Discussion Papers No. 532, March 2008
Statistics Norway, Research Department

Terje Skjerpen

Engel elasticities, pseudo-maximum likelihood estimation and bootstrapped standard errors

A case study

Abstract:
Estimation of standard errors of Engel elasticities within the framework of a linear structural model formulated on two-wave panel data is considered. The complete demand system is characterized by measurement errors in total expenditure and by latent preference variation. The estimation of the parameters as well as the standard errors of the estimates is based on the assumption that the variables are normally distributed. Considering a concrete case it is demonstrated that normality does not hold as a maintained assumption. In the light of this standard errors are estimated by means of bootstrapping. However, one obtains rather similar estimates of the standard errors of the Engel elasticities no matter whether one sticks to classical normal inference or perform non-parametric bootstrapping.

Keywords: Engel elasticities, standard errors, classical normal theory, bootstrapping

JEL classification: C13; C14; C15; C33; D12

Acknowledgement: I thank Erik Biørn for useful comments. Any remaing errors or shortcomings are my own.

Address: Terje Skjerpen, Statistics Norway, Research Department. E-mail: terje.skjerpen@ssb.no

Discussion Papers comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:
http://www.ssb.no
http://ideas.repec.org/s/ssb/dispap.html

For printed Discussion Papers contact:

Statistics Norway
Sales- and subscription service
NO-2225 Kongsvinger

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

Introduction

Maximum likelihood estimators of parameters in models based on the normality of disturbances assumption often retain the consistency property even if normality does not hold, but statistical inference may be influenced. In this paper this issue is considered in a concrete case, namely by considering the analysis of Aasness *et al*. (1993), who applied a structural equation modeling framework in the analysis of a consumer demand system.[1] Their inference was based on classical normal theory. The focus is on Engel elasticities within a linear consumer demand system consisting of five commodities, characterized by measurement errors in total expenditure and the presence of unobserved preference heterogeneity. We compare standard errors of Engel elasticity estimates calculated in two different ways. The first one, which builds upon the assumption of normality, employs the delta method, whereas the second one is based upon bootstrapping.[2] The main conclusion is that there is only a very modest deviation in the estimated standard errors when using the two approaches.

The rest of the paper is organized in the following way. In Section 1 we provide notation and give the specification of the econometric model. Section 2 gives a very short description of the data. Besides, results related to univariate normality tests are reported. Section 3 is devoted to estimation issues. In Section 4 we compare the estimated standard errors obtained by using different approaches. Some concluding remarks are offered in Section 5.

1. Modeling framework

Consider the following complete system of linear Engel curves specified for panel data with 2 replications, 5 commodities and 408 households

(1) $$\eta_{ith} = a_{it} + b_i \xi_{th} + c_{i1} z_{1h} + c_{i2} z_{2h} + \mu_{ih};\ i = 1,...,5;\ t = 1,2;\ t = 1,2; h = 1,...,408.$$

In (1) η_{ith} and ξ_{th} denote, respectively, latent expenditure at constant prices on consumption category *i* and total latent expenditure in period *t* by household *h*. To account for demographic effects we add two time-invariant observed demographic variables, namely z_{1h} and z_{2h} which represent the number

---

[1] All their calculations were conducted using the software program LISREL. For a later version than the one used by Aasness *et al.* (1993) cf. Jöreskog and Sörbom (1996).

[2] For general description of bootstrapping cf. Efron and Tibshirani (1993), Hall (1994) and Horowitz (2001). For bootstrapping within the framework of structural equation models cf. Stine (1990) and Yung and Bentler (1998).

of children and number of adults, respectively. The variable μ_{ih} captures unobserved household specific heterogeneity and may be associated with commodity specific preferences attached to commodity *i* by household *h*. Lastly, time-specific intercepts are allowed for.

The adding-up conditions, $\sum_{i=1}^{5}\eta_{ith} = \xi_{th}\ (t = 1,2; h = 1,...,408)$, imply that $\sum_{i=1}^{5}\mu_{ih} = 0\,(h = 1,...,408)$ and the following parameter restrictions

$$\sum_{i=1}^{5} a_{it} = 0; \sum_{i=1}^{5} b_i = 1; \sum_{i=1}^{5} c_{ij} = 0;\ t = 1,2;\ j = 1,2.$$

For later use we define the following symbols

$$a_t = \left[a_{1t}, a_{2t}, a_{3t}, a_{4t}\right]',\ t = 1,2,$$

$$b = \left[b_1, b_2, b_3, b_4\right]',$$

$$c_j = \left[c_{1j}, c_{2j}, c_{3j}, c_{4j}\right]',\ j = 1,2,$$

$$\mu_h = \left[\mu_{1h}, \mu_{2h}, \mu_{3h}, \mu_{4h}\right]',\ h = 1,...,408.$$

The first- and second-order moments of μ_h are given by

$$E(\mu_h) = 0$$

and

$$E(\mu_h \mu_h') = \Sigma_{\mu\mu} = \begin{bmatrix} \sigma^2_{\mu_1\mu_1} & & & \\ \sigma^2_{\mu_2\mu_1} & \sigma^2_{\mu_2\mu_2} & & \\ \sigma^2_{\mu_3\mu_1} & \sigma^2_{\mu_3\mu_2} & \sigma^2_{\mu_3\mu_3} & \\ \sigma^2_{\mu_4\mu_1} & \sigma^2_{\mu_4\mu_2} & \sigma^2_{\mu_4\mu_3} & \sigma^2_{\mu_4\mu_4} \end{bmatrix}.$$

We have the following measurement error model that links the observed expenditures of the consumption categories to their latent counterparts

(2) $$y_{ith} = \eta_{ith} + v_{ith}; i = 1,...,5; t = 1,2; h = 1,...,408,$$

where y_{ith} represents observed expenditure on consumption category *i* in period *t* by household *h*, and where v_{ith} is interpreted as measurement error. Let

$$v_{th} = \left[v_{1th}, v_{2th}, v_{3th}, v_{4th}, v_{5th}\right]', \; t = 1,2,$$

and

$$v_h = \left[v_{1h}', v_{2h}'\right]'.$$

The first- and second-order moments of this vector are given by

$$Ev_h = 0$$

and

$$E\left(v_h v_h'\right) = I_2 \otimes Diag\left[\sigma_{v_1v_1}, \sigma_{v_2v_2}, \sigma_{v_3v_3}, \sigma_{v_4v_4}, \sigma_{v_5v_5}\right].$$

Thus measurement errors in different periods are assumed to be uncorrelated, and so are the measurement errors of the different consumption categories in the same period. Besides the measurement errors are assumed to be homoskedastic.

From adding-up it follows that

$$x_{th} = \sum_{i=1}^{5} y_{ith} = \xi_{th} + \sum_{i=1}^{5} v_{ith} : \; t = 1,2, \; h = 1,...,408.$$

The total latent expenditure variable in the two periods, i.e., ξ_{1h} and ξ_{2h}, are specified as

$$\xi_{1h} = \chi_h + u_{1h}$$

and

$$\xi_{2h} = q_{02} + q_2\left(\chi_h + u_{2h}\right).$$

We refer to χ_h as the permanent component of total latent expenditure and u_{th} (t = 1,2) as the volatile component of total latent expenditure.

For later use it is convenient to introduce the vector

$$\kappa_h = \left[\chi_h, z_{1h}, z_{2h}, u_{1h}\ u_{2h}\right]'.$$

We assume that the two first order moments of κ_h are given by

$$E\kappa_h = \mu_\kappa = \left[\mu_\chi, \mu_{z_1}, \mu_{z_2}, 0, 0\right]'$$

and

$$\Sigma_{\kappa\kappa} = E\left[(\kappa_h - \mu_\kappa)(\kappa_h - \mu_\kappa)'\right] = \begin{bmatrix} \sigma^2_{\chi\chi} & & & & \\ \sigma^2_{z_1\chi} & \sigma^2_{z_1z_1} & & & \\ \sigma^2_{z_2\chi} & \sigma^2_{z_2z_1} & \sigma^2_{z_2z_2} & & \\ 0 & 0 & 0 & \sigma^2_{uu} & \\ 0 & 0 & 0 & 0 & \sigma^2_{uu} \end{bmatrix}.$$

We also define

$$\Sigma^a_{\kappa\kappa} = \begin{bmatrix} \sigma^2_{\kappa\kappa} & & \\ \sigma^2_{z_1\kappa} & \sigma^2_{z_1z_1} & \\ \sigma^2_{z_2\kappa} & \sigma^2_{z_2z_1} & \sigma^2_{z_2z_2} \end{bmatrix}$$

for later use.

The observed purchases of the different commodities can be viewed as indicators of latent total expenditure. In addition, two additional indicator variables of total latent expenditure are utilized

(3) $$w_{kth} = d_{kt} + e_k\xi_{th} + f_{k1}z_{1h} + c_{k2}z_{2h} + \lambda_{kh} + \varepsilon_{kth};\ k = 1,2; t = 1,2; h = 1,...,408.$$

In (3) w_{kth} denotes the observed value of income measure k (k=1,2) in period t of household h.

As in the consumer demand system we allow for time-specific intercepts. Latent total expenditure and the number of adults occur on the right hand side of (3). The latent variable λ_{kh} takes care of unobserved heterogeneity across households and ε_{kth} is a genuine error term.

We introduce some more notation

$$d_t = \left[d_{1t}, d_{2t}\right]',\ t = 1,2,$$

$$e = \left[e_1, e_2\right]',$$

$$f_k = \left[f_{k1}, f_{k2}\right]',\ k = 1,2,$$

$$F = \left[f_1\ f_2\right],$$

$$\lambda_h = \left[\lambda_{h1}, \lambda_{h2}\right]',$$

$$\varepsilon_{th} = \left[\varepsilon_{1th}, \varepsilon_{2th}\right]^{/}, t = 1,2.$$

$$\varepsilon_h = \left[\varepsilon_{1h}^{/}, \varepsilon_{2h}^{/}\right]^{/}.$$

The first- and second-order moments of λ_h are given by

$$E(\lambda_h) = 0$$

and

$$E\left(\lambda_h \lambda_h^{/}\right) = \Sigma_{\lambda\lambda} = \begin{bmatrix} \sigma^2_{\lambda_1\lambda_1} & \\ \sigma^2_{\lambda_2\lambda_1} & \sigma^2_{\lambda_2\lambda_2} \end{bmatrix},$$

whereas the first and second order moments of ε_h are given by

$$E(\varepsilon_h) = 0$$

and

$$E\left(\varepsilon_h \varepsilon_h^{/}\right) = I_2 \otimes \Sigma_{\varepsilon\varepsilon},$$

where

$$\Sigma_{\varepsilon\varepsilon} = \begin{bmatrix} \sigma^2_{\varepsilon_1\varepsilon_1} & \\ \sigma^2_{\varepsilon_2\varepsilon_1} & \sigma^2_{\varepsilon_2\varepsilon_2} \end{bmatrix}.$$

Let us define the following observation vectors

$$y_{th} = \left[y_{1th}, y_{2th}, y_{3th}, y_{4th}, y_{5th}\right]^{/}, t = 1,2,$$

$$w_{th} = \left[w_{1th}, w_{1th}\right]^{/}, t = 1,2,$$

$$z_h = \left[z_{1h}, z_{2h}\right]^{/},$$

$$y_h = \left[y_{1h}^{/}, y_{2h}^{/}\right]^{/},$$

$$w_h = \left[w_{1h}^{/}\ w_{2h}^{/}\right]^{/},$$

$$y_h^{*} = \left[y_h^{/}, w_h^{/}, z_h^{/}\right]^{/}.$$

Formally we may now write the vector equation for the whole observation vector of household h as

(4) $$y_h^* = \mu_{y^*} + B_\kappa(\kappa_h - \mu_\kappa) + B_\mu \mu_h + B_\nu \nu_h + B_\lambda \lambda_h + B_\varepsilon \varepsilon_h.$$

The explicit expressions for the vector μ_y^* and the matrices $B_\kappa, B_\mu, B_\nu, B_\lambda, B_\varepsilon$ are reported in Appendix A. We assume that the vectors $(\kappa_h - \mu_\kappa), \mu_h, \nu_h, \lambda_h, \varepsilon_h$ are all uncorrelated with each other. Hence we may write the first- and second-order moments of y_h^* as

$$E(y_h^*) = \mu_{y^*}$$

$$E\left\{(y_h^* - \mu_{y^*})(y_h^* - \mu_{y^*})^/\right\} = \Sigma_{y^*y^*} = B_\kappa \Sigma_{\kappa\kappa} B_\kappa^/ + B_\mu \Sigma_{\mu\mu} B_\mu^/ + B_\nu (I_2 \otimes \Sigma_{\nu\nu}) B_\nu^/ + B_\lambda \Sigma_{\lambda\lambda} B_\lambda^/ + B_\varepsilon (I_2 \otimes \Sigma_{\varepsilon\varepsilon}) B_\varepsilon^/.$$

The vectors $y_1^*, y_2^*, \cdots, y_{408}^*$ are assumed to be stochastically independent.

We next define three vectors with parameters which together constitutes all the parameters entering the first- and second-order theoretical moments. These are

$$\theta_{12} = \left[b^/, c_1^/, c_2^/, e^/, f_1^/, f_2^/, q_2\right]^/,$$

$$\theta_2 = \left[\left(vech\Sigma_{\kappa\kappa}^a\right)^/, \sigma_{uu}^2, \left(vech\Sigma_{\mu\mu}\right)^/, \sigma_{\nu_1\nu_1}^2, \sigma_{\nu_2\nu_2}^2, \sigma_{\nu_3\nu_3}^2, \sigma_{\nu_4\nu_4}^2, \sigma_{\nu_5\nu_5}^2, \left(vech\Sigma_{\lambda\lambda}\right)^/, \left(vech\Sigma_{\varepsilon\varepsilon}\right)^/\right]^/,$$

$$\theta_1 = \left[a_1^/, a_2^/, \mu_\chi, q_{02}, d_1^/, d_2^/, \mu_{z_1}, \mu_{z_2}\right]^/.$$

The theoretical first-order moments are functions of θ_1 and θ_{12}, whereas the second-order theoretical moments are functions of θ_{12} and θ_2.

2. Data and univariate tests of skewness, excess kurtosis and non-normality

The data set is from the years 1975-1977 and are formally treated as a balanced panel data set with two observations for each of 408 observational units, i.e. households. The two data sources are the Norwegian Surveys of Consumer Expenditure and tax files. Altogether there are nine observable variables and seven of these are two-dimensional. Five of the variables are purchase expenditures in constant prices of the following commodities: (i) Food, tobacco and beverages (y_1), (ii) Clothing and footwear (y_2), (iii) Housing, fuel and furniture (y_3), (iv) Travel and recreation (y_4) and (v) Other goods and services ($y5$). Together these variables cover purchases of all goods and services. Furthermore we include two income variables, which we refer to as Income measure 1 and Income measure 2. Income measure 1 (w_1) is "Taxable income for the central government tax assessment minus taxes", whereas Income measure 2 (w_2) is "Income base used for calculating social security premiums and pension rights in the public social security system". The two last variables are, respectively, the number of children (z_1) and the number of adults (z_2) in the households. These are time invariant variables. In each of the two period the upper tail distribution of the two-dimensional variables have been moderately winsorized. The first- and second-order empirical moments are reported in Aasness et al. (1993, pp. 1419-1421).

In the next section the log-likelihood functions depend on the first- and second-order empirical moments.

Let

$$\bar{y}_{it.} = \frac{1}{408}\sum_{h=1}^{408} y_{ith}, i = 1,...,5, t = 1,2 ,$$

$$\bar{w}_{jt.} = \frac{1}{408}\sum_{h=1}^{408} w_{jth}, j = 1,2, t = 1,2 ,$$

$$\bar{z}_{k.} = \frac{1}{408}\sum_{h=1}^{408} z_{kh}, k = 1,2 ,$$

$$\bar{y}_t = \left[\bar{y}_{1t.}, \bar{y}_{2t.}, \bar{y}_{3t.}, \bar{y}_{4t.}, \bar{y}_{5t.}\right]', t = 1,2,$$

$$\bar{w}_{t.} = \left[\bar{w}_{1t.}, \bar{w}_{2t.}\right]', t = 1,2, \text{ t=1,2,}$$

$$\bar{z} = \left[\bar{z}_{1.}, \bar{z}_{2.}\right]'.$$

The vector of empirical means is then given by

$m = \left[\bar{y}_{1.}', \bar{y}_{2.}', \bar{w}_{1.}', \bar{w}_{2.}', \bar{z}'\right]'$. The empirical covariance matrix is given by $S = \frac{1}{H}\sum_{h=1}^{H}\left(y_h^* - m\right)\left(y_h^* - m\right)'$.

We test whether each of the variables $y_{11},\ldots,y_{51}, y_{12},\ldots,y_{52}, w_{11}, w_{21}, w_{12}, w_{22}, z_1, z_2$ are normally distributed. The test statistic for non-normality, which is asymptotically chi-square distributed with two degrees of freedom, is additive in two components each being chi-square distributed with one degree of freedom. They are functions of the sample skewness and kurtosis, respectively. For a further description and discussion of the test statistics cf. Davidson and MacKinnon (1993, pp. 568-569) and

Hall and Cummins (2005, p. 271). The test results are reported in Table D1 in Appendix D. For most of the variables the hypothesis that they are normally distributed is clearly rejected. Generally both skewness and excess kurtosis contribute to the rejection. Thus one cannot claim that the variables have been drawn from a normal distribution, at least not marginally.

3. Normality, maximum likelihood and sufficient statistics

Under the assumption that y_h is normally distributed maximum likelihood estimation is implemented by minimizing the following fit function with respect to θ_{12}, θ_2 and θ_1 (cf. Jöreskog et al., 2000, p. 7)

(5)
$$\begin{aligned} L_{12}\left(\theta_1, \theta_{12}, \theta_2; m, S\right) = \log\left\|\Sigma_{y^*y^*}(\theta_{12},\theta_2)\right\| + tr\left(S\left[\Sigma_{y^*y^*}(\theta_{12},\theta_2)\right]^{-1}\right) - \log\left\|S\right\| - 16 \\ +\left(\mu_{y^*}(\theta_1,\theta_{12}) - m\right)^{/}\left[\Sigma_{y^*y^*}(\theta_{12},\theta_2)\right]^{-1}\left(\mu_{y^*}(\theta_1,\theta_{12}) - m\right). \end{aligned}$$

In the model we are considering there is perfect fit of the theoretical first-order moments, which means that all the information in the first-order empirical moments is used to estimate the parameters in θ_1.[3] This means that S is a sufficient statistic for θ_{12} and θ_2 and that maximum likelihood estimates of these parameters vectors are obtained by minimizing the following fit function

(6)
$$L_2\left(\theta_{12}, \theta_2; S\right) = \log\left\|\Sigma_{y^*y^*}(\theta_{12},\theta_2)\right\| + tr\left(S\left[\Sigma_{y^*y^*}(\theta_{12},\theta_2)\right]^{-1}\right) - \log\left\|S\right\| - 16.$$

Under non-normality the estimators of θ_{12}, θ_2 and θ_1 have status as pseudo-maximum likelihood estimators, which are consistent estimators. However the estimated standard errors based on normality theory may be biased. In light of this, we estimate standard errors using bootstrapping.

4. Bootstrapped standard errors of Engel elasticities

The model specified in Section 2 corresponds to the base model E3P3C1M1[4] in the nomenclature of Aasness et al. (1993). In Appendix C we report the parameter estimates, which corresponds to those reported in Aasness et al., op. cit.

[3] In Appendix C we demonstrate how the estimates of parameters in θ_1 are obtained in a second round after having obtained estimates of the parameters in the vectors θ_2 and θ_{12}.

[4] Sometimes they apply a shorter form of this name.

Let the budget shares evaluated at the mean values of total latent expenditure, the demographic variables and the latent preference variables be defined as

$$\rho_i\left(\theta_{12},\theta_1\right)=0.5\frac{\left(a_{i1}+b_i\mu_\chi+c_{i1}\mu_{z1}+c_{i2}\mu_{z2}\right)}{\mu_\chi}+0.5\frac{\left(a_{i1}+b_i(q_{02}+q_2\mu_\chi)+c_{i1}\mu_{z1}+c_{i2}\mu_{z2}\right)}{q_{02}+q_2\mu_\chi},\ i=1,...,5.$$

The two additive terms correspond to the first and second period, respectively.

We define the Engel elasticities as

$$E_i=\frac{b_i}{\rho_i\left(\theta_{12},\theta_1\right)},i=1,...,5.$$

The Engel-elasticities can be estimated by plugging in the ML-estimates of the parameters occurring in the expression of E_i. When it comes to standard errors of Engel elasticities we compare the results based on two different methods. In the first case the standard errors are calculated utilizing normal distribution theory and the delta method, which involves a first order linearization of the expression for E_i (cf. Kmenta, 1997, p. 486).[5] As an alternative we employ non-parametric bootstrapping, i.e., we draw R new samples with replacement from the empirical distribution and minimize the fit function (5) each time.

The bootstrapped Engel elasticities in replication r are, hence, given by

$$E_i^{[r]}=\frac{b_i^{[r]}}{\rho_i^{[r]}},i=1,...,5,$$

where

$$\rho_i^{[r]}=0.5\frac{\left(a_{i1}^{[r]}+b_i^{[r]}\mu_\chi^{[r]}+c_{i1}^{[r]}\mu_{z1}^{[r]}+c_{i2}^{[r]}\mu_{z2}^{[r]}\right)}{\mu_\chi^{[r]}}+0.5\frac{\left(a_{i2}^{[r]}+b_i^{[r]}(q_{02}^{[r]}+q_2^{[r]}\mu_\chi^{[r]})+c_{i1}^{[r]}\mu_{z1}^{[r]}+c_{i2}^{[r]}\mu_{z2}^{[r]}\right)}{q_{02}^{[r]}+q_2^{[r]}\mu_\chi^{[r]}}.$$

Bootstrapped standard errors of the estimated Engel elasticities are then obtained by calculating the empirical standard deviations of the bootstrapped Engel elasticities, i.e.,

$$\hat{\sigma}_{\hat{b}_i}^{boot}=\sqrt{\frac{1}{R-1}\sum_{r=1}^{R}\left(\rho_i^{[r]}-\bar{\rho}_i^{boot}\right)^2},$$

[5] These calculations have been done in TSP 4.5 (cf. Hall and Cummins, 2005).

where

$$\overline{\rho}_i^{boot} = \frac{1}{R}\sum_{r=1}^{R}\rho_i^{[r]}.$$

In the numerical calculations the number of replications, R, have been set to 100, 1000 and 10000, respectively. The results are reported in Table 1 below. In the second column we report the estimates of the Engel-elasticities. The estimated standard errors based on the normal distribution and the delta method are reported in the third column. The three last columns in Table 1 contain standard errors obtained by bootstrapping with 100, 1000 and 10 000 replications, respectively. All the four sets with standard errors are rather similar. There are some changes in the standard errors when increasing from 100 to 1000 replications, but only minor changes when increasing from 1000 to 10 000 replications. If we compare the standard errors obtained by using 10 000 replications with those obtained using the delta method, we see that the former method produces higher standard errors for "Food, beverages and tobacco", "Housing, fuel and furniture" and "Travel and recreation" and lower standard errors for "Clothning and footwear" and "Other goods and services". The conclusions that "Food, beverages and tobacco" is a necessary and that "Other goods and services" is a luxury good go through in all four cases.

Table 1. Pseudo-maximum likelihood estimates of Engel elasticities and different measures of standard errors

Commodity	Estimate[a]	Standard errors			
		Delta method[b]	Number of bootstrap replications[c]		
			100	1,000	10,000
Food, beverages tobacco	0.632	0.048	0.046	0.051	0.051
Clothing and footwear	1.143	0.107	0.097	0.099	0.099
Housing, fuel and furniture	1.079	0.068	0.072	0.074	0.075
Travel and recreation	1.098	0.073	0.065	0.075	0.078
Other goods and services	1.381	0.106	0.105	0.099	0.100

[a] This is the Engel elasticities reported in Table VII in Aasness *et al.* (1993).

[b] The estimated standard errors of the Engel elasticties reported in Table VII in Aasness *et al.* (1993) only accounted for the estimation uncertainty in the marginal budget shares, i.e., in the estimates of the b_i-parameters.

[c]The means of the bootstrapped Engel-elasticities are very close to the quasi-maximum likelihood estimates.

5. Conclusions

A complete set of linear Engel-curves where total expenditure is assumed to be contaminated by measurement error has been estimated using a structural equation modeling framework minimizing a fit function which is the optimal one under normality. However, normality is rejected by formal testing. In such a case the procedure can be labeled pseudo-maximum likelihood estimation, which is believed to yield consistent estimates of the parameters but may involve biased estimates of the standard errors. In the light of this we have calculated standard errors by bootstrapping. From an economic point of view focus is often on Engel elasticities. In our case the elasticity is not a parameter, but a function of a set of variables. We consider the case where the Engel elasticties are evaluated at the expected value of the variables of which they are functions. In this point the Engel elasticities are non-linear functions of parameters in the model. We calculate standard errors by two different methods. The first is based on normality and application of the delta method, whereas the second is based on bootstrapping. It turns out that the deviations between the estimated standard errors of the Engel elasticities are rather modest. For instance, the classification of the consumption categories as luxury and necessities is not influenced.

The above results may be related to the fact that some literature in the structural equation modeling tradition shows that, asymptotically, estimation of standard errors of parameter estimates can proceed as if the observed variables were normally distributed even if this is not the case, cf. for instance the contribution by Satorra (1990, 1992) and Satorra and Bentler (1990). This is an interesting area for further work.

References

Aasness, J., Biørn, E. and T. Skjerpen (1993): Engel Functions, Panel Data and Latent Variables. *Econometrica*, **61,** 1395–1422.

Browne, M.W. and A. Shapiro (1988): "Robustness of Normal Theory Methods in the Analysis of Linear Latent Variate Models. *British Journal of Mathematical and Statistical Psychology*, **41**, 193−208.

Davidson, R. and J.G. MacKinnon (1993): *Estimation and Inference in Econometrics*. Oxford: Oxford University Press.

Efron, B. and R.J. Tibshirani (1993): *Introduction to the Bootstrap*. London: Chapman and Hall.

Hall, B.H. and C. Cummins (2005): TSP Reference Manual. Version 5.0. Palo Alto, CA: TSP International.

Hall, P. (1994): Methodology and Theory for the Bootstrap. Chapter 39 in Engle R. and D.L. McFadden (Eds.) *Handbook of Econometrics*. *Volume 4*. Amsterdam: Elsevier, pp. 2341−2381.

Horowitz, J.L. (2001): The Bootstrap. Chapter 52 in Heckman, J. J. and E. Leamer (Eds.): *Handbook of Econometrics*, *Volume 5.* Amsterdam: Elsevier, pp. 3160−3228.

Jöreskog, K.G. and D. Sörbom (1996): *LISREL® 8: User'sReference Guides*. Lincolnwood, IL: Scientific Software International, Inc.

Jöreskog, K.G., Sörbom, D., Du Toit, S. and M. Du Toit (2000): *LISREL® 8: New Statistical Features*. Lincolnwood, IL: Scientific Software International, Inc.

Kmenta, J. (1997): *Elements of Econometrics*. Second edition. Ann Arbor, MI: University of Michigan Press.

Mooijaart, A. and P.M. Bentler (1991): Robustness in Normal Theory Statistics in Structural Equation Models. *Statistica Neerlandica*, **45**, 159−171.

Satorra, A. (1990): "Robustness Issues in Structural Equation Modeling: A Review of Recent Developments". *Quality & Quantity*, **24**, 367−386.

Satorra, A. (1992): "Asymptotic Robust Inferences in the Analysis of Mean and Covariance Structures. In P.V. Marsden (Ed.), *Sociological Methodology* 1992. Vol. 22. Oxford: Blackwell, pp. 249−278.

Satorra, A. and P.M. Bentler (1990): "Model Conditions for Asymptotic Robustness in the Analysis of Linear Relations". *Computational Statistics & Data Analysis*, **10**, 235−249.

Stine, R. (1990): An Introduction to Bootstrap Methods: Examples and Ideas. *Sociological Methods & Research*, **18**(2/3), 243−291.

Yung, Y.-F. and P.M. Bentler (1998): "Bootstrapping Techniques in Analysis of Mean and Covariance Structures". In Marcoulides, G.A. and R.E. Schumacker (Eds.): *Advanced Structural Equation*

Modeling: Issues and Techniques. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers, pp. 195–226.

Appendix A

Explicit expressions for the matrices $B_\kappa, B_\mu, B_\nu, B_\lambda, B_\varepsilon$ and the vector μ_{y^*}.

We start out with defining some vectors and matrices which will be employed below.

Let ι_j denote a column vector with *j* elements, all being equal to 1, I_j the identity matrix of order *j*, 0_j a quadratic matrix of order zero in which all elements are equal to zero and $0_{i\times j}$ (where $i\neq j$) a matrix with *i* rows and *j* columns where all elements are equal to zero. We define the following vectors and matrix

$$a_t^* = \left[a_t^{/}, -\iota_4^{/} a_t\right]^{/}, t = 1,2,$$

$$b^* = \left[b^{/}, 1 - \iota_4^{/} b\right]^{/},$$

$$c_j^* = \left[c_j^{/}, -\iota_4^{/} c_j\right]^{/}, j = 1,2,$$

$$C^* = \left[c_1^* \; c_2^*\right].$$

We partition the *B*-matrices in the following way:

$$B_m = \left[B_m^{(1)/} \; B_m^{(2)/} \; B_m^{(3)/}\right]^{/}, \text{ where } m = \kappa, \mu, \nu, \lambda, \varepsilon .$$

The number of rows in these 3 submatrices are 10, 4 and 2 for all values of m, whereas the number of columns differ. We now specify all the submatrices:

$$B_\kappa^1 = \begin{bmatrix} b^* & c_1^* & c_2^* & b^* & 0_{5\times 1} \\ q_2 b^* & c_1^* & c_2^* & 0_{5\times 1} & q_2 b^* \end{bmatrix},$$

$$B_\kappa^2 = \begin{bmatrix} e & f_1 & f_2 & e & 0_{2\times 1} \\ q_2 e & f_1 & f_2 & 0_{2\times 1} & q_2 e \end{bmatrix},$$

$$B_\kappa^3 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix},$$

$$B_\mu^1 = \iota_2 \otimes I^*,$$

where

$$I^* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & -1 & -1 & -1 \end{bmatrix},$$

$$B_\mu^2 = 0_4 ,$$

$$B_\mu^3 = 0_{2\times 4} ,$$

$$B_\nu^1 = I_{10} ,$$

$$B_\nu^2 = 0_{4\times 10} ,$$

$$B_\nu^3 = 0_{2\times 10} ,$$

$$B_\lambda^1 = 0_{10\times 2} ,$$

$$B_\lambda^2 = \iota_2 \otimes I_2 ,$$

$$B_\lambda^3 = 0_2 ,$$

$$B_\varepsilon^1 = 0_{10\times 4} ,$$

$$B_\varepsilon^2 = I_4 ,$$

$$B_\varepsilon^3 = 0_{2\times 4} ,$$

The population covariance matrix is now given by

$$\Sigma_{y^* y^*} = \begin{bmatrix} B_\kappa^1 \Sigma_{\kappa\kappa} B_\kappa^{1\prime} + B_\mu^1 \Sigma_{\mu\mu} B_\mu^{1\prime} + B_\nu^1 (I_2 \otimes \Sigma_{\nu\nu}) B_\nu^{1\prime} & B_\kappa^1 \Sigma_{\kappa\kappa} B_\kappa^{2\prime} & B_\kappa^1 \Sigma_{\kappa\kappa} B_\kappa^{3\prime} \\ B_\kappa^2 \Sigma_{\kappa\kappa} B_\kappa^{1\prime} & B_\kappa^2 \Sigma_{\kappa\kappa} B_\kappa^{2\prime} + B_\lambda^2 \Sigma_{\lambda\lambda} B_\lambda^{2\prime} + B_\varepsilon^2 (I_2 \otimes \Sigma_{\nu\nu}) B_\varepsilon^{2\prime} & B_\kappa^2 \Sigma_{\kappa\kappa} B_\kappa^{3\prime} \\ B_\kappa^3 \Sigma_{\kappa\kappa} B_\kappa^{1\prime} & B_\kappa^3 \Sigma_{\kappa\kappa} B_\kappa^{2\prime} & B_\kappa^3 \Sigma_{\kappa\kappa} B_\kappa^{3\prime} \end{bmatrix}$$

We partition μ_{y^*} as

$$\mu_{y^*} = \left[\mu_{y^*}^{1\prime}, \mu_{y^*}^{2\prime}, \mu_{y^*}^{3\prime} \right]^\prime ,$$

where the three subvectors contain 10, 4 and 2 elements, respectively. They are given, in partitioned form, by

$$\mu_{y^*}^1 = \begin{bmatrix} a_1^* + b^*\mu_\chi + c_1^*\mu_{z_1} + c_2^*\mu_{z_2} \\ a_2^* + q_2 b^*\mu_\chi + c_1^*\mu_{z_1} + c_2^*\mu_{z_2} \end{bmatrix},$$

$$\mu_{y^*}^2 = \begin{bmatrix} d_1 + e\mu_\chi + f_1\mu_{z_1} + f_2\mu_{z_2} \\ d_2 + q_2 e\mu_\chi + f_1\mu_{z_1} + f_2\mu_{z_2} \end{bmatrix},$$

$$\mu_{y^*}^3 = \begin{bmatrix} \mu_{z_1} \\ \mu_{z_2} \end{bmatrix}.$$

Appendix B

Estimation of the parameters occurring only in the first-order moments

The theoretical first order moments may be written as

(7) $$\mu_{y^*} = G(\theta_{12})\theta_1.$$

Let us partition the G-matrix in the following way

$$G = \begin{bmatrix} G^1_{10\times 16} \\ G^2_{4\times 16} \\ G^3_{2\times 16} \end{bmatrix}.$$

The submatrices defining G are then given as

$$G^1_{10\times 16} = \begin{bmatrix} I^* & 0_{5\times 4} & b & 0_{5\times 1} & 0_{5\times 2} & 0_{5\times 2} & C^* \\ 0_{5\times 4} & I^* & q_2 b^* & q_{02} b^* & 0_{5\times 2} & 0_{5\times 2} & C^* \end{bmatrix},$$

$$G^2_{4\times 16} = \begin{bmatrix} 0_{2\times 4} & 0_{2\times 4} & e & 0_{2\times 1} & I_2 & 0_2 & F \\ 0_{2\times 4} & 0_{2\times 4} & q_2 e & q_{02} e & 0_2 & I_2 & F \end{bmatrix}$$

and

$$G^3_{2\times 16} = \begin{bmatrix} 0_{2\times 4} & 0_{2\times 4} & 0_{2\times 1} & 0_{2\times 1} & 0_{2\times 1} & 0_{2\times 1} & I_2 \end{bmatrix}.$$

If we invert (7) we obtain

$$\theta_1 = \left[G(\theta_{12})\right]^{-1} \mu_{y^*}.$$

We can estimate μ_{y^*} by m, and the ML-estimator of θ_1 is

$$\hat{\theta}_1 = \left[G\left(\hat{\theta}_{12}\right)\right]^{-1} m,$$

where $\hat{\theta}_{12}$ denotes the ML-estimator of θ_{12}. This means, in contrast to what is the case for the second-order moments, that there is a perfect fit as far as the theoretical first-order moments are concerned.

Estimates of the parameters in the econometric model and estimated standard errors based on normal theory

Table C1. Estimates of the parameters in the vector θ_{12} and estimated standard errors based on normal theory[a]

Parameter	Estimate	Standard error
b_1	0.162	0.012
b_2	0.122	0.011
b_3	0.268	0.016
b_4	0.343	0.021
γ_{11}	0.907	0.148
γ_{21}	0.109	0.132
γ_{31}	-0.330	0.192
γ_{41}	-0.492	0.253
γ_{12}	0.569	0.236
γ_{22}	0.048	0.210
γ_{32}	-1.526	0.307
γ_{42}	1.122	0.408
q_2	1.104	0.030
e_1	0.514	0.053
e_2	1.110	0.100
f_{11}	-1.384	0.673
f_{21}	-0.121	1.270
f_{12}	9.474	1.066
f_{22}	11.111	2.011

[a] Most of these results are reproduced from tables V and VII in Aasness *et al.* (1993). The relevant column in Table V is the one labeled 'Base model'.

Table C2. Estimates of the parameters in the vector θ_2. Estimates of standard errors based on normal theory[a]

Parameter	Estimate	Standard error
$\sigma^2_{\chi\chi}$	380.015	33.679
$\sigma^2_{z_1\chi}$	8.797	1.402
$\sigma^2_{z_2\chi}$	10.003	1.091
$\sigma^2_{z_1z_1}$	1.579	0.111
$\sigma^2_{z_2z_1}$	0.079	0.057
$\sigma^2_{z_2z_2}$	0.827	0.058
σ^2_{uu}	15.149	4.596
$\sigma^2_{\mu_1\mu_1}$	6.228	0.839
$\sigma^2_{\mu_2\mu_1}$	-0.204	0.512
$\sigma^2_{\mu_3\mu_1}$	-0.804	0.756
$\sigma^2_{\mu_4\mu_1}$	-4.938	1.107
$\sigma^2_{\mu_2\mu_2}$	3.014	0.732
$\sigma^2_{\mu_3\mu_2}$	-2.194	0.692
$\sigma^2_{\mu_4\mu_2}$	-0.759	0.930
$\sigma^2_{\mu_3\mu_3}$	7.735	1.458
$\sigma^2_{\mu_4\mu_3}$	-4.103	1.519
$\sigma^2_{\mu_4\mu_4}$	10.324	2.570
$\sigma^2_{\nu_1\nu_1}$	9.819	0.719
$\sigma^2_{\nu_2\nu_2}$	13.146	0.934
$\sigma^2_{\nu_3\nu_3}$	26.914	1.963
$\sigma^2_{\mu_4\mu_4}$	89.017	6.161
$\sigma^2_{\nu_5\nu_5}$	5.316	0.395
$\sigma^2_{\lambda_1\lambda_1}$	192.570	16.677
$\sigma^2_{\lambda_2\lambda_1}$	276.534	27.716
$\sigma^2_{\lambda_2\lambda_2}$	721.531	58.804
$\sigma^2_{\varepsilon_1\varepsilon_1}$	57.440	4.441
$\sigma^2_{\varepsilon_2\varepsilon_1}$	53.574	5.587
$\sigma^2_{\varepsilon_2\varepsilon_2}$	92.816	9.128

[a] Most of these results are reported in tables IV, V and VI in Aasness *et al.* (1993). The relevant columns in table IV, V and VI are those labeled M1, 'Base model' and P3, respectively. The notation in Table V differs slightly from the one used in the current paper.

Table C3. Estimates of the parameters in the vector θ_1. Estimates of standard errors based on normal theory[a]

Parameter	Estimate	Standard error
μ_χ	39.750	1.151
q_{02}	-1.165	1.527
μ_{z1}	0.804	0.062
μ_{z2}	2.225	0.045
a_{11}	2.125	0.475
a_{21}	-0.674	0.428
a_{31}	2.908	0.620
a_{41}	-3.785	0.818
a_{21}	1.650	0.484
a_{22}	-0.991	0.436
a_{32}	2.795	0.632
a_{42}	-2.891	0.834
d_{11}	-2.291	2.138
d_{21}	-13.565	4.025
d_{12}	0.078	2.176
d_{22}	-13.483	4.097

[a] Most of these results are reported in tables V and VII in Aasness *et al.* (1993). The relevant column in Table V is the one labeled E3P3M1C1.

Testing for skewness, excess kurtosis and normality

Table D1. Univariate tests of skewness, excess kurtosis and normality

Variable	Skewness[a]		Excess kurtosis[b]		Normality	
	Statistic	p-value	Statistic	p-value	Statistic	p-value
y_{11}	42.007	<0.00000	2.816	0.093	44.823	<0.00000
y_{12}	27.735	<0.00000	0.003	0.956	27.738	<0.00000
y_{21}	138.406	<0.00000	49.352	<0.00000	187.758	<0.00000
y_{22}	131.439	<0.00000	28.965	<0.00000	160.405	<0.00000
y_{31}	149.174	<0.00000	96.111	<0.00000	245.286	<0.00000
y_{32}	116.387	<0.00000	37.824	<0.00000	154.211	<0.00000
y_{41}	154.740	<0.00000	63.570	<0.00000	218.310	<0.00000
y_{42}	130.413	<0.00000	40.585	<0.00000	170.998	<0.00000
y_{51}	218.484	<0.00000	201.658	<0.00000	420.142	<0.00000
y_{52}	135.468	<0.00000	68.929	<0.00000	204.397	<0.00000
w_{11}	4.798	0.028	0.014	0.905	4.812	0.090
w_{12}	8.056	0.004	1.369	0.242	9.425	0.009
w_{21}	4.957	0.026	6.793	0.009	11.750	0.003
w_{22}	5.204	0.023	7.040	0.008	12.244	0.002
z_1	283.128	<0.00000	528.550	<0.00000	811.678	<0.00000
z_2	72.379	<0.00000	42.137	<0.00000	114.516	<0.00000

[a] Cf. formula (16.41) of Davidson and MacKinnon (1993, p. 568).

[a] Cf. formula (16.42) of Davidson and MacKinnon (1993, p. 569).