Jones, A.M. (1989): A Double-Hurdle Model of Cigarette Consumption, in: Journal of Applied Econometrics, Vol. 4, No 1, 23-39

Lechner, M. (1992): Some Specification Tests for Static Limited Dependent Variable Models Estimated on Panel Data. Working Paper, Institut für Volkswirtschaftlsehre und Statistik, Universität Mannheim und ZEW, Mannheim.

Maddala, G. S. (1983): Limited-Dependent and Qualitative Variables in Econometrics; New York.

Malhotra, N. (1984): The Use of Linear Logit Models in Marketing Research, in: Journal of Marketing Research, Vol XXI, 20-31.

McFadden, D. (1980): Econometric Models for Probabilistic Choice among Products. In: Journal of Business, Vol.53, No.3, 13-29.

McFadden, D. (1984): Econometric Analysis of Qualitative Response Models. In: Griliches, Z., Intriligator, M.D. (Ed.): Handbook of Econometrics, Volume II,.1395-1457.

McKelvey, R.; Zavoina, W.(1975): A Statistical Model for the Analysis of Ordinal Level Dependent Variables. In: Journal of Mathematical Sociology 4, 103-120.

Ronning, G. (1991): Mikroökonometrie. Springer Verlag Heidelberg.

Schneider, H. (1993): Binary Discrete Choice Under Asymmetric Restrictions. In: *Statistical Papers*, Vol. 34, 43-57.

Thursby, J., Schmidt, P. (1977): Some Properties of Test Specification Errors in a Linear Regresson Model. In: Journal of the American Statistical Association, Vol. 72,.635-641.

Veall, M.,Zimmermann, K. (1990): Evaluating Pseudo $R^{2}$´s for Binary Probit Models. Working Paper, University of Munich.

White, H. (1982): Maximum Likelihood Estimation of Misspecified Models. In: Econometrica,Vol. 50, Nr. 1, 1-25.

White, H. (1983): Corrigendum. In: Econometrica, Vol. 51, 513.

Wittink, D.R., P. Cattin (1989): Commercial Use of Conjoint Analysis: An Update. In: Journal of Marketing, Vol. 7, 91-96.

Zabel, J. (1993): The Relationship between Hours of Work and Labor Force Participation in Four Models of Labor Supply Behaviour. In: Journal of Labor Economics, Vol. 11, No. 2, 387-416.

Zeger, S., Liang, K. (1986): Longitudinal Data Analysis for Discrete and Continuous Outcomes. In : Biometrics Vol. 42, 121-130.

## References


Amemiya, T. (1975): Qualitative Response Models. In Annals of Economic and Social Measurement, 4/3, 1975, 363-372.

Amemiya, T. (1985): Advanced Econometrics. Oxford.

Arminger, G. (1995): Specification and Estimation of Mean Structures: Regression Models. In: Arminger, G., Clogg, C., Sobel, M. (Ed.): Handbook of Statistical Modeling for the Social and Behavioral Sciences,. 77-177.

Backhaus, K.; Erichson,B.; Plinke, W.;Weiber, R. (1990): Multivariate Analysemethoden; New York.

Berndt, E., Hall, B., Hall, R., Hausman, J. A. (1974): Estimation and Inference in Nonlinear Structural Models. In: Annals of Economic and Social Measurement 3, 653-666.

Blundell, R.; Meghir, C.(1987): Bivariate Alternatives to the Tobit Model. In: Journal of Econometrics, Vol.34, 179-200.

Chamberlain, G. (1982): Multivariate Regression Models for Panel Data. In: Journal of Econometrics, Vol. 18, 5-46.

Chamberlain, G. (1984). Panel Data. In: Griliches, Z., Intriligator, M.D.(Ed..), Handbook of Econometrics, Vol. II. Amsterdam, 1248-1318

Chintagunta, P. (1992): Estimating a Multinomial Probit Model of Brand Choice Using the method of Simulated Moments, in: Marketing Science, Vol. 11, No. 4, 386-407.

Chintagunta, P. (1994): Heterogeneous Logit Model Implications for Brand Positioning, in: Journal of Marketing Research, Vol. XXXI, 304-311.

Chintagunta, P.; Dipak, Jain, C.; Vilcassim, N. (1991): Investigating Heterogeneity in Brand Preferences in Logit Models for Panel Data. In: Journal of Marketing Research, Vol.28, November 1991, 417-428.

Cragg, J.G. (1971): Some Statistical Models for Limited Dependent Variables with Applications to the Demand for Durable Goods. In: Econometrica, 39, 829-844.

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

Elrod, T.; Keane, M. (1995): A Factor-Analytic Probit Model for Representing the Market Structure in Panel Data, in: Journal of Marketing Research, Vol. XXXII, 1-16.

Fader, P.; Lattin, J. (1993): Accounting for Heterogeneity and Nonstationarity in a Cross-Sectional Model of Consumer Purchase Behavior, in: Marketing Science, Vol. 12, No. 3, 304-317.

Gourieroux, C., Monfort, A. Trognon, A. (1984): Pseudo Maximum Likelihood methods: Theory. In : Econometrica, Vol. 52, 681-720.

Green P.; Srinivasan, V. (1978): Conjoint Analysis in Consumer Research. In: The Journal of Consumer Research, Vol. 5, 103-122.

Gruca, T.; Sudharshan D. (1991): Equilibrium Characteristics of Multinomial Logit Market Share Models. In: Journal of Marketing Research, Vol.XXVIII, November 1991,.480-482

Gupta, S., Chintagunta P., Kaul, A., Wittink R. (1996): Do Household Scanner Data Provide Representative Inferences From Brand Choices: A Comparison With Store Data. In: Journal of Marketing Research, Nr. 4, Vol. 33, 383-398

Heckman, J.; Willis, R. (1976): Estimation of a Stochastic Model of Reproduction: An Econometric Approach. In: Terleckyj, N.(Ed.): Household Production and Consumption, 99-138.

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

Huber, P.J. (1967): The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions, Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol.1, University of California Press, Berkeley and Los Angeles

Grammig J., Hujer R., Löwenbein, O. (1997): Mikroökonometrische Modellierung der Markenwahl bei diskontinuierlichem Kaufverhalten. Forthcoming in: Allgemeines Statistisches Archiv

Hujer, R., Grammig, J., Fryns, H., Herterich, R. (1996): Preisfindung und optimale Marketingstrategien für neue pharmazeutische Produkte. In: Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung, Heft 3.

$$\hat{S} = \frac{1}{N \cdot K} \cdot \frac{\partial Q\left(\hat{\theta}\right)}{\partial \theta'} \hat{A}\left(\hat{\theta}\right)^{-1} \frac{\partial s\left(\hat{\theta}\right)'}{\partial \theta}$$
$$\left[ \frac{\partial s\left(\hat{\theta}\right)}{\partial \theta'} \hat{\Omega} \frac{\partial s\left(\hat{\theta}\right)'}{\partial \theta} \right]^{-1} \frac{\partial s\left(\hat{\theta}\right)}{\partial \theta'} \hat{A}\left(\hat{\theta}\right)^{-1} \frac{\partial Q\left(\hat{\theta}\right)}{\partial \theta}$$

with:

$$\hat{A}\left(\hat{\theta}\right) = \frac{1}{K \cdot N} \sum_{k=1}^{K} \sum_{i=1}^{N} \frac{\partial^2 \ln L_i\left(\hat{\theta}\right)}{\partial \theta \, \partial \theta'}$$

$$\hat{\Omega} = \hat{A}\left(\hat{\theta}\right)^{-1} \cdot \hat{B}\left(\hat{\theta}\right)^{-1} \cdot \hat{A}\left(\hat{\theta}\right)^{-1}$$

$$\hat{B}\left(\hat{\theta}\right) = \frac{1}{K \cdot N} \left\{ \sum_{i=1}^{N} \sum_{k=1}^{K} \frac{\partial \ln L_i\left(\hat{\theta}\right)}{\partial \theta} \cdot \frac{\partial \ln L_i\left(\hat{\theta}\right)}{\partial \theta'} \right\} \qquad \text{(A-3)}$$

Under the null the generalised score statistic is asymptotically $\chi^2(q)$. Under $H_A$ $\hat{S} \to \infty$, if $N \to \infty$.

### A.2.2 RESET Tests

We use RESET type tests (**Re**gression **S**pecification **E**rror **T**est) (Ramsey 1969, Thursby, Schmidt 1977) to test the specification of the three equation censored model. As suggested by Thursby, Schmidt (1977) we use higher order polynomials of the mean functions to test the specifications. Generally, we have

$$V_{ik}^{*} = \mu_{Vk} + \sum_{l=2}^{M} \beta_l \mu_{Vk}^{l} + \varepsilon_{Vik}$$
$$U_{ik}^{*} = \mu_{Uk} + \sum_{l=2}^{M} \beta_l \mu_{Uk}^{l} + \varepsilon_{Uik}$$
$$D_{ik}^{*} = \mu_{Dk} + \sum_{l=2}^{M} \beta_l \mu_{Dk}^{l} + \varepsilon_{Dik} \qquad \text{(A-4)}$$

A RESET-Test, that includes a second order polynomial of the mean is a RESET2. Including the second and third order polynomial of the mean function we have a RESET23, which can be used as a test of normality. In (A-4) we have the general formulation of a RESET23..M-Test. To conduct the test, we formulate the null that the additional parameters are zero and compute the generalised score statistic (A-3).

## A.2 Generalised Score Testing

### A.2.1 General Idea

White (1982,1983) showed that conventional Score- and Wald statistics take on wrong values in the presence of misspecification of the probability model. The generalised Score- and Wald-test statistics derived by White (1982,1983) take on the correct size in the presence of misspecification, too. In order to test hypotheses using the generalised Score and Wald statistics the null is expressed in a system of homogeneous equations, representing restrictions on the parameters:

$$\begin{aligned} H_0&: s(\theta) = 0 \\ H_A&: s(\theta) \neq 0. \end{aligned} \tag{A-1}$$

$\theta$ is the vector of parameters, $r$ the number of parameters in $\theta$ The vector valued function $s(\theta)$ is then a function from $R^r \rightarrow R^q$, where $q$ is the number of rows in $s(\theta)$.

We assume that $s(\theta)$ contains no redundant restrictions, the rank of the Jacobi-matrix $\frac{\partial s(\theta)}{\partial \theta'}$ is equal to $q$. The estimation of $\theta$ can be done under the restrictions $s(\theta) = 0$ or unrestricted.

The parameter vector that maximises (11) under the restrictions $s(\theta) = 0$

$$\max_{\theta,\lambda} \sum_{i=1}^{N} \ln L_i(\theta) + \lambda' \cdot s(\theta) \tag{A-2}$$

is defined as $\hat{\theta}$ .
If the parameter estimate $\hat{\theta}$ available, then we can use the generalised Score statistic

The model contains a latent equation that explains the qualitative decision to purchase a product and that controls simultaneously the observable volume and preference score decision. Unlike the single equation Tobit model, the restrictive assumptions of how the qualitative and quantitative consumer decision are linked together are relaxed.

We place the application of the model in a context of a market entry study for a medicament after a fundamental change in the German health insurance system. After the German Health reform, cost awareness was very much stressed by imposing budget constraints on the prescription budget of the physicians. Comparing the results with the Tobit formulation, we find that the Tobit model´s implicit parameter restrictions can lead to significant different results and policy implications. The application shows that the model's ability to distinguish prescription probability, purchase volume and preference scores gives greater insight in these different aspects of consumer behaviour.

# Appendix

## A.1 Goodness of Fit Measures

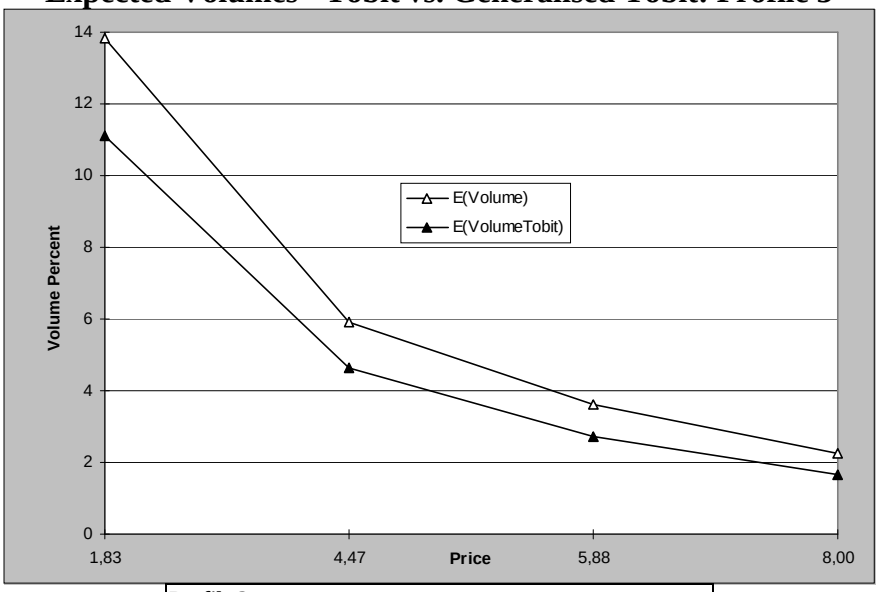
### A.1.1 McKelvey-Zavoina´s R squared for three equation generalised Tobit

$$R^2_{MZ}(V^*) = \frac{\sum_{k=1}^{K}\left(\hat{\mu}_{VK} - \frac{1}{K}\sum_{k=1}^{K}\hat{\mu}_{VK}\right)^2}{\sum_{k=1}^{K}\left(\hat{\mu}_{VK} - \frac{1}{K}\sum_{k=1}^{K}\hat{\mu}_{VK}{}_{V}\right)^2 + K\sigma_V^2} \cdot 100 \text{ (Volume Propensity)}$$

$$R^2_{MZ}(U^*) = \frac{\sum_{k=1}^{K}\left(\hat{\mu}_{UK} - \frac{1}{K}\sum_{k=1}^{K}\hat{\mu}_{VK}\right)^2}{\sum_{k=1}^{K}\left(\hat{\mu}_{UK} - \frac{1}{K}\sum_{k=1}^{K}\hat{\mu}_{VK}\right)^2 + K\sigma_U^2} \cdot 100 \text{ (Preference Propensity)}$$

$$R^2_{MZ}(D^*) = \frac{\sum_{k=1}^{k}\left(\hat{\mu}_{DK} - \frac{1}{K}\sum_{k=1}^{K}\hat{\mu}_{DK}\right)^2}{\sum_{k=1}^{k}\left(\hat{\mu}_{DK} - \frac{1}{K}\sum_{k=1}^{K}\hat{\mu}_{DK}\right)^2 + K} \cdot 100 \text{ (Choice Propensity)}$$

**Figure 7:**
**Expected Volumes - Tobit vs. Generalised Tobit: Profile 3**



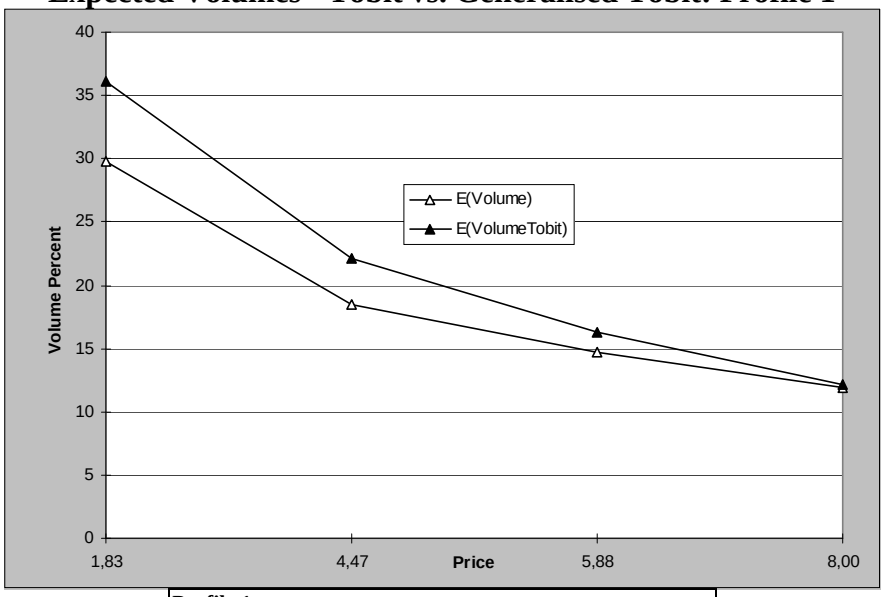

| **Profile 3** | |
|---|---|
| Cartilage Protection: | yes |
| Efficacy | **less than reference product** |
| Manufacturer: | Stellapharma |
| Side effects | **more than reference product** |

From profile 2 we move to profile 3 by changing the attribute level of efficacy to „less than reference product“. This would present the worst case for the manufacturer when cartilage protection would be associated with „more side effects“ and „less efficacy“ by the physicians. Figure 7 shows that expected volume sharply declines compared to profile 2 for all price levels. When raising the price level from the *1,83 DM* to *4,47 DM* expected volume falls from 13.8 to 5.9 percent. Again, the Tobit model now seems to underestimate expected volume. From figures 5 to 7 we learn that the decline in volume when moving from the lowest price level to one higher level is the sharper the less favourable attribute levels are present in the profile.

## 4. Summary and Conclusion

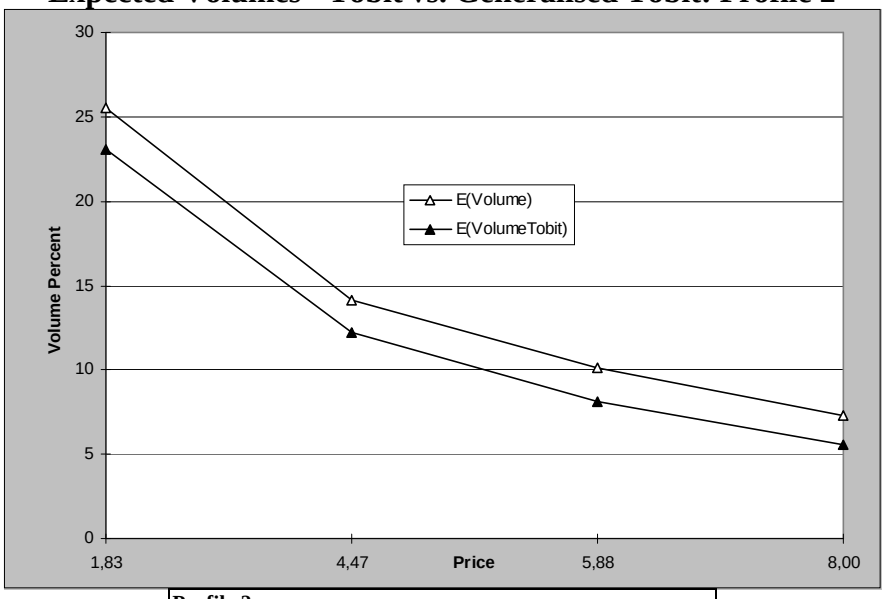
Our study was motivated by the drawbacks of traditional Conjoint Analysis to take into account other aspects of consumer behaviour than consumer preferences. By placing a microeconometric threshold model in a Conjoint-typical full profile design, we are able to model the qualitative decision to purchase a product, the volume decision and preference scores in a censored three equation model.

**Figure 5:**
**Expected Volumes - Tobit vs. Generalised Tobit: Profile 1**




| **Profile 1** | |
|---|---|
| Cartilage Protection: | yes |
| Efficacy | as reference product |
| Manufacturer: | N.N |
| Side effects | same as reference product |

**Figure 6:**
**Expected Volumes - Tobit vs. Generalised Tobit: Profile 2**




| **Profile 2** | |
|---|---|
| Cartilage Protection: | yes |
| Efficacy | as reference product |
| Manufacturer: | N.N |
| Side effects | more  than reference product |

effects". For all possible combinations the negative effect of „more side effects" will clearly dominate. Examples will be given below.

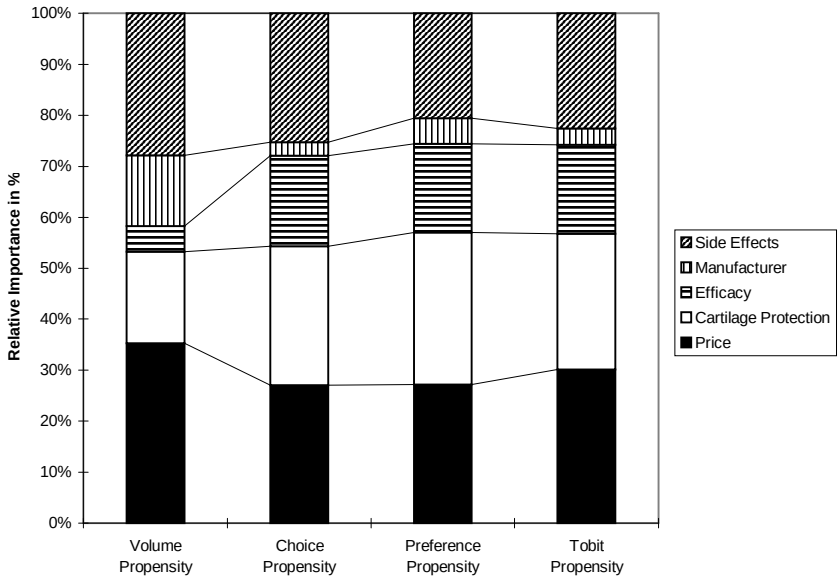
Especially the reversed effect of „more side effects" shows that the distinction between choice and volume propensity in the generalised model is indeed helpful, because the Tobit specification does not allow to distinguish such effect. Only one latent equation is estimated, and the result show a unambiguously negative effect of „more side effects" on the latent equation (see table 3).

Figures 5 to 7 illustrate the effects of price level changes on expected volume for distinct profile levels. Using the formula for $E(V_{ik})$ in the Tobit formulation (see equation (5)) and using the estimated parameters from table 3, we also plot the expected volume estimates from the Tobit model.

Profile 1 in figure 5 represents the „best case" for the manufacturer of the new medicament: A product with accepted cartilage protection, accepted efficacy and as much side effects as the reference product. The expected volume percent drop from 30 % to 20 % when changing the price level from *1,88 DM* to *4,47 DM*. At the high price end, the expected volume percent is about 12 %. The expected volume computed from the Tobit results lie above the volumes implied by the general model, especially for the low price levels.

Profile 2 is identical to profile 1 with the exception that „more side effects" are present (a likely case for the new medicament). We see that the negative impact on volume clearly dominates with expected volumes 25,5 %, 14,1 % 10,1 % and 7,3 % associated with the different price levels. Note that the Tobit model appears to overestimate the negative effect of „more side effects" on volume compared to the generalised model.

**Figure 4:**
**Comparison of Relative Importance**




When comparing the volume propensity equation with the others more significant differences appear. Attempting an interpretation of the latent volume propensity, one has to consider that the censoring rules in equation (9) determine the observability of positive outcomes $V_{ik}$ and $U_{ik}$. The results show that in the volume propensity equation only the price label *1,83 DM* increases the volume propensity significantly. Price, however is now the relatively most important attribute for volume: A high volume propensity can best be reached by a low price level. Cartilage protection also has a statistically significant positive effect on volume propensity. The efficacy labels (less and equal than reference) have no effect on the basic volume propensity. The lack of significance of the parameters is reflected in the low Pseudo R-squared. The mean propensity of around 25 % is not very much affected by different attribute levels.

Note that the levels of the attribute „side effects" have a reversed impact in the volume propensity equation. To interpret this result one has to consider the interaction of choice propensity and the censored observables, $V_{ik}$ and $U_{ik}$. In order to estimate the effect of different attribute levels on expected volume, $\mathrm{E}(V_{ik})$, we use the formula in equation 12 and replace the parameters by the estimated coefficients in table 4, i.e. $\hat{\mathrm{E}}(V_{ik}) = \hat{\mu}_{Vk} \cdot \Phi(\hat{\mu}_{Dk}) + \hat{\sigma}_{D,V}\phi(\hat{\mu}_{Dk})$. We can distinguish three effects of „more side effects" on the expected volume: By decreasing $\Phi(\hat{\mu}_{Dk})$, the effect of „more side effects" on $\hat{\mathrm{E}}(V_{ik})$ is unambiguously negative, through $\hat{\mu}_{Vk}$, however, it is positive. The effect of „more side effects" on $\hat{\sigma}_{D,V}\phi(\hat{\mu}_{Dk})$ depends on whether $\hat{\mu}_{Dk}$ is greater (negative) or less than zero (positive) before changing the level of the attribute „side

**Table 3 Estimation Results**[5]

| | Generalised Model | | | | | | Tobit | |
|---|---|---|---|---|---|---|---|---|
| | **Choice Propensity** | | **Volume Propensity** | | **Preference Propensity** | | | |
| **Variable** | param. | t-value | param. | t-value | param | t-value | param. | t-value |
| Constant | -1.323 | -14.17 | 24.73 | 8.06 | -18.48 | -4.93 | -38.93 | 14.45 |
| Daily therapy cost 1.83 DM | 0.701 | 9.88 | 8.23 | 4.54 | 23.12 | 11.76 | 25.59 | 13.65 |
| Daily therapy cost 4.47 DM | 0.272 | 3.85 | 0.56 | 0.33 | 8.13 | 3.95 | 8.66 | 4.41 |
| Daily therapy cost 8.00 DM | -0.264 | -3.10 | 0.80 | 0.40 | -2.87 | -1.09 | -7.14 | -3.14 |
| Cartilage Protection ensured | 0.970 | 16.90 | 4.20 | 2.87 | 28.49 | 14.85 | 29.04 | 18.51 |
| More side effects than reference product | -0.597 | -6.96 | 6.36 | 2.71 | -11.81 | -4.19 | -15.70 | 5.39 |
| Less side effects than reference product | 0.304 | 5.27 | -0.13 | -0.10 | 7.83 | 4.84 | 8.81 | 7.40 |
| Efficacy as reference product | 0.635 | 10.23 | 1.16 | 0.88 | 16.53 | 8.31 | 18.99 | 11.53 |
| Manufacturer: Stellapharma | 0.017 | 0.24 | -2.42 | -1.34 | -1.16 | -0.54 | -0.02 | -0.01 |
| Manufacturer: Lunapharma | -0.077 | -0.99 | -3.25 | -1.58 | -4.81 | -2.03 | -3.49 | -1.73 |
| **Variances and Correlations** | | | | | | | | |
| Std.Dev. | 1 | | 20.12 | 28.70 | 25.01 | 33.43 | 27.87 | 42.08 |
| Correlation of errors with choice errors | | | -0.47 | -10.67 | 0.85 | 25.41 | | |
| **McKelvey-Zavoina** | 39.8 % | | 5.2 % | | 43.2 % | | | |
| | $\chi(2)$ | p-value | $\chi(2)$ | p-value | $\chi(2)$ | p-value | | |
| **Reset 23** | 4.33 | 0.11 | 3.22 | 0.20 | 2.33 | 0.31 | | |

Since all variables enter the model as dummy variables, it is possible to compute the relative importance of each attribute by comparing the relative size of the parameters for each latent equation (figure 4).

[5] t-values computed using heteroscedasticity consistent variance covariance matrix. The implementation of the maximisation of (11) was done using the routines in the MAXLIK library of the econometric computer language GAUSS. Procedures to built the log likelihood (11) and its gradients are available from the authors upon request. We used the routines in the econometric package LIMDEP for the estimation of the Tobit model. The test procedures were implemented in GAUSS. They are available from the authors upon request

### 3.4 Estimation Results

We present the estimation results of the estimated generalised three equation Tobit model together with the results of a simple Tobit estimation in table 3 The variables are the attribute levels from table 1 coded as dummy variables that equal 1 if the attribute level is present in the profile and 0 else. For each of the attributes one level is chosen as a reference.

In each equation the price level *5,88 DM* was chosen as the reference category. The estimation shows that a price level of *1,83 DM* implies a significant higher choice propensity and a price level *8,00 DM* a significant lower choice probability compared to the reference category. Using the formula for the choice probability, $P(D_{ik}=1)=P(\varepsilon_{Dik}>-\mu_{Dk})=\Phi(\mu_{Dk})$ we estimate the prescription probability[4] for a profile „A" ( *cartilage protection ensured, more side effects than reference, less efficacy than reference, manufacturer N.N.* ) at a price level of *1,83 DM* to be 40,1 % and for the same profile at a price level of *8,00 DM* to be 11,2 %.

Cartilage protection significantly raises prescription propensity. For profile „A" at a price level of *4,47 DM* the estimated prescription probability is 24,9 %, if cartilage protection is lacking, the estimated probability drops to 5 %. Side effects are very important in both directions. The estimated 24,9 % prescription probability for profile „A" at price level *4,47 DM* is raised to 58,8 %, if there are less side effects than for the reference product. Lack of efficacy reduces prescription probability significantly. The estimated prescription probability for Profile „A" at price level *4,47 DM* is 48,2 % if efficacy is considered as large as for the reference product instead of less. The three levels of the attribute „manufacturer" have no significant different impact on choice probabilities. Preference propensity falls with higher price levels, lack of cartilage protection, more side effects and lack of efficacy. The manufacturer label „Lunapharma" has a negative impact on preference. The attribute relatively most important for preferences is „cartilage protection". Compared to choice propensity, the side effects are somewhat less important. In terms of relative importance of attributes, preference and choice propensity are not so much different.

[4] Choice and prescription probability are from now on used as synonyms

**Figure 2:**
**Percentages of Preference Scores and Volume Values in 10 % bins**

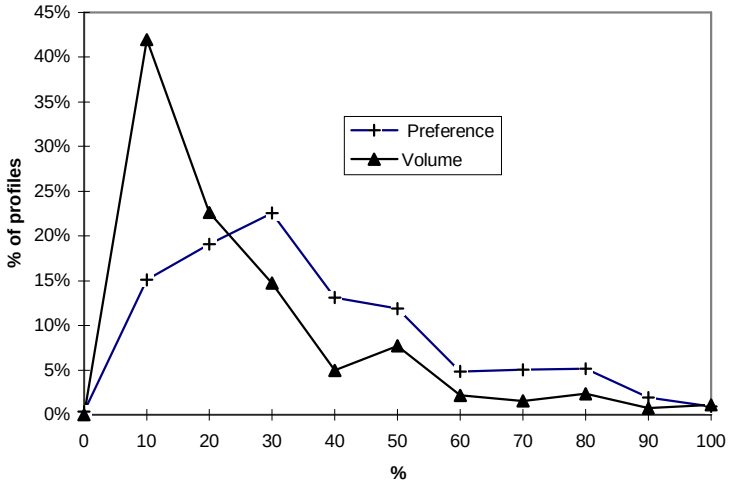
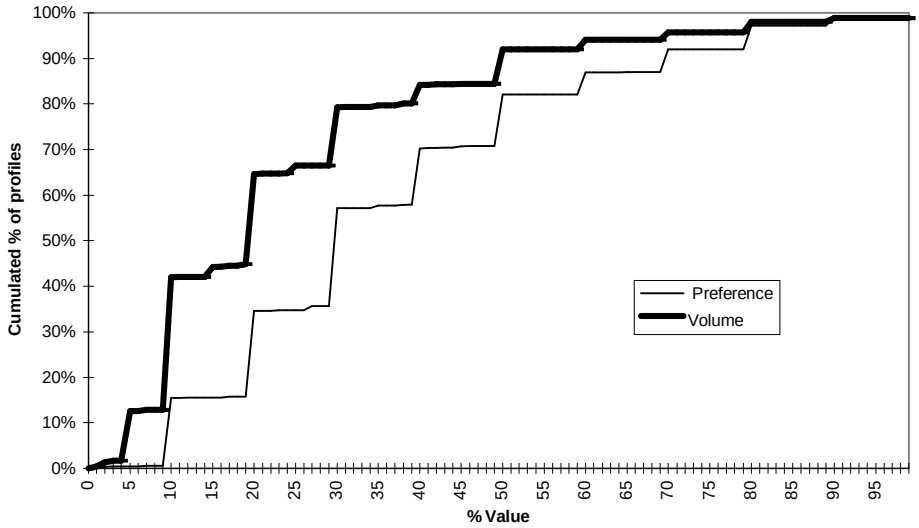



**Figure 3:**
**Cumulated Percentages of Preference Scores and Volume Values in 1 % bins**




The histogram in figure 2 shows that the shape of the distribution of volumes and preferences is quite different. Figure 3 shows the cumulated percentages (in 1 % bins), illustrating the differences between volume and preferences in another way. For example, 60 % of the evaluated profiles have a preference score of 40 % or less, but 60 % of the evaluated profiles have less than 20 % of prescription volume.

According to the questionnaire layout presented in section 2.1, the 124 physicians had to decide for each of the 20 profiles whether he/she would prescribe a medicament with that profile, and if yes, to which extent: The percentage of arthrosis patients for which the product would be prescribed provides the volume information for the model. For the chosen profiles, a preference score had to be given, too. This preference range was defined from a 0 to 100 % equivalence to an „ideal“ product. The descriptive information in table 2 shows that from the profiles evaluated by the physicians, on average 42 % were accepted and would be prescribed. For those profiles, the average preference score value is 36.4 %, and the average volume (patients for which the product would be prescribed) is 24.1 %.

**Table 2: Sample descriptives**

| | Mean | Std.Dev. | Min. | Max. |
|---|---|---|---|---|
| Physician´s Age in Years | 43.1 | 7.6 | 25 | 59 |
| Profile Preference Score | 36.4 | 20.4 | 0 | 100 |
| Profile Prescription Volume | 24.1 | 21.7 | 1 | 100 |
| Total Sample Individuals | 124 | | | |
| Share of Internists | 71% | | | |
| Share of practitioners | 29% | | | |
| Share of male physicians | 85% | | | |
| Profiles evaluated | 2480 | | | |
| Profiles accepted | 42% | | | |

In figure 2 we arranged the values given for the volume and preference information in 10 % bins and computed percentages of observations.

## 3.2 Factorial Design

The study layout or factorial design was developed in a way to be able to incorporate in the model the institutional aspects described in section 3.2. Product profiles represent combinations of the attribute levels in table 1. The four price levels span the price range on the market as well as price levels discussed by the manufacturer before market entry. The market leading product with well accepted efficacy and given side effect was chosen as a reference product. The attribute level of the attribute „manufacturer“ represent well known medicament manufacturers (we use aliases here) and a no name manufacturer (N.N.)

**Table 1: Factorial design**

| **Attribute** | **Attribute Level** |
|---|---|
| Price<br>(Daily therapy cost) | 1.83 DM<br>4.47 DM<br>5.88 DM<br>8.00 DM |
| Cartilage Protection | Provided<br>Not provided |
| Side effects | More than reference product<br>As much as reference product<br>Less side than reference product |
| Efficacy | Efficacy as reference product<br>Efficacy significantly less than reference product |
| Manufacturer | N.N.<br>Stellapharma<br>Lunapharma |

From the possible 144 products profiles that are possible by combining the attribute levels, we extracted 20 profiles of a reduced design using the algorithms described in section 2.1.

## 3.3 Description of the data

The sample of physicians consist of 124 physicians that had to assess each of the 20 profiles in the reduced design. Demographic characteristics of the sample individuals are given in table 2

# 3. Empirical Analysis

## 3.1 Background of the Study

The data for the application of the model developed in the previous sections come from a joint market entry study conducted by the Institute for Statistics and Econometrics, University Frankfurt, IMS GmbH, Frankfurt, and Heumann Pharma, Nürnberg.[3]. Consumers of traditional arthrosis medicaments bear the risk of cartilage damage as a severe side effect. A newly designed medicament that was planned to be introduced in the German market in summer 1994 aimed to avoid this risk by using an added substance that ensured cartilage protection. Before, a physician could ensure cartilage protection only by the prescription of an additional drug.

Two problems were associated with the new medicament: Firstly, it was controversially discussed whether the additional substance ensuring cartilage protection would cause less efficacy of the arthrosis drug. Some studies indicated that efficacy was indeed lessened, others could not find a negative effect. Secondly, the question whether the added substance for cartilage protection might cause additional side effects was undecided among pharmacologists. As a consequence, the market success of the product was considered very much depending on the prescribing physicians' opinion about the effect on efficacy and side effects caused by the cartilage protective substance.

Prices of arthrosis drugs on the market did significantly vary at the time of the study. Low priced generics, i.e. no name products exploiting expired patents on substances, competed against high priced brand products. The possible market entry price range for the new medicament was thus very wide. Another aspect complicated product pricing: Shortly before the study was carried out, the German public health system had undergone a severe reform. One of the fundamental changes implied by the German health reform in 1993-1994 (GHR) was that physicians' cost awareness was very much sharpened: Physicians now faced a rigid budget for prescriptions for their patients that were members of the public health insurance.

[3] The market entry study was carried out in a joint project of Heumann Pharma, Nürnberg, IMS GmbH, Frankfurt and the Institute for Statistics and Econometrics, Frankfurt University. The data used in our paper is the unmodified original data, also used by Hujer et al. (1996). We have been asked to modify the labels in the original questionnaire (not the data): We changed the price labels and the labels for the manufacturers compared to the original questionnaire.

$$\log L(\theta) = \sum_{i=1}^{N}\sum_{k=1}^{K} \log\left[1 - \Phi\left(\mu_{DK}\right)\right] \cdot \left(1 - D_{ik}\right) +$$

$$\log\left[\Phi\left(\frac{\mu_{VK} - \mathrm{E}\left(\tilde{\varepsilon}_{Dik}\right)}{\mathrm{Var}\left(\tilde{\varepsilon}_{Dik}\right)}\right) \cdot \frac{1}{\sigma_V \sigma_U} \phi_{BV}\left[V_{ik} - \mu_{VK}, U_{ik} - \mu_{UK}, \rho_{V,U}\right]\right] \cdot D_{ik}$$

with

$$\mu_{DK} = \alpha_{0D} + \sum_{j=1}^{J}\sum_{m=1}^{M_{j-1}} \alpha_{jmD} x_{jm}$$

$$\mu_{VK} = \alpha_{0V} + \sum_{j=1}^{J}\sum_{m=1}^{M_{j-1}} \alpha_{jmV} x_{jm}$$

$$\mu_{UK} = \alpha_{0U} + \sum_{j=1}^{J}\sum_{m=1}^{M_{j-1}} \alpha_{jmU} x_{jm}$$

$$\theta = \left(\alpha_{0D}, \alpha_{0U}, \alpha_{0V}, \alpha_{11D}, \ldots, \alpha_{11V}, \ldots, \alpha_{11U}, \ldots, \sigma_V, \sigma_U, \rho_{DV}, \rho_{DU}, \rho_{UV}\right),$$

$$\rho_{D,V} = \frac{\sigma_{D,V}}{\sigma_D \sigma_V}, \rho_{D,U} = \frac{\sigma_{D,U}}{\sigma_D \sigma_U}, \rho_{V,U} = \frac{\sigma_{V,U}}{\sigma_V \sigma_U},$$

$$\tilde{\varepsilon}_{ik1} = \left[\varepsilon_{Dik} \middle/ \begin{pmatrix} \varepsilon_{Vik} = V_{ik} - \mu_{VK} \\ \varepsilon_{Uik} = U_{ik} - \mu_{UK} \end{pmatrix}\right]$$

$$\mathrm{E}\left(\tilde{\varepsilon}_{Dik}\right) = \frac{1}{\sigma_{\varepsilon_V}^2 \cdot \sigma_{\varepsilon_U}^2 - \sigma_{V,U}} \left(\left(V_{ik} - \mu_{DK}\right)\left(\sigma_U^2 \sigma_{D,V} - \sigma_{D,U}\sigma_{D,V}\right)\right.$$
$$\left. + \left(U_{ik}^{*} - \mu_{UK} - z_{i2}'\pi_2\right)\left(\sigma_V^2 \sigma_{D,U} - \sigma_{V,U}\sigma_{D,U}\right)\right)$$

$$\mathrm{Var}\left(\tilde{\varepsilon}_{Dik}\right) = 1 - \frac{1}{\sigma_V^2 \cdot \sigma_U^2 - \sigma_{V,U}} \left(\sigma_U^2 \sigma_{D,V} - \sigma_{D,U}\sigma_{D,V}\right)\sigma_{D,V}$$
$$+ \left(\sigma_V^2 \sigma_{D,U} - \sigma_{V,U}\sigma_{D,V}\right)\sigma_{D,U}. \qquad (11)$$

$\phi_{\mathrm{BV}}(\,)$ denotes the density function of the bivariate standard normal distribution. Maximisation of $\log L(\theta)$ yields estimates for the parameter vector $\theta$. The expectation for volumes is given by

$$\mathrm{E}\left(V_{ik}\right) = \mu_{VK} \cdot \Phi\left(\mu_{DK}\right) + \sigma_{D,V}\phi\left(\mu_{DK}\right) \qquad (12)$$

In order to test the model assumption and model specifications, we use following statistics: To assess the goodness of fit of our latent variable specification, we compute McKelvey-Zavoina´s Pseudo R-squared for each of the latent equations. To test the normality assumption and correct specification we compute generalized score test statistics (White 1982,1983) and perform RESET tests. Details of the derivation of the test statistics and goodness of fit statistics are given in the appendix

The third latent variable is the preference score that individual *i* attaches to profile *k*.

$$U_{ik}^{*} = \alpha_{0U} + \sum_{j=1}^{J} \sum_{m=1}^{M_j - 1} \alpha_{jmU} x_{jm} + \varepsilon_{ikU} \tag{8}$$

The censoring rules which translate the latent variables $D_{ik}^{*}$, $V_{ik}^{*}$, $U_{ik}^{*}$ into their observable counterparts $D_{ik}$ - a dummy variable that takes on 1 profile *k* is chosen by individual *i* and 0 else -, $V_{ik}$ -the observable purchase volume- , and $U_{ik}$ - the observed preference score - are as follows

$$\begin{aligned} D_{ik} &= 1 \text{ if } D_{ik}^{*} > C_A \\ &= 0 \text{ else} \\ V_{ik} &= V_{ik}^{*} \text{ if } D_{ik}^{*} > C_B \\ &= 0 \text{ else} \\ U_{ik} &= U_{ik}^{*} \text{ if } D_{ik}^{*} > C_C \\ &= 0 \text{ else.} \end{aligned} \tag{9}$$

Rearranging the profiles that have been chosen by consumers and those that were not we can write the joint probability of the sample realisation to be

$$\prod_{0} \left[ P\left(D_{ik} = 0\right) \right] \prod_{1} \int_{0}^{\infty} f\left(\varepsilon_{ikD}, \varepsilon_{ikV}, \varepsilon_{ikU}\right) d\varepsilon_{ikD} \tag{10}$$

with $f(\cdot)$ as the joint density of the random vector $\varepsilon_{ik} = \left(\varepsilon_{ikD}, \varepsilon_{ikV}, \varepsilon_{ikU}\right)$ . Assuming that $\varepsilon_{ik}$ is independent, identical, trivariate normal, $\varepsilon_{ik}$ , ~ N(0,$\Sigma$), with

$$\Sigma = \begin{bmatrix} 1 & \sigma_{DV} & \sigma_{D,U} \\ \sigma_{D,V} & \sigma_{V}^{2} & \sigma_{V,U} \\ \sigma_{D,U} & \sigma_{V,U} & \sigma_{U}^{2} \end{bmatrix}$$

and setting $C_A = C_B = C_C = 0$, we have the following operational log likelihood:

Zabel (1993) has criticised this assumption for being too restrictive[1]: The product attributes that might lead a consumer to purchase a product in the first place might not be the same as the ones that will lead to a high consumption volume. Status considerations, for example, might lead to a purchase despite a relative high price but a low consumption in the presence of high variable costs [2].

An example will be presented in our empirical analysis where we study the prescription behaviour of physicians after a cost cutting health reform. In this environment, a physician might well prescribe a new medicament with appropriate features despite a relative high price, but the amount of his prescriptions might be small, given the budget constraints imposed by the health reform.

The ability to distinguish the factors that influence market chances of a product in terms of reaching enough consumers and sales volumes is thus an important feature for a model for consumer behaviour. Following Zabel´s (1993) idea of formulating a generalised Tobit model, we develop an econometric model that is able to model purchase probability, purchase volume and preferences simultaneously but that does not impose the restrictive assumptions of the Tobit model. We specify three equations with latent endogenous variables that describe consumer behaviour:

The latent index $D_{ik}^{*}$ controls the unobservable propensity of consumer *i* to accept profile *k* as a product he/she wants to consume. This propensity depends on the attributes of profile *k* and a set of parameters, $\alpha_{0D}, \alpha_{jmD}$ and an unobservable random variable $\varepsilon_{ikD}$ that measures unobserved heterogeneity over profiles and individuals:

$$D_{ik}^{*} = \alpha_{0D} + \sum_{j=1}^{J} \sum_{m=1}^{M_j - 1} \alpha_{jmD} x_{jm} + \varepsilon_{ikD} \tag{6}$$

The notation of attribute levels and weight parameters is chosen as in (1). A second equation describes the consumer's propensity to consume a certain volume of profile *k.* Again, this latent variable depends on a set of parameters, the profile attributes and a random error term:

$$V_{ik}^{*} = \alpha_{0V} + \sum_{j=1}^{J} \sum_{m=1}^{M_j - 1} \alpha_{jmV} x_{jm} + \varepsilon_{ikV} \tag{7}$$

[1] Zabel´s (1993) analysis is placed in the context of the labour supply decision.
[2] Think for example of the consumption of cellular telephone services

$V_{ik}^*$: Purchase propensity of consumer i for profile k

$\alpha_o$, $\alpha_{jm}$: Coefficients

$x_{jm} = \begin{cases} 1, & \text{if attribute level m of attribute j is present in profile k} \\ 0 & \text{else} \end{cases}$

$J$: Number of attributes

$M_j$: Number of attribute levels of attribute j

$\varepsilon_{ik}$: Stochastic error term

The unobservable purchase propensity translates into observable (in the sense of being available from the questionnaire) purchase volume $V_{ik}$ with the following threshold model:

$$\begin{aligned} V_{ik} &= V_{ik}^* \text{ if } V_{ik}^* > C \\ V_{ik} &= 0 \text{ if } V_{ik}^* \le C \end{aligned} \tag{2}$$

Identification can be assured by setting $C = 0$, leading to the likelihood function for a sample of $N$ individuals assessing $K$ profiles:

$$L = \prod_{i=1}^{N} \prod_{k=1}^{K} \left[ \mathrm{P}\left(V_{ik}^* \le 0\right) \right]^{1-Y_{ik}} \left[ \mathrm{f}\left(V_{ik} \mid V_{ik}^* > 0\right) \cdot \mathrm{P}\left(V_{ik}^* > 0\right) \right]^{Y_{ik}} \tag{3}$$

$Y_{ik}$ is a dummy indicator that equals one if profile $k$ is accepted for purchase by consumer $i$ and zero else. Assuming a normal distribution for $\varepsilon_{ik} \sim N(0,\sigma)$, and define $\mu_k = \alpha_0 + \sum_{j=1}^{J} \sum_{m=1}^{M_j - 1} \alpha_{jm} x_{jm}$, we obtain the Tobit Likelihood

$$L = \prod_{i=1}^{N} \prod_{k=1}^{K} \left( \Phi\left(-\mu_k \cdot \sigma^{-1}\right) \right)^{1-Y_{ik}} \cdot \left( \sigma^{-1} \cdot \phi\left( \left(V_{ik} - \mu_k\right) \cdot \sigma^{-1} \right) \right)^{Y_{ik}} \tag{4}$$

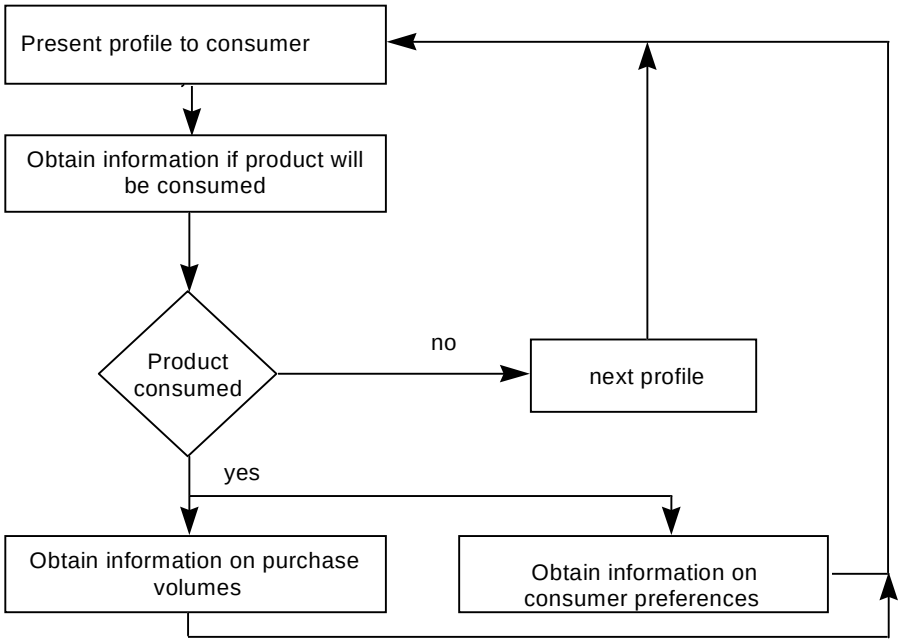
with $\Phi(\cdot)$ and $\phi(\cdot)$ defining the cumulative density and the density of the standard normal respectively.

Note the Tobit model's idiosyncrasy that the expected volume of purchases

$$\mathrm{E}\left(V_{ik}\right) = \left(\mu_k\right) \cdot \Phi\left(\mu_k \cdot \sigma^{-1}\right) + \sigma\varphi\left(\mu_k \cdot \sigma^{-1}\right) \tag{5}$$

and the purchase probability $\mathrm{P}\left(u_{ik}^* > 0\right)$ are dependent on a single set of explanatory variables and the same parameter vector.

**Figure 1:**
**Questionnaire Filtering**




## 2.2 Econometric Model and Specification Testing

In a revealed preference framework, the usage of probabilistic choice models is a constantly refined state of the art technology (Chintagunta, Jain, Vilcassim 1991, Chintagunta 1992,1994, Elrod, Keane 1995). The basic assumption of a probabilistic choice model is that an individual chooses one alternative among a given choice set. This alternative is the one that gives him/her maximum utility, given the attributes of the choices. The approach in a full profile environment is different. Here, the consumer is presented a set of product profiles and he/she is not restricted to choose only one. One consumer might choose all profiles for consumption, another individual none. Another point is that probabilistic choice models focus on purchase probabilities (and underlying preferences of course). Volume considerations are not part of the model.

Econometric models for censored dependent variables provide a more appropriate tool for studies using profile design data. Hujer et al. (1996) show the application of a censored regression (Tobit) model. Their idea is to introduce a unobservable variable $V_{ik}^{*}$ that denotes the purchase propensity of consumer *i* for a product profile *k* . This profile is described by a set of *J* attributes with $M_j$ attribute levels:

$$V_{ik}^{*} = \alpha_0 + \sum_{j=1}^{J} \sum_{m=1}^{M_j - 1} \alpha_{jm} x_{jm} + \varepsilon_{ik} \qquad (1)$$

from the products already established in the market. So will the behaviour of the consumers.

The profile method helps to circumvent this drawbacks by explicitly providing a choice set of product profiles that may include descriptions of hypothetical but feasible products. The idea is to set up a factorial design by defining the relevant attributes of a product (price, layout, ect.) and the attribute levels ( prices, different layout, etc.). Note that a certain amount of a priori information has to be provided: Usually, a selection of the presumably most important product attributes requires a thorough discussion and a considerable amount of market knowledge. After having determined the factorial design, the choice set for an individual is now perfectly controllable and the profile attributes are available without missing values or errors in variables.

Having determined the factorial design, the complete set of possible alternatives is reduced to a smaller set of profiles which are then presented to a sample of consumers. Algorithms and tables for reducing symmetrical and asymmetrical designs into orthogonal arrays were laid down by Addelman (1962) and Placket, Burman (1946). The algorithms are implemented in statistical software packages, e.g. SAS or SPSS.

The traditional Conjoint method requires the consumer to evaluate the profiles (stimuli) of the reduced design according to a given preference scale or by ranking the alternatives. Hence, the traditional Conjoint method focuses on preferences, which, however, represent only one possible way to analyse consumer behaviour. In our approach, we use the questionnaire that gathers information in a different way. Figure 1 presents the routing of the consumer through the questionnaire.

Figure 1 shows that we collect qualitative information (product choice yes/no) as well as quantitative information on purchase volumes and preference scores. In order to utilise this information one has to employ an econometric model that is prepared to deal with this combination of qualitative and censored information.

In our paper, we try to avoid the drawbacks of the Conjoint method and the restrictiveness of the Tobit approach by formulating an econometric model that is able to model purchase probability, purchase volume and preferences simultaneously. The model is an extension of Zabel´ s (1993) generalised Tobit with an additional preference equation in a full profile setup.

The applicability of the approach is tested in an empirical study using a full profile design of medicaments evaluated by a sample of German physicians. This example is particularly useful for the context of this paper, since it looks at the markets for medicaments after the German Health Reform (GHR). After GHR, German physicians´ cost awareness was sharpened by introducing fixed budgets for prescriptions for patients that are members of the public health insurance. Since the physician is the one who decides whether to prescribe the medicament in the first place and to what extend, the distinction of a physician's product preferences and the prescription probability and prescription volume for a certain medicament profile become especially important.

In the following sections we proceed as follows: In section two, we first develop the questionnaire layout that is needed to gather the information for model. After that, the econometric model, the likelihood function and model tests are presented. Section three contains the empirical application, the estimation results and some sensitivity analyses. We conclude in section four with a summary of the main results.

## 2. Study Design and Econometric Modelling

### 2.1 Full Profile vs. Revealed Preference Approach and Questionnaire Layout

There are two distinct ways to model consumer behaviour empirically. One is the revealed preference approach where records of purchases from household panel surveys are used for econometric modelling. Since the recruited panel members usually provide a large set of household characteristics like household composition, income, demographics, etc., and new information technology facilitates gathering of detailed consumer behaviour data, the database that can be collected contains rich information on consumer choices.

For market entry analyses, however, the revealed preference has its drawbacks. At the stage of designing, initial pricing and market entry of product innovations, data on observed purchases do not help, since the new product may differ in important attributes

## 1. Introduction

The empirical analysis of consumer behaviour on micro level can be conducted by using records on purchases of members of household panels („revealed preference“ approach) or by developing an experimental design where profiles of products have to be evaluated by a random sample of households or individuals („profile“ approach). The revealed preference approach has the advantage of providing detailed information on consumer level, especially when the study is designed as a panel. Panel econometrics can then be utilised to model dynamic consumer behaviour. Recent microeconometric studies that use revealed preferences from consumer panels include Gupta et al. (1996) and Grammig et al. (1997).

The second approach that implies setting up an experimental design and obtaining information on product profiles is favourable if predictions of the market acceptance and market potential of product innovations are needed. The profile approach has traditionally been the domain of the Conjoint method, a standard tool for market researchers (Wittink, Cattin 1989).

The Conjoint method has the idiosyncrasy that it focuses exclusively on consumer preferences. Modelling preferences, however, is only one possible way to look at consumer behaviour and in some cases not the most important one. Other aspects of consumer behaviour, i.e. purchase probability and purchase volume, are not considered. Although the derivation of market shares from preference values is an exercise often applied in - and demanded from - Conjoint analyses, the method to do this projection is not stringent. Every textbook on market research includes the lemma that preferences have to exceed certain thresholds should they actually turn into observable purchases. A threshold model, however, is not integral part of the traditional Conjoint approach.

Recent studies try to widen the narrow focus of traditional Conjoint Analysis by combining microeconometric threshold models and traditional full profile settings. Employing a Tobit Model (Amemiya 1984,1985, Ronning 1991), Hujer et al. (1996) were able to model purchase probability and purchase volume in a consistent framework. The drawback of this approach is that by using a Tobit model, certain implicit restrictions have to be considered. Although purchase probability and purchase volume are modelled in a single framework, there is an assumption of the Tobit approach that has been criticised for being to restrictive (Zabel 1993). It is the assumption that the purchase probability and the decision to purchase a certain amount of a product depend on the same set of parameters.

# ABSTRACT

We propose the application of a three equation generalised Tobit to model different aspects of consumer behaviour in a full profile study design. The model takes into account that consumer behaviour can be measured by preference scores, purchase probability and purchase volume. We aim to avoid the drawbacks of traditional conjoint analysis where the latter two aspects are disregarded. Starting from a full profile design, we develop the appropriate questionnaire layout, the econometric model, the likelihood function and tests. The model is applied in a market entry study for an innovative medicament after a reform of Germany´s public health system in 1993-1994.

Keywords: generalised tobit, market entry study

JEL Classification: C35,M31,L65

# Modelling Consumer Behaviour in a Profile Design Using a Three Equation Generalised Tobit Model

**by**

**Joachim Grammig and Reinhard Hujer**

**Department of Economics**
**Johann Wolfgang Goethe-University, Frankfurt am Main**

**March 1997**

The paper is based on a joint project of the Insitute for Statistics and Econometrics, University Frankfurt, IMS GmbH, Frankfurt and Heumann Pharma, Nürnberg. We thank W. Niedermeier, Heumann Pharma, and H. Fryns, IMS GmbH, for providing us the permission to use the data for our study and K.-O. Maurer and M. Wellner, University Frankfurt, for helpful comments. K.-O. Maurer also provided the algorithms for setting up the dataset used for model estimation.

Joachim Grammig, Reinhard Hujer
Department of Economics
Johann Wolfgang Goethe-University
Mertonstraße 17
D-60054 Frankfurt am Main

Phone: +49 (0)69 798-28115 (Hujer) or 22893 (Grammig)
Fax: +49 (0)69 798-23673

E-Mail: grammig@wiwi.uni-frankfurt.de, hujer@wiwi.uni-frankfurt.de

WWW: http://www.wiwi.uni-frankfurt.de/professoren/empwifo/