8068	1.5569
dem1	-1.1555	0.2555	-0.8690	0.0772	-0.8810	0.0756	-0.2829	0.2005	-0.1345	0.0608	-0.1532	0.0596
dem2	0.2774	0.3070	0.0908	0.0914	0.2172	0.0904	1.5280	0.2402	1.1543	0.0717	1.2389	0.0711
dem3	1.4608	0.4721	-0.0923	0.1344	0.1353	0.1136	2.8281	0.3696	1.4890	0.1055	1.9102	0.0893
dem4	0.3611	0.3699	-0.2609	0.1807	0.1213	0.1731	1.8030	0.2903	0.6093	0.1419	0.7579	0.1361
Age	5.1383	5.6318	5.0725	5.6255	0.0813	0.2761	5.8055	4.4160	4.0819	4.4213	-1.5989	0.2162
Coho	3.8645	5.4355	3.3471	5.4325	0.6688	0.2650	8.9751	4.2611	7.0359	4.2686	-1.1328	0.2074
Gend	-1.1168	2.9716	-0.7735	2.9667	0.7394	0.1654	-8.2606	2.3290	-9.4720	2.3304	-0.9689	0.1300
Inac	0.2513	0.1806	0.2156	0.1801	0.2021	0.1755	1.4360	0.1407	1.4223	0.1406	1.1921	0.1371
Texp* Texp	-0.8635	0.1069	-0.8617	0.1063	--	--	-0.4242	0.0815	-0.4432	0.0810	--	--
dem1*dem1	0.1119	0.0484	--	--	--	--	0.1499	0.0380	--	--	--	--
dem2*dem2	-0.0015	0.0755	--	--	--	--	-0.0604	0.0589	--	--	--	--
dem3*dem3	-0.5157	0.1624	--	--	--	--	-0.1416	0.1271	--	--	--	--
dem4*dem4	-0.0929	0.0886	--	--	--	--	-0.2161	0.0696	--	--	--	--
dem1*dem2	0.0664	0.0932	--	--	--	--	0.3796	0.0729	--	--	--	--
dem1*dem3	-0.0982	0.1146	--	--	--	--	-0.3945	0.0898	--	--	--	--
dem1*dem4	0.3212	0.2716	--	--	--	--	-0.5231	0.2127	--	--	--	--
dem2*dem3	-0.1332	0.1323	--	--	--	--	-0.3690	0.1035	--	--	--	--
dem2*dem4	0.0069	0.1974	--	--	--	--	-0.2778	0.1540	--	--	--	--
dem3*dem4	-0.5559	0.2162	--	--	--	--	-0.1446	0.1693	--	--	--	--
Age*Age	-0.2677	0.2854	-0.2784	0.2849	--	--	-0.3516	0.2237	-0.2726	0.2239	--	--
Coho*Coho	-0.1110	0.2657	-0.0849	0.2655	--	--	-0.6025	0.2081	-0.5111	0.2085	--	--
Gend*Age	-0.0116	0.2988	-0.0695	0.2985	--	--	0.6721	0.2342	0.7454	0.2345	--	--
Gend*Coho	0.3339	0.2712	0.3227	0.2710	--	--	0.7623	0.2126	0.8719	0.2128	--	--
Age*Coho	-0.4476	0.5425	-0.3951	0.5422	--	--	-0.8282	0.4251	-0.6350	0.4259	--	--
west*rur	-4.0380	0.3184	--	--	--	--	0.3991	0.2509	--	--	--	--
west*dens	-3.3823	0.2665	--	--	--	--	1.3851	0.2101	--	--	--	--
west*city	-2.6256	0.3412	--	--	--	--	0.7631	0.2690	--	--	--	--
cent*rur	-3.5159	0.4236	--	--	--	--	0.9002	0.3338	--	--	--	--
cent*dens	-3.5474	0.3003	--	--	--	--	1.7863	0.2368	--	--	--	--
cent*city	-2.6626	0.3908	--	--	--	--	0.8880	0.3081	--	--	--	--
nor*rur	-4.6370	0.4257	--	--	--	--	2.2750	0.3354	--	--	--	--
nor*dens	-4.0608	0.3341	--	--	--	--	2.0842	0.2634	--	--	--	--
east*rur	-3.1976	0.2718	--	--	--	--	0.9387	0.2142	--	--	--	--
east*dens	-2.1077	0.2253	--	--	--	--	0.8395	0.1776	--	--	--	--
west	--	--	-1.4625	0.1573	-1.4686	0.1576	--	--	0.3047	0.1244	0.2831	0.1248
cent	--	--	-1.4858	0.1908	-1.5228	0.1911	--	--	0.7272	0.1508	0.7117	0.1512
nor	--	--	-1.9292	0.2406	-1.9228	0.2410	--	--	1.3441	0.1902	1.3367	0.1908
rur	--	--	-2.3197	0.2036	-2.5061	0.2025	--	--	0.5270	0.1608	0.4102	0.1602
dens	--	--	-1.5572	0.1711	-1.6368	0.1711	--	--	0.8153	0.1353	0.7700	0.1354

Table 7 B: Coefficient estimates and standard errors in selected non-linear models. Smokers												
	Cigarettes						Handrolling tobacco					
	QQQQ		QLQL		LLLL		QQQQ		QLQL		LLLL	
	Coef	St.err.	Coef	St.err.	Coef	St.err.	Coef	St.err.	Coef	St.err.	Coef	St.err.
Const	-7.5960	50.1460	-12.4108	50.1022	-3.3399	3.7285	-104.1250	35.9460	-86.9061	36.0046	17.3912	2.6813
Texp	11.8150	0.5410	11.7237	0.5371	7.5922	0.2754	1.1760	0.3910	1.1873	0.3884	-0.1204	0.1984
Pcig	-14.3700	4.4800	-14.1633	4.4832	-13.9358	4.4995	7.7920	3.2290	7.2697	3.2364	6.7095	3.2445
Phan	8.3960	3.8140	8.1855	3.8165	8.4209	3.7417	-4.9600	2.7480	-4.7628	2.7545	-6.9771	2.6975
dem1	-2.2220	0.4290	-1.1835	0.1321	-1.1211	0.1306	-0.8660	0.3070	0.0824	0.0948	0.1206	0.0938
dem2	-1.2360	0.5500	-0.5896	0.1549	-0.3988	0.1537	1.6010	0.3950	1.1608	0.1113	1.2333	0.1105
dem3	0.1210	0.8750	-0.9737	0.2389	-0.6994	0.2042	3.5330	0.6270	1.8998	0.1717	2.2694	0.1468
dem4	-1.2940	1.1000	-0.8905	0.3431	-0.5240	0.3335	2.9580	0.7890	1.0661	0.2466	0.9263	0.2397
Age	4.4820	9.9460	5.6365	9.9369	1.5231	0.4941	19.6800	7.1300	16.6026	7.1408	-1.3425	0.3554
Coho	1.1650	9.5750	1.4016	9.5733	1.8856	0.4704	24.1630	6.8640	21.2816	6.8802	-1.2654	0.3384
Gend	2.6920	5.5910	4.5481	5.5706	1.6336	0.3025	-8.7550	4.0090	-11.1698	4.0048	-1.3048	0.2175
Inac	-0.3630	0.3140	-0.4081	0.3135	-0.4745	0.3066	1.7310	0.2260	1.6918	0.2257	1.4269	0.2207
Texp* Texp	-1.5650	0.1710	-1.5389	0.1697	--	--	-0.4610	0.1240	-0.4803	0.1232	--	--
dem1*dem1	0.1310	0.0830	--	--	--	--	0.2120	0.0600	--	--	--	--
dem2*dem2	0.1480	0.1260	--	--	--	--	0.0360	0.0910	--	--	--	--
dem3*dem3	-0.5850	0.2820	--	--	--	--	-0.1440	0.2020	--	--	--	--
dem4*dem4	0.1040	0.3700	--	--	--	--	-0.4010	0.2650	--	--	--	--
dem1*dem2	0.2330	0.1580	--	--	--	--	0.6520	0.1130	--	--	--	--
dem1*dem3	0.2760	0.1910	--	--	--	--	-0.1440	0.1370	--	--	--	--
dem1*dem4	0.5080	0.4540	--	--	--	--	-0.9930	0.3260	--	--	--	--
dem2*dem3	0.1020	0.2320	--	--	--	--	-0.6180	0.1670	--	--	--	--
dem2*dem4	0.4510	0.3490	--	--	--	--	-0.4910	0.2510	--	--	--	--
dem3*dem4	-0.3850	0.4350	--	--	--	--	-0.3230	0.3120	--	--	--	--
Age*Age	-0.2090	0.5080	-0.2953	0.5069	--	--	-1.0160	0.3640	-0.8479	0.3643	--	--
Coho*Coho	0.1270	0.4690	0.1231	0.4691	--	--	-1.4010	0.3360	-1.2674	0.3372	--	--
Gend*Age	-0.2710	0.5680	-0.4271	0.5657	--	--	0.7770	0.4070	1.0021	0.4067	--	--
Gend*Coho	-0.0230	0.5120	-0.1795	0.5111	--	--	0.7200	0.3670	0.9036	0.3674	--	--
Age*Coho	-0.1900	0.9590	-0.2234	0.9588	--	--	-2.2170	0.6880	-1.9178	0.6892	--	--
west*rur	-7.9930	0.6000	--	--	--	--	2.3310	0.4290	--	--	--	--
west*dens	-6.8010	0.4760	--	--	--	--	3.2430	0.3400	--	--	--	--
west*city	-4.7180	0.6120	--	--	--	--	2.0470	0.4380	--	--	--	--
cent*rur	-6.9740	0.7720	--	--	--	--	2.1540	0.5530	--	--	--	--
cent*dens	-7.4840	0.5220	--	--	--	--	2.8230	0.3740	--	--	--	--
cent*city	-5.0440	0.6860	--	--	--	--	2.0590	0.4900	--	--	--	--
nor*rur	-9.4780	0.7400	--	--	--	--	3.7440	0.5300	--	--	--	--
nor*dens	-8.4080	0.5770	--	--	--	--	3.3830	0.4130	--	--	--	--
east*rur	-6.2710	0.4880	--	--	--	--	2.1120	0.3490	--	--	--	--
east*dens	-4.2340	0.4000	--	--	--	--	1.7610	0.2860	--	--	--	--
west	--	--	-2.8434	0.2855	-2.7959	0.2861	--	--	1.2908	0.2048	1.2419	0.2054
cent	--	--	-3.1794	0.3336	-3.2260	0.3344	--	--	1.0642	0.2393	1.0523	0.2401
nor	--	--	-4.0737	0.4136	-4.0682	0.4147	--	--	1.7678	0.2966	1.7419	0.2977
rur	--	--	-4.7943	0.3677	-5.1296	0.3663	--	--	1.3778	0.2639	1.2836	0.2630
dens	--	--	-3.4367	0.3034	-3.5573	0.3035	--	--	1.4480	0.2177	1.4102	0.2179

Tabel 8 A: Coefficient estimates and standard errors in selected linear models. All households												
	Cigarettes						Handrolling tobacco					
	LLLL		LLLO		LLOO		LLLL		LLLO		LLOO	
	Coef	St.err.	Coef	St.err.	Coef	St.err.	Coef	St.err.	Coef	St.err.	Coef	St.err.
Const	2.5752	2.0986	2.4185	2.1111	3.8490	0.6443	14.9728	1.6472	14.8930	1.6503	4.8102	0.5019
Texp	4.5085	0.1606	4.9702	0.1583	5.1058	0.1562	0.1401	0.1248	0.0500	0.1226	0.0063	0.1211
Pcig	-6.2321	2.4453	-6.1453	2.4591	-5.1915	2.4373	3.6194	1.8938	3.5158	1.8969	2.3030	1.8820
Phan	3.0235	2.0081	3.0357	2.0195	3.6705	1.9979	-3.8068	1.5569	-3.7531	1.5595	-5.0590	1.5426
dem1	-0.8810	0.0756	-1.0886	0.0751	-0.8788	0.0688	-0.1532	0.0596	-0.0996	0.0590	0.1450	0.0542
dem2	0.2172	0.0904	-0.0316	0.0898	0.0025	0.0869	1.2389	0.0711	1.3009	0.0703	1.4589	0.0681
dem3	0.1353	0.1136	-0.0758	0.1135	-0.6467	0.0960	1.9102	0.0893	1.9729	0.0889	1.8180	0.0754
dem4	0.1213	0.1731	-0.1403	0.1733	-1.3631	0.1211	0.7579	0.1361	0.8198	0.1358	0.5489	0.0954
Age	0.0813	0.2761	-0.1286	0.2773	--	--	-1.5989	0.2162	-1.5180	0.2163	--	--
Coho	0.6688	0.2650	0.5151	0.2663	--	--	-1.1328	0.2074	-1.0558	0.2076	--	--
Gend	0.7394	0.1654	0.7961	0.1662	--	--	-0.9689	0.1300	-0.9639	0.1301	--	--
Inac	0.2021	0.1755	0.1809	0.1765	--	--	1.1921	0.1371	1.2249	0.1373	--	--
west	-1.4686	0.1576	--	--	--	--	0.2831	0.1248	--	--	--	--
cent	-1.5228	0.1911	--	--	--	--	0.7117	0.1512	--	--	--	--
nor	-1.9228	0.2410	--	--	--	--	1.3367	0.1908	--	--	--	--
rur	-2.5061	0.2025	--	--	--	--	0.4102	0.1602	--	--	--	--
dens	-1.6368	0.1711	--	--	--	--	0.7700	0.1354	--	--	--	--

Tabel 8 B: Coefficient estimates and standard errors in selected linear models. Smokers												
	Cigarettes						Handrolling tobacco					
	LLLL		LLLO		LLOO		LLLL		LLLO		LLOO	
	Coef	St.err.	Coef	St.err.	Coef	St.err.	Coef	St.err.	Coef	St.err.	Coef	St.err.
Const	-3.3399	3.7285	-2.9967	3.7849	7.1645	1.1723	17.3912	2.6813	17.2375	2.6912	8.4713	0.8353
Texp	7.5922	0.2754	8.4332	0.2737	8.6010	0.2705	-0.1204	0.1984	-0.3089	0.1951	-0.4690	0.1929
Pcig	-13.9358	4.4995	-13.8143	4.5614	-11.9995	4.5369	6.7095	3.2445	6.5735	3.2544	5.3608	3.2391
Phan	8.4209	3.7417	8.2736	3.7935	10.6637	3.7374	-6.9771	2.6975	-6.8203	2.7058	-8.5079	2.6684
dem1	-1.1211	0.1306	-1.5633	0.1307	-1.5146	0.1188	0.1206	0.0938	0.2608	0.0929	0.3432	0.0845
dem2	-0.3988	0.1537	-0.9361	0.1535	-1.0456	0.1512	1.2333	0.1105	1.3935	0.1091	1.4592	0.1075
dem3	-0.6994	0.2042	-1.2107	0.2055	-1.8329	0.1668	2.2694	0.1468	2.4182	0.1460	2.5268	0.1186
dem4	-0.5240	0.3335	-1.1204	0.3369	-2.3544	0.2288	0.9263	0.2397	1.0837	0.2394	1.5646	0.1626
Age	1.5231	0.4941	1.1088	0.5010	--	--	-1.3425	0.3554	-1.1947	0.3564	--	--
Coho	1.8856	0.4704	1.5150	0.4770	--	--	-1.2654	0.3384	-1.1254	0.3393	--	--
Gend	1.6336	0.3025	1.8644	0.3063	--	--	-1.3048	0.2175	-1.3740	0.2177	--	--
Inac	-0.4745	0.3066	-0.5039	0.3109	--	--	1.4269	0.2207	1.4542	0.2214	--	--
west	-2.7959	0.2861	--	--	--	--	1.2419	0.2054	--	--	--	--
cent	-3.2260	0.3344	--	--	--	--	1.0523	0.2401	--	--	--	--
nor	-4.0682	0.4147	--	--	--	--	1.7419	0.2977	--	--	--	--
rur	-5.1296	0.3663	--	--	--	--	1.2836	0.2630	--	--	--	--
dens	-3.5573	0.3035	--	--	--	--	1.4102	0.2179	--	--	--	--

Table 9: Engel and Cournot elasticities at sample mean for different models								
	All households				Smokers			
	Cigarettes		Handrolling		Cigarettes		Handrolling	
Model	elast	st.err.	elast	st.err.	elast	st.err.	elast	st.err.
Engel elasticities:								
QLQL	1.4819	0.0627	0.1422	0.0287	1.2956	0.0573	0.0483	0.0258
QLLL	1.4570	0.0515	0.1555	0.0321	1.2646	0.0479	0.0510	0.0282
LLLL	1.2298	0.0331	0.0305	0.0185	1.0538	0.0306	-0.0133	0.0161
LLLO	1.3558	0.0335	0.0109	0.0181	1.1706	0.0311	-0.0342	0.0159
LLOL	1.2608	0.0329	0.0191	0.0182	1.0786	0.0304	-0.0316	0.0159
LOLL	1.1730	0.0307	0.2823	0.0173	0.9569	0.0280	0.1457	0.0150
LLOO	1.3915	0.0334	0.0014	0.0179	1.1939	0.0310	-0.0519	0.0157
LOLO	1.2491	0.0311	0.2648	0.0172	1.0174	0.0286	0.1328	0.0149
LOOL	1.2086	0.0293	0.3991	0.0164	0.9545	0.0272	0.1477	0.0145
LOOO	1.2833	0.0297	0.3827	0.0163	1.0061	0.0278	0.1363	0.0145
Elasticity with respect to the price of cigarettes:								
QLQL	-1.6663	0.7867	0.7886	0.4335	-1.9387	0.7419	0.7277	0.3894
QLLL	-1.5969	0.7686	0.7818	0.4880	-1.8637	0.7183	0.7370	0.4267
LLLL	-1.7000	0.8057	0.7884	0.4979	-1.9343	0.7511	0.7427	0.4315
LLLO	-1.6763	0.8102	0.7658	0.4987	-1.9176	0.7614	0.7277	0.4328
LLOL	-1.3864	0.7975	0.5022	0.4936	-1.6279	0.7468	0.5923	0.4293
LOLL	-1.6703	0.8066	1.0239	0.5034	-1.9261	0.7511	0.8990	0.4362
LLOO	-1.4149	0.8021	0.5013	0.4944	-1.6657	0.7572	0.5932	0.4306
LOLO	-1.6922	0.8115	1.0148	0.5047	-1.9573	0.7628	0.8992	0.4383
LOOL	-1.3109	0.7985	0.6572	0.5012	-1.5782	0.7468	0.7015	0.4351
LOOO	-1.3557	0.8035	0.6667	0.5026	-1.6315	0.7587	0.7126	0.4375
Elasticity with respect to the price of handrolling tobacco:								
QLQL	0.7946	0.6691	-0.5876	0.3753	1.1205	0.6374	-0.4768	0.3402
FLLL	0.7740	0.6457	-0.8169	0.4107	1.1326	0.6105	-0.7648	0.3628
LLLL	0.8248	0.6769	-0.8292	0.4190	1.1688	0.6384	-0.7723	0.3669
LLLO	0.8281	0.6807	-0.8175	0.4197	1.1485	0.6472	-0.7550	0.3680
LLOL	1.0454	0.6691	-1.1331	0.4144	1.5681	0.6294	-0.9804	0.3618
LOLL	0.6945	0.6777	-0.8754	0.4238	1.0538	0.6384	-0.8029	0.3709
LLOO	1.0003	0.6729	-1.1012	0.4150	1.4802	0.6380	-0.9415	0.3629
LOLO	0.6825	0.6819	-0.8610	0.4249	1.0096	0.6485	-0.7793	0.3728
LOOL	1.0205	0.6699	-1.3163	0.4208	1.5915	0.6293	-1.1272	0.3666
LOOO	1.0088	0.6740	-1.2895	0.4219	1.5415	0.6392	-1.0965	0.3686

Table 10 A: Coefficients based on different estimators. Model LLLL. All households observed twice								
	Cigarettes				Handrolling tobacco			
	OLS	FGLS	Between	Within	OLS	FGLS	Between	Within
Const	2.7254	3.6918	--	--	12.2885	14.1672	--	--
Texp	4.4514	3.0790	5.4267	1.9137	0.1957	0.4960	-0.0113	0.7200
Pcig	-0.3180	-0.5273	2.0425	-0.9775	2.6063	0.6107	6.0380	0.2097
Phan	-1.1152	0.2962	-3.9928	2.0402	-4.9177	-2.2907	-8.4542	-1.4507
dem1	-0.8053	-0.5489	-0.9036	0.6240	-0.2616	-0.0638	-0.3062	0.4517
dem2	-0.1265	0.2398	-0.3177	1.2096	1.1199	1.1262	1.1385	1.0176
dem3	0.2680	0.3498	0.1472	0.9527	1.9124	1.5954	2.0473	0.7813
dem4	0.1500	0.1395	0.1223	0.6722	0.6844	0.8203	0.6522	1.4767
Age	-0.2170	-0.2865	-0.1268	-3.2166	-1.0719	-1.2413	-0.8988	-1.1640
Coho	0.4278	0.2191	0.5580	-3.1062	-0.5862	-0.9058	-0.3537	-1.4583
Gend	1.0965	0.3056	1.4081	-1.6830	-1.1033	-0.9040	-1.1986	-0.5019
Inac	0.4616	0.2701	0.5809	0.2681	0.9739	0.1785	1.3453	-0.4772
west	-0.3601	-0.4835	-0.3078	--	-0.0142	-0.0677	0.0003	--
cent	-1.0526	-1.1723	-0.9881	4.4441	0.5703	0.5803	0.5578	1.2815
nor	-1.3980	-1.4120	-1.3974	--	1.1051	1.1094	1.1003	--
rur	-2.6450	-3.0990	-2.3795	0.7350	0.3531	0.3254	0.3217	-0.4895
dens	-1.5736	-1.8460	-1.4226	1.5524	0.7047	0.7437	0.6723	1.0997

Table 10 B: Coefficients based on different estimators. Model LLLL. Smokers observed twice								
	Cigarettes				Handrolling tobacco			
	OLS	FGLS	Between	Within	OLS	FGLS	Between	Within
Const	-2.0146	-0.9511	--	--	18.2089	19.5033	--	--
Texp	7.2254	5.0593	8.9013	2.9858	-0.0444	0.7170	-0.5315	1.0672
Pcig	-1.2188	-0.8787	2.8368	-1.6024	6.6339	3.4269	11.2747	0.7405
Phan	-0.2504	0.6929	-4.7734	3.3623	-8.7523	-5.7099	-13.1472	-3.0777
dem1	-0.7932	-0.6374	-0.8831	0.9653	-0.0479	0.0481	-0.0873	0.7062
dem2	-0.8522	-0.2492	-1.2192	1.8544	0.9465	0.9935	0.9339	1.5895
dem3	-0.6992	-0.2111	-1.0408	1.6534	2.1270	1.8495	2.3010	1.3652
dem4	-0.7518	-0.5441	-0.9238	1.2782	0.8373	1.1249	0.7259	3.0043
Age	0.8453	0.6613	0.9854	-6.0093	-1.2611	-1.3587	-1.1077	-2.4092
Coho	1.1188	0.9866	1.2231	-5.7333	-1.0672	-1.2641	-0.8492	-2.8772
Gend	2.2888	0.8454	2.9432	-2.7961	-1.6478	-1.1879	-1.8092	-0.8300
Inac	0.0594	0.2818	0.0597	0.5817	0.9529	0.2164	1.4068	-0.6556
west	-0.7575	-0.9046	-0.6882	--	0.5833	0.5772	0.5967	--
cent	-2.2074	-2.4007	-2.0959	--	1.0227	1.0530	1.0059	--
nor	-3.0415	-3.0348	-3.0553	--	0.8863	0.8775	0.8922	--
rur	-4.9141	-5.5949	-4.4516	-1.2920	1.2870	1.3508	1.2187	-2.6235
dens	-3.0779	-3.5315	-2.8082	--	1.4200	1.5209	1.3649	--

Appendix. Proof of (3) and (4)

Let $\phi(\cdot)$ and $\Phi(\cdot)$ denote the marginal density function and the cumulative distribution function, respectively, of the standardized normal distribution. Let $\Psi(\cdot,\cdot)$ denote the cumulative distribution function of the standardized binormal distribution with coefficient of correlation ρ. In Rosenbaum (1961, p. 406) [see also Maddala (1983, p. 368)] it is shown that

$$\mathsf{E}(u_g|u_g > c, u_j > d) = \frac{\phi(c)[1-\Phi(d^*)] + \rho\phi(d)[1-\Phi(c^*)]}{P(u_g > c, u_j > d)}.$$

where u_g, u_j $(g, j = C, H)$ are the disturbances in (1) and

$$c^* = \frac{c - \rho d}{(1-\rho^2)^{\frac{1}{2}}}, \qquad d^* = \frac{d - \rho c}{(1-\rho^2)^{\frac{1}{2}}}.$$

In a similar way, we find

$$\text{(A.1)} \qquad \begin{aligned} \mathsf{E}(u_g|u_g < c, u_j < d) &= -\frac{\phi(c)\Phi(d^*) + \rho\phi(d)\Phi(c^*)}{P(u_g < c, u_j < d)}, \\ \mathsf{E}(u_g|u_g < c, u_j > d) &= -\frac{\phi(c)[1-\Phi(d^*)] - \rho\phi(d)\Phi(c^*)}{P(u_g < c, u_j > d)}, \\ \mathsf{E}(u_g|u_g > c, u_j < d) &= \frac{\phi(c)\Phi(d^*) - \rho\phi(d)[1-\Phi(c^*)]}{P(u_g > c, u_j < d)}, \\ & \qquad g, j = C, H;\ g \neq j. \end{aligned}$$

Since

$$\begin{aligned} P(u_g < c, u_j < d) &= \Psi(c,d), \\ P(u_g < c, u_j > d) &= \Phi(c) - \Psi(c,d), \qquad g, j = C, H;\ g \neq j, \\ P(u_g > c, u_j < d) &= \Phi(d) - \Psi(c,d), \end{aligned}$$

we then define

$$\text{(A.2)} \qquad \begin{aligned} \lambda(c) &= \mathsf{E}(u_g|u_g < c) = -\frac{\phi(c)}{\Phi(c)}, \\ \mu_{LL}(c,d;\rho) &= \mathsf{E}(u_g|u_g < c, u_j < d) = -\frac{\phi(c)\Phi(d^*) + \rho\phi(d)\Phi(c^*)}{\Psi(c,d)}, \\ \mu_{LG}(c,d;\rho) &= \mathsf{E}(u_g|u_g < c, u_j > d) = -\frac{\phi(c)[1-\Phi(d^*)] - \rho\phi(d)\Phi(c^*)}{\Phi(c) - \Psi(c,d)}, \\ \mu_{GL}(c,d;\rho) &= \mathsf{E}(u_g|u_g > c, u_j < d) = \frac{\phi(c)\Phi(d^*) - \rho\phi(d)[1-\Phi(c^*)]}{\Phi(d) - \Psi(c,d)}, \\ & \qquad g, j = C, H;\ g \neq j. \end{aligned}$$

The moment generating function of the truncated multinormal distribution is derived and discussed in Tallis (1961); see also Amemiya (1974, section 2).

From (1) and (2), using (A.2), we derive the following expresions for the conditional expectations of the latent expenditures with one threshold

$$\text{(A.3)} \qquad \begin{aligned} \mathsf{E}(y_C^*|x, y_C^* > 0) &= x\beta_C - \sigma_C \mathsf{E}(u_C|u_C < x\beta_C/\sigma_C) &= x\beta_C - \sigma_C\lambda_C, \\ \mathsf{E}(y_H^*|x, y_H^* > 0) &= x\beta_H - \sigma_H \mathsf{E}(u_H|u_H < x\beta_H/\sigma_H) &= x\beta_H - \sigma_H\lambda_H, \end{aligned}$$

where

$$\lambda_g = \lambda(x\beta_g/\sigma_g), \qquad g = C, H,$$

and with two thresholds

$$\text{(A.4)} \qquad \begin{aligned} \mathsf{E}(y_C^*|x, y_C^* > 0, y_H^* > 0) &= x\beta_C - \sigma_C \mathsf{E}(u_C|u_C < x\beta_C/\sigma_C, u_H < x\beta_H/\sigma_H) \\ &= x\beta_C - \sigma_C\mu_{LLCH}, \\ \mathsf{E}(y_H^*|x, y_H^* > 0, y_C^* > 0) &= x\beta_H - \sigma_H \mathsf{E}(u_H|u_C < x\beta_C/\sigma_C, u_H < x\beta_H/\sigma_H) \\ &= x\beta_H - \sigma_H\mu_{LLHC}, \\ \mathsf{E}(y_C^*|x, y_C^* > 0, y_H^* < 0) &= x\beta_C - \sigma_C \mathsf{E}(u_C|u_C < x\beta_C/\sigma_C, u_H > x\beta_H/\sigma_H) \\ &= x\beta_C - \sigma_C\mu_{LGCH}, \\ \mathsf{E}(y_H^*|x, y_H^* > 0, y_C^* < 0) &= x\beta_H - \sigma_H \mathsf{E}(u_H|u_C > x\beta_C/\sigma_C, u_H < x\beta_H/\sigma_H) \\ &= x\beta_H - \sigma_H\mu_{LGHC}, \end{aligned}$$

where

$$\mu_{LLgj} = \mu_{LL}(x\beta_g/\sigma_g, x\beta_j/\sigma_j, \rho), \quad \mu_{LGgj} = \mu_{LG}(x\beta_g/\sigma_g, x\beta_j/\sigma_j, \rho), \qquad g, j = C, H;\ g \neq j.$$

Since $y_g = \max[y_g^*, 0]$, $(g = C, H)$ [cf. (2)], (A.3) and (A.4) can be restated in terms of the observed expenditures as

$$\text{(A.5)} \qquad \begin{aligned} \mathsf{E}(y_C|x, y_C > 0) &= x\beta_C - \sigma_C\lambda_{LC}, \\ \mathsf{E}(y_H|x, y_H > 0) &= x\beta_H - \sigma_H\lambda_{LH}, \end{aligned}$$

and

$$\text{(A.6)} \qquad \begin{aligned} \mathsf{E}(y_C|x, y_C > 0, y_H > 0) &= x\beta_C - \sigma_C\mu_{LLCH}, \\ \mathsf{E}(y_H|x, y_H > 0, y_C > 0) &= x\beta_H - \sigma_H\mu_{LLHC}, \\ \mathsf{E}(y_C|x, y_C > 0, y_H = 0) &= x\beta_C - \sigma_C\mu_{LGCH}, \\ \mathsf{E}(y_H|x, y_H > 0, y_C = 0) &= x\beta_H - \sigma_H\mu_{LGHC}. \end{aligned}$$

Using (A.2), (A.5), and the law of iterated expectations, we find that

$$\text{(A.7)} \qquad \begin{aligned} \mathsf{E}(y_C|x) &= \mathsf{E}(y_C|x, y_C > 0)\Phi_C + 0(1 - \Phi_C) &= x\beta_C\Phi_C + \sigma_C\phi_C, \\ \mathsf{E}(y_H|x) &= \mathsf{E}(y_H|x, y_H > 0)\Phi_H + 0(1 - \Phi_H) &= x\beta_H\Phi_H + \sigma_H\phi_H, \end{aligned}$$

where $\Phi_g = \Phi(x\beta_g/\sigma_g)$, $\phi_g = \phi(x\beta_g/\sigma_g)$ $(g = C, H)$, Φ_C and Φ_H being the marginal smoking probabilities for cigarettes and handrolling tobacco, respectively. An alternative way of deriving (A.7) is to use (A.2), (A.6), and the law of iterated expectations. This gives

$$\text{(A.8)}\qquad \begin{aligned} \mathsf{E}(y_C|x) &= \mathsf{E}(y_C|x, y_C>0, y_H>0)\Psi_{CH} \\ &\quad +\mathsf{E}(y_C|x, y_C>0, y_H=0)(\Phi_C-\Psi_{CH})+0(1-\Phi_C) \\ &= x\beta_C\Phi_C+\sigma_C\phi_C, \\ \mathsf{E}(y_H|x) &= \mathsf{E}(y_H|x, y_H>0, y_C>0)\Psi_{CH} \\ &\quad +\mathsf{E}(y_H|x, y_H>0, y_C=0)(\Phi_H-\Psi_{CH})+0(1-\Phi_H) \\ &= x\beta_H\Phi_H+\sigma_H\phi_H, \end{aligned}$$

which completes the proof of (3).

The probability of being a smoker, *i.e.*, of consuming at least one tobacco commodity can be written as

$$\text{(A.9)}\qquad \begin{aligned} \Phi_S &= P(y_C+y_H>0) \\ &= (\Phi_C-\Psi_{CH})+(\Phi_H-\Psi_{CH})+\Psi_{CH} \quad = \quad \Phi_C+\Phi_H-\Psi_{CH}, \end{aligned}$$

where Ψ_{CH}, $\Phi_C-\Psi_{CH}$, and $\Phi_H-\Psi_{CH}$ are the probabilities of smoking, respectively, both commodities, cigarettes only, and handrolling tobacco only. Now, the law of iterated expectations implies

$$\text{(A.10)}\qquad \begin{aligned} \mathsf{E}(y_C|x) &= \mathsf{E}(y_C|x, y_C+y_H>0)\Phi_S+0(1-\Phi_S) &= \mathsf{E}(y_C|x, y_C+y_H>0)\Phi_S, \\ \mathsf{E}(y_H|x) &= \mathsf{E}(y_H|x, y_C+y_H>0)\Phi_S+0(1-\Phi_S) &= \mathsf{E}(y_H|x, y_C+y_H>0)\Phi_S. \end{aligned}$$

From (A.7) and (A.10) we get

$$\text{(A.11)}\qquad \begin{aligned} \mathsf{E}(y_C|x, y_C+y_H>0) &= x\beta_C\frac{\Phi_C}{\Phi_S}+\sigma_C\frac{\phi_C}{\Phi_S}, \\ \mathsf{E}(y_H|x, y_C+y_H>0) &= x\beta_H\frac{\Phi_H}{\Phi_S}+\sigma_H\frac{\phi_H}{\Phi_S}, \end{aligned}$$

where Φ_S is given by (A.9). This completes the proof of (4).

References

Amemiya, T. (1974): Multivariate Regression and Simultaneous Equation Models when the Dependent Variables are Truncated Normal. *Econometrica*, 42, 999 – 1012.

Amemiya, T. (1984): Tobit Models: A Survey. *Journal of Econometrics*, 24, 3 – 61.

Amemiya, T. (1985): *Advanced Econometrics.* Cambridge, Mass.: Harvard University Press.

Avery, R.B. (1977): Error Components and Seemingly Unrelated Regressions. *Econometrica*, 45, 199 – 209.

Baltagi, B.H. (1980): On Seemingly Unrelated Regressions with Error Components. *Econometrica*, 48, 1547 – 1551.

Baltagi, B.H., and Li, Q. (1992): A Monotonic Property for Iterative GLS in the Two-Way Random Effects Model. *Journal of Econometrics*, 53, 45 – 51.

Becker, G.S., and Murphy, K.M. (1988): A Theory of Rational Addiction. *Journal of Political Economy*, 96, 675 – 700.

Biørn, E. (1994): Moment Estimators and the Estimation of Marginal Budget Shares from Household Panel Data. *Structural Change and Economic Dynamics*, 5, 133 – 154.

Biørn, E. (1999): Estimating Regression Systems from Unbalanced Panel Data: A Stepwise Maximum Likelihood Procedure. Department of Economics, University of Oslo, Memorandum No. 20/1999.

Biørn, E., and Jansen, E.S. (1982): *Econometrics of Incomplete Cross-Section/Time-Series Data: Consumer Demand in Norwegian Households 1975 – 1977.* Oslo: Statistics Norway.

Blundell, R., and Meghir, C. (1987): Bivariate Alternatives to the Tobit Model. *Journal of Econometrics*, 34, 179 – 200.

Breusch, T.S. (1987): Maximum Likelihood Estimation of Random Effects Models. *Journal of Econometrics*, 36, 383 – 389.

Chaloupka, F. (1991): Rational Addictive Behavior and Cigarette Smoking. *Journal of Political Economy*, 99, 722 – 742.

Chaloupka, F., and Warner, K.E. (1999): The Economics of Smoking. NBER Working paper No. 7047.

Cragg, J.G. (1971): Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durables Goods. *Econometrica*, 39, 829 – 844.

Deaton, A. (1990): Price Elasticities from Survey Data. Extensions and Indonesian Results. *Journal of Econometrics*, 44, 281 – 309.

Deaton, A., and Irish, M. (1984): Statistical Models for Zero Expenditures in Household Budgets. *Journal of Public Economics*, 23, 59 – 80.

Godfrey, L.G. (1988): *Misspecification Tests in Econometrics.* Cambridge: Cambridge University Press.

Goldberger, A.S. (1981): Linear Regression After Selection. *Journal of Econometrics*, 15, 357 – 366.

Greene, W.H. (1981): On the Asymptotic Bias of the Ordinary Least Squares Estimator of the Tobit Model. *Econometrica*, 49, 505 – 513.

Heckman, J.J. (1976): The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. *Annals of Economic and Social Measurement*, 5, 475 – 492.

Heckman, J.J. (1990): Varieties of Selection Bias. *American Economic Review*, 80 (Papers and Proceedings), 313 – 318.

Hsiao, C. (1986): *Analysis of Panel Data.* Cambridge: Cambridge University Press.

Jones, A.M. (1989): A Double Hurdle Model of Cigarette Consumption. *Journal of Applied Econometrics*, 4, 23 – 39.

Keen, M. (1986): Zero Expenditures and the Estimation of Engel Curves. *Journal of Applied Econometrics*, 1, 277 – 286.

Krishnakumar, J. (1996): Simultaneous Equations. Chapter 9 in: Mátyás, L. and Sevestre, P. (eds.): *The Econometrics of Panel Data. Handbook of the Theory with Applications.* Dordrecht: Kluwer.

Labeaga, J.M. (1999): A Double-Hurdle Rational Addiction Model with Heterogeneity: Estimating the Demand for Tobacco. *Journal of Econometrics*, 93, 49 – 72.

Labeaga, J.M., Preston, I., and Sanchis-Llopis, J.A. (1998): Children and Demand Patterns: A Panel Data Approach. Paper presented at the Eighth International Conference on Panel Data, Göteborg, June 1998.

Leung, S.F., and Yu, S. (1996): On the Choice Between Sample Selection and Two-Part Models. *Journal of Econometrics*, 72, 197 – 229.

Lluch, C. (1974): Expenditure, Savings and Habit Formation. *International Economic Review*, 15, 786 – 97.

Maddala, G.S. (1983): *Limited-Dependent and Qualitative Variables in Econometrics.* Cambridge: Cambridge University Press, 1983.

Muellbauer, J. (1988): Habits, Rationality and Myopia in the Life Cycle Consumption Function. *Annales d'économie et de statistique*, No. 9, 47 – 70.

Oberhofer, W., and Kmenta, J. (1974): A General Procedure for Obtaining Maximum Likelihood Estimates in Generalized Regression Models. *Econometrica*, 42, 579 – 590.

Olsen, R.J. (1980): Approximating a Truncated Normal Regression with the Method of Moments. *Econometrica*, 48, 1099 – 1105.

Powell, J.L. (1984): Least Absolute Deviation Estimation for the Censored Regression Model. *Journal of Econometrics*, 25, 303 – 325.

Rosenbaum, S. (1961): Moments of a Truncated Bivariate Normal Distribution. *Journal of the Royal Statistical Society* (Series B), 23, 405 – 408.

Tallis, G.M. (1961): The Moment Generating Function of the Truncated Multi-Normal Distribution. *Journal of the Royal Statistical Society* (Series B), 23, 223 – 229.

USDHHS (1989): Reducing the Health Consequences of Smoking. 25 Years of Progress. A Report for the Surgeon General, 1989. US Departement of Health and Human Services. Public Health Service, Centers for Disease Control, Centers for Chronic Disease Prevention and Health Promotion. Office on Smoking and Health. DHHS publication No. (CDC) 89-8411, 1989.

Wangen, K.R., and Aasness, J. (2001): Demand for Cigarettes and Handrolling Tobacco – A Time Series Analysis of Cross Section Elasticities. Under preparation, Statistics Norway.

Wangen, K.R., and Biørn, E. (2001): Prevalence and Composition of Tobacco Consumption – A Discrete Choice Analysis Based on Panel Data. Under preparation, Statistics Norway.

Recent publications in the series Discussion Papers

203 J. Aasness and L. Belsby (1997): Estimation of Time Series of Latent Variables in an Accounting System: Petrol Consumption of Norwegian Households 1973-1995

204 A. Rygh Swensen (1997): Change in Regime and Markov Models

205 K. Nyborg and I. Spangen (1997): Cost-Benefit Analysis and the Democratic Ideal

206 L. Belsby and J.F. Bjørnstad (1997): Modelling and Estimation Methods for Household Size in the Presence of Nonresponse: Applied to The Norwegian Consumer Expenditure Survey

207 K.O. Aarbu and T.O. Thoresen (1997): The Norwegian Tax Reform; Distributional Effects and the High-Income Response

208 T. Hægeland and T.J. Klette (1997): Do Higher Wages Reflect Higher Productivity? Education, Gender and Experience Premiums in a Matched Plant-Worker Data Set

209 J. Gjerde, S. Grepperud and S. Kverndokk (1998): Optimate Climate Policy under the Possibility of a Catastrophe

210 T. Eika and K.A. Magnussen (1998): Did Norway Gain from the 1979-85 Oil Price Shock?

211 K.O. Aarbu and J.K. MacKie-Mason (1998): Why Some Corporations Pay More Tax than Necessary

212 R. Aaberge (1998): UMP Unbiased Tests for Multiparameter Testing Problems with Restricted Alternatives

213 M. Søberg (1998): "EPA's New Emissions Trading Mechanism: A Laboratory Evaluation" – A Comment

214 K. Nyborg (1998): Non-Verifiable Emissions, Voluntary Agreements, and Emission Taxes

215 H. C. Bjørnland (1998): Economic Fluctuations in a Small Open Economy - Real versus Nominal Stocks

216 L.C. Zhang (1998): Post-Stratification and Calibration – A Synthesis

217 R. Aaberge and Y. Zhu: The Pattern of Household Savings during a Hyperinflation. The Case of Urban China in the Late 1980s

218 L.C. Zhang (1998): A Note on Post-stratification when Analyzing Binary Survey Data Subject to Non-response

219 T.J. Klette and J. Møen (1998): From Growth Theory to Technology Policy – Coordination Problems in Theory and Practice

220 L.C. Zhang (1998): Synthetic Graphical Modelling of Binary Panel Data Subject to Non-response. A Case Study

221 J.K. Dagsvik (1998): Choice among Lotteries when Preferences are Stochastic

222 J.K. Dagsvik (1998): Nonparametric Identification of Discrete Choice Models

223 L. Brubakk and J.K. Dagsvik (1998): Consumer Demand and Unobservable Product Attributes

224 E. Holmøy (1998): Potential Aggregate Gains from Improved Sectoral Efficiency: General Equilibrium Estimates for Norway

225 B.H. Vatne and J.K. Dagsvik (1998): Estimation of Generalized Extreme Value Models by a Max-spectral Representation

226 I. Svendsen (1998): Rational Expectations in Price Setting – Tests Based on Norwegian Export Prices

227 K.A. Brekke and R.B. Howarth (1998): The Social Contingency of Wants: Implications for Growth and the Environment

228 K.A. Brekke and E. Moxnes (1998): Do Models Improve Fishery Management? Empirical Evidence from a Experimental Study

229 J.K. Dagsvik, Yu Zhu and R. Aaberge (1998): A Framework for Empirical Modelling of Consumer Demand with Latent Quality Attributes

230 R. Aaberge, U. Colombino and S. Strøm (1998): Social Evaluation of Individual Welfare Effects from Income Taxation: Empirical Evidence Based on Italian Data for Married Couples

231 R. Nesbakken (1998): Residential Energy Consumption for Space Heating in Norwegian Households. A Discrete-Continuous Choice Approach

232 R. Nesbakken (1998): Price Sensitivity of Residential Energy Consumption in Norway

233 M. Søberg (1998): Uncertainty and International Negotiations on Tradable Quota Treaties

234 J.K. Dagsvik and L. Brubakk: Price Indexes for Elementary Aggregates Derived from Behavioral Assumptions

235 E. Biørn, K.-G. Lindquist and T. Skjerpen (1998): Random Coefficients and Unbalanced Panels: An Application on Data from Norwegian Chemical Plants

236 K. Ibenholt (1998): Material Accounting in a Macroeconomic Framework. Forecast of waste generated in manufacturing industries in Norway

237 K-G. Lindquist (1998): The Response by the Norwegian Aluminium Industry to Changing Market Structure

238 J.K. Dagsvik, A.S. Flaatten and H. Brunborg: A Behavioral Two-Sex Model

239 K.A. Brekke, R.B. Howarth and K. Nyborg (1998): Are there Social Limits to Growth?

240 R. B. Howarth and K. A. Brekke (1998): Status Preferences and Economic Growth

241 H. Medin, K. Nyborg and I. Bateman (1998): The Assumption of Equal Marginal Utility of Income: How Much Does it Matter?

242 B. Bye (1998): Labour Market Rigidities and Environmental Tax Reforms: Welfare Effects of Different Regimes

243 B.E. Naug (1999): Modelling the Demand for Imports and Domestic Output

244 J. Sexton and A. R. Swensen (1999): ECM-algorithms that converge at the rate of EM

245 E. Berg, S. Kverndokk and K.E. Rosendahl (1999): Optimal Oil Exploration under Climate Treaties

246 J.K. Dagsvik and B.H. Vatne (1999): Is the Distribution of Income Compatible with a Stable Distribution?

247 R. Johansen and J.K. Dagsvik (1999): The Dynamics of a Behavioral Two-Sex Demographic Model

248 M. Søberg (1999): Asymmetric information and international tradable quota treaties. An experimental evaluation

249 S. Grepperud, H. Wiig and F.A. Aune (1999): Maize Trade Liberalization vs. Fertilizer Subsidies in Tanzania: A CGE Model Analysis with Endogenous Soil Fertility

250 K.A. Brekke and Nils Chr. Stenseth (1999): A Bio-Economic Approach to the study of Pastoralism, Famine and Cycles. Changes in ecological dynamics resulting from changes in socio-political factors

251 T. Fæhn and E. Holmøy (1999): Welfare Effects of Trade Liberalisation in Distorted Economies. A Dynamic General Equilibrium Assessment for Norway

252 R. Aaberge (1999): Sampling Errors and Cross-Country Comparisons of Income Inequality

253 I. Svendsen (1999): Female labour participation rates in Norway – trends and cycles

254 A. Langørgen and R. Aaberge: A Structural Approach for Measuring Fiscal Disparities

255 B. Halvorsen and B.M. Larsen (1999): Changes in the Pattern of Household Electricity Demand over Time

256 P. Boug (1999): The Demand for Labour and the Lucas Critique. Evidence from Norwegian Manufacturing

257 M. Rege (1999): Social Norms and Private Provision of Public Goods: Endogenous Peer Groups

258 L. Lindholt (1999): Beyond Kyoto: CO_2 permit prices and the markets for fossil fuels

259 R. Bjørnstad and R. Nymoen (1999): Wage and Profitability: Norwegian Manufacturing 1967-1998

260 T.O. Thoresen and K.O. Aarbu (1999): Income Responses to Tax Changes – Evidence from the Norwegian Tax Reform

261 B. Bye and K. Nyborg (1999): The Welfare Effects of Carbon Policies: Grandfathered Quotas versus Differentiated Taxes

262 T. Kornstad and T.O. Thoresen (1999): Means-testing the Child Benefit

263 M. Rønsen and M. Sundström (1999): Public Policies and the Employment Dynamics among new Mothers – A Comparison of Finland, Norway and Sweden

264 J.K. Dagsvik (2000): Multinomial Choice and Selectivity

265 Y. Li (2000): Modeling the Choice of Working when the Set of Job Opportunities is Latent

266 E. Holmøy and T. Hægeland (2000): Aggregate Productivity and Heterogeneous Firms

267 S. Kverndokk, L. Lindholt and K.E. Rosendahl (2000): Stabilisation of CO_2 concentrations: Mitigation scenarios using the Petro model

268 E. Biørn, K-G. Lindquist and T. Skjerpen (2000): Micro Data On Capital Inputs: Attempts to Reconcile Stock and Flow Information

269 I. Aslaksen and C. Koren (2000): Child Care in the Welfare State. A critique of the Rosen model

270 R. Bjørnstad (2000): The Effect of Skill Mismatch on Wages in a small open Economy with Centralized Wage Setting: The Norwegian Case

271 R. Aaberge (2000): Ranking Intersecting Lorenz Curves

272 J.E. Roemer, R. Aaberge , U. Colombino, J, Fritzell, S.P. Jenkins, I. Marx, M. Page, E. Pommer, J. Ruiz-Castillo, M. Jesus SanSegundo, T. Tranaes, G.G.Wagner and I. Zubiri (2000): To what Extent do Fiscal Regimes Equalize Opportunities for Income Acquisition Among citizens?

273 I. Thomsen and L.-C. Zhang (2000): The Effect of Using Administrative Registers in Economic Short Term Statistics: The Norwegian Labour Force Survey as a Case Study

274 I. Thomsen, L.-C. Zhang and J. Sexton (2000): Markov Chain Generated Profile Likelihood Inference under Generalized Proportional to Size Non-ignorable Non-response

275 A. Bruvoll and H. Medin (2000): Factoring the environmental Kuznets curve. Evidence from Norway

276 I. Aslaksen, T. Wennemo and R. Aaberge (2000): "Birds of a feather flock together". The Impact of Choice of Spouse on Family Labor Income Inequality

277 I. Aslaksen and K.A. Brekke (2000): Valuation of Social Capital and Environmental Externalities

278 H. Dale-Olsen and D. Rønningen (2000): The Importance of Definitions of Data and Observation Frequencies for Job and Worker Flows - Norwegian Experiences 1996-1997

279 K. Nyborg and M. Rege (2000): The Evolution of Considerate Smoking Behavior

280 M. Søberg (2000): Imperfect competition, sequential auctions, and emissions trading: An experimental evaluation

281 L. Lindholt (2000): On Natural Resource Rent and the Wealth of a Nation. A Study Based on National Accounts in Norway 1930-95

282 M. Rege (2000): Networking Strategy: Cooperate Today in Order to Meet a Cooperator Tomorrow

283 P. Boug, Å. Cappelen and A.R. Swensen (2000): Expectations in Export Price Formation: Tests using Cointegrated VAR Models

284 E. Fjærli and R. Aaberge (2000): Tax Reforms, Dividend Policy and Trends in Income Inequality: Empirical Evidence based on Norwegian Data

285 L.-C. Zhang (2000): On dispersion preserving estimation of the mean of a binary variable from small areas

286 F.R. Aune, T. Bye and T.A. Johnsen (2000): Gas power generation in Norway: Good or bad for the climate? Revised version

287 A. Benedictow (2000): An Econometric Analysis of Exports of Metals: Product Differentiation and Limited Output Capacity

288 A. Langørgen (2000): Revealed Standards for Distributing Public Home-Care on Clients

289 T. Skjerpen and A.R. Swensen (2000): Testing for long-run homogeneity in the Linear Almost Ideal Demand System. An application on Norwegian quarterly data for non-durables

290 K.A. Brekke, S. Kverndokk and K. Nyborg (2000): An Economic Model of Moral Motivation

291 A. Raknerud and R. Golombek: Exit Dynamics with Rational Expectations

292 E. Biørn, K-G. Lindquist and T. Skjerpen (2000): Heterogeneity in Returns to Scale: A Random Coefficient Analysis with Unbalanced Panel Data

293 K-G. Lindquist and T. Skjerpen (2000): Explaining the change in skill structure of labour demand in Norwegian manufacturing

294 K. R. Wangen and E. Biørn (2001): Individual Hetero-geneity and Price Responses in Tobacco Consumption: A Two-Commodity Analysis of Unbalanced Panel Data