

---

## INTRODUCTION

In modern societies, private consumption, purchasing and using up goods and services have become constitutive elements of social life. Political, institutional and individual actions are oriented towards consumption. Therefore, information on consumption is very important for monitoring and analyzing societal development. In the last decade, sociology and psychology have started to pay attention to consumption and consumer behavior. Societal problems like the BSE crisis and issues like sustainable development demand more knowledge on causal mechanisms underlying consumption patterns.

Behavior patterns in modern societies with multiple options to act cannot be easily understood using simple social categories (see the controversy on individualization, *FRIEDRICH* 1998). Social reality is transformed by individually organized lifestyle as well by differentiated household structures. This seems to be acknowledged by social sciences at least in the context of theoretical perspectives. There is a lack of empirical information, especially longitudinal information on the dynamics of consumer behavior. Therefore social and economic research needs also longitudinal and detailed data on product demand by consumers.

In Germany, there are two large standard data-generation programs which provide data and information for the social sciences:

- the Allbus Survey and
- the German Socio Economic Panel.

These datasets offer data on opinions and economic and demographic features of the households, but they are only marginally related to consumption. Actually, no standard data base for analyzing consumption of food or other goods and services is available. Consumption data collected and provided by the German National Bureau of Statistics and specific economic institutes proved to be good reference data, but they are not sufficient for analyzing the social contexts of consumption. Nevertheless, until now there has been no initiative to establish standard data programs focused on consumption.

In contrast, non-academic market research institutes provide information and data services on consumption on a very high level. They run consumer panels, continuously producing longitudinal behavioral data on the consumption of individuals and households. As these services are aimed at supporting companies in marketing their specific products, consumer panels produce detailed information on the demand for specific consumer goods.

Actually, they are continuously recording the purchase event history of product flow into the household over years. Combined with socio-economic and demographic information on the household as well as surveys on attitudes of specific reference persons in the household, the consumer panel data program seems to be of high value for social science consumer analysis. Unfortunately, standard access to the original data was not possible recently. This resulted in only little systematic social and economic research with consumer panel data.

To open the data to public access, Georgios Papastefanou and Peter Schmidt contacted GfK<sup>1</sup> in 1997 to establish to an agreement for using the data for scientific analysis. GfK was interested in cooperation and provided the data. At the beginning of 1999 ZUMA organized an interdisciplinary working group, whose task was to examine the potential as well as the manageability of the data. Members of the group were Georgios Papastefanou, Peter Schmidt, Yasemin Boztug, Lutz Hildebrandt, Axel Börsch-Supan, Matthias Fengler, Joachim Winter, Frank Faulbaum, Petra Stein, John Thøgersen, Corinna Willhöft, Uwe Fachinger, Norbert Schneider, Doris Rosenkranz, Kerstin Hartmann, Ulrich Oltersdorf, Bärbel Geisel, Hartmut Lüdtkke, Dieter Ohr. The department of Income and Consumption at ZUMA organized and systematized the data and delivered them to the members of the working group. In October 1999 the working group members met again for a conference. They reported results and discussed the shortcomings of the data with participants from social, market and economic research. In the present special edition of “ZUMA-Nachrichten” a selection of these analyses is presented to demonstrate the scientific and practical value of this set of consumer panel data. Table 1 provides an overview of the different approaches of the analyses.

The first paper is written by *HANS-GEORG PRESTER*, who presents an overview on the production of consumer panel data at GfK. He achieves transparency by documenting methods and data collection procedures which are driven by the ultimate goal of providing a useful information service for industrial customers. He gives first hand information on the data collection programs at GfK. We also get a clear insight into specific strategies of analysis which might be interesting für social research. The paper expresses a deep interest in intensifying the cooperation between socio-economic and market research.

The paper of *HARTMUT LÜDTKE* and *JÖRG SCHNEIDER* starts by considering how behavioral life style patterns determine meanings of personal and social communication and exclusion. They examine the data for life style strategies that can also be observed in purchases of everyday consumer products. By including indicators of social position and

---

1 Gesellschaft für Konsumforschung, Nürnberg

attitudinal expression they further differentiate behavioral patterns along subjective and social meanings. One of the central findings is that social position and mentality correlates with fast moving consumer good purchases like food, beverages and hygiene products.

Table 1: Overview of the different approaches

<b>Name and Titel</b>	<b>Dependent Variable</b>	<b>Independent Variable</b>	<b>Method</b>	<b>Results</b>
<i>ROGER BERGER:</i> Consumer panel data and rational choice based theories of myopic habit formation	- coffee - beer - yogurt	- habit stock - price - income	- regression analysis	- OLS-regression of habit stock to logarithmic quantity of consumed good - OLS-regression of habit stock, logarithmic price and the logarithmic income on the logarithmic quantity of consumed: beer, coffee, yogurt - elasticities of demand for habit stock, price and income: beer, coffee, yogurt
<i>YASEMIN BOZTUG/LUTZ HILDEBRANDT:</i> Nonparametric modeling of buying behavior in fast moving consumer goods markets	- promotion	- loyalty - price	- multinomial logit model	- different model types for parametric and nonparametric choice models - estimation results of the continuous variables made by backfitting - estimation results of the continuous variables made by marginal integraion

Table 1: Overview of the different approaches (continued)

<b>Name and Titel</b>	<b>Dependent Variable</b>	<b>Independent Variable</b>	<b>Method</b>	<b>Results</b>
<i>MATTHIAS FENGLER</i> <i>/JOACHIM WINTER:</i> Price-setting and price-adjustment behavior for fast-moving consumer goods	- ground coffee	- time	- frequencies	- distribution of prices for the three fast-moving consumption goods - psychological pricing points for the three fast-moving consumption goods
<i>UWE FACHINGER:</i> Continuity or discontinuity in the decision of the application of the income of private households?	- roasted pure coffee - universal detergent - yeast	- household income - age of the housekeeping person	- product-limit estimation - proportional hazard model - cox regression	- results of the Kaplan-Meier estimation for subgroups of WGS 12 - results of the Kaplan-Meier estimation for subgroups "with caffeine" of WGS 12 - results of the Kaplan-Meier estimation for subgroups of universal detergent - the median household income in age classes - results of the Kaplan-Meier estimation for subgroups of yeast
<i>HARTMUT LÜDTKE</i> <i>/JÖRG SCHNEIDER:</i> Can patterns of everyday consumption indicate lifestyles?	- beverages - food - personal hygiene	- social position - householder's age, education, occupation - size of building, household, residence - number of children	- cluster analysis - analysis of variance	- consumer lifestyle typology

Table 1: Overview of the different approaches (continued)

<b>Name and Titel</b>	<b>Dependent Variable</b>	<b>Independent Variable</b>	<b>Method</b>	<b>Results</b>
<i>DIETER OHR:</i> Purchasing healthy food in Germany	- percentage of expenditures for healthy food	- education - age - children under 14 - income - east vs. west Germany	- cross sectional and panel approaches	- healthy food in the GfK consumer data - household expenditures for healthy food - relative frequencies for the percentage of healthy food articles in the GfK data - attitudes on nutrition - the purchase of healthy food and its determinants - a multivariate model for the purchase of healthy food
<i>GEORGIOS PAPASTEFANOU:</i> Household income level and liquor consumption	- Amount of alcoholic beverage per quarter	- purchase frequency - income position of the household - household structure	- regression analysis	- effects of income position and its change on the amount of purchased liquors in the first quarter of 1995 - effects of income position and change on purchased quantity of liquor in the first quarter of 1995 by: educational status of household head, employment status of housekeeping person

Table 1: Overview of the different approaches (continued)

<b>Name and Titel</b>	<b>Dependent Variable</b>	<b>Independent Variable</b>	<b>Method</b>	<b>Results</b>
<p><i>NORBERT F. SCHNEIDER / DORIS ROSENKRANZ / KERSTIN HARTMANN:</i>                      Living arrangements and consumption. Plead for a lifestyle orientated market research.</p>	<p>- crispbread</p>	<p>- Living arrangements                      - occupational status                      - age</p>	<p>- Analysis of variance</p>	<p>- significant differences between forms of living with and without children                      - significant differences between forms of living and age                      - significant differences between forms of living by occupational status                      - significant differences between selected forms of living by age and occupational status</p>
<p><i>JOHN THÖGERSEN:</i>                      Is a sustainable consumption pattern gradually emerging in Germany?</p>	<p>- product groups with environment-relevant packaging attributes</p>	<p>- attitude data</p>	<p>- factor analysis</p>	<p>- product groups with environment-relevant packaging attributes                      - the propensity to choose environment-friendly options in the studied product groups in first and second half of 1995                      - the correlation between general environmental concern and the propensity to choose an environment-friendly version of various convenience goods                      - average correlations between propensities to choose environment-friendly options in different categories</p>

*JOHN THÖGERSEN*'s paper deals with the problem of sustainable development and its foundation in individual behavior patterns. Environmentally sound behavior is based on the consumption of fast moving consumer goods for household maintenance. In his main question Thögersen asks if there is a mechanism of spillover, where a proenvironmental purchasing disposition (expressed in the purchasing of specific goods) is transferred to other environmentally relevant products. Using product information on packaging of purchased products, he also analyses whether spillover is supported by specific attitudes and values.

The analysis of *DIETER OHR* concentrates on aspects of health in the context of a high public sensitivity towards health risks in modern food production. He asks if purchasing of food products related to health issues is influenced by attitudes towards nutrition and health. He finds consistency between purchasing and attitudes towards healthy nutrition, even if the involvement is on a medium level. Differentiating the expenses for healthy food products according to age, education and region underlines that social position, mediated by socialization and cognitive competence, determines everyday life.

*MATTHIAS FENGLER* and *JOACHIM WINTER* start from a critical evaluation of theoretical economic models on price formation and examine the dynamics and variation of retail prices. They analyze the purchases of coffee, since it is a product bought frequently and having changing prices. Their results show the importance of psychological price thresholds for price settings.

*NORBERT SCHNEIDER*, *DORIS ROSENKRANZ* and *KERSTIN HARTMANN* deal with the social differentiation of household structures. They focus on crisp bread as a product and look for its demand in various types of households by taking into account marriage status, household size, age of household head and net household income.

*ROGER BERGER* performs a trial to exploit the special informational value of consumer panel data, namely the continuous and total measurement of purchases over a period of one year. This kind of data makes it possible to model habit effects on purchase decisions, which he reports for purchases of yoghurt, coffee and beer.

Dynamic, longitudinal information is also used by *UWE FACHINGER*, who is interested in risk evaluation by consumers. He argues that individual risk evaluation should be best detected when considering continuity of purchasing behavior. The available information on precise timing of a purchase allows him to apply hazard rate models on purchases of coffee, detergents and yeast.

Another type of longitudinal information, changes in net household income is used by *GEORGIOS PAPASTEEANOU*. In his paper he examines the hypothesis that there is more demand for alcoholic beverages (with more than 20 Vol. alcohol) in higher income classes than in

lower classes. By separating income changes from income level effects on the amount of liquors bought and by differentiating this covariation by household structures, he finds empirical support for different socio-psychological functions of alcoholic beverages consumption.

Finally, *YASEMIN BOZTUNG* and *LUTZ HILDEBRANDT* focus on methodological issues of modelling consumer behavior by discrete decision models. They show advantages and disadvantages of nonparametric and semi-parametric approaches as well as an application of the semiparametric approach to specific consumer panel data.

With this volume ZUMA is documenting its initiative to improve scientific access to special longitudinal data for social and economic research. The initial working group did a very good job in showing the research potential of the consumer panel data from various points of view. So, as our main effort was aimed at supporting the consumer related social and economic research, we now want to open the data to the scientific community. One can get the ZUMA version of the GfK ConsumerScan Panel data by just sending in the signed terms of agreement (this document is printed in the Annex, but also available on the attached CD-ROM). A summary description of the data sets and some explanation for matching transformations of the data is also included in the Annex. Detailed information on the data as well as on the survey documents is available on the CD-ROM and at [http://www.gesis.org/Dauerbeobachtung/Einkommen/Daten/Gfk/gfk\\_index.htm](http://www.gesis.org/Dauerbeobachtung/Einkommen/Daten/Gfk/gfk_index.htm).

The initiative to open access to consumer panel data for scientific research has been taken by ZUMA, but it would not have been fruitful without the special motivation of the participants of the working group, the constant support of Wolfgang Twardawa (GfK), the patient assistance of Jolantha Müllner in organizing the symposium, and the untiring efforts of Antje Springer and Matthias Fleck in helping produce this volume. We wish to thank all of them for their invaluable contribution to a new service for social and economic research.

### **Editors**

Georgios Papastefanou  
Peter Schmidt  
Axel Börsch-Supan  
Hartmut Lüdtke  
Ulrich Oltersdorf

### **References**

Jürgen Friedrichs (Ed.) 1998, *Die Individualisierungs-These*. Opladen: Leske + Budrich.



---

# CONSUMER PANEL RESEARCH OF GfK

*HANS-GEORG PRESTER*

## **Preliminary Remarks**

There is still reluctance of commercial market research and scientific research to deal with each other. A black and white view on this topic shows the following outlines:

On the one hand it is doubted that a profit oriented institute is always willing and able to fulfill the standards of scientific research. Focus on the price of research and the informational interests of the clients are seen as possible dangers which may reduce quality of research. Sometimes the commercial institutes contribute to this view, when they hide the details of their research. But the reasons for this covering are often that a) the clients want to be better informed than their competitors and b) the institutes want to secure their commercial success by protecting themselves against copies of their research instruments.

On the other hand you can hear complains about scientific research: it were slow, hair-splitting and would not help to provide solutions for the challenges of business. Besides the difficulties to get resources for scientific research, probably the rules of academic career („publish or perish“) are the breeding ground of these complains. Publications about research with complex multivariate methods in established academic fields help academic career, but this is seldom compatible with research about specific business problems.

All the better that despite these obstacles ZUMA and GfK have begun to cooperate. ZUMA has shown interest for the data coming from consumer panel research and GfK has delivered panel data. The following notes will therefore give an impression about GfK, consumer panel research, and the panel data.

## **1. Who is GfK?**

The GfK group (Gesellschaft für Konsumforschung) was established in 1934. In universities, business associations and industry a view disseminated, which Wilhelm Vershofen, one of the founders of GfK formaluted as follows: „In present days the insight is more and more accepted, that the consumer – in the sense of the end buyer - is the economic

factor of prime importance. His attitudes, his habits and his purchasing decisions determine the success of all products which are produced for the market, i.e. for sale.“ Consumer panel data help branded goods manufacturers and retailers to understand consumers' product preferences, attitudes and behavior patterns, and to tailor their goods and services to changing consumer behavior.

In the beginning of GfK secondary statistics (e.g. maps about regional differences in the spending power of consumers) and consumer, retailers and manufacturer surveys about consumption played the key role. The date of birth of a real consumer panel was 1957 (1000 households). This panel was continuously increased in the following years and later complemented by an individual panel..

Today the GfK Group includes 14 German subsidiaries and in addition 90 subsidiaries located in 43 countries worldwide. It's core business is the collection, analysis and interpretation of trading and consumer information on behalf of brand manufacturers, retailers and services companies and the media. With total revenues of DM 760 millions (1999) the GfK Group worldwide is one of the leading companies in a market research market of some DM 30 billion. These revenues were achieved with about 4,100 staff in 4 business divisions:

- Consumer Tracking (consumer panel research and mail panel)
- Non-Food Tracking (retail audits for consumer technology products – principally in the IT, telecommunications, consumer electronics, household goods, photographic and Do it Yourself products sectors)
- Media (TV audience ratings, quantitative and qualitative data on the use of other electronic and print media)
- Ad Hoc Research (information services about different stages of the marketing decision making process, e.g. product test, advertising pre-tests, monitoring of the use of advertising media, test market simulation etc.)

In 1999 the most outstanding event for GfK was its flotation. Since September 23, GfK has been quoted on the Frankfurt Stock Exchange. Since March 2000 GfK is a member of the M-Dax.

## 2. What is Consumer Panel Research?

### 2.1 The Methods of Consumer Panel Research

In consumer panel research data about purchasing behavior are continuously collected with the same instrument from the same consumers (for an overview about market research with panels see: Günther et al., 1998)

The sample may be *households* (household panel) or *individuals* (individual panel). Ideally always the same consumers should be in the sample, but in reality people finish their participation after a while (panel mortality) and have to be replaced in the sample. In an established and well managed consumer panel continuous reporting over a whole year of about 70% to 80% of the sample can be achieved.

The collected data concern the *purchases of the consumers* in certain product fields. Usually data about so called FMCG (fast moving consumer goods, like food, beverages, products for washing and cleaning etc.) but also about the purchase of durables and services (e.g. lending of videos, visits to the cinema) are collected. The main informations captured, are the date of the purchase, the shop, the identification of the item purchased, the amount bought, and the price paid. A household panel is the most economic type of sample, because one person in the household (usually the housekeeping person) has the task to report the purchases of all members of the household. In practice this works well as long as items are concerned, which where bought by the housekeeping person itself or which she uses. A broad product area of items necessary for housekeeping (e.g. washing, cleaning and cooking) can be covered with this method. But for products which are mostly individually bought and consumed, an individual panel delivers more reliable data. In GfK Germany data about articles for personal care and confectionery are therefore collected in a sample of individuals. In the meantime the individual panel is a combination of a household sample and an individuals sample, where the purchases of housekeeping persons are taken from an individualised household sample and the purchases of non-housekeeping persons are taken from an individual sample.

In addition to the purchase data, *sociodemographic* informations are collected, e.g. size of the household, region, age of the housewife etc. It is also possible to measure consumption relevant *attitudes* under certain restrictions. These restrictions are, that the work load on the panel members must not exceed certain limits and that questions have to be avoided which influence the purchasing behavior. Informations which allow for the identification of panel members cannot be used in reporting and analysis, but of course name and adress are necessary for the administration of the panels, and their processing follows the regulations of the data processing act.

In consumer panel research the traditional *method of data collection* is the *paper & pencil diary* with prepared sheets for household's input. Usually the panel members get the diary for a quarter of a year in advance, and send back the sheets weekly (or monthly) in a stamped addressed envelope. For this purpose the sheets in the diary are organized by weeks (or months), so that for all weeks (months) the same set of forms is contained in the diary. In all cases the day of the purchase is registered and the panel members are instructed to enter their items at the same day they purchased them.

A special challenge of this method is the precise identification of the article bought. The panel members therefore had to write down brand name, manufacturer, packsize and other features of an article, but in the mean time the collection of the EAN number, which is printed as a *barcode* on most packages, is preferred whenever possible. Because the EAN is just a numbering of products and not a code system for article features, the institute has to administrate an EAN master file, which contains the translation of the numbers in product features and their values. This has to be done by the institute itself because at the moment there is no general EAN master file which specifies product descriptions in a depth required for market research.

In the paper & pencil method the EAN number is written down (it is always printed below the barcode), but the possibility to capture barcodes electronically led to a revolution in consumer panel research by using hand held scanners instead of paper & pencil diaries in the households. During 1997 GfK switched its main household panel to *electronic data collection* (Electronic Diary). Each household got a mobile hand held device which can be used to scan the barcodes and also has a display and a keyboard. These features allow asking questions and offer possibilities to select informations like the shop from an option list. The number of packages bought and the price paid are key entered. Also non bar-coded products can be entered by means of an electronic questionnaire which asks for the product features.

The purchase data entered are stored in the device and automatically transferred to the institute via modem and phone every weekend (free of charge for the panel members).

Electronic data capturing with the present technology of handheld scanners comes to its limits, when products are reported which usually have no EAN number, e.g. fresh food or textiles (too laborious for the panel members if much details are required). Therefore GfK still runs traditional paper & pencil panels for these special product fields.

## 2.2 Panel Maintenance and Quality Control

The objectives of panel maintenance in an established panel are to motivate the panel members to stay in the panel and to secure a good quality of reporting. Furthermore those who finished participation have to be replaced.

The consumer panels of GfK Germany are quota samples which are adjusted by weighting according to regions, household size, age, gender, townsizes etc.

When *recruiting* panel members importance is attached to a broad regional variance and the inclusion of random elements into the recruitment procedures. A recruitment mix of interviewers (quota by region x age and region x household size), mailings and CATI (random addresses / phone numbers stratified by regions and townsizes) is used. It must not be forgotten that it is much more difficult to recruit panel members for continuous reporting than respondents for a single shot survey.

To motivate panel members a mix of *incentives* is used. They get brochures and news letters several times a year, which show them on a generalized level what GfK does with the data they reported, birthday cards, and can choose a gift of about 70.- DM after one year of regular reporting. Furthermore there are prize draws every quarter containing a car (VW Golf), dream holidays for two persons including a generous pocket money, two day trips to Nürnberg including a visit to GfK, and gifts of money. A special role plays the telephone hotline who is not only available for questions of the panel members but also actively calls them when quality problems in reporting are recognized.

*Quality control* of the panels is done on two levels. On the one hand the quality of reporting is tracked continuously and bad reporters are given their notice to leave the panel, on the other each single purchase reported has to be controlled for plausibility.

To control the *reporting quality of panel members*, only respondents who fulfill certain requirements are included in the part of the total sample used for reporting. At the end of each month the response behavior is controlled whether data were delivered at all and whether the amount purchased in total is in a plausible range around the mean of all households of same size. Only panel members fulfilling these criteria are included in the reporting mass of the passed month. Regularly the panel members are checked for their long term response rates, if necessary they get a notice to leave the panel.

This system of monthly selecting the good responders as the base for the reports (monthly mass) has consequences for the analyses. For facts like volume, value and market shares the monthly masses can be added (full mass), but for facts where the buying history is important (e.g. penetration) only those panel members which reported continuously during all interesting periods (continuous mass) are the basis of the analysis. The

*continuous reporters* are therefore the most valuable part of a consumer panel, because they allow for dynamic analyses of purchasing behavior like brand switching etc.

In *controlling the single purchase acts* the intention is to recognize errors which the panel members made when entering the purchase. Most of these errors are careless mistakes when entering prices (e.g. DM 1.08 instead of DM 10.80), but these deviations can be recognized by automatic check procedures (comparison with the average price of the article) and corrected to a plausible value.

### 2.3 Consumer Panels offered by GfK

In 1999 GfK Germany offered the following main consumer panels for FMCG, which are summarized under the brand name „*ConsumerScan*“.

<i>Panel</i>	<i>Universe</i>	<i>Sample Size</i>	<i>Data Collection</i>	<i>Product Fields</i>
ConsumerScan Households	Private German Households	12.000 households	Electronic	Fast moving consumer goods (food, beverages, washing, cleaning, personal care, pet food, other products for housekeeping)
ConsumerScan Fresh Food	Private german households	5.000 households	Paper & Pencil	Fresh products (vegetables, fruits, meat, fish) Alcoholic beverages
ConsumerScan Individuals	Persons living in private german households age 10 years and older	10.000 individuals	Paper & Pencil	Confectionery, personal care, drugs Communication and services (e.g. magazines, cinema, video, lotteries etc.)

Beside some other special panels, additionally under the brand name „ConsumerScope“ a household panel of 20.000 households is offered for mail surveys. The mailpanel is used for monthly surveys which can be focussed on target groups and for the data collection about the purchases of durables (like household appliances and consumer electronics).

Beginning in 2000 for personal care and confectionary combinations of the purchases of housekeeping persons in the household panel with the purchases of individuals in the individual panel are offered (ConsumerScan Health and Beauty, ConsumerScan Confectionery). The combined samples are individualized and made proportional according to housekeeping (yes/no) and other individual characteristics by weighting procedures.

## 2.4 The Insights coming from Consumer Panel Research

The previous chapters may have formed an idea of that consumer panels have to be continuously cared for. So it should be no surprise that they are costly instruments. They deliver a *continuous stream of purchasing data* about the same product fields.

The data are sold via long term contracts (e.g. for one year up to three years). This contracts specify the rhythm of report delivery (e.g. monthly, quarterly, etc.), the product fields reported, the facts (e.g. penetration, volume, market share, prices), the extent of consulting (e.g. yearly presentations of the market) and budgets for special analyses with the panel data, not covered by regular reporting.

The *clients* of the consumer panel are the brand manufacturers but also retail companies and associations. The insights from consumer panel data for these clients can be summarized as follows:

- *Market Tracking*: How is my brand developing compared to the market and other brands?  
Here especially time series of market shares, the success of innovations and the development of markets are analyzed.
- *Target Groups*: Who purchases my brand?  
In which segments are my brands / other brands successful?
- *Brand Health Check*: What are the strengths and weaknesses of my brand compared to other brands?  
Compared to competitive brands, benchmarks like the penetration rate, intensity, repeat buying etc. are analysed.

- *Evaluation of Marketing Policy*: What effects have changes in the marketing mix? It is important to learn how changes in the four P's (product, place, price, promotion and people) of marketing mix effect the market share.
- *Category Management*: How can store assortments, product introduction and promotion be optimized? The behavior of the customers is analysed to define categories, evaluate their performance and make recommendations about optimal store assortments.
- *Forecast*: Which benchmarks will a product introduction achieve? The question for each manufacturer who launches a new product in the market is, whether his product will achieve a market share necessary for commercial success. Consumer panel data help to forecast whether the product will be a success or a flop by analyzing the consumer reactions in the phase of the introduction.

These points illustrate that the purchase data themselves are in the centre of interest. The next chapter will show how consumer panel data are used in market research.

### 3. Methods of Analyzing Consumer Panel Data

#### 3.1 Benchmarks for Market Evaluation derived from Consumer Panel Data

Some characteristics became established in consumer panel research which are in principle tautological decompositions of volume, value or market shares. But the components react differently – this is the practical experience – to marketing decisions.

These characteristics are described in the following section:

Be „X“ a certain product and „Pg“ the product group to which X belongs, furthermore:

N	:= sample size of the panel
Nb(Pg)	:= number of buyers of the product group in the panel
Nb(X)	:= number of buyers of X in the panel
Vol(Pg)	:= volume of the product group bought in the panel
Vol(X)	:= volume of X bought in the panel
Vol(Pg/X)	:= volume of the product group, bought by the purchasers of X



Then it can be derived:

Penetration of X	Pen(X)	:= Nb(X)/N
Relative Penetration of X	Rpen(X)	:= Nb(X)/Nb(Pg)
Market share of X (basis volume)	MSVol(X)	:= Vol(X)/Vol(Pg)
Volume of X per Household	VolpHH(X)	:= Vol(X)/N
Volume of X per Buyer of X	VolpB(X)	:= Vol(X)/Nb(X)
Brand Share amongst Brand Buyer	BSBB(X)	:= Vol(X)/Vol(Pg/X)
Product Group Vol. per X Buyer	VolpB(Pg/X)	:= Vol(Pg/X)/Nb(X)
Intensity Factor of X	Ifactor(X)	:= VolpB(Pg/X)/VolpB(Pg)

Penetration, Relative penetration, Market Share and BSBB are usually reported as percentages. Instead of Volume per household usually Volume per 100 households is reported.

Immediately plausible is the following *basic formula of market decomposition*:

$$\text{VolpHH}(X) = \text{Pen}(X) * \text{VolpB}(X)$$

That means that the volume bought by a product depends on the circle of buyers which can be reached and the intensity of buying. One gets more insights into market dynamics if one tracks and compares the single components because they can be influenced by different policies. A high penetration rate can be achieved by a broad target group, high distribution rate (the optimum is, if one can buy the product in every shop), much advertising to make the product well known, and many promotions (e.g. price promotions). But if the buyers of X are intensive buyers (compared to the buyers of other brands), depends whether loyalty to the product can be induced. E.g. by optimal adaption of the product to the demands of the target group, high quality of the product, good value for money and an attractive product image.

The basic formula can be decomposed further on, by replacing Volume per Buyer with the product of Buying Acts per Buyer times Volume per Buying Act. Instead of Volume also Value or Number of Packages can be taken.

Not so intuitively plausible, but easily to verify by inserting the definitions is the *Parfitt-Collins decomposition of Market Share*:

$$\text{MSVol}(X) = \text{Rpen}(X) * \text{BSBB}(X) * \text{Ifactor}(X)$$

The Relative Penetration of X indicates how many of the potential buyers of X are reached. The brand share amongst brand buyers describes the loyalty of the brand buyers to the brand and the intensity factor is a measure, whether the X-buyers are more or less intensive product groups buyers than an average product group buyer. The formula says

that a product gains a high market share if it reaches a high percentage of the potential buyers of this product, if the buyers are loyal to the product and buy intensively in the product group.

Especially in brand evaluation the strength and the weaknesses of brands become evident by comparing these type of characteristics. This means that consumer panel data have its own value in benchmarking, even if they are not used for causal or dynamic analyses.

Typically these characteristics are calculated on an aggregated level, but they may serve as a guideline for social research for the construction of *different dependent variables on the individual level*. E.g. the following characteristics can be calculated for each panel member:

- A dummy variable whether the panel member bought in the product group or not, indicating interest in the product group.
- Inside the buyers of the product group: the volume of the product group a panel member bought, indicating the intensity of demands for the product group.
- Inside the buyers of the product group: a dummy variable whether a panel member bought product X, indicating who is interested at all in X.
- Inside the buyers of product X: the share of the product group which is allotted to product X, indicating the loyalty to X.

Instead of only looking to the amount of X bought, one could use the dependent variables above, which indicate different aspects of the buying act. Probably one will detect, that different causal factors are relevant on the different levels.

### **3.2 Special Analyses about the Dynamics of Purchasing Behavior**

At least a short impression should be given about special analyses of consumer panel data which refer to the dynamic aspect of panel data. Typically two time periods or the buying acts before and after the purchase of a certain product are compared.

Without describing the established dynamic analyses (brand switching, gain and loss, etc.) in detail two examples may give an impression what usually is analysed.

Example 1: A retail company A does not offer a certain brand Y any more. What will the purchasers do, which previously bought brand Y in A? How many of them stay in the shops of A, but buy a competitive product of Y? How many of them go to another shop and are loyal to Y? And what are the effects on the value share of Y? These questions can be answered by comparing the purchases of the buyers which bought Y in A before and after the policy change of A.

Example 2: A manufacturer recognizes, that his Brand Z loses market share. That means that purchasers of Z switch to other brands (loss of purchasers) or that they buy at least other brands in higher proportions than before (sinking loyalty). So the question is: Which other brands gain how much by the losses of Z. A „gain and loss“ analysis answers these questions by comparing two time slices.

It is also possible to create variables on the household level which describe the dynamics of purchasing behavior (e.g. brand switching) and analyse them (e.g. with survival analysis).

## 4. The Data Delivered

The GfK sold an extensive data set to ZUMA under the following restrictions:

- Characteristics which allow to identify brands and manufacturers were removed from the data
- Characteristics which allow the identification of shops were removed from the data

In doing this GfK fulfills it's obligations against it's clients not to allow others an evaluation of brands, manufacturers or retail companies.

The data are *household panel data from 1995* (at this time: paper & pencil diary). In this year the total sample was divided into two subsamples of same size and same structure. Data for the product groups partly were collected in both panels or only in one of the two subsamples, so that depending on the product group the sample size differs.

In total the structural data and the purchase data of 9064 households were delivered, which reported at least the whole year of 1995 continuously. These data stem from the unadjusted sample, which is disproportional in household size and will have other disproportionalities, because panel mortality is not random (e.g. it is higher for young people). Before making projections to all German households an adjustment is necessary.

The *household characteristics and attitude data* are organised in a way, that for each household a data record containing the informations in a fixed format was produced.

The household informations cover:

- Region
- Townsize
- Size of the household
- Number of children below 14 years in the household
- Age of the housekeeping person

- Net income of the household
- And other informations about profession, housing conditions, equipment with household appliances, etc.

The attitude data cover the answers on 74 items of the areas of life style, eating habits and environmental consciousness. Also the scales constructed from these items and the results of cluster analyses of the households on the base of the scales were delivered.

In total the *purchases* of 81 product groups are reported. 36 are covered in both panels, the others in one of the two subsamples. The purchase data are organized in a flat file containing one record per household per buying act of a certain product. The contents of the purchase records are especially:

- The household number, which allows to connect the purchase data with the household characteristics
- The date of the purchase
- The type of shop where the product was bought
- A product description consisting of a product group number, product type and several product group dependent features like type of packaging, kind of product etc.
- The number of items bought from the same article (at this day in the same shop)
- The cumulated volume of these items in grams, milliliter or pieces.
- The cumulated value of these items in Pfennigen (0,01 DM).

When analysing the data, please be careful with the following issues:

- Package size can be calculated by dividing volume by nr. of items bought.
- Price per package is calculated as value divided by nr. of items bought.
- Price per unit is calculated as value divided by volume.
- Because volume units are measured in different units depending on the product group, it is recommended to use value, when combining several product groups.
- Inside some product groups also different units may be used, e.g. milliliter for fluids, grams for powder, pieces for tabs.

Two ways of data organisation may help to analyse the data with standard tools. One is to load household structures and purchase data into two database tables and relate them by the household number. The other is to add the household information to each purchase data record.

Since ZUMA showed interest in analysing the effects of changes in household structures and attitudes on purchasing behavior, also the household characteristics and the attitude measures of 1994 and 1996 were delivered as far as they were enquired. The household characteristics and the attitude questions are updated once a year, so the exact date when a change in these features happened cannot be detected.

## 5. Final Comments

How far these data may lead to interesting insights for social science and vice versa, how far GfK gains knowledge from the analyses of the data cannot be finally settled at the moment. The hope of GfK is, that new methods of analysing this type of data may emerge from this cooperation. Some subjective remarks may conclude these notes:

- Purchasing behavior is a central part of peoples everyday-life and also a central part of our economic system. It's worthwhile to analyse it.
- Consumer panel data are the best information about consumer behavior available at the present. Behavior is recorded as close to it's occurrence as possible.
- The dynamic character of panel data delivers interesting material – even when the type of behavior does not interest the analyst – for developing models of analysing these type of data.
- The data may help to test hypotheses about social behavior as far as the interesting variables are operationalised.

The first results are contained in this publication. It's a start to interchange ideas and data for intensifying public private partnership.

## Contact

*Dr. Hans-Georg Prester  
Operations Consumer Panel  
GfK AG, Bereich Panel Services  
Nordwestring 101  
90309 Nürnberg  
hans-georg.prester@gfk.de*

## References

Günther, M., Vossebein, U. & Wildner, R. (1998): *Marktforschung mit Panels: Arten – Erhebung – Analyse – Anwendung*. Wiesbaden: Gabler.

# **CAN PATTERNS OF EVERYDAY CONSUMPTION INDICATE LIFESTYLES? A SECONDARY ANALYSIS OF EXPENDITURES FOR FAST MOVING GOODS AND THEIR SOCIAL CONTEXTS**

*HARTMUT LÜDTKE AND JÖRG SCHNEIDER*

## **1. Lifestyles and patterns of consumption**

Lifestyles are patterns of everyday behaviour acquired by actors following their interests and preferences in the long term, framed by values, beliefs and other contexts of social meaning (Hartmann, 1999, Hradil, 1992, Konietzka, 1995, Lüdtke, 1989, 1995a, 1996, Mitchell, 1983, Müller, 1992, Spellerberg, 1996). In the process of presenting elements of their lifestyles in settings of communication and social exchange, actors are forming groups of symbolic affiliation with similar others, respectively symbolic distance from dissimilar others. As a result, collective forms of lifestyle are produced in a process of comparing, exchanging, and accepting signs of individual lifestyles by the relevant actors in social situations. From an actor's point of view, the meaning of lifestyle depends on three functions (Lüdtke, 1989, 1995a): (a) as a memory of everyday routines providing self-assurance in orientation and action, (b) as a link between personal and social identity, (c) as a medium as well as an outcome of social affiliation, respectively distinction.

According to this concept lifestyles are markers of the "expressive" or (following Max Weber) consumptive dimension of social inequality crossing the vertical dimension of status and power (Hartmann, 1999: 20 ff., Hradil, 1987, Lüdtke, 1989, Spellerberg, 1996). Their theoretical position is one on the mesolevel, placed between individual forms of living and small networks of these and, on the other side, macrostructures like classes, strata, large milieus or subcultures. That's why lifestyle affiliation as an exogenous vari-

able can better serve for explaining or predict behaviour and attitudes as endogenous variables than socio-economic variables of traditional description do.

In keeping with their expressive, semiotic meaning in social exchange it seems to be conclusive to operationalize lifestyles, using interview or observation data on individuals, by indicators of overt behaviour or, in our own term, of "performance". They include categories of relevant others in situations and artifacts as concomitants of everyday action. Because of their functioning as contexts and means of self-representation, lifestyles are usually inquired upon variables concerning performance in symbolically relevant parts of living, like home comfort, outfit and personal hygiene, leisure and sports, holidays, media, cultural activities, food, modes of eating and drinking. Although no theoretical consent among researchers has been achieved so far, how to select the most significant indicators of style performance from an open-ended list of opportunities, most of the life style studies carried out in recent years were founded on data from various parts of living mentioned above.

Having identified lifestyles as patterns of performance the analysis is continued by assigning typical characteristics of social position (e. g. status, income, gender, age, household structure and other resources or restrictions) as well as indicators of mentality (e. g. value-orientations, attitudes, self-concepts) to them. In this way additional information becomes available for better interpreting the subjective meaning of styles and for better explaining their generation in the process of socialization and social allocation of the representatives of style in question (Lüdtke, 1989, 1995a, 1996).

Among numerous researchers in this field, it was Sobel (1983) who chose a rather thrifty, but exclusively performance-related, way of constructing lifestyle types using data of 18.000 respondents in the US-Survey of Consumer Expenditures from 1972/73: logarithmic measurements of expenditures for 17 classes of goods. He found, by means of factor analyses, four patterns: two on a high, respectively "normal" level of "prestige", one of "high life" (focusing entertainment, pop, liquor), and one of "home life" (focusing family living, camping, health, sports, tv). Although Sobel dealt with only a narrow section of lifestyle research he demonstrated that data about consumer goods and expenditures can be suitable indicators of lifestyle patterns. Following his approach within an enlarged context we attempted to reconstruct more informative lifestyle types basing the GfK data. Although the kinds of goods available to us do represent only some trivial and highly transient elements of lifestyle, it can be shown that they, in connection with variables of social position and mentality; open our view of several, highly distinct as well as stable, stylistic patterns. They may signify something "more" than the obvious point of fast moving goods picked up by consumers.

## 2. Problems and hypotheses

Looking for a theoretically fruitful consumer typology we have to face the problem that the GfK data, being available for secondary analysis, do represent only an imbalanced selection of performing indicators. We are therefore running a high risk expecting to find lifestyle patterns of that kind which can serve for marking *complex* forms of living which we have our eyes on. From the point of a positive turn of things our first assumption is:

(1) *Classifying data about preferences and expenditures for fast moving goods, in connection with circumstances of shopping, makes it possible to clearly separate consumer groups of high homogeneity and distinctiveness. They adequately represent the performance level of lifestyles.*

If this hypothesis proves true and with respect to the fact that lifestyles develop within an individual's given frame of attitudes and restrictions, the following assumption seems plausible:

(2) *There are sets of specific variables of social position as well as of mentality corresponding to each of the performance subgroups. They are indicated by showing above or below average values in a group. Having identified lifestyle types by patterns of performance, social position, and mentality, these groups are, as a result, located in some general dimensions of lifestyles known from former studies (Lüdtke, 1995b, Spellerberg, 1996): traditionalism versus modernism, home- versus public-centeredness (radius of action), activity/creativity versus passivity/reception, and one or two less remarkable dimensions of economic and cultural resources at disposal.*

The shopping data noted by the diarists reach over one year. As far as we may assume that an individual does realize his or her special lifestyle by the set of all relevant acts at least over this time, we also can say: If consumers allocated their purchases over this period by chance there wouldn't appear any typical pattern of performance in one's purchases done within a shorter span of time. Constancy of lifestyles can therefore be tested by means of comparing the respective ties of diarists to analogous types separated at several cross-section analyses, though the span of one year seems to be a rather short panel length. If allocation by chance was fictitious, we may expect that the respondents performed structurally similar shopping behaviours at each time of analysis. To cut the periods of comparison not too short we decided to separate data of the quarters in 1995, thus far our next assumption is:

(3) *Constancy of one's lifestyle performance is given if there is no remarkable reduction in strength of the association between the pattern displayed in one quarter and that or those displayed in the following quarter(s), beginning with the first of them. This is*



*equivalent to the observation that most of the diarists are tied to an analogous lifestyle type each across all the quarters.*

As a consequence of our second hypothesis concerning the mentality dimension of lifestyles and according to the restricted meaning of fast moving goods as well as to the given small number of mental indicators, measured by GfK once a year, we assume:

*(4) Values on the scales of „General Life Orientation“ and „Orientation Towards Nourishment“ prove effective predictors of a diarist’s style affiliation, i. e. ties to analogous subgroups across the quarters.*

According to the theoretical claim that lifestyle ties should prove fruitful exogenous variables for the explanation of value orientations or other forms of cognition and attitudes, measured independently from indicators explicitly applied to describe mentality (Hartmann, 1999, Lüdtke, 1995a, 2000), our last hypothesis is:

*(5) The lifestyle types reconstructed at each quarter will be closely associated with affiliates’ values on scales of „Ecological Consciousness“.*

### **3. Data and sample**

Our study is based on the GfK fast moving goods data from the 1995 panel: a combination of data from shopping diaries (currently kept) and questionnaires (once in that year). The following variables from the diaries were accumulated over the year as well as over each quarter (January to March, April to June etc.), so as to getting one total and four partial sets of data to be independently analysed at the further steps:

- (a) 50 sorts of goods as parts of 3 main groups<sup>1</sup>: beverages (22), food (21)<sup>2</sup>, and personal hygiene (7);
- (b) for each sort, respectively act, noted in the diary: number of purchases, quantity per entry, price per unit, time of purchase, location (small retailers, discounters, supermarkets); all quantity and expenditure data were transformed into values per capita, holding sizes of household constant;

---

<sup>1</sup> A fourth group: household articles (e. g. cleanser, paper towels), were not included because of their low affinity to „symbolic significance“ of use.

<sup>2</sup> Unfortunately goods like fruits, vegetables, meat, sweets, preserved food, bread, cakes and pastries had not been included in the investigation.

(c) social position and resources: householder's age, formal education, and occupation; size of building, of residence, of household; number of children below 18 years; car(s), garden and domestic appliances property (microwave, video, dish-washer)<sup>3</sup>;

(d) mentality: 7 Likert scales of „General Life Orientation” meaning tendencies like objectives, basic attitudes or self-concepts: traditionalism, innovativeness, seeking stimulating experiences, nostalgia, mistrust of anything new, convenient cookery, striving for high quality of things; 13 Likert scales of „Orientations Towards Nourishment”<sup>3</sup>: slimness, healthiness, natural products, curiosity to learn, German food, convenience food, homely fare, full food, delightfulness, freshness, proprietary articles, vitamins and minerals, un-critical way of nourishment;

(e) three scales (varimax rotated, principal components, explaining 49.5 % of variance) of Ecological Consciousness<sup>3</sup> gained from 13 statements concerning attitudes towards conservation: Factor loadings above 0.50

(1)	<i>Disapprobation of Conservational Intervention</i>	
	There is being made too much a fuss about environment	0.75
	There is enough conservation at the moment	0.70
	Cars aren't an environmental problem	0.68
	No possibility of conservation in the household	0.67
	I don't worry about the environmental harmfulness of products	0.57
	Conservation is a matter of state and industry instead of citizen	0.54
(2)	<i>Conservation at One's Personal Disposal</i>	
	Paying attention to pollution-free products at shopping	0.76
	Expending more money for pollution-free wrapping	0.68
	Buying less harmful products than in the past	0.61
	Taking limitations on oneself for the benefit of conservation	0.59
(3)	<i>Conservation as a Macrolevel Concern</i>	
	No more industry	0.83
	The whole of nature not only environment has to be aided	0.75

(f) „sensitivity to price” (Preisbewusstsein): a statement to be judged along three values: non/low/predominant/high.

The Likert scales mentioned were taken over from the original GfK data set available, the more so as their internal consistencies seemed to be sufficient. The data according to (a) and (b) were reduced, by means of factor analysis, down to 7 principal components of

---

3 Unfortunately data of income were not available in the 1995 set.

beverages, 6 such of food, and one such of goods for personal hygiene, showing contributions of variance each between 47 and 51 %. Four further factors (explaining 84 % of variance) were identified indicating general dimensions of shopping: quantity and wideness of goods bought; standard versus special prices paid; avoiding discounters versus preferring supermarkets; preferring small retailers. In this way we defined 18 variables indicating shopping, respectively „lifestyle” performance to serve as active variables for the subsequent cluster analysis of subgroups.

Obviously a small number of diarists were characterized by some extremely imbalanced tendencies of shopping or making entries, for instance by noting lots of very few goods or minimal quantities of many goods. In order not to reconstruct artificial patterns of shopping, we decided to „iron out” those data according to two criteria: a) eliminating such goods which were bought by only 10 % or less of the diarists only once during the year, and b) holding the maximum quantities per capita constant over the year at the level of the respective 90<sup>th</sup> percentile. Finally, the parallelized data sets from diaries and questionnaires refer to a net sample size of  $N = 4426$  consumers.

## 4. Analyses and results

### 4.1 Lifestyle clusters in total

Taking the first step we carried out a cluster analysis<sup>4</sup> including the 18 yearly variables defined above. An optimal solution was found with 15 clusters involving a reduction of 33.2 % of total variance. 40 % of the similarities between clusters are higher than 0.20 with a maximum amount of 0.48 – what we may judge all in all as just good enough. Describing the subgroups by the active indicators of performance (PER), those of social position (POS) and those of mentality (MEN)<sup>5</sup> we find the following clear-cut profile of lifestyle types<sup>6</sup>.

---

4 By application of the CONCLUS model by Bardeleben (1991) which is an iterative, maximum-likelihood approach towards minimization of variance within clusters. Ist main criteria of goodness are:  $\text{Eta}^2$  (proportion of variance between clusters to total variance), consistency of a cluster (on the analogy to  $\text{Eta}^2$ ), and similarities between clusters (a correlation coefficient).

5 Including the subscales of „General Life Orientation“ (GLO) and „Orientation Towards Nourishment“ (OTN).

6 The values are scores of standard deviation from total mean (= 0). Only those „active“ (PER) variables with absolute values above 0.32 are given, „passive“ variables (POS, MEN) however with values above 0.24. Variables not mentioned indicate that there is a roughly average amount or category in the respective dimension.

*Cluster 1: Single or small households with wideley spreading consumption (5.5 %, consistency = -0.05)*

PER	quantity and wideness of goods bought	2.02
	standard products for daily dental and personal hygiene	1.44
	sauces, mustard, chips	1.23
	coffee and tinned milk	1.10
	liquor	1.02
	juices, lemonades, preserved milk, chocolate	0.88
	milk products	0.74
	fresh milk	0.53
	high-fat curd	0.35
POS	household size	-0.39
	number of children below 18	-0.35
	domestic appliances	0.33
MEN	no characteristics	

This pattern is characterized by a wide range of sorts of goods bought that seems to be typical for single or couple households (children hardly ever) located in the middle class. High level of liquor consumption and of appliances property may indicate their affinity to a “high life” style (Sobel 1983). The amount of consistency near zero is caused by the extremely high value of the first indicator of general shopping behaviour in relation to the other variables.

*Cluster 2: Young families of working parents tending to convenience food (8.4 %, consistency = 0.34)*

PER	juices, lemonades, preserved milk, chocolate	1.34
	sauces, mustard, chips	0.78
	milk products	0.54
	standard products for daily dental and personal hygiene	0.42
	cheese, low-fat curd, yoghurt	-0.49
	coffee and tinned milk	-0.40
	fresh milk	-0.36
POS	household size	1.08
	number of children below 18	-1.04
	age of housholder	-0.86
	domestic appliances	0.48
	car(s)	0.36
	working householder	0.32
MEN	GLO: high quality of things	-0.43
	OTB: proprietary articles	-0.45
	healthiness	-0.42
	freshness	-0.38
	full food	-0.35
	natural products	-0.39
	convenience food	0.34
sensitivity to price	0.32	

These middle class group of young families with at least one child are, through all PER, POS and MEN dimensions, characterized by a rather unobtrusive, convenience-orientated way of consumption. The pattern seems to reflect some habits and attitudes strongly affected by experiencing everyday stress of lacking time or money.

*Cluster 3: Liquor drinkers (7.0 %, consistency = 0.34)*

PER	liquor	1.82
	sauces, mustard, chips	0.57
	cheese, low-fat curd, yoghurt	0.51
	quantity and wideness of goods bought	0.44
	non-carbonated mineral water in returnable bottles	-0.44
	instant whips, sour cream, low-fat curd	-0.38
	cereals	-0.38
POS	formal education	(0.23)
MEN	OTN: curiosity	0.33
	delightfulness	0.30

This group unites diarists with a relatively luxurious and open-minded (curious) way of consumption that culminates in a high extent of liquor drinking. Affiliants are at a slightly above average level of formal education.

*Cluster 4: Large young families with well-equipped homes and gardens (5.9 %, consistency = 0.25)*

PER	fresh milk	1.80
	cereals	1.54
	standard products for daily dental and personal hygiene	0.50
	juices, lemonades, preserved milk, chocolate	0.47
	sauces, mustard, chips	0.44
	quantity and wideness of goods bought	0.40
	instant whips, sour cream, low-fat curd	0.35
	coffee and tinned milk	-0.43
	non-carbonated mineral water in returnable bottles	-0.42
POS	number of children	1.19
	household size	1.12
	age of householder	-0.62
	domestic appliances	0.59
	car(s)	0.41
	size of building	-0.35
	garden	0.30
MEN	OTN: uncritical way of nourishment	-0.39
	homely fare	-0.38
	slimness	-0.32
	GLO: traditionalism	-0.33

This pattern is similar to cluster 2 differing from that by living outside a city center (small home, garden) and a higher affinity to fresh and natural products such as fresh milk and cereals as well as to a more critical way of nourishment.

*Cluster 5: Extremely reduced amount of shopping (13.4 %, consistency = 0.56)*

This pattern seems to be rather a methodological artifact than a consumer “style”: it unites cases showing negative values in 9 dimensions of goods and a very low level of quantity and wideness (-1.04), additionally with absence of positional and mental features.

*Cluster 6: Single or couple, elderly health fans (5.7 %, consistency = 0.30)*

PER	cheese, low-fat curd, yoghurt	1.39
	instant coffee, teas	0.96
	cereals	0.56
	high-fat curd	-0.93
	sausages, mustard, chips	-0.79
	juices, lemonades, preserved milk, chocolate	-0.50
POS	age of householder	0.35
	household size	-0.32
MEN	OTN: full food	0.67
	healthiness	0.42
	natural products	0.41
	freshness	0.34
	slimness	0.31



Cluster 7: *Young consumers with low-critical attitude making use of special offers (2.6 %, consistency = 0)*

PER	standard versus special prices paid	-2.42
	juices, lemonades, preserved milk, chocolate	0.68
	sausages, mustard, chips	0.61
	milk products	0.58
	standard products for daily dental and personal hygiene	0.56
	supermarkets versus discounters	0.56
	instant coffee, teas	-0.43
	cheese, low-fat curd, yoghurt	-0.38
	coffee and tinned milk	0.37
	preferences for small retailers	-0.36
	beverages in disposable packs	0.36
POS	age of householder	0.26
MEN	GLO: nostalgia	-0.31
	OTN: sensitivity to price	0.32
	uncritical way of nourishment	0.28

This profile of characteristics speaks for itself making sense of the heading as our interpretation of this pattern. Lack of consistency is caused once more by the extremely high first value.

*Cluster 8: Consumers of preserved products in non-urban home preferring local products (4.9 %, consistency = 0.16)*

PER	instant whips, sour cream, low-fat curd	1.86
	instant coffee, teas	1.30
	quantity and wideness of goods bought	0.63
	juices, lemonades, preserved milk, chocolate	0.60
	milk products	0.56
	standard products for daily dental and personal hygiene	0.44
	fresh milk	-0.43
	super markets versus discounters	-0.37
	cheese, low-fat curd, yoghurt	-0.38
POS	size of residence	-0.39
	size of building	-0.35
	garden	0.33
MEN	OTN: German food	0.27
	GLO: convenient cookery	-0.25

The most interesting feature of this pattern seems to be an “unconsistent” relationship between the preference for ready-to-serve and instant products on the one hand and a tendency against convenient cookery respectively towards German food.

*Cluster 9: Elderly consumers with limited mobility and selective taste (7.3 %, consistency = 0.47)*

PER	preference for small retailers (i.e., located in the neighbourhood)	2.05
	preference for supermarkets versus discounters	0.44
	quantity and wideness of goods bought	-0.84
	sub-average values in 7 dimensions of goods	
POS	car property	-0.63
	domestic appliances	-0.44
	age of householder	0.28
MEN	GLO: high quality of things	0.33
	nostalgia	0.32
	OTN: proprietary articles	0.31
	sensitivity to price	-0.45

*Cluster 10: Single elderly coffee drinkers at a lower level of education (mostly widows?) (6.5 %, consistency = 0.34)*

PER		coffee and tinned milk	1.40
		cereals	-0.67
		instant coffee, teas	-0.52
		sauces, mustard, chips	0.52
		milk products	-0.47
		preferences for small retailers	-0.34
POS		household size	-0.41
		number of children	-0.39
		age of householder	0.37
		formal education	-0.26
MEN	GLO:	convenient cookery	-0.40
	OTN:	convenient food	-0.32
		homely fare	0.30
		uncritical way of nourishment	0.30

*Cluster 11: Drinkers of noncarbonated water (6.4 %, consistency = 0.27)*

PER		noncarbonated mineral water in returnable bottles	1.95
		preferences for supermarkets versus discounters	0.43
		fresh milk	0.36
		beverages in disposable packs	-0.34
		standard products for daily dental and personal hygiene	0.32
POS		no characteristics	
MEN		sensitivity to price	-0.24
	OTN:	homely fare	-0.25

This pattern seems to be one of consumers quite open to issues of environmental conservation.

*Cluster 12: Urban consumers of beverages in disposable packs (6.7 %, consistency = 0.31)*

PER	beverages in disposable packs	2.28
	preferences for supermarkets versus discounters	-0.41
	juices, lemonades, preservable milk, chocolate	-0.34
POS	size of building	0.57
	garden	-0.41
	domestic appliances	-0.38
	size of residence	0.37
	car(s)	-0.33
MEN	no characteristics	

These urban consumers, less equipped with cars or domestic appliances, moreover showing features of performance and mentality at an average level, appear as producers of rubbish to an extreme extent by having drinks. Therefore we expected them to have attitudes of relative indifference towards environmental conservation – an assumption that was not confirmed by data (see sections 4.2 and 4.5).

*Cluster 13: Senior citizens with high standards of consumption and traits of traditionalism (4.0 %, consistency = 0.20)*

PER		preference for small retailers	2.03
		coffee and tinned milk	0.87
		instant whips, sour cream, low-fat curd	0.67
		quantity and wideness of goods bought	0.65
		cereals	-0.42
		milk products	0.38
		preference for supermarkets versus discounters	0.33
POS		age of householder	0.67
		household size	-0.45
		number of children	-0.43
		working householder	-0.37
		domestic appliances	-0.27
MEN		sensitivity to price	-0.40
	GLO:	traditionalism	0.38
		nostalgia	0.32
		high quality of things	0.31
	OTN:	proprietary articles	0.36
		curiosity	-0.34

*Cluster 14: Young working consumers at a moderate level of shopping behaviour (10.5 %, consistency = 0.53)*

PER		preferences for supermarkets versus discounters	1.20
		quantity and wideness of goods bought	-0.91
		sub-average values in 7 dimensions of goods	
POS		age of householder	0.21
		working householder	0.26
MEN		sensitivity to price	-0.26

*Cluster 15: Single or couple, health-oriented, senior citizens (5.2 %, consistency = 0.18)*

PER	high-fat curd	1.77
	cheese, low-fat curd, yoghurt	1.12
	instant coffee, teas	0.75
	quantity and wideness of goods bought	0.70
	juices, lemonades, preservable milk, chocolate	-0.52
	milk products	-0.45
	fresh milk	0.42
	standard products for daily dental and personal hygiene	0.36
	liquor	0.35
	coffee and tinned milk	0.34
	saucers, mustard, chips	-0.34
POS	age of householder	0.64
	household size	-0.48
	number of children	-0.44
	working householder	-0.31
MEN	GLO: high quality of things	0.30
	OTN: healthiness	0.42
	freshness	0.38
	natural products	0.36
	proprietary articles	0.29
	German food	0.28

In contrast to the slightly similar cluster 6 affiliates to this type are elder on average, and they show a higher standard and variety of consumption.

## 4.2 General lifestyle dimensions of clusters

In accordance with our first hypothesis we achieved to find a set of well-discriminated, homogenous consumer groups represented by clusters. Having been connected with characteristics of social position and mentality, they reflect a more complex typology of consumption as well. To interpret them to be “entire lifestyles” however would sound like a considerable speculation with respect to the narrow section of lifestyle indicators having been at our disposal. In spite of this limitation, an indirect approach of testing how far these types are empirical, though special, manifestations of latent lifestyles stands to reason, namely arguing this way:

They are located in general dimensions of action, i.e., basic orientations of individuals in the process of developing and maintaining their lifestyles.

They are constant over time as lifestyles, in the sense of biographical constructions, do (see section 4.3).

Their performative “core”, i.e., association with a cluster, can be predicted by mental traits, here: “general life orientations” and “orientation towards nourishment”, because lifestyle performance is affected by affiliates’ preferences and self-concepts (see section 4.4).

They are effective predictors of “environmental consciousness” because individual’s cognitions and value orientations of such kind are formed and stabilized by the context of practising and experiencing his or her lifestyle (see section 4.5 involving a reverse logic of analysis).

Our starting point is the first problem. It is, by inspecting the results of classification, simple to duplicate that at least six clusters are placed in two dimensions of action: radius of action with the poles of high life (clusters 1 and 3) versus home life (clusters 4,8,9), and modernism (cluster 4) versus traditionalism (clusters 9, 13). This seems not exhausting but an interesting hint. In addition to it ten of our fifteen lifestyle types can – without exaggerating our interpretation – be related to Sobel’s (1981, 1983) general lifestyles: nos 13, 15 to “high prestige”, nos 2, 7, 14 to “normal level of prestige”, nos 1, 3 to “high life”, and nos 4, 8, 9 to “home life”.

By use of the informations given from similarities between clusters we extended our view. We ran an analysis of multidimensional scaling based on these measurements with application of the SPSS procedure ALSCAL. The 15\*14/2 similarities were reduced to three dimensions as summerized in table 1.



Table 1: Dimensions of cluster similarities ( $N = 105$ ), stress = 0.204

Stimulus Name: Cluster	Stimulus Coordinates for Dimension		
	1	2	3
CL1: Single or small households with wideley spreading consumption	2.29	-.02	.78
CL2: Young families of working parents tending to convenience food	.49	1.32	-.38
CL3: Liquor drinkers	.26	-.21	1.48
CL4: Large young families with well-equipped homes and gardens	.81	.63	-1.5
CL5: Extremely reduced amount of shopping	-1.75	.14	-.27
CL6: Single or couple, elderly health fans	-.18	-1.05	-1.22
CL7: Young consumers with low-critical attitude making use of special offers	.54	1.56	1.04
CL8: Consumers of preserved products in non-urban home preferring local products	1.34	-.77	-.58
CL9: Elderley consumers with limited mobility and selective taste	-1.96	-.74	.24
CL10: Single elderly coffee drinkers at a lower level of education (mostly widows?)	-.06	-.18	1.23
CL11: Drinkers of noncarbonated water	.03	.50	-1.18
CL12: Urban consumers of beverages in disposable packs	-.93	1.34	.36
CL13: Senior citizens with high standards of consumption and traits of traditionalism	.08	-1.54	.67
CL14: Young working consumers at a moderate level of shopping behaviour	-1.70	.42	-.27
CL15: Single or couple, health-oriented, senior citizens	.73	-1.41	-.39

Taking the most prominent coordinates into account the following interpretation of the findings seems to hit the point:

Dimension 1:

*Quantity and wideness of consumption, associated with socioeconomic status*

high	←	clusters 1, 8	versus	clusters 9, 14	→	low
------	---	---------------	--------	----------------	---	-----

(cluster 5 as an artifact)

Dimension 2:

*Quality and modernity of taste and design of products and wrapping, associated with age*

ready to serve convenient uncritical young	←	clusters 2, 7, 12	versus	clusters 6, 13, 15	→	fresh, natural high quality healthy elderly
--	---	-------------------	--------	--------------------	---	--

Dimension 3:

*Attitudes towards conservation*

indifferent rejecting naive	←	clusters 3, 7, 10	versus	clusters 4, 6, 11	→	pro reflexive active
-----------------------------------	---	-------------------	--------	-------------------	---	----------------------------

For all that these classifications of clusters to general dimensions of orientation and action suggest slight or moderate correspondences with dimensions of “genuine” lifestyles. Obviously, there is given some evidence of our first assumption, although some *different* reductions of types to general dimensions presented themselves, leaving undecided how to interpret a homogeneous theoretical meaning of them.

### 4.3 Constancy of types throughout the year

In order to analyse the constancy of our typology it was necessary first of all to find equivalent sets of clusters for all quarters of 1995 to be valid reflections of the yearly ones at the same time. Because of some tendencies of diarists not to allocate shopping events in a homogenous way over the year we had to reduce the number of indicators of perform-

ance: twelve of the eighteen original ones remained being present in the quarters approximately to the same extent. These variables were measured, and compared, as structurally identical factors (principal components), by a procedure of construct validating as it were. It wouldn't furnish any more relevant information to describe these, 12 time 4, factors here in detail. After having constructed then four equivalent 15-clusters-sets<sup>7</sup> of the quarters, in the same way we did for the yearly data, we were able to analyse the strenghts of association between each two cluster sets appearing in crosstables. We used two measures indicating the proportional reduction in error when predicting the association of cases with clusters of one period by the clusters of a previous one: Guttman's Lambda and the coefficient of uncertainty UC. Table 2 gives a summary of the results. There is strong evidence of relatively high consistency of clusters over two, even three and four periods confirming our third hypothesis.

Table 2: Dependencies of cluster sets on those of previous quarters\* (N = 4426)

		dependent		
		Q 2	Q 3	Q 4
independent	Q 1	.31 .26	.28 .24	.28 .22
	Q 2		.30 .24	.29 .22
	Q 3			.29 .24

\*Asymtetric coefficients of Lambda (first value) and UC (second value). All analogous coefficients of contingency > .80!

Another, more detailed test of constancy is the way of predicting association of cases with clusters by the 12 variables of performance in accordance with the logic we followed above. Analysis of discriminance is an adequate model for doing so. The results of it are summarized in table 3: any percentage of "correct" prediction does amount at least seven times of a "hit" by chance (= 6.7 %). They obviously indicate high constancy across the quarters likewise.

---

7 All amounts of Eta<sup>2</sup> gained by these partitious are between 0.42 and 0.44.

Table 3: Percentages of “correct hits” when associations with clusters are predicted by variables of performance (analyses of discriminance,  $N = 4426$ )

		15 clusters per quarter			
		Q 1	Q 2	Q 3	Q 4
12 variables per quarter	Q 1	93.2	50.0	46.7	44.9
	Q 2		94.8	47.4	45.5
	Q 3			93.4	46.2
	Q 4				93.1

Hardly worthy of note that all measurements in this passage are highly significant ( $p < 0.001$ ) with respect to the null hypotheses.

#### 4.4 Predicting performance types by variables of mentality

Understanding lifestyles evokes the problem of explaining covert behaviour (e.g. patterns of consumption) by overt behaviour (e.g. attitudes towards general objectives or goals of way of life): How strong and asymmetrical is an expected relationship between a set of attitudes and a set of behaviours theoretically associated with them? From a psychological point of view attitudes indicate beliefs about an object or action concerning the outcome that is expected from action (Ajzen 1988). From a sociological point they indicate how an actor normatively frames his definition of the situation (Esser 1999). As far as we know, attitude-behaviour connections are rather weak, as a rule, because attitudes refer to a wider scope of objectives than behaviour does, and social control as well as control beliefs have additional effects on action.

The case of lifestyle is, over and above that, a special variant of this problem having two aspects: a) The performative core of a lifestyle represents neither a particular class of action nor a concrete class of situations but a complex pattern of generalizations of several actions and situations, respectively experiences gained by and in them. Any prediction of performance type by attitudes must involve therefore more fuzziness than one of particular action. b) Lifestyle-related attitudes are to be taken not only as independent variables of one’s lifestyle performance but as reflections of one’s former actions and learning by success just as much. Looking at it from both aspects, we may expect even lower associations between attitudes and lifestyle patterns than we might do in the common case.

To test our fourth assumption two analyses of discriminance were carried through taking the 15 lifestyle clusters as dependent groups.

The 7 scales of General Life Orientation as independent variables permit a correct classification of 15.5 % of all respondents. This means a 2.3 fold of a random classification (6.7 %). The best predictions refer to the following groups: Extremely reduced amount of shopping (no 5): 49.4 % correct classifications; young families of working parents tending to convenience food (no 2): 32.3 %.

The 13 scales of Orientation towards Nourishment permit a correct classification of 21.2 % of cases, i.e. a 3.2 fold of random. The most exactly predicted groups are here: young families of working parents tending to convenience food (no 2): 45.1 %; Extremely reduced amount of shopping (no 5): 29.6 %; Single elderly coffee drinkers at a lower level of education (no 10): 27.9 %; Single or couple, elderly health fans (no 6): 22.4 %.

These results proved statistically significant with respect to moderate levels of various indicators of power and fit. The striking means of particular scales per group are shown in the description of groups in section 4.1, and, referring to the exploratory character of this paper, there is no need for a detailed comment here.

As a result, we can now specify our fourth assumption in this way: In spite of all theoretical restrictions mentioned above, indicators of mentality prove to be moderately effective predictors of a diarist's style affiliation, at least in a few dimensions of general life orientation and orientation towards nourishment. The power of style-discrimination is greater with the latter scales, according to the higher affinity between the operationalization of types by means of food and beverage indicators and the particular dimensions of attitudes.

## **4.5 Ecological consciousness as explanandum**

Being associated with a type of lifestyles means to be involved in a special set of habits and frames. It has been stepwise acquired and assimilated in a long period of one's biography, and will direct one's orientations and behaviours in a long-term perspective as well. Consequently, actual lifestyle performance must prove true as an explanans of actor's general beliefs, value-orientations, or central life interests in a more fruitful way than social class or other features of macro-structural position do (Lüdtke 1955a, Georg 1998, Spellerberg 1996).

According to this presumption, and in a special sense of our fifth hypothesis, we expected to find, among the 15 types of shopping (life-)styles, some that would show values significantly above or below average on the scales of orientation towards conservation. We did so since

conservation as an object to attitudes is, no doubt, an important issue for a consumer's way of using resources and accepting restrictions, even at using fast moving goods.

In agreement with this idea<sup>8</sup>, 5 subgroups are accentuated in the dimension of

*Disapprobation of Conservational Interventions:*

above average: no 7 (young consumers with low-critical attitude, 0.26); no 10 (single elderly coffee-drinkers, 0.28);

below average: no 4 (large young families with well-equipped homes and gardens, -0.49); no 6 (single or couple, elderly health fans, -0.27); no 11 (drinkers of noncarbonated water, -0.27);

and 2 subgroups in the dimension of

*Conservation at One's Personal Disposal:*

above average: no 6 (see above, 0.30);

below average: no 2 (young families of working parents tending to convenient food, -0.29), whereas there are no substantial differences in the dimension of conservation as a Macrolevel Concern.

One-factorial analyses of variance carried through with lifestyles as (treatment) groups falsify the null-hypothesis supposing no differences of conservational orientations between the groups, and that with respect to each dimension. The results are summarized in table 4.

---

<sup>8</sup> See also ALSICAL dimension 2 in table 1.

Table 4: ANOVA Ecological Consciousness by lifestyles

Scale	source of variation	sum of squares	df	mean of squares	F	p
Disapprobation of Conservational Interventions	between groups	178.977	14	12.784	13.281	.000
	within groups	4246.023	4411	.963		
	total	4425.000	4425			
Conversation at One's Personal Disposal	between groups	101.690	14	7.264	7.411	.000
	within group	4323.310	4411	.980		
	total	4425.000	4425			
Conservation as a Macrolevel Concern	between groups	50.755	14	3.625	3.656	.000
	within groups	4374.245	4411	.992		
	total	4425.000	4425			

## 5. Conclusion

Our initial question: do shopping styles as patterns of everyday consumption prove substantial indicators of lifestyles, can be answered in the affirmative now, though we have to concede some restrictions. Our findings clearly show that even the consumption of fast moving goods, hardly symbolizing an expressive function for the customers at first sight, are imbedded in broader lifestyles. After having joined 15 plain clusters of shopping performance with respect to selected goods from the fields of food, beverages and personal hygiene to their co-varying features of social position and mentality, we found most of them to represent meaningful patterns of conduct of life: normatively framed, and socially adjusted types of habits and routines. Considering the narrow selection of variables of consuming behaviour available from the GfK data, a clear-cut congruence of shopping styles with lifestyles differentiated in recent social research could not be verified, though. Variables of social position like age, size of household and economic level seem to be more effective in the differentiation of styles than mental variables, which is, in a way, a selective effect by the shopping data themselves. Our five presumptions largely stood the test, however we see fit to modify or specify them in a few points:

The shopping styles can be read as simplified copies or segments of complex lifestyles in so far as their performative cores are associated with several variables of social position and mentality – an indication of the structural as well as motivational integration of the behaviour patterns represented by the clusters.

The fifteen style types could be reduced to a few dimensions of orientation of action. In doing so we found some, though not exhausting, correspondences with the general life-style typology by Sobel (1981, 1983), as well as with the theoretical dimensions of orientation modernism and radius of action. Three further dimensions were found by an additional ALSCAL analysis of the similarities between clusters: quantity and wideness of consumption (indicating the SES axis), quality and modernity of taste and design (similar to “modernism”), and attitudes towards conservation.

It seems to be arbitrary so far to attempt, only by means of inductive interpretation, a theoretical integration of these different co-ordinates of the general meanings of shopping, respectively life styles. This could be a challenge for further investigation into the ways of the reproduction of lifestyles by consumption.

In applying the life style approach to the explanation of shopping behaviour we postulated it to be highly constant over time. If shopping patterns are copies of lifestyles, single shopping acts of the members of a lifestyle group must not necessarily be repeated within a certain period of time, but a certain constancy of an individual consumer’s association with a particular pattern or style must be given therein. This theoretical postulate proved true, at least for the period of one year: We found remarkable measurements of constancy of style associations throughout the year, by means of stepwise comparing the data per quarter and of analyses of discriminance done quarterly predicting identical groups at the time by the shopping variables as well.

Characteristics of mentality proved to be rather effective predictions of style associations, and that all the more the closer the attitudes in question are to the behaviours, i.e. buying food, beverages or goods for personal hygiene. Thus, the dimensions of orientation towards nourishment permitted somewhat better predictions than those of general life orientation.

Reversing our logic of analysis we considered value orientations of a specific kind to be outcomes of lifestyle practice. In doing so, our style types proved to be moderately effective explanantia of ecological consciousness as such kind of value orientations, with this, those dimensions of orientation that are closer to personal concern, are clearly dependent on life style, whereas differences in attitudes on the macrolevel of conservational policy are not substantial.

If only the effects of socio-economic variables on style association, we did not quantify more detailed in this paper, would be taken into account, our analysis might be exposed to the objection of being trivial: it is notorious in the field of research on consumption after all that consuming patterns are dependent on given resources/restrictions in a certain way.



Therefore, we put up the theoretical relevance of our contribution for discussion by the following statements:

Patterns of shopping behaviour on the microlevel are imbedded in general lifestyles, and this becomes apparent even in the area of fast moving goods. Further refining and specifying relationships of such kind might improve the definition of target groups and market segments in everyday consumption.

The explanation of shopping patterns according to the lifestyle approach might enlarge our knowledge about the development and framing of consumers' preferences – probably a useful auxiliary service of sociology to the micro-economics of consumption.

## Contact

*Prof. Dr. Hartmut Lüdtke and Jörg Schneider*  
*Institut für Soziologie, Philipps-Universität Marburg*  
*Ketzerbach 11*  
*35032 Marburg*  
*Tel. 28 24584 (Prof. Lüdtke)*  
*Tel. 28 24588 (Mr. Schneider)*  
*Fax 28 28978*  
*luedtkeH@mail.uni-marburg.de*  
*schneidj@mail.uni-marburg.de*

## References

- Ajzen, Icek, 1988: *Attitudes, Personality and Behavior*. Milton Keynes: Open University Press.
- Bardeleben, Hans, 1991, *CONCLUS-PC. Programmbeschreibung*. Gießen.
- Esser, Hartmut, 1999, *Soziologie. Spezielle Grundlagen. Bd. 1: Situationslogik und Handeln*. Frankfurt a.M./New York: Campus.
- Georg, Werner, 1998, *Soziale Lage und Lebensstil. Eine Typologie*. Opladen: Leske+Budrich.
- Hartmann, Peter H., 1999, *Lebensstilforschung: Darstellung, Kritik und Weiterentwicklung*. Opladen: Leske+Budrich.
- Hradil, Stefan, 1987, *Sozialstrukturanalyse in einer fortgeschrittenen Gesellschaft*. Opladen: Leske+Budrich.

Hradil, Stefan, 1992, Alte Begriffe und neue Strukturen. In: ders. (ed.), *Zwischen Bewusstsein und Sein. Die Vermittlung 'objektiver' Lebensbedingungen und 'subjektiver' Lebensweisen*. Opladen: Leske+Budrich, 15-55.

Konietzka, Dirk, 1995, *Lebensstile im sozialstrukturellen Kontext*. Opladen: Westdeutscher Verlag.

Lüdtke, Hartmut, 1989, *Expressive Ungleichheit. Zur Soziologie der Lebensstile*. Opladen: Leske+Budrich.

Lüdtke, Hartmut, 1995a, *Zeitverwendung und Lebensstile. Empirische Analysen zu Freizeitverhalten, expressiver Ungleichheit und Lebensqualität in Westdeutschland*. Marburger Beiträge zur Sozialwissenschaftlichen Forschung 5. Marburg: Philipps-Universität.

Lüdtke, Hartmut, 1995b, *Vier Dimensionen von Lebensstilen. Zur Anwendung der Cluster- und Korrespondenzanalyse*. *Angewandte Sozialforschung* 19, Heft 1, 77-92.

Lüdtke, Hartmut, 1996, Methodenprobleme der Lebensstilforschung. Probleme des Vergleichs empirischer Lebensstiltypologien und der Identifikation von Stilpionieren. In: Otto Schwenk (ed.), *Lebensstil zwischen Sozialstrukturanalyse und Kulturwissenschaft*. Opladen: Leske+Budrich, 139-164.

Lüdtke, Hartmut, 2000, Konsum und Lebensstile. In: Doris Rosenkranz, Norbert F. Schneider (eds.), *Konsum. Soziologische, ökonomische und psychologische Perspektiven*. Opladen: Leske+Budrich, 117-132.

Mitchell, Arnold, 1983, *The Nine American Lifestyles. Who we are and where we're going*. New York: Macmillan.

Müller, Hans-Peter, 1992, *Sozialstruktur und Lebensstile. Der neuere theoretische Diskurs über soziale Ungleichheit*. Frankfurt/M.: Suhrkamp.

Sobel, Michael E., 1981, *Lifestyle and Social Structure. Concepts, definitions, analyses*. New York: Academic Press.

Sobel, Michael E., 1983, *Lifestyle Expenditures in Contemporary America. Relations between stratification and culture*. *American Behavioral Scientist* 26: 521-533.

Spellerberg, Annette, 1996, *Soziale Differenzierung durch Lebensstile. Eine empirische Untersuchung zur Lebensqualität in West- und Ostdeutschland*. Berlin: Edition Sigma.

---

# **IS A SUSTAINABLE CONSUMPTION PATTERN GRADUALLY EMERGING IN GERMANY? GRADUAL CHANGES IN THE SUSTAINABILITY OF THE CONSUMPTION PATTERN ANALYZED BY MEANS OF A CONSUMER PANEL**

*JOHN THØGERSEN*

## **1. Introduction**

When ways to activate a sustainable consumption pattern are discussed in the research literature and in political debates, implicitly or explicitly one or both of two main roads to this end is/are generally envisioned: the road of the small steps, where one may hope that one step leads to the next, and the detour via the main road of influencing general values and norms, where one hopes that the new values and norms lead to more reasoned lifestyle changes. Currently, the behavioural sciences are not able to provide scientifically based answers as to which road might be preferable, or whether a combination should be striven for. The main reason is that most research on environment and consumer behaviour has concentrated on studying behaviour within narrowly defined sectors. We even lack scientific knowledge about how individuals' propensities to behave in an environmentally friendly way in different domains or situations are related to each other. The international literature contains studies that report positive correlations between people's propensities to behave in an environmentally friendly way across domains (e.g., Berger, 1997) as well as studies that fail to find such correlations (e.g., Stern & Oskamp, 1987).

There are good arguments both for and against a hypothesis about interrelated propensities to behave in an environmentally friendly way in different situations. It is possible that environmentally friendly behaviour often requires knowledge or a willingness to seek information that are more prevalent in some people than in others. If this is true, these people would more often than others, everything else being equal, behave in an environmentally friendly way in any specific domain. A positive relationship across domains may

also have causes of a psychological nature. The hypothesis that environmentally friendly behaviour has a tendency to "spill over" into other behavioural domains (Frey, 1993; Thøgersen, 1999b) is backed by a group of psychological theories (especially Balance Theory and Dissonance Theory) that claim that we have a need to avoid inconsistencies in our beliefs, attitudes, and behaviours (Thøgersen, 1999a). Bem's (1972) Self-Perception Theory of attitude formation gives additional support to this hypothesis. This theory predicts that if a person starts behaving environmentally friendly in one area (e.g., separating one's waste for recycling), that person's attitudes and self-image might change in a way that increase his or her preparedness to behave environmentally friendly in other areas.

However, other arguments support the view that correlations between behaviours in different situations or domains are small, non-existing, or perhaps even negative. Environmentally friendly behaviour in specific settings is sometimes interpreted as an attempt of making amends (e.g., Wenke, 1993). It is claimed that people to some degree perform relatively easy environmentally friendly behaviours (like participation in a municipal recycling program) in order to make it easier to reject performing more demanding or costly behaviours (like biking to work instead of going by car) (Halkier, 1997). Some psychological theories also suggest mechanisms that may block the "spillover" between behavioural domains, for example Schwartz's (1977) Norm Activation Theory. This theory predicts that when the personal costs of behaving in a way that primarily benefits others or the society at large are perceived to be too high people tend, as a defence reaction, to post-rationalise the situation. People may neutralise the moral attitude or norm dictating pro-social behaviour by denying that continuing their current behaviour has any serious consequences or by denying their own responsibility for solving the problems produced by their current behaviour (Schwartz, 1968, 1973, 1977; Schwartz & Howard, 1980).

The primary objective of this paper is to detect whether, due to psychological mechanisms like those mentioned above interacting in a virtuous circle or for other reasons, environmentally friendly behaviours are spreading to more and more areas of the consumption pattern. An additional objective is to determine whether such a virtuous circle is facilitated if the individual possesses certain attitudes or values.

One may reasonably expect that spillover be facilitated if two product groups share the same environmentally relevant characteristic. More generally, spillover should be facilitated more the more common characteristics the product groups are perceived to have, i.e., the more likely it is that they are somehow categorised in the same mental category

(Thøgersen, 1999a; Thøgersen & Ölander, 1999). Hence, a third objective is to determine whether spillover is facilitated if the behaviours in question share common characteristics.

In order to reach these objectives, a research design that deviates from most previous studies of consumer behaviour in the environmental field is needed. Since the research questions concern evolution through time (changes in attitudes and behaviour), panel data are needed. By means of panel data it is possible to investigate whether consumers who behave environmentally friendly in area A today are more likely than others to behave environmentally friendly in area B tomorrow, which would indicate that a virtuous circle is operating. The study reported in this paper is based on household panel data from GfK (see specifications below). This data set makes it possible to study whether consumers who buy environment friendly products in one product category (e.g., beverages in returnable bottles) at one point in time are more likely than those who do not buy environment friendly products in other product categories (e.g., paper goods made of recycled fibres) later. Further, by classifying the consumption areas covered by the data according to shared characteristics; it is possible to see whether consistency in environment-friendly buying is facilitated by such similarities. In addition, attitude data connected with the panel data makes it possible to investigate whether certain attitudes or values facilitate environmentally friendly changes in behaviour. Hence the data set allows us to answer the questions implied by the objectives stated above with regard to this particular setting.

## **2. Data and method**

### *Panel data*

The empirical part of this study is based on a large consumer panel, GfK's Household Panel no. 1 for 1995, consisting of more than 4.000 German households who have filled out detailed shopping diaries on a continuous basis for a full year.<sup>1</sup> The panel data covers 57 product groups, each consisting of a number of specific products. In order to simplify the analysis and condense changes in the shopping pattern during the year covered by the data, the panel data set was divided in two, covering the first and the last six months respectively. Obviously, under normal circumstances deep and radical changes cannot be expected to happen over such a short time period. However, due to the fairly large sample size it is possible to detect the embryonic signs of changes that may significantly mark consumption patterns of the future. This paper represents a first attempt to use this type of data for a study of this type. Hence, in order not to complicate matters unnecessarily only

---

<sup>1</sup> I am grateful to the Zentrum für Umfragen, Methoden und Analysen (ZUMA) in Mannheim, Germany, for making this data set available to me.

simple descriptive and correlation analysis tools are used. Broadly spoken, this panel is limited to packaged goods and beverages, including dairy products, but excluding fresh vegetables and meat. Three types of environment-relevant product characteristics are reported in the shopping diaries:

For two product groups - kitchen rolls and toilet paper - it is reported whether the product is "environment-friendly" or not. These products are coded as environment-friendly if their packaging is marked with an environment-label ("Der Blaue Engel") or with information stating that it is made of recycled fibres or is unbleached.

For one additional product group - coffee and tea filters - it is registered whether the product is bleached or unbleached (bleaching being environmentally harmful).

Environment-relevant packaging attributes are reported for 28 product groups. This category is by far the most heterogeneous and the only one of the three that needs further introduction.

### ***Packaging data***

A number of packaging attributes are perceived by consumers as environmentally relevant (Bech-Larsen, 1996; Thøgersen, 1996). Particularly, the German regulation as of 1991, mandating that packaging waste should be separately collected by producers (or their representatives), is likely to have increased the attention of German consumers towards this issue. However, it is not at all obvious what packaging attributes consumers perceive to be environment-friendly, and studies show that many German consumers are uncertain about this issue (e.g., Thøgersen, 1996). Thøgersen (1996) found that in 1992 the packaging material was the most important cue to a cream cheese packaging's environment-friendliness,<sup>2</sup> and that glass was considered the most environmentally friendly material followed by paper and cardboard, while plastic and aluminium were considered environmentally unfriendly. However, the degree to which consumers associate packaging material with environment-friendliness is bound to depend on a number of conditions, particularly the packaging traditions of the product group and the presence or absence of more unequivocal cues to environment-friendliness. As regards tradition, it is hard to imagine that consumers associate environment-friendliness to glass wine-bottles since glass bottles is the traditional and dominant way of packaging wine.<sup>3</sup> More unequivocal

---

<sup>2</sup> None of the more unequivocal cues to environment-friendliness mentioned below were available for cream cheese at this time.

<sup>3</sup> If anything, the glass packaging here - as in a number of other cases - is presumably associated with high quality since wines that are filled in alternative packagings are generally of an inferior quality.

cues may be an environmental label, a "packaging free" system, a refill-system, or a return-and-reuse system (with or without a deposit). Again, the degree to which these cues are perceived as environment-relevant presumably depends on the packaging tradition within the product category (and which other associations the packaging form holds). In Germany, the presence of the "Green Dot" on most any packaging may have removed the importance of labels as a cue to the environment-friendliness of the packaging and may have made consumers even more uncertain about how to (and the need to) distinguish between packagings on this dimension. Table 1 lists the product groups in the GfK Household Panel 1 1995 where possible environment-relevant packaging attributes were registered. Below it will be explored whether German consumers perceive these product characteristics, as well as those mentioned earlier, as environment-relevant.

Information about possible environment-relevant packaging attributes are available for 28 of the 57 product groups covered by the Panel (not including wine and alcoholic beverages and a few other products nearly exclusively marketed in glass bottles or jars). The most common attribute is glass packaging (14 cases). In some of the included cases a glass packaging is probably not associated with environment-friendliness. Particularly, in cases where there are both one-way glass and returnable packaging options, it is assumed from the outset that only returnable bottles are perceived as environment-friendly. In 7 cases - all in the group of detergents and cleaners - it is registered whether or not a concentrated product is bought. Concentration clearly has implications for the amount of packaging, but may be preferred for other reasons as well (which is the case for most of these attributes). Perhaps the most unequivocal environmental attributes in the set are returnable (beverages and dairy products) and "refill" (detergents and cleaners and instant coffee) packagings.

Analytically, one may treat the environmental characteristics as dichotomous variables. However, one may suspect that it makes a difference with regard to the likelihood of spillover of environmentally concerned packaging choice between product groups whether or not the two groups share the same environmentally relevant characteristic (such as refill or concentrated). Still, in order to make the analyses more manageable, different kinds of returnable bottles (i.e., glass and plastic), different kinds of concentrated (referring to different levels of concentration as compared with the traditional standard), and different kinds of non-returnable glass packaging (i.e., bottles and jars) are aggregated.

Table 1: Product groups with environment-relevant packaging attributes

Product category	Label	Loose	Refill	Returnable	Glass	Concentrated
Windows, carpet and WC cleaners	W0					1
Tomato pure	W1				1	
Mayonnaise etc.	W2				1	
Detergent, low temp	W3		1			
Washing-up liquid	W4		1			1
Rough cleaner	W5					
Milk	W8	1		1	1	
Softener	W10		1			1
Tea	W18				1	
Universal cleaner	W21		1			1
Mustard	W23				1	
Floor maintenance	W28					1
Bath additives	W29		1			1
Potato products	W30	1				
Dessert	W31			1	1	
Beer	W33			1	1	
Juice	W36			1	1	
Soft drinks	W46			1	1	
Soft cheese	W47	1				
Linen starch etc.	W48		1			
Cream	W50			1	1	
Coffee whitener	W51				1	
Air freshener	W52		1			
Quark	W73			1	1	
Yoghurt	W78			1	1	
Hard cheese	W81	1				
Special cleaners	W82		1			
Mineral water	W84			1	1	
Number of cases		4	9	9	14	7

### *Attitude data*

In October every year, members of GfK's consumer panels fill out a questionnaire with a large number of background, consumption related, and attitudinal questions. The questionnaire given to Household Panel no. 1 in October 1995 contained, among other things, 14 items measuring various aspects of environmental attitudes. Appropriate attitudinal



indices for the present study have been extracted from these 14 items by means of exploratory factor analysis and item analysis.

The exploratory factor analysis using all environment-related attitude items as input produced a three-factor solution (after applying the Kaiser-criterion). However, the three factors only account for 48% of the variance in the items, indicating that either the sample holds very heterogeneous perceptions of environmental issues or the instrument as a whole is not well suited to capture the common perceptions that exist.<sup>4</sup> All three factors, but particularly Factor 2, capture items that seem to reflect environmental concern in general. In addition, Factor 1 captures items seemingly reflecting denial.

Based on these factors, item analysis was performed with the aim of maximising the internal reliability of a general environmental concern and a denial construct. A three-item indicator seems to be the most reliable measure of general environmental concern, the three items being "when I buy cosmetics and household products I'm very attentive towards their environment-friendliness," "the protection of nature is more important than continued economic growth," and "I'm prepared to pay higher prices for environment-friendlier packaging." Cronbach's alpha for this instrument is .69, which is acceptable.

The items that seems suitable for measuring the inclination to deny responsibility for solving environmental problems ("denial" for short) are "the government and industry should take the first steps to protect the environment, not the ordinary citizens," "cars have nothing to do with environmental problems," and "I cannot do much to protect the environment in my household." Cronbach's alpha for a scale based on these items is only .48. However, the three items may still reflect a latent denial construct, but may do so in a formative (of various types of denial) rather than a reflective way. Indeed, the items arguably reflect three types of denial (of responsibility for, consequences of, and ability to solve environmental problems) that individuals may differ in their propensity to use. However, what really matter for behaviour is not which type or denial is preferred, but rather whether or not individuals have a propensity to fall back on denial in order to avoid making unpleasant sacrifices. As should be expected, there is a significant negative correlation ( $r = -.31, p < .001$ ) between denial and environmental concern.

---

4 The factor analysis results can be acquired from the author.

### 3. Analyses and results

#### 3.1 Purchase frequency

This study is limited to the purchase of what is generally perceived as fast moving, "non-durable" or convenience goods. Still, the shopping diaries reveal that the frequency of buying these items varies a lot, both among consumers and among product categories (see Table 2). On average, these consumers buy milk nearly once a week and some does it every day. At the other end of the spectrum are a variety of household chemicals that the average household hardly buys once a year. Obviously, the volume of consumption and, hence, also of resource use, waste, and a number of other environmental impacts are positively correlated with purchase frequency, meaning that high-frequency product categories also should be the target of the highest concern. At the same time, the likelihood of developing a habitual shopping pattern increases with the behavioural frequency (Ouellette & Wood, 1998), meaning that environmentally harmful (but also environmentally beneficial) behaviour patterns are more difficult to change the higher the behaviour frequency. I will return to this issue later.

#### 3.2 Consumers' propensity to choose environment-friendly

German consumers' propensity to choose the environment-friendly option within a product category is measured as the proportion of purchases in the product category that have one of the environment-friendly attributes mentioned earlier. For most products, only one environment-relevant attribute is registered (environmental label, unbleached, returnable bottle, concentrated, etc.), but in three cases two attributes are registered (concentrated as well as refill packaging). In these cases, the propensities to choose products with each of these characteristics are calculated separately. Table 3 shows the propensity to choose environment-friendly regarding each of the studied product categories in the first and the last half of 1995. The stability of each propensity over the year, measured as Pearson's  $r$  and including only consumers who purchased the product in both halves of the year, is also shown.

German consumers' average propensity to choose (what is here assumed to be) the most environment-friendly option varies tremendously among these 31 product types, from (practically) 0 (potato products in lose weight and desserts in a returnable glass packaging) to about 90% (unbleached coffee and tea filters) of the purchases. The variation is practically unchanged between the first and the second half of 1995, and so is the rank order of the propensities to choose environment-friendly in the different product groups (Spearman's rank correlation = .95).

Table 2: *Buying frequencies<sup>1</sup> in the studied product groups in first and second half of 1995. N = 4426*

Product category	Label	Mean t1	Std Dev t1	Min t1	Max t1	Mean t2	Std Dev t2	Min t2	Max t2
Floor maintenance	W28	0.30	0.81	0	18	0.24	0.69	0	11
Air freshener	W52	0.31	1.17	0	19	0.30	1.14	0	17
Special cleaners	W82	0.35	0.83	0	9	0.38	0.85	0	10
Linen starch etc.	W48	0.40	0.99	0	19	0.37	0.98	0	17
Kitchen rolls	W32	1.17	1.85	0	23	1.20	1.87	0	20
Rough cleaner	W5	1.22	1.88	0	44	1.18	1.86	0	47
Detergent, low temp	W3	1.33	1.86	0	24	1.20	1.72	0	21
Filter paper	W75	1.45	1.63	0	25	1.41	1.59	0	26
Windows, carpet, and WC cleaners	W0	1.47	1.99	0	23	1.45	1.97	0	21
Softener	W10	1.71	2.55	0	27	1.60	2.46	0	29
Mustard	W23	1.72	2.11	0	23	1.68	2.11	0	22
Bath additives	W29	1.91	2.68	0	42	2.02	2.77	0	37
Washing-up liquid	W4	1.98	2.23	0	25	1.94	2.13	0	26
Potato products	W30	2.16	3.59	0	52	1.93	3.49	0	52
Mayonnaise etc.	W2	2.30	3.30	0	40	2.29	3.27	0	39
Universal cleaner	W21	2.33	2.57	0	29	2.23	2.44	0	28
Tomato pure	W1	2.82	3.88	0	50	2.71	3.82	0	50
Toilet paper	W99	3.65	3.65	0	47	3.68	3.63	0	51
Dessert	W31	3.97	7.55	0	83	3.58	7.02	0	67
Tea	W18	4.53	5.93	0	71	4.23	5.54	0	69
Coffee whitener	W51	6.00	6.64	0	46	5.89	6.61	0	44
Beer	W33	7.02	10.44	0	96	7.28	10.65	0	153
Soft drinks	W46	7.96	13.91	0	150	8.88	14.53	0	182
Cream	W50	8.52	9.06	0	75	8.20	8.91	0	78
Mineral water	W84	9.68	10.15	0	130	10.04	10.31	0	115
Juice	W36	10.73	14.16	0	168	9.80	12.93	0	133
Hard cheese	W81	11.07	13.20	0	114	10.79	12.71	0	103
Quark	W73	11.24	11.11	0	112	10.13	10.56	0	110
Yoghurt	W78	14.29	15.48	0	156	12.66	14.29	0	154
Soft cheese	W47	16.63	15.77	0	208	15.92	15.38	0	208
Milk	W8	23.01	23.27	0	189	23.15	23.99	0	209

<sup>1</sup> The buying frequency is the number of purchases in the product category in six months.

Table 3: *The propensity to choose environment-friendly options in the studied product groups in first and second half of 1995*

Product category	Label <sup>1</sup>	Mean t1	Std Dev t1	N t1	Mean t2	Std Dev t2	N t2	r t1-t2	N
Potato products	W30	0.00	0.04	2380	0.00	0.05	2278	0.58	1835
Dessert	W31	0.00	0.04	2318	0.00	0.03	2181	0.42	1827
Quark	W73	0.01	0.06	3929	0.01	0.04	3818	0.30	3667
Cream	W50	0.02	0.11	3625	0.02	0.11	3557	0.67	3347
Windows, carpet, and	W0	0.03	0.15	2585	0.05	0.19	2598	0.20	1975
Tea	W18	0.04	0.15	3258	0.06	0.18	3170	0.49	2784
Air freshener	W52	0.04	0.17	479	0.04	0.17	472	0.03 <sup>2</sup>	247
Floor maintenance	W28	0.05	0.22	897	0.05	0.20	776	0.33	362
Bath additives	w29	0.05	0.18	2713	0.07	0.20	2751	0.55	2170
Linen starch etc.	W48	0.05	0.20	1016	0.06	0.21	908	0.54	544
Yoghurt	W78	0.08	0.20	3830	0.08	0.20	3783	0.72	3584
Soft cheese	W47	0.09	0.18	4059	0.09	0.18	4004	0.68	3894
Milk	W8	0.09	0.23	4097	0.08	0.22	4075	0.86	3987
Special cleaners	W82	0.11	0.29	974	0.11	0.29	1079	0.48	504
Coffee whitener	W51	0.13	0.29	3333	0.13	0.29	3285	0.73	3006
Washing-up liquid	W4R	0.14	0.29	3283	0.13	0.28	3264	0.49	2740
Kitchen rolls	W32	0.15	0.32	2244	0.14	0.31	2284	0.46	1723
Detergent, low temp	W3	0.15	0.30	2492	0.20	0.35	2327	0.52	1756
Rough cleaner	W5	0.21	0.35	2393	0.29	0.39	2412	0.38	1685
Juice	W36	0.28	0.35	3746	0.29	0.36	3706	0.68	3436
Hard cheese	W81	0.30	0.36	3746	0.28	0.35	3722	0.78	3516
Softener	W10C	0.31	0.39	2573	0.05	0.20	2499	0.29	2029
Softener	W10R	0.34	0.40	2573	0.37	0.41	2499	0.53	2029
Toilet paper	W99	0.41	0.42	3784	0.41	0.42	3784	1.00	3784
Mayonnaise, etc.	W2	0.49	0.42	2720	0.53	0.42	2662	0.55	2182
Washing-up liquid	W4C	0.50	0.43	3283	0.60	0.41	3264	0.46	2740
Universal cleaner	W21R	0.51	0.42	3380	0.61	0.42	3366	0.53	2881
Tomato pure	W1	0.59	0.39	2952	0.61	0.40	2888	0.49	2403
Soft drinks	W46	0.61	0.40	3040	0.57	0.40	3204	0.70	2728
Universal cleaner	W21C	0.65	0.41	3380	0.45	0.41	3366	0.33	2881
Mustard	W23	0.68	0.41	2903	0.68	0.41	2843	0.62	2280
Beer	W33	0.77	0.35	3276	0.76	0.35	3309	0.71	2956
Mineral water	W84	0.86	0.30	3911	0.86	0.29	3952	0.83	3754
Filter paper	W75	0.91	0.26	2802	0.90	0.26	2760	0.51	2330
Average		0.28		2882	0.28		2848	0.54	2458

<sup>1</sup> The suffix R in a label means returnable or refillable packaging. C means concentrated.<sup>2</sup> Not significant at  $p < .05$ . All others are significant at  $p < .001$ .

The data set makes it possible to calculate two summary indicators that may be used to evaluate changes in the overall propensity to choose environment-friendly options when buying the studied products. One may count the number of product groups where individual consumers have chosen an environment-friendly option in the covered time periods. This measure can be used as an indicator of the *breadth* of environment-friendly buying which is central to the spillover ideas mentioned in the introduction. The average breadth indicator for the first six months is 10.05 and for the last six months it is 10.77. The increase is marginal, but statistically significant ( $t = 19.44$ , d.f. = 4425,  $p < .001$ ).

Further, one may calculate the simple average of the propensities to choose environment-friendly in the studied products groups. This measure may serve as an indicator of the *depth* of environment-friendly buying in the areas where people choose environment-friendly. The depth indicator for the first and second half of 1995 is shown in Table 3. According to this indicator, the overall propensity to choose environment-friendly has not changed during this year.

### 3.3 The stability of environment-friendly choice

The stability of environment-friendly choice in each product group is calculated as the Pearson correlation between individual consumers propensity to choose the most environment-friendly option in the first and the second half-year. In all cases except one, correlations between propensities to choose the environment-friendly option are significant and positive. The exception is the most infrequently bought item in the set, air freshener, where the environment-friendly attribute is a refill container. Besides this case, correlations vary from weak ( $r = .29$ , concentrated textile softener) to perfect ( $r = 1$ , environment-labelled toilet paper). The average stability is .54. Hence, the propensity to choose environment-friendly consumer goods (or not) seems to be characterised more by stability than by change. This is hardly a surprise, especially when the time-span is as short as in the present case.

Behavioural stability, particularly in areas like the one studied here, is typically attributed to habituation (e.g., Ouellette & Wood, 1998), although it is acknowledged that it may have other causes as well (such as stable preferences, cf., e.g., Ajzen, 1991). Habituation depends on the behaviour being performed at a high frequency (Ouellette & Wood, 1998). Hence, if (part of) the reason why practically all correlations reported in Table 3 are significant and positive is that consumers tend to buy whatever they buy (environment-friendly or not) habitually, the correlation between past and future behaviour should depend on the purchase frequency. This prediction may be tested by analysing the correlation between purchase frequency, reported in Table 2, and the stability of environment-friendly choice (the correlation between environment-friendly choice in the first and

second half of 1995), reported in Table 3. The correlation between average purchase frequency (calculated for the whole year) and the stability of environment-friendly choice in the product categories is indeed significant and positive ( $r = .57$ , d.f. = 30,  $p = .001$ ), thus supporting the suggestion that habitualisation is among the causes of the stability of the propensity to choose environment-friendly (or not). However, that it is also partly due to stable preferences can be seen by comparing the stabilities in those product areas where the environment-friendly purchase frequency is significantly correlated with the index of environmental concern (average stability .58, see Table 4) and those where it is not (average stability .50). The difference is statistically significant,  $p < .001$ . Hence, the stability of the propensity to buy environment-friendly is highest in areas where the environment-friendly attribute is recognised as such.

### 3.4 The perceived environmental relevance of these choices

To the degree that consumers' choice of an environment-friendly option is at least partly voluntary, it matters whether or not they perceive differences between options to be environment-relevant. Whether they do that in the cases studied here is not registered in either GfK's panel data or in the accompanying survey. However, if we accept the assumption that consumers are more likely to favour an option that they perceive as environment-friendly the more environmentally concerned they are, the correlation between a measure of environmental concern and the propensity to choose a product with an alleged environment-friendly attribute can be used as a rough indicator of whether or not the attribute is conceived as environment-relevant by consumers (see Table 4).

In absolute terms, none of the correlations that are reported in Table 4 are impressive. This is as should be expected when a general attitude measure is correlated with such specific behaviour measures (Ajzen & Fishbein, 1977; Weigel & Newman, 1976). In order to illustrate the importance of correspondence for the attitude-behaviour correlation, an aggregate index of the propensity to choose environment-friendly in these product areas has been calculated. The last row of Table 4 shows that this index correlates more strongly with environmental concern than any of the individual items.<sup>5, 6</sup>

---

5 The purchase of convenience goods is still a very narrow behavioural category compared to the measure of environmental concern. Hence the fairly low correlation.

6 In order to further substantiate the claim that the propensity to choose the options in focus here is related to their perceived environmental relevance, the correlation between the denial index and the aggregate propensity was also calculated. The correlation is statistically significant with the expected direction ( $r = .12$ ,  $p < .001$ ).

Table 4: *The correlation between general environmental concern<sup>1</sup> and the propensity to choose an environment-friendly version of various convenience goods*

Product area	Label	r	N	p
Softener	W10K	0.049	1902	0.034
Tomato pure	W1	0.027	1902	0.244
Soft drinks	W46	0.011	3272	0.539
Softener	W10R	0.009	3410	0.590
Tea	W18	0.004	3208	0.804
Detergent, low temp	W3	-0.010	2873	0.594
Potato products	W30	-0.011	2635	0.559
Rough cleaner	W5	-0.012	2934	0.525
Quark	W73	-0.022	3815	0.174
Windows, carpet and WC cleaners	W0	-0.023	2990	0.210
Linen starch etc.	W48	-0.025	1308	0.361
Coffee whitener	W51	-0.026	3369	0.127
Bath additives	w29	-0.029	3084	0.108
Dessert	W31	-0.032	2510	0.104
Mayonnaise etc.	W2	-0.034	3008	0.066
Special cleaners	W82	-0.038	1448	0.153
Mineral water	W84	-0.043	3832	0.008
Mustard	W23	-0.045	3244	0.011
Cream	W50	-0.045	3587	0.007
Washing-up liquid	W4R	-0.048	3556	0.004
Hard cheese	W81	-0.051	3691	0.002
Beer	W33	-0.056	3387	0.001
Juice	W36	-0.057	3751	0.000
Kitchen rolls	W32	-0.068	2636	0.000
Universal cleaner	W21R	-0.072	3612	0.000
Washing-up liquid	W4K	-0.074	3556	0.000
Universal cleaner	W21K	-0.075	3612	0.000
Floor maintenance	W28	-0.077	1229	0.007
Filter paper	W75	-0.077	3018	0.000
Soft cheese	W47	-0.084	3896	0.000
Milk	W8	-0.087	3904	0.000
Air freshener	W52	-0.102	659	0.009
Yogurt	W78	-0.106	3761	0.000
Toilet paper	W99	-0.147	3774	0.000
All product areas		-0.149	4121	0.000

<sup>1</sup> Measured on an index going from 3 (highest) to 15 (lowest).

Significant, but weak correlations found in studies based on large samples are often discounted based on the argument that had the sample size not been large they would not have reached significance. This is an appropriate careful way of reasoning in cases where the correlation has a weak theoretical backing. However, in cases, such as the present, where there are sound theoretical reasons to expect a correlation, and also that the correlation is bound to be weak, significant, but weak, correlations cannot be discounted with reference to the sample size. On the other hand it is a strong message when a correlation does not reach significance in spite of a large sample size, which is the case in 14 of the 34 cases analysed in Table 4. In one additional case, the correlation is significant, but with the wrong sign. Hence, in 15 cases product attributes that are technically environment-relevant seem not to be perceived as such by the consumers. Many of these cases are infrequently bought household chemicals and speciality foods, but there are also more frequently bought products among them, such as soft drinks, tea, and coffee whitener. In 19 cases is the correlation between the choice of an environment-friendly option and environmental concern statistically significant and in the right direction.

### **3.5 The breadth and depth of buying environment-friendly in areas perceived as environment-relevant**

If only those product areas where the propensity to choose environment-friendly is significantly correlated with environmental concern are counted in, the average breadth indicator (mentioned above) for the first six months is 7.06 and for the last six months it is 6.74. The decrease is marginal, but statistically significant ( $t = -10.37$ , d.f. = 4425,  $p < .001$ ). If the same procedure is used for calculating the depth indicator (mentioned above), it falls from .36 in the first to .34 in the second half of 1995. The difference is not statistically significant ( $z = .898$ ), however.

### **3.6 Consistency in environment-friendly purchases across products**

The key question, raised when motivating this study, is whether consumers are consistent in their propensity to choose environment-friendly options in different product groups. It was suggested that, due to spillover, consistency should be expected to increase over time. In addition it was suggested that the (increase in) consistency may depend on perceived similarities between the product (or behavioural) categories and on whether or not the consumer holds certain environment-relevant attitudes. In order to test these propositions, correlation analyses have been performed focusing on propensities to choose environment-friendly options in different product categories. The focus is not on the individual bivariate correlations, but on aggregates produced by averaging over groups of products, or more precisely over bivariate correlations between environment-friendly purchases of products in the group. Such analyses have been made for the total sample as well as for



sub-samples, where the score on the environmental concern and denial indexes have been used to split the sample (see Table 5).

Table 5: Average correlations between propensities to choose environment-friendly options in different product categories in the first and last six months of 1995

	All		Products where propensity to buy environment-friendly correlates with environmental concern									
			All		Low concern		High concern		Low denial		High denial	
	t1	t2	t1	t2	t1	t2	t1	t2	t1	t2	t1	t2
Total	0.03	0.03	0.04	0.04	0.03	0.02	0.05	0.04	0.04	0.05	0.04	0.03
Packaging												
Loose	0.18	0.20	0.49	0.48	0.45	0.46	0.51	0.50	0.51	0.50	0.46	0.47
Return	0.11	0.11	0.15	0.15	0.12	0.13	0.16	0.16	0.16	0.17	0.13	0.12
Refill	0.04	0.04	0.03	0.02	0.06	0.02	0.01	0.00	0.00	0.02	0.06	0.01
Glass	0.00	0.01										
Concentrated	0.01	0.02	0.02	-0.01	0.03	0.00	0.02	-0.01	0.03	-0.01	0.01	0.00
Green label	0.14	0.17	0.14	0.17	0.21	0.17	0.14	0.17	0.20	0.19	0.14	0.15
Product category												
Detergents and cleaners	0.04	0.03	0.07	0.04	0.08	0.05	0.06	0.04	0.07	0.03	0.07	0.06
Cheese	0.49	0.48	0.49	0.48	0.45	0.46	0.51	0.50	0.51	0.50	0.46	0.47
Dairy products	0.14	0.16	0.19	0.19	0.12	0.13	0.21	0.22	0.23	0.22	0.12	0.14
Beverages	0.22	0.22	0.20	0.19	0.19	0.18	0.20	0.20	0.20	0.21	0.19	0.17
Paper goods	0.09	0.09	0.09	0.09	0.11	0.08	0.10	0.09	0.13	0.11	0.08	0.07
Dressings	0.00	0.00										

Note: Differences between correlations larger than NN are significant,  $p < .05$ .

Focusing first at the most aggregate level, Table 5 shows the average of correlations between different propensities to buy environment-friendly, calculated both for all covered cases and for only those cases where the propensity is significantly correlated with environmental concern. Consistent with the suggestion that consistency depends on environmental concern and the perceived environmental relevance of the options, the average correlation is highest in the latter case and only in this case it is (marginally) statistically significant (d.f.  $\approx 2,300$ ). However, the difference between the correlations is not significant at this level of aggregation. Neither does the aggregate correlation change between

the two time periods. Hence, judged at this level of aggregation the behaviour of German consumers seems neither to be particularly consistent, nor to be in a process towards a higher level of consistency. On this level of aggregation, when the sample is split according to environmental concern or denial (in those scoring below and those scoring above the sample mean of each index), differences between sub-samples have the expected sign (high concern and low denial seeming to facilitate greater consistency), but they are not statistically significant.

In most cases, the consistency becomes quite a bit higher when the focus is narrowed to product categories that share more common characteristics. Exceptions are a category where the common characteristic is the material of the packaging (glass) and a category of speciality foods (dressings). In neither of these categories is more than one purchase behaviour significantly related to environmental concern. Consistency in choice is also absent regarding the purchase of refills and concentrated products (two options that are limited to detergents, cleaners, and the like). Those among these purchase behaviours that are related to environmental concern tend to be so relatively weak.

The strongest average correlation among packaging choices regarding different products concerns “loose weight” and the strongest one among environment-friendly choices within product categories concerns (different types of) cheese. When only purchase behaviours that are significantly correlated with environmental concern are counted in, the loose weight and the cheese categories boil down to the same thing. Obviously, there are a lot of common characteristics among the purchases of different kinds of cheese in loose weight, including that they are often bought in the same speciality store or speciality counter in the supermarket. Table 4 shows that the propensity to choose cheese in loose weight is correlated with environmental concern, but of course, other factors may play an even bigger role. Average correlations are also relatively high for the choice of returnable packaging and for choosing beverages and dairy products (other than cheese) in environment-friendly packaging. The same is the case for the choice of environment-labelled products and paper goods (the environment-labelled products plus filters). In all these cases (where it is relevant) except one, the average correlation improves when behaviours that are not significantly correlated with environmental concern are removed. Further, in all of these cases, except for the purchase of paper goods and environment-labelled products (i.e., also part of the paper goods category), environmental concern makes a significant difference (except for beverages) for the level of consistency in behaviour. The inclination to denial makes an equally significant difference in the same cases and also for the purchase of environment-labelled products and paper goods. However, no tendency to increased consistency (or the reverse) over time can be detected in either of the categories or (sub-) samples.

## **4. Summary and implications**

The key question motivating this paper is whether or not environmentally friendly behaviours are spreading to more and more areas of the consumption pattern. The answer to this question, and answers to follow up questions about conditions for and determinants of such a tendency, are of utmost importance for strategies aimed at facilitating a more sustainable consumption pattern. The conditions in focus in this paper are individuals' possession of certain attitudes or values and the degree to which product or behavioural categories share the same environmentally relevant characteristics.

The main conclusions are:

There are mixed evidence as to whether the breadth of environment-friendly buying by German consumers expanded or contracted over the studied time period (one year). In all cases, the change was marginal. Overall, the depth or intensity of environment-friendly buying inside product areas did not change.

The propensity of German consumers to shop in an environment-friendly way is highly stable within most product areas. The stability is positively correlated with the purchase frequency of a product, indicating that it is at least partly due to habit. However, the stability also depends on the perceived environmental relevance of the choice, indicating that it is partly due to stable preferences.

The propensity of German consumers to shop in an environment-friendly way is far from consistent across (convenience) product areas overall. However, the consistency is higher within areas that seem to be perceived as environment-relevant and more so among purchases that share many than among those sharing few common characteristics.

Environmental concern seems to facilitate and an inclination to denial seems to reduce consistency.

The level of consistency did not change over the studied time period.

As regards the key question, the cross-sectional and the time-series evidence from this study points in opposite directions. The study finds no clear signs of environment-friendly behaviour spreading to more areas of the consumption pattern over time. Here, the overwhelming evidence points towards stability rather than change. On the other hand, the cross-sectional evidence indicates that under the right conditions consumers tend to be consistent in their propensity to shop in an environment-friendly way. Hence, it follows that environment-friendly behaviour must have spread between different areas of the consumption pattern when these conditions were present. The conditions identified to influence consistency – and, hence, the spread or spillover of a propensity to shop in an

environment-friendly way – are perceived environmental relevance, common characteristics (i.e., perceived similarity) between areas, environmental concern, and an inclination to denial (negative influence).

Although seeming contradictory, these results can be reconciled in the conclusion that a tendency for environment-friendly behaviour to spread to more and more areas of the consumption pattern apparently exists, but under conditions similar to those existing in Germany in the middle of the 1990's it is so weak and slow that it is not traceable in time-series data covering a period of just one year. From this follows a number of implications. Research-wise it suggests that an attempt should be made to replicate this study using a longer timeframe. Policy-wise attempts should be made to identify conditions that facilitate and conditions that hamper the spillover of environment-friendly behaviour. This study suggests that consumers insufficient understanding of the environmental relevance of choices they make, and perhaps also of the similarity of environmentally relevant characteristics across product and behavioural areas, are among the barriers to sustainable consumption. Such barriers could be targeted with specific educational programmes. The study also suggests that more basic attitudes and personality traits influence consumers propensity to adopt a more sustainable consumption pattern. Specifically, this propensity is facilitated by a high level of environmental concern and hampered by an inclination to deny ones responsibility. Attitudes and personality traits like these are presumably built over a long period of time and should be targeted by a broad based long run education and socialization effort at all levels of society, but with the family and in the education sector being key agents of change.

## **Contact**

*John Thøgersen  
Department of Marketing  
Aarhus School of Business  
Haslegaardsvej 10  
DK-8210 Aarhus V  
Tel 8948 6440  
Fax 8615 3988  
John.Thogersen@mar.hha.dk*

## References

- Ajzen, I. (1991). The Theory of planned behavior. *Organizational Behavior and Decision Processes*, 50, 179-211.
- Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, 84(5), 888-918.
- Bech-Larsen, T. (1996). Danish consumers' attitudes to the functional and environmental characteristics of food packaging. *Journal of Consumer Policy*, 19, 339-363.
- Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 6, pp. 1-62). New York: Academic Press.
- Berger, I. E. (1997). The demographics of recycling and the structure of environmental behavior. *Environment and Behavior*, 29, 515-531.
- Frey, B. S. (1993). Motivation as a limit to pricing. *Journal of Economic Psychology*, 14, 635-664.
- Halkier, B. (1997, 28. januar). Nemt at tørre ansvaret af på forbrugerne. *Information*, pp. 3.
- Ouellette, J. A., & Wood, W. (1998). Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychological Bulletin*, 124(1), 54-74.
- Schwartz, S. H. (1968). Words, deeds, and the perception of consequences and responsibility in action situations. *Journal of Personality and Social Psychology*, 10, 232-242.
- Schwartz, S. H. (1973). Normative explanations of helping behavior: A critique, proposal, and empirical test. *Journal of Experimental social psychology*, 9, 349-364.
- Schwartz, S. H. (1977). Normative influence on altruism. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, Vol. 10 (pp. 221-279). New York: Academic Press.
- Schwartz, S. H., & Howard, J. A. (1980). Explanations of the moderating effect of responsibility denial on the personal norm-behavior relationship. *Social Psychology Quarterly*, 43(4), 441-446.
- Stern, P. C., & Oskamp, S. (1987). Managing scarce environmental resources. In D. Stokols & I. Altman (Eds.), *Handbook of environmental psychology* (Vol. 2, pp. 1043-1088). New York: Wiley.
- Thøgersen, J. (1996). The demand for environmentally friendly packaging in Germany (MAPP Working Paper 30): Department of Marketing, The Aarhus School of Business.
- Thøgersen, J. (1999a). Cognitive dissonance and the development of a sustainable consumption pattern. Submitted for review.

Thøgersen, J. (1999b). Spillover processes in the development of a sustainable consumption pattern. *Journal of Economic Psychology*, 20, 53-81.

Thøgersen, J., & Ölander, F. (1999). Danske forbrugeres oplevelse af forskelle og ligheder mellem miljørelevante aktiviteter (Working Paper 1999-3). Århus: Marketing og Miljø Forskningsgruppen, Institut for Markedsøkonomi, Handelshøjskolen i Århus.

Weigel, R. H., & Newman, L. S. (1976). Increasing attitude-behavior correspondence by broadening the scope of the behavioral measure. *Journal of Personality and Social Psychology*, 33(6), 793-802.

Wenke, M. (1993). Konsumstruktur, Umweltbewußtsein und Umweltpolitik. Eine makro-ökonomische Analyse des Zusammenhanges in ausgewählten Konsumbereichen. Berlin: Duncker & Humblot.

---

# **PURCHASING HEALTHY FOOD IN GERMANY. AN EMPIRICAL ANALYSIS OF ITS ATTITUDINAL AND SOCIO-ECONOMIC ANTECEDENTS**

*Dieter Ohr*

**I**n diesem Beitrag wird der Kauf gesunder Nahrungsmittel untersucht und geprüft, welchen Einfluß dabei ernährungsbezogene Einstellungen einerseits, sozioökonomische Merkmale andererseits auf das Kaufverhalten ausüben. Datengrundlage der empirischen Analysen sind zwei Stichproben deutscher Haushalte ( $n_1=4426$ ,  $n_2=4638$ ). Es zeigt sich, daß Haushalte, in denen gesunde Ernährung auf der Einstellungsebene bedeutsam ist, auch tatsächlich signifikant mehr gesunde Nahrungsmittel kaufen. Neben dem Einfluß ernährungsbezogener Einstellungen lassen sich auch einige interessante sozioökonomische Differenzierungen nachweisen. So werden mehr gesunde Nahrungsmittel konsumiert bei höherer Bildung, niedrigerem Lebensalter, wenn kleine Kinder im Haushalt leben und in westdeutschen Haushalten, vergleicht man diese mit ostdeutschen Haushalten.

**T**his paper examines the purchase of healthy food in Germany, in particular the influence which two classes of determinants may exert, that is attitudes concerning nutrition and socio-economic characteristics. The empirical analyses are based on two samples of German households ( $n_1=4426$ ,  $n_2=4638$ ). It is shown that attitudes regarding the health aspect of nutrition do in fact exert a significant influence on the purchase of healthy food. Households where beliefs and preferences favoring a healthy nutrition predominate are also more likely to purchase healthy food products. In addition to nutritional attitudes, socio-economic differences play also a role in explaining the decision to buy healthy food. Its purchase is more likely with higher education, lower age, when small children are living in the household, and, for households in West-Germany in comparison to East-Germany.

## 1. Why Consumers Choose Healthy Food: Arguments and Research Hypotheses<sup>1</sup>

During the last decades the general public has become more sensitive with respect to its food. Several food scandals, for instance pollution or even toxic substances in food articles, set the stage for a still growing consciousness that there might be a close link between nutrition and health. More recently, the so-called 'mad cow disease' marked another culmination point in a long chain of food scandals. To be sure, an increasing number of media reports on food scandals does not necessarily mean that the quality of food currently is in fact worse than three or four decades ago. Rather, a growing number of reported food scandals may in part also be due to the fact that the media system has changed substantially during the last decades and competition has increased. But whether or not the greater awareness of consumers concerning healthy nutrition is based on an accurate perception of real-world developments, the increase in public sensitiveness can hardly be disputed and will most certainly affect consumer behavior.

There is another important reason why the link between health and nutrition could have become more prevalent in public debates and in the consumers' minds. Quite apart from food scandals we have observed a long-term secular development towards a higher well-being of the western, industrialized societies. As in the course of this process the economic situation of most citizens in the developed societies has improved, the demands which are put on nutrition have risen because health is, in microeconomic terminology, a superior good. Citizens thus develop clear beliefs and intense preferences on how and what to eat and drink. These beliefs and preferences could, then, in turn be reflected in the citizens' consumer choices. Therefore, beliefs and preferences – in short: attitudes – on nutrition should have a clear and relevant effect on the purchase of food articles. In particular, consumers with intense nutritional attitudes towards a pro health direction should choose healthy food articles to a higher degree than consumers without such attitudes. Whether and to what degree this is in fact the case will be analyzed in this article.

Postulating that attitudes on nutrition have a strong impact on the purchase of healthy food implies that a consumer's choice in the small local shop or the supermarket is at least partly motivated by conscious consideration and the desire to act accordingly. Conversely, the act of purchasing food is neither completely spontaneous and largely based on the momentary mood nor a mere execution of habits without any conscious *involvement* (cf.

---

<sup>1</sup> For their help in coding the food articles I would like to thank *Stephanie Bous* and *Petra Streitz* as well as *Omer Gersten* for his support in translating the often very specific concepts for the food articles into English.



Kroeber-Riel 1992 for the involvement concept in consumer research). Purchasing food in the literature is rather conceived as behavior with an intermediate degree of involvement and with a certain degree of mental control. Conceiving food purchasing in this way may also imply that long-standing habits are followed if these habits are connected with stable attitudes (Kroeber-Riel 1992: 169).

It is, first, an aim of this article to test and to examine the postulated relationship between nutritional attitudes and the purchase of healthy food articles. Secondly, it will also be investigated in a more exploratory manner which socio-demographic variables are able to account for the purchase of healthy food.

## 2. Cross-sectional and Panel Approaches to Analyze the Purchase of Healthy Food

A cross-sectional approach is a first, straightforward method to shed some light on the relationship between nutritional attitudes and purchasing healthy food. When using this approach it is assumed that current attitudes on nutrition as well as current socio-economic positions are causally antecedent to food purchase:

$$Y_{i,t} = \beta_0 + \beta_1 \cdot A_{i,t} + \dots + \beta_2 \cdot SE_{i,t} + \dots + \varepsilon_{i,t} \quad (1)$$

In equation (1)  $A_{i,t}$  indicates a particular attitude on nutrition at time  $t$ ,  $SE_{i,t}$  stands for a typical variable out of a set of socio-economic variables, and  $Y_{i,t}$  represents a measure for the purchase of healthy food.

In many cases it is theoretically as well as empirically promising to include the former level of the dependent variable into the model equation:

$$Y_{i,t} = \beta_0 + \beta_1 \cdot A_{i,t} + \dots + \beta_2 \cdot SE_{i,t} + \dots + \beta_3 \cdot Y_{i,t-1} + \varepsilon_{i,t} \quad (2)$$

The effect of the  $Y$ -variable measured at  $t-1$  on the same variable at time  $t$  can designate how stable purchase behavior is over time. A large effect, for example, might indicate that the behavior is habitualized for the most part. Alternatively, the variable can also be understood as a proxy measure for those determinants of purchase behavior not taken account of in the model. Since most often past behavior is an excellent predictor of behavior

in the future, explained variance in model (2) is likely to be much higher compared to model (1). If nutritional attitudes and socio-economic variables are important predictors for purchase behavior in  $t$  both types of variables will also be strongly correlated with purchase behavior in  $t-1$ . Hence, the effects being estimated for these variables with model (2) will most certainly be much smaller than those in model (1). Which kind of estimates better reflect the 'true' effects is difficult to decide in general and strongly depends on the researcher's theoretical model. It is safe to say, however, that the coefficients for attitudes and socio-economic variables are likely to be somewhat overestimated in model (1) and somewhat underestimated in model (2).

Attitudes on nutrition may change during a time period. These changes may be due to normal maturation in the course of the life-cycle but can also go back to so-called period effects, for example media reports covering the contamination of food products. Attitudinal changes may, then, in turn induce changes in consumer expenditures on food. To test this presumption, one can make use of model (3):

$$\Delta Y_i = \beta_0 + \beta_1 \cdot \Delta A_i + \dots + \beta_2 \cdot \Delta SE_i + \dots + \beta_3 \cdot Y_{i,t-1} + \varepsilon_{i,t} \quad (3)$$

Here,  $\Delta A_i$ , for example, indicates the attitudinal change from  $t-1$  to  $t$ . In most applications of model (3) the effect of the lagged endogenous variable,  $Y_{t-1}$ , will be negative. This negative effect, then, designates that an already high level of purchases in the preceding period is followed by rather small increases in the following period.

Equation (3) represents the common view which states that, generally, attitudinal changes produce behavioral changes. As far as purchase behavior is concerned, attitude changes regarding nutrition would be causally prior to changes in purchasing food. Yet, the competing hypothesis which postulates just the opposite causal direction is plausible as well: Consumers might change their purchase habits, as a result of TV commercials for instance, and may consequently adjust their attitudes. Seen in this perspective,  $\Delta Y_i$  would be causally antecedent to  $\Delta A_i$ . Thus, one can plausibly argue for both causal directions.

In principle, each of the three approaches mentioned could be conducted with the *ZUMA datasets of the GfK household and consumer panel 1995* (henceforth in short: GfK data). In practical terms, however, this proves to be difficult at present, since only two different time points of measurement are currently available at the household level: Socio-demographic data were gathered in January 1994 and in January 1995, attitudes on nutrition in October 1994 and October 1995. Data for purchase behavior are available for the whole

year of 1995. In order to work with the longitudinal models (2) or (3) one could, for instance, examine the change in purchase behavior, if any, between the first and the last months in 1995. Following this or a similar procedure would have its drawbacks, however. First, changes in nutritional attitudes during a period of only twelve months are likely to be quite small and can hardly be distinguished from measurement error. Secondly, with an interval as short as one year previous purchase behavior,  $Y_{t-1}$ , is very likely to be an almost perfect predictor of current purchase behavior, thus leaving little room for competing explanatory variables. Thirdly, both types of changes, that is changes in purchase behavior and changes in nutritional attitudes, would have some overlap in the time periods. Hence, in causal terms it would be impossible to clarify the intriguing question of whether changes in purchase behavior do in fact follow attitudinal changes or vice versa. As soon as more panel waves and longer time spans between the measurement points are available in the future, it will be fruitful to expand the empirical analyses to models (2) and (3). Given the data constellation presently at hand, it appears to be most sensible to start analyzing the demand of healthy food with an approach which is very similar to a cross-sectional approach such as model (1). It has been possible, though, to make use of the fact that the purchase data and the attitudinal and socio-economic data have been gathered at different time periods. So, in order to measure purchase behavior with respect to healthy food, the time period from *october 1995 to december 1995* was chosen. Therefore, purchase behavior during this period will be accounted for by socio-economic variables which were collected in *january 1995* and nutritional attitudes gathered in *october 1995*.

### **3. Measuring the Purchase of Healthy Food and its Determinants**

To define 'healthy' food and, conversely, food which is regarded as not being healthy is not an easy task. Even scientific evidence on the relationship between health and nutrition does not seem to be conclusive as the sometimes contradictory food recommendations in newspapers and TV magazines give an illustration of. One should not, however, overstate this point. Although nutritional suggestions in the media do not always point into the same direction, the overall picture emerging for the attentive consumer is fairly clear. Broadly speaking, from the perspective of healthy nutrition some quite simple rules on what one ought to eat and drink are available: It seems to be conducive to one's health if the food does not contain too much fat, if it does not include too much sugar, if nutrition is – more or less – wholesome, if not too much alcohol is consumed regularly etc. Despite these commonly accepted rules, it may still appear as somewhat naïve to classify a particular food product as healthy or unhealthy. To be sure, most food articles are not healthy

or unhealthy as such. Fairly often this is rather a matter of degree. And, still more important, it crucially depends on how much is consumed of a given food product: One glass of red wine, for example, may be conducive to one's health, two liters each day most certainly not. While it is often difficult to distinguish between healthy and non-healthy food with respect to a particular food product at a certain point in time, the situation is different for the *aggregate of all food articles* which are consumed *during a longer period*. Here it seems plausible to assume that consumers who are conscious with respect to their nutrition behavior will attempt to increase the share of healthy food in their total food plan.

There is a wide variety of food products in the GfK data sets (see Table A1 in the appendix). In order to classify the various food articles as 'healthy', the 'nutrition rules' mentioned above have been applied among others. First, it has been assumed that it is harmful to one's health if alcohol is consumed to a relevant degree. Therefore, drinks such as beer, wine, or whisky have been classified into the non-healthy category. Secondly, the food articles have been examined according to in how far *sugar* is an important ingredient. Hence, food articles such as cakes or cookies have also been coded as non-healthy. As a third criterion the *fat* content has been employed. When it comes to cheese, for instance, only those types of cheese with a fat content below 45 percent have been considered as being healthy where 45 percent is, approximately, the average value for soft and hard cheese. Fourthly, food articles have not been designated as healthy food when food additives or preservatives could have been part of the product. In addition to these negative selection criteria positive criteria are important as well. Food products which are commonly regarded as belonging to wholesome nutrition, for instance yoghurt, curd, or wholemeal rice, have been put to the healthy category.

To be sure, this classification can be no more than a rough approximation of those classification schemes which supposedly predominate in the general population. Other researchers might arrive at other decisions. On the whole, however, the food articles which have been classified as healthy should represent the picture in the German population reasonably well.

When we speak, henceforth, of food and healthy food an important qualification has to be made. The food articles in the GfK data sets do neither cover the whole range of food in general nor the whole range of healthy food in particular. For example, fresh meat, fish, and vegetables do not belong to the consumer data sets at present. When interpreting the empirical findings below, this constraint has to be taken into account. This said, it should hold nonetheless, however, that consumers with clear beliefs and intense preferences pro healthy nutrition should, *ceteris paribus*, purchase more of those food products which

have been regarded as furthering the health. In Table 1 the food articles which have been classified as healthy are documented.

*Table 1: Healthy food in the GfK consumer data*

Milk	unprocessed milk, full-cream milk, soy drinks, soured milk,
Poultry	e.g. legs, breast, wing, escalope, but not inwards oder giblets and
Teas	e.g. medicinal tea, green tea, herb fruit tea
Potato products	wholemeal dumpling, wholemeal potato fritter, wholemeal
Fruit juices	fruit juice, nectar juice, vegetable juice
Crispbread	e.g. soy crispbread, sunflower crispbread, wholemeal wheat
Non-alcoholic drinks	e.g. energy drinks
Cheese	(1) soft cheese such as camembert, melted cheese, blue mold
Rusk	e.g. wholemeal rusk, tea rusk, diet rusk
Yoghurt	e.g. natural yoghurt, fruit yoghurt, drink yoghurt, spice yoghurt
Curd, meals of curd	maximum fat content: 20 percent
Meal	wheat flour, wholemeal wheat, wholemeal, rye flour, rye
Mineral water	
Rice	e.g. Basmati, parboiled longseed, broken rice, natural rice,
Cereals	e.g. rolled oats, cereal bar, wheat bran, linseed

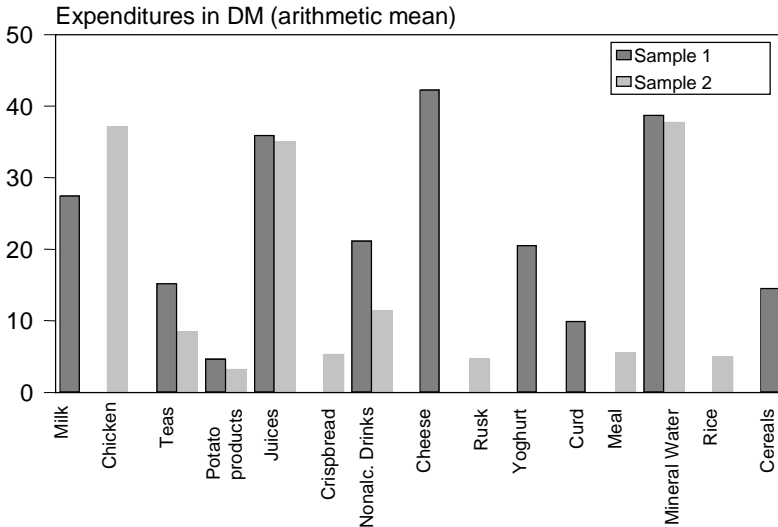
Figure 1 depicts how the expenditures for healthy food are distributed in both GfK samples.<sup>2</sup> As the food articles in both samples are not the same but overlap only partly, the average expenditures from october to december 1995 are shown for both samples. The dark grey bars represent the expenditures for food groups in sample 1, the lighter grey bars stand for food groups in sample 2.<sup>3</sup>

When we compare the distribution of healthy food articles between both GfK samples, it appears that sample 1 is somewhat better suited to represent food articles which are commonly regarded as healthy. Food articles such as milk, cheese, yoghurt, and curd are not part of sample 2 but of sample 1. Hence, all in all, food articles which in the public and the media are often declared as an essential part of wholesome nutrition and as being especially conducive to one's health are more likely to be found in sample 1.

<sup>2</sup> GfK sample 1 contains 4426 households, GfK sample 2 4638 households.

<sup>3</sup> Expenditures over a certain time period combine, of course, two aspects. First, they mirror the price which had to be payed for a food article. Secondly, they represent the amount which was purchased during the period. Whereas, for instance, the relatively high expenditure value for cheese is mainly due to the price level for cheese, expenditures for mineral water are largely a function of the amount.

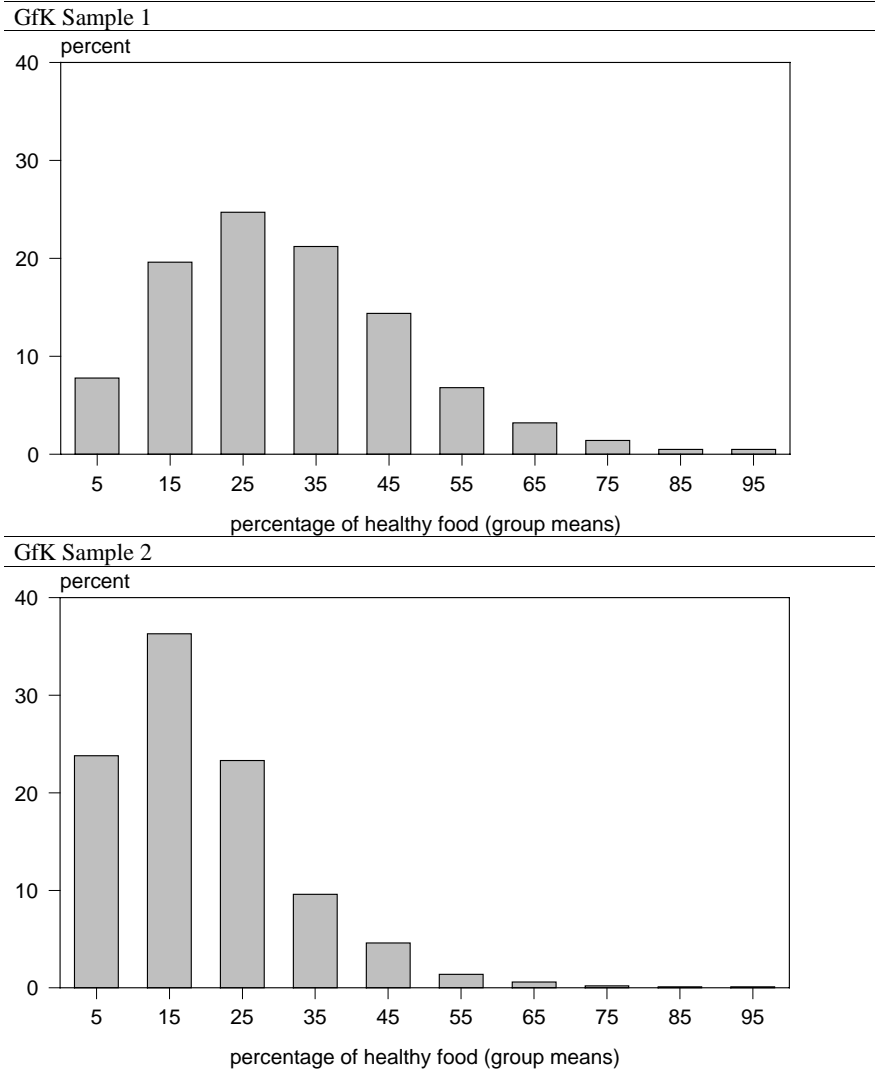
Figure 1: Household Expenditures for Healthy Food



Period: October to December 1995. Data Base: Consumer Data, GfK.

In principle, one could plausibly conceive the *absolute expenditures* for healthy food (in DM) as the explanandum of an empirical analysis. In this case, a household's economic restrictions – for instance its monthly income or household size – would exert a very strong effect on consumer expenditures. Conversely, nutritional attitudes would very likely have only a marginal influence on absolute expenditures. Moreover, according to the notion favored in this article, purchasing food implies the more or less conscious consideration of which kind of food to choose, be it healthy food or food regarded as not so healthy. From this point of view, it seems to be more convincing to look at how important the expenditures for healthy food are *in comparison* to the total food expenditures. Hence, the percentage of expenditures for healthy food with the total expenditures for food as the percentage base will be an adequate measure for nutritional behavior which is motivated by health consciousness. Figure 2 shows histograms for both GfK samples which indicate how the criterion variable of the following empirical analyses, *the percentage of expenditures for healthy food articles with respect to all food articles*, is distributed.

Figure 2: Relative Frequencies for the Percentage of Healthy Food Articles in the GfK Data (Expenditures Healthy Food Articles/Expenditures all Food Articles)



As can be seen in the figure, healthy food is in fact an important component in German households' food plans. Almost half of the households in sample 1, for instance, spend at least 30 percent of their total food expenditures on healthy food. Similar to Figure 1 it is evident here as well that both samples contain fairly different food articles. Whereas the mean percentage for healthy food amounts to 30.6 percent in GfK sample 1, the corresponding percentage equals only 18.9 percent in GfK sample 2.

Let us now come to the potential determinants of the relative expenditures for healthy food. In particular, which kind of *nutritional attitudes* can be expected to determine consumer choices? Certainly, the well-known *principle of correspondence in the level of specificity* does apply here (Fishbein/Ajzen 1975: 369). Therefore, consumer attitudes should affect consumer behavior the greater the more they correspond to the goods involved and the better they specifically refer to the goods in question.

Four attitude scales on nutrition were finally chosen. Although they are not optimally suited in each respect to the consumer data at hand they proved to be adequate overall – in terms of their content as well as in terms of their scale quality. Question wording, Cronbach's  $\alpha$  as a measure of scale quality, and the arithmetic means for each item are documented in Table 2.

The four scales represent different aspects of the respondents' beliefs and preferences on nutrition. By means of the first scale, the *preference for healthy food*, an overall high priority of health with respect to one's nutritional behavior is measured. If the food articles documented in Table 1 are in fact perceived as being healthy by the consumers a fairly strong relationship between attitudes and purchase behavior can be expected. Given the fact that fresh food is not an important part in the consumer data at hand, one should not expect huge effects for the *preference for natural food* as well as for the *preference for fresh food*<sup>4</sup> on purchase behavior. However, both scales measure an interested and positive attitude towards healthy nutrition as the positive correlation between each scale and the preference for healthy food shows (pearson correlation .481 between healthy food preference and natural food preference; .194 between healthy food and fresh food). At least weak effects on the purchase of healthy food should thus be possible nonetheless. The expectation is much more positive for the fourth attitude scale which refers to the *preference for wholesome food*. Here the question wording is almost optimally suited to the food products actually bought. As a wide variety of wholesome food articles is avail-

---

4 The arithmetic means for the second and third item of the 'fresh food scale' refer to their original values. For the further analyses the items have been recoded into the positive, pro-fresh direction. Computation of Cronbach's alpha is also based on the recoded items.



able in the GfK data sets, the correspondence between attitudes and behavior is quite high. Given this, one should legitimately expect a strong relationship between this particular attitude scale and the percentage of healthy food.

Table 2: *Attitudes on Nutrition*

	Item mean
<i>Preference for healthy food (Cronbach's <math>\alpha = 0.720</math>)</i>	
In my household I eat mild food.	3.1
I pay attention to what I eat and drink because I have to take care of my health.	3.5
I avoid all food that can damage my health.	3.5
<i>Preference for natural food (Cronbach's <math>\alpha = 0.773</math>)</i>	
I decline products that have preservatives.	3.3
When buying food, I look for food without additives.	3.5
<i>Preference for fresh food (Cronbach's <math>\alpha = 0.597</math>)</i>	
I always buy fresh food instead of, for example, preserved or frozen food.	3.3
Preserved food does taste as good as fresh food.	2.3
I can hardly imagine to cook without prepared foods.	3.0
<i>Preference for wholesome food (Cronbach's <math>\alpha = 0.682</math>)</i>	
We live by consuming wholesome foods.	2.5
We eat vegetarian.	1.5
Meals with grain assume a greater and greater part of our total food plan.	2.0
For each item answers from 1 = I strongly disagree to 5 = I strongly agree.	

Number of households: 8406 (GfK sample 1 + GfK sample 2)

The item means in Table 2 clearly demonstrate that health consciousness in Germany with respect to nutrition has reached a fairly high level. Each of the first three items which explicitly stress the health aspect of nutrition is approved by the respondents on the aver-

age. Furthermore, consumers prefer fresh food to preserved food and decline food additives. Of course, both types of preferences can in part also be matters of taste but they reflect as well health considerations. Mixed motives may also exist for the *preference for wholesome food*. To eat vegetarian, for instance, can be ethically inspired (cf. v. Ziehlberg/v. Alvensleben 1998). It also includes a health connotation, however (pearson correlation .387 between healthy food preference and wholesome food preference).

In addition to the nutritional attitudes I expect furthermore some substantive effects of the *socio-economic position* on purchasing healthy food. Choosing between healthy and unhealthy food requires information which is often difficult to obtain. Acquiring and processing the relevant information on nutrition will certainly be the more easier the better one's education and cognitive competence. Therefore, *education* is one of the socio-economic determinants to be examined. Moreover, *generational influences* could have an impact on nutrition practices. Public discussion about nutrition and health has intensified during the last two or three decades. Hence, younger generations should give higher priority to healthy food than older generations if we assume that individuals' attitudes are more moldable during their formative years. It thus appears reasonable to break down the purchase of healthy food according to age. In the cross-sectional design used in this article age automatically indicates the birth cohort as well.

Socialization is again the principal idea behind the distinction between *East- and West-Germany*. Consumers in East-Germany vs. West-Germany have grown up in radically different societies – politically as well as culturally. It can be surmized, then, that their preferences concerning healthy food will differ as well.

For *families with small children* it is expected that they have a greater percentage of healthy food since they are obliged each day to reflect about the best nutrition for their children. Finally, it will also be tested whether *household income* is related to the purchase of healthy food.

In operational terms the socio-demographic variables are defined straightforwardly. Education is measured by means of the main earner's formal school degree. Age is the age in years of the person leading the household. Income represents the overall net household income per month, and, lastly, number of children under 14 years simply reflects the factual numbers.

#### **4. Empirical Findings**

Let us at first take a look at the bivariate relationships between the dependent variable, that is the percentage of expenditures for healthy food, and its potential determinants.

Tables 3.1 and 3.2 show the results for sample 1 and sample 2, respectively. To begin with sample 1, strong and significant relationships between all four *attitude scales on nutrition*<sup>5</sup> and the expenditures for healthy food can be detected. As the Eta-value of .205 shows, the *preference for wholesome food* has by far the strongest relationship with the purchase of healthy food among all four scales.

Table 3.1: *The Purchase of Healthy Food and its Determinants - Bivariate Relationships*

GfK Sample 1

Nutritional Scales	Percentage of healthy food (arithmetic means)					Eta	p
	1	2	3	4	5		
Healthy food	28.90	28.11	29.37	32.47	33.94	.123	.000
Natural food	27.37	28.28	28.42	31.49	35.11	.157	.000
Fresh food	24.98	28.39	29.75	32.09	34.43	.125	.000
Wholesome food	27.98	30.11	34.22	38.96	43.22	.205	.000
Education	low	middle	high				
	29.35	30.59	32.28			.079	.000
Age (birth cohort)	19-24	25-34	35-44	45-59	60+		
	35.57	33.54	31.59	28.96	30.02	.100	.000
Children under 14	no	yes					
	29.98	32.91				.074	.000
Income (DM)	-1000	1000-1999	2000-2999	3000-3999	4000+		
	31.99	30.23	29.38	31.05	31.84	.059	.005
East- vs. West-Germany	east	west					
	26.70	31.65				.120	.000

<sup>5</sup> The attitude scales on nutrition were built as the arithmetic means of the items which belong to the respective concept. For the bivariate analyses (Tables 3.1 and 3.2) digits after the decimal point have been rounded for the attitude scales in order to get whole digit values.

Table 3.2: *The Purchase of Healthy Food and its Determinants - Bivariate Relationships*

GfK Sample 2

	<i>Percentage of healthy food (arithmetic means)</i>					Eta	p
	1	2	3	4	5		
Nutritional Scales							
Healthy food	15.48	17.99	19.02	19.17	19.97	.064	.001
Natural food	17.01	17.09	18.43	18.77	21.25	.104	.000
Fresh food	18.02	17.06	17.63	20.42	22.13	.141	.000
Wholesome food	17.89	18.54	20.95	21.04	23.21	.100	.000
Education	low	middle	high				
	18.61	18.85	19.36			.026	.207
Age (birth cohort)	19-24	25-34	35-44	45-59	60+		
	20.22	20.18	19.67	18.74	18.09	.062	.001
Children under 14	no	yes					
	18.37	20.80				.081	.000
Income (DM)	-1000	1000- 1999	2000- 2999	3000- 3999	4000+		
	17.28	17.62	18.92	19.47	19.44	.056	.006
East- vs. West- Germany	east	west					
	16.27	19.61				.105	.000

Therefore, beliefs and preferences on nutrition appear to play an important role when it comes to explaining the purchase of healthy food.

Regarding the socio-economic variables, *education* and *age* are both related to the percentage of healthy food as has been hypothesized. With higher education the tendency to choose healthy food rises. Conversely, higher age groups, that is birth cohorts which did not experience the intense public debates on nutrition and health during their formative years, purchase somewhat less healthy food than younger age groups. As was expected, *households with children under 14* buy more healthy food than households without or with older children. As far as the monthly *income* is concerned, group differences are statistically significant. However, the Eta-value is small (.059) and the arithmetic means for the percentage of healthy food do not increase – or decrease – monotonically with

income. Therefore, an interpretation for the relationship between income and healthy food seems to be difficult.

There is, finally, a clear and marked divide between the former communist *East-Germany* and the *West-German Länder*. Households in East-Germany significantly purchase much less healthy food products than their western counterparts. The difference adds up to five percentage points which can be called a fairly large effect. One might speculate that the Eastern part of Germany is still – five years after the German unification at the time when the GfK surveys were conducted – more traditional in a cultural sense. Another interpretation which does not contradict the first one refers to the lower consumption experience of the East-German population. If we take together the effect of age (or birth cohort) and the difference between East- and West-Germany we get some indications that socialization could in fact be an important mechanism in explaining nutritional preferences.

For sample 2 (Table 3.2) the overall pattern of the relationships is fairly similar to sample 1. The associations are lower, however. Probably the main reason can be found in the different composition of both GfK samples regarding the food products.<sup>6</sup> As already mentioned, GfK sample 1 is more representative of food commonly regarded as healthy in general and wholesome in particular.

In the last step of the analysis a multivariate model for the percentage of healthy food is estimated for GfK sample 1.<sup>7</sup> Since the dependent variable's measurement level is clearly metric, a linear regression approach is the suitable method. In addition to the determinants in Tables 3.1 and 3.2, two further potential explanatory variables have been included, household size and the product price paid for healthy food.<sup>8</sup> Together with household

**6** As the items which measure the preference for wholesome food include an indicator for eating vegetarian, it is, of course, not optimal to construct the percentage of healthy food for GfK sample 2 with poultry as an element (see Figure 1). However, the relationship is only marginally influenced by the expenditures for poultry. If it is excluded from the calculation of the percentage, Eta increases only slightly from .100 to .119.

**7** Codings for the regression analysis are as follows: *education* (1: low (Hauptschule); 2: middle (Realschule); 3: high (Abitur+)); *age* (1: - 19 years; 2: 20 - 24 years; 3: 25 -29 years; 4: 30 - 34 years; 5: 35 - 39 years; 6: 40 - 44 years; 7: 45 - 49 years; 8: 50 - 54 years; 9: 55 -59 years; 10: 60 - 64 years; 11: 65 - 69 years; 12: 70 years +); *children under 14* (0: no child under 14 in the household; 1: 1 child or more under 14); *income per month 1994* (DM 250, 750, 1125, 1375, 1750, 2250, 2750, 3250, 3750, 4250, 4750, 5250, 5750); *household size 1994* (number of persons, 1 to 10); *East-West-Germany* (0: East-Germany; 1: West-Germany); *payed price for healthy food* (in DM); *attitude scales on nutrition* (1: strongly disagree to 5: strongly agree).

**8** There is, of course, no average price for those households who did not buy anything in the healthy food category. For these households the overall average price for healthy goods has been

income, both variables did not prove to be significant, though. Hence, the three variables have been dropped from the regression equation and the model has been reestimated. In order to assess how important the nutritional attitudes are if taken together, the change in explained variance has also been computed for these variables.

Table 4: *A Multivariate Model for the Purchase of Healthy Food; (OLS-Regression; GfK sample 1)*

	Unstandardized regression coefficients	Standardized regression coefficients
<i>Socio-economic Variables</i>		
Education	1.538 (5.48)	.084
Age (birth cohort)	-.586 (-5.39)	-.105
Children under 14	1.552 (2.22)	.040
Income per month	n.s.	
Household Size	n.s.	
East-West-Germany	5.043 (8.03)	.122
Payed Price for healthy food	n.s. <sup>a</sup>	
<i>Nutritional Attitudes</i>		
Healthy food	1.414 (4.24)	.081
Natural food	.821 (2.86)	.052
Fresh food	1.068 (3.74)	.060
Wholesome food	2.755 (8.84)	.151
Intercept	11.410 (7.13)	
$\Delta R^2$ when including nutritional attitudes	.059	
$R^2$	.088	
Number of Households	4100	

T-Values in parentheses. a: not significant.

In comparison to the bivariate analyses the pattern is very similar. The preference for wholesome food is still the most important predictor if the standardized coefficient is used as the measure of effect strength. Preference for healthy food, natural food, and fresh food also have significant impacts although their effects are clearly weaker in the multivariate

substituted. As each household can be regarded as a price taker, this assumption seems to be fairly reasonable.

case. In comparison to the whole variance explained, the increase in explained variance makes up more than half when the four attitude scales are added to the socio-economic predictor variables. This strong effect of nutritional attitudes is another indication for the notion that purchasing food is neither a spontaneous act nor a simple execution of habits unconnected to beliefs and preferences.

Again we find a fairly strong effect for the distinction between households in East- and West-Germany. Even after controlling for nutritional attitudes and the other socio-economic determinants, the difference in purchasing healthy food between East- and West-Germany is relevant and highly significant.

## 5. Conclusion

In the preceding analysis a partial model for the purchase of healthy food has been developed and tested. It is a partial model since some important explanatory variables, such as purchase habits, have not yet been included. Therefore, the main focus in evaluating the findings presented should not be on overall variance explained but rather on the pattern of effects. Strong effects of nutritional attitudes on purchase behavior support the conception stated at the outset that buying food is a type of behavior which is characterized by an *intermediate degree of involvement* (Kroeber-Riel 1992: 169). Viewed from another perspective, these findings demonstrate that beliefs and preferences regarding healthy nutrition do not reflect mere wishful thinking or social desirability on the respondents' side without any behavioral consequences. Consumers do not "preach water – to others – and drink wine for themselves" as the famous proverb would suggest. They rather act according to their overt beliefs and preferences.

Which kind of food is chosen appears to depend as well on factors such as socialization and cognitive capabilities. The negative effect of age and the strong impact of the East-West-German dichotomy are both compatible with a socialization hypothesis whereas the estimated influence of education may indicate that purchasing food requires the ability to acquire and process often complex information on nutrition.

## Contact

Dieter Ohr  
Universität zu Köln  
Institut für Angewandte Sozialforschung  
Greinstraße 2  
50939 Köln  
Dieter.Ohr@uni-koeln.de

## References

- Baudrillard, Jean, 1970, *La société de consommation. Ses mythes, ses structures*. Paris: Denoël.
- Felser, Georg, 1997, *Werbe- und Konsumentenpsychologie*. Eine Einführung. Stuttgart: Schäffer-Poeschel.
- Finkel, Steven E., 1995, *Causal Analysis with Panel Data*. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-105. Thousand Oaks, CA: Sage.
- Fishbein, Martin, Icek Ajzen, 1975, *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Reading, Mass.: Addison-Wesley.
- Fricke, Andreas, 1996, Eine Kaufverhaltensanalyse bei Öko-Produkten unter besonderer Berücksichtigung des Kohortenkonzepts. In: *Jahrbuch der Absatz- und Verbrauchsforschung* 4/96, 372-400.
- Kessler, Ronald C., David F. Greenberg, 1981, *Linear Panel Analysis. Models of Quantitative Change*. New York: Academic Press.
- Kroeber-Riel, Werner, 1992, *Konsumentenverhalten*. München: Franz Vahlen. 5th ed.
- Kutsch, Thomas, 2000, Konturen einer Ernährungssoziologie. In: Doris Rosenkranz, Norbert F. Schneider (ed.), *Konsum. Soziologische, ökonomische und psychologische Perspektiven*. Opladen: Leske+Budrich, 149-168.
- Papastefanou, Georgios, 2000, Struktur und Wandel des Güterkonsums privater Haushalte 1993 und 1993. In: Doris Rosenkranz, Norbert F. Schneider (ed.), *Konsum. Soziologische, ökonomische und psychologische Perspektiven*. Opladen: Leske+Budrich, 265- 282.
- Trommsdorff, Volker, 1993, *Konsumentenverhalten*. Stuttgart: W. Kohlhammer. 2nd ed.
- Ziehlberg, Roland v., Reimar v. Alvensleben, 1998, Die Bedeutung ethischer Motive beim Kauf von Lebensmitteln am Beispiel fair gehandelten Kaffees. In: *Jahrbuch der Absatz- und Verbrauchsforschung* 2/98, 201-216.



---

## Appendix

*Table A1: Food Articles in the GfK consumer data set*

---

Ketchup/spicy sauce/ noodle sauce (1)
Mayonnaise/ remoulade/ salad dressing (2)
Canned vegetables/ tomato paste (7)
Milk (8)
Roasted coffee beans (12)
Instant coffee (13)
Salts (14)
Poultry (16)
Frozen food (excluding untreated poultry) (17)
Teas (18)
Cocoa (19)
Alcohols (20)
Fats (22)
Mustard/ horse-radish (23)
Ready-made cakes (24)
Champagne (26)
Apple-cider (27)
Potato products (30)
Ready-made pudding/ ready-made desserts (31)
Beer (33)
Vermouth/ aperitif (34)
Wine/ mulled wine (35)
Fruit juices (36)
Canned sour food/ sauerkraut/ red cabbage (37)
Crispbread (38)
Coffee products (42)
Fine/ready-made/buffet style salads (43)

---

---

Sherry/ port (44)  
Pudding - / dessert powder (45)  
Non-alcoholic drinks with carbonic acid (e.g. energy drinks) (46)  
Soft cheese (e.g. mozzarella) (47)  
Cream (50)  
Canned milk, coffee cream, coffee whitener (51)  
Cakes and cookies for autumn and winter (54)  
Vinegar (57)  
Cookies (64)  
Rusk (68)  
Dried ready-made meals (i.e. pizza, spaghetti) (71)  
Curd and meals of curd (73)  
Savory biscuits (74)  
Instant ready-made meals (76)  
Canned ready-made meals (77)  
Yoghurt (78)  
Meal (79)  
Hard cheese (e.g. Leerdamer, Gouda etc.) (81)  
Baking-powder/ vanilla/ leaven (83)  
Mineral water (84)  
Icecream (86)  
Baking mix (for cakes) (87)  
Rice (89)  
Cereals (90)  
Pastas (91)  
Complete ready-made meals in package (93)  
Cocktail products (e.g. peanuts, chips) (94)

---

The numbers in parentheses refer to the product group in the GfK data sets.

# PRICE-SETTING AND PRICE-ADJUSTMENT BEHAVIOR FOR FAST-MOVING CONSUMER GOODS

MATTHIAS FENGLER AND JOACHIM WINTER\*

## 1. Introduction

When analyzing aggregate dynamics, economists traditionally assume that prices adjust instantaneously in response to changes of economic conditions. In that case, relative prices, i.e., the ratios of prices of different goods, act as a signal of the relative scarcity of goods. Flexible prices are a crucial condition for an efficient allocation of resources in an economy. In reality, however, the simplifying assumption that prices are flexible need not hold: wages are stipulated in long-term contracts, and prices of many consumption goods and services change only infrequently.<sup>1</sup> If prices do not adjust fully and instantaneously in response to changes of economic conditions, one speaks of *price rigidities* (see Carlton, 1986). Based on casual empiricism, one might expect that many prices are rigid.<sup>2</sup>

In this paper, we present new empirical evidence that helps to understand price-setting behavior in retail markets and its aggregate implications. Using data from the GfK Consumer Panel 1995 (a large-scale household survey conducted by the *Gesellschaft für Konsumforschung*, Nürnberg), we investigate the dynamics and dispersion of prices for fast-moving consumer goods. We focus on one frequently purchased consumption good that exhibited substantial price dynamics during 1995, ground coffee. Our empirical

---

\* We are grateful to Axel Börsch-Supan, Annette Giering, Annette Köhler and Nicole Koschate for helpful discussions. The GfK Consumer Panel 1995 was provided by the Zentrum für Umfragen, Methoden und Analysen (ZUMA), Mannheim.

1 We treat wages as prices and services as goods.

2 One might ask whether by this strict definition there are *any* flexible prices at all. Under certain circumstances one can justify that in financial markets prices of stocks are fully flexible, i.e., that they adjust instantaneously in response to changes in economic conditions (e.g., to new information on expected profits of a company). However, it is beyond the scope of this paper to discuss this question.

results show that psychological pricing points affect both static price setting and adjustment to cost shocks significantly. We interpret our findings as evidence for substantial rigidity of prices in German retail.

The remainder of this paper is structured as follows. In the next section, we present some basic concepts used in the analysis of price setting and price adjustment behavior, and we review the central questions of current empirical research in this area. In section 3, we discuss whether price data obtained from the GfK Consumer Panel are suitable to answer these empirical questions. We then focus on one explanation for price rigidities, psychological pricing points in section 4. We sketch other empirical issues that can be explored with data from the GfK Consumer Panel in section 5. Section 6 provides a summary of our results.

## **2. Theoretical and empirical analysis of price setting and price adjustment behavior**

For economists, the phenomenon of rigid prices is associated with several important questions. First, it is necessary to document whether and to what extent observable prices are rigid. Once the existence of price rigidities is established, the reasons why prices do not adjust fully and instantaneously need to be explored, and finally, the aggregate effects of rigid prices need to be examined. These issues have been, and they continue to be, areas of active research in macroeconomics over the last decades (see Blanchard, 1990).

Rigid prices play a major role in Keynesian models of aggregate economics. Whereas in frictionless economies, it is difficult to justify active policy measures, frictions (such as rigid prices) open the possibility for policy makers to influence the economy's aggregate performance positively. In the earlier Keynesian literature, most theoretical models of aggregate economics invoked rigid prices by assumption; this was often motivated by long-term employment contracts or the inertia of prices of goods and services. However, just *assuming* rigid prices seems to be unsatisfactory from a methodological point of view and, even worse, turned out to be inconsistent with economic theory. It is obvious that firms could increase profits if they adjusted prices to their optimal level more often. Without any further arguments, rigid prices cannot be derived by optimizing behavior of economic agents. This was one of the main criticisms of traditional Keynesian models, and the existence and importance of price rigidities became one of the battlefields of the Keynesians and the neoclassic school.

Having realized the importance of price rigidities, especially in the aggregate economic analysis from the Keynesian perspective, it is not surprising that in the 80s, economists

developed a large number of models that try to derive price rigidities consistently from a microeconomic optimization framework. This change of focus also revealed a drawback of earlier empirical research in Keynesian economics, the lack of disaggregated micro data (see Danziger, 1987). While on the theoretical side, this research program has come to an end, there is still only little convincing empirical evidence on price rigidities. Only recently, some advances using disaggregated price data have been made.<sup>3</sup>

At this point, another drawback of traditional macroeconomic analysis (not only of the Keynesian approach) emerged – the missing link between the microeconomic activities of individual agents (firms and households) and the dynamics of those economic aggregates macroeconomists are actually interested in, i.e., the *aggregation problem*. With respect to price rigidities, for example, it is not clear whether and how the inflexibility of prices at the micro level influences the dynamics of the aggregate price level, i.e., inflation (e.g., Caplin and Spulber, 1987). These issues have not yet been fully resolved, neither theoretically nor empirically, and they are an active area of current research in macroeconomics (see Bryan and Cecchetti, 1999, for a recent review of this literature).

In this paper, we discuss how the price data obtained from the GfK Consumer Panel can be used to approach some of the empirical questions raised above. We identified four issues which could be addressed, although by the nature of the data, the analysis is restricted to price setting and price adjustment for fast-moving consumption goods:

- the empirical relevance of psychological pricing points,
- the empirical relevance of (fixed) costs of price adjustment,
- the aggregate impact of micro-level price rigidities,
- the interaction of micro-level price dynamics, price dispersion, the aggregate price level and inflation.

In the remainder of this paper, we focus on the first issue, that is the empirical relevance of psychological pricing points for individual price setting, and we discuss whether they can explain price rigidities. The other topics are sketched only briefly; a more thorough discussion of the theoretical literature and additional empirical results can be found in Fengler (2000).

---

<sup>3</sup> Köhler and Winter (1993) provide an extensive summary of this literature. Major contributions include Rotemberg (1982), Cecchetti (1986), Kashyap (1995), Blinder et al. (1996), Slade (1998, 1999) and, for Germany, Köhler (1996).

### 3. Obtaining micro-level price data from the GfK Consumer Panel

In this paper, we use price data based on individual transactions that were obtained from the 1995 wave of the GfK Consumer Panel. This data set was designed for household demand analysis from a marketing perspective. For many issues in applied household analysis, however, researchers face the problem that the socio-demographic and socio-economic characteristics of the panel households (e.g., income and employment status) are recorded only once a year. Although the consumption data are sampled with high frequency, the corresponding household information exhibits little variation over time. Within-household event studies are hardly possible because events such as changes of household composition or income cannot be related to changes to consumption patterns observed in the transactions data. However, long-term issues of demand behavior, such as the role of households' attitudes, can be analyzed quite well as long as between-household comparisons are sufficient.

In this paper, we focus exclusively on the price data available in the GfK Consumer Panel and ignore all other information such as household characteristics. For every transaction (i.e., the purchase of an individual product), the data set provides very detailed information such as product classification, brand, size, type of retailer, and last but not least, its purchase value. We can therefore extract daily price data for a vast number of fast-moving consumption goods covering the year 1995. From this perspective, the GfK Consumer Panel constitutes a unique source of high-frequency, micro-level price data.

For the analysis of price adjustment, there is one important drawback which stems from the fact that the data are recorded on a transaction basis from the household perspective. Transaction prices can only be traced to four different types of retailers, but not to a *specific* retailer (such as a specific grocery store), and we cannot construct consistent time series of prices quoted by unique retailer. Specifically, price changes are not directly observable. Therefore, our analysis is restricted to the dynamics of the whole *distribution* of prices over time.<sup>4</sup> Since these distributions of prices are available daily on the level of individual products (the lowest level of aggregation possible), we can still address some of the empirical questions posed in section 2 from the perspective of German retail markets, such as the relevance of psychological pricing points. We return to the other issue below in section 5.

---

4 The only way to construct a time series of individual prices is to identify those households which purchase a given product with high frequency at the same type of retailer and then to impose the assumption that this is indeed the same retailer all the time; see Fengler (2000).

#### 4. Empirical results on psychological pricing points

Prices that account for the subjective price perception of consumers are very popular among retailers of consumption goods (see Monroe, 1983, and Wiswede, 1995). To explain the wide-spread use of such prices, it is usually argued that consumers react only little in response to price changes within a certain price range, but react strongly when the limits of these price ranges are violated in either direction. These limits, the so-called pricing points, are typically associated with even prices (such as 1 DM, 5 DM, 100 DM etc.). Actual prices are set just below these limits. Therefore, especially with fast-moving consumption goods, we observe odd prices (e.g., prices of 49 Pf. or 99 Pf., 4.99 DM or 89 DM) very frequently, but rarely even prices. In the remainder of this paper, we refer to these prices as *psychological pricing points*; a related term used in the literature is *focal point pricing*. For Germany and other countries, the relevance of focal prices has been documented extensively, and we report additional evidence below.<sup>5</sup>

Whereas the existence of pricing points is generally accepted, the traditional motivation for them is subject to controversial debate. In field experiments, Diller and Brielmaier, 1996, do not find any evidence that switching from odd to even prices results in significant demand reactions of consumers (which would be an implication of the traditional model). Therefore, the psychological motivation for focal prices is questionable, and Diller and Brielmaier conclude “that psychological pricing points are presumably rooted more deeply in the brains of researchers and managers than in those of consumers” (1996: 709, our translation). In her survey study, Köhler (1996) finds that psychological pricing points are not relevant for price-setting in the manufacturing sector. Similar results have been obtained by Blinder et al. (1998) who conducted a survey among U.S. companies. It is therefore still an open question why psychological pricing points are so important, and more theoretical research is needed (see Wedel and Leeflang, 1998, for a recent theoretical contribution).

Since psychological pricing points can be documented empirically, economists should be interested in their aggregate impact. When those economic variables which determine prices change continuously, actual prices will be adjusted either too early or too late in most cases and therefore deviate from optimal prices most of the time which implies price rigidities. Therefore, in addition to confirming that retailers prefer psychological pricing

---

5 We do not attempt to summarize this vast literature. Recent evidence for Germany is provided by Dahlhäuser (1996), Diller and Brielmaier (1996), Müller-Hagedorn and Zielke (1998) and Gedenk and Sattler (1998). For a recent study based on American price data see Huston and Kamdur (1996).

points when *setting* prices, we also need to investigate price *adjustment*, that is how prices are adapted after changes in the economic environment.

The GfK Consumer Panel does not only provide the opportunity to document the empirical relevance of psychological pricing points for a large number of fast-moving consumption goods, but because of its time dimension, we can also investigate the effect of psychological pricing points on price adjustment. We now present preliminary empirical evidence on these issues; for further details, see Fengler and Winter (2000b). We concentrate on three products (ground coffee, milk, and butter); see tables 1 and 2a to 2c.

Table 1: *Distribution of prices for three fast-moving consumption goods*

	Number of observations	Mean	Standard deviation	Min	Max
Ground coffee, vacuum packed, 500g	14247	8,21	1,00	5,98	14,98
Milk, 1l	4926	1,02	0,09	0,49	1,98
Butter, 250g	33732	1,74	0,16	0,89	3,69

Source: GfK Consumer Panel 1995 (coffee: commodity group 12, article ID 24199; milk: commodity group 08, article ID 15109; butter: commodity group 22, article ID 43990); own calculations.

With respect to coffee, 86 percent of the 14,247 purchases registered in the GfK Consumer Panel 1995 may be classified as being subject to psychological focal pricing. Besides the “classic” pricing points ending in 49 Pf., 98Pf. or 99 Pf., we also treat prices such as 6.66 DM, 7.77 DM and 8.88 DM as “psychological”. For example, 7.77 DM is observed in approximately 10 percent of all purchases.<sup>6</sup> Moreover, the distribution exhibits a large dispersion (spread and variance). The smallest price observed is 5.98 DM, the most expensive is 14.98 DM. This is partially due to extensive quality and product discrimination in the coffee market, but also to the dynamics of coffee prices in 1995; we return to this issue below.

<sup>6</sup> The fractions of prices below 6.00 DM and over 10.00 DM are not reported since they are negligible.



Table 2a: Psychological pricing points for ground coffee

Price	Number of obs.	Relative frequency	Rank
Ground coffee, vacuum packed, 500g			
6,49	59	0,00	
6,66	158	0,01	
6,98	187	0,01	
6,99	1410	0,10	4
7,49	911	0,06	6
7,77	1429	0,10	3
7,98	295	0,02	8
7,99	3834	0,27	1
8,49	681	0,05	7
8,88	51	0,00	
8,98	191	0,01	
8,99	1827	0,13	2
9,49	153	0,01	
9,98	107	0,01	
9,99	911	0,06	5
<i>Other</i>	2043	0,14	
Sum	14247	1,00	

Source: GfK Consumer Panel 1995 (coffee: commodity group 12, article ID 24199); own calculations.

Table 2b: Psychological pricing points for milk

Price	Number of obs.	Relative frequency	Rank
Milk, 1l			
0,89	125	0,03	4
0,99	3504	0,71	1
1,09	844	0,17	2
1,19	232	0,05	3
Other	221	0,04	
Sum	4926	1,00	

Source: GfK Consumer Panel 1995 (milk: commodity group 08, article ID 15109); own calculations.

Table 2c: Psychological pricing points for ground butter

Price	Number of obs.	Relative frequency	Rank
<b>Butter, 250g</b>			
1,49	1226	0,04	6
1,59	2241	0,07	3
1,69	15195	0,45	1
1,79	7139	0,21	2
1,89	2036	0,06	5
1,99	2043	0,06	4
Other	3852	0,11	
Sum	33732	1,00	

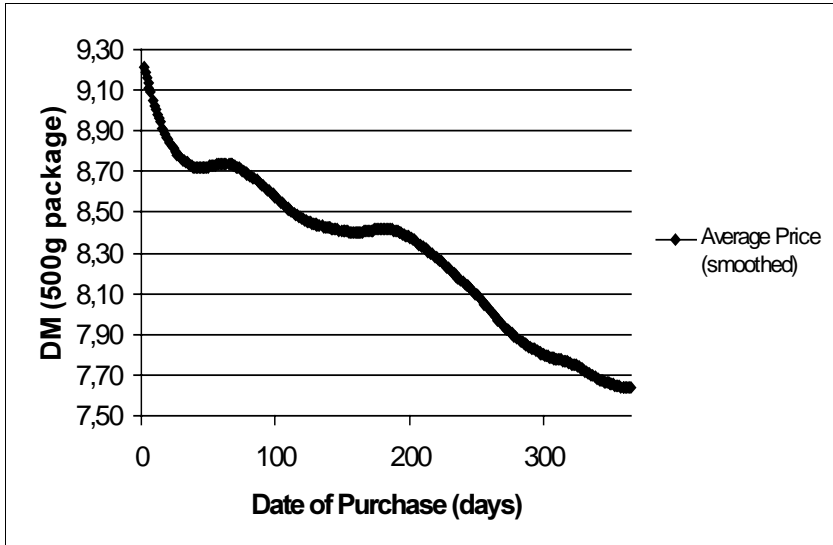
Source: GfK Consumer Panel 1995 (butter: commodity group 22, article ID 43990); own calculations.

Prices of milk exhibit less spread, but psychological pricing points dominate even more. Approximately 96 percent of all purchases show only four different prices, and 71 percent belong to only one price (99 Pf.). In contrast to coffee, the average price of milk remained constant over the entire year. With regard to butter, there were also only little changes in the average price over the year, but price dispersion is higher which might be due to quality and price differentiation. Again, we confirm the dominance of psychological pricing points (about 89 percent of observed purchases).

These results establish, once again, the importance of psychological pricing points and focal pricing in the German retail market (in this case for three grocery products of daily use). However, they only refer to the distribution of prices over the entire year. This static approach might mask actual price dynamics. In our dynamic analysis, we concentrate on one of the products presented above, precisely: on ground coffee. We choose this product because the average retail price exhibited an almost dramatic decline during 1995 (see figure 1), and because production costs are rather well known – the price of green coffee is determined on the world market, and it is essentially the same for all German coffee roasters.<sup>7</sup>

<sup>7</sup> Almost all ground coffee sold in Germany is also roasted in Germany; the proportion of imported coffee is negligible (see Feuerstein, 1999).

Figure 1: Retail Coffee Prices



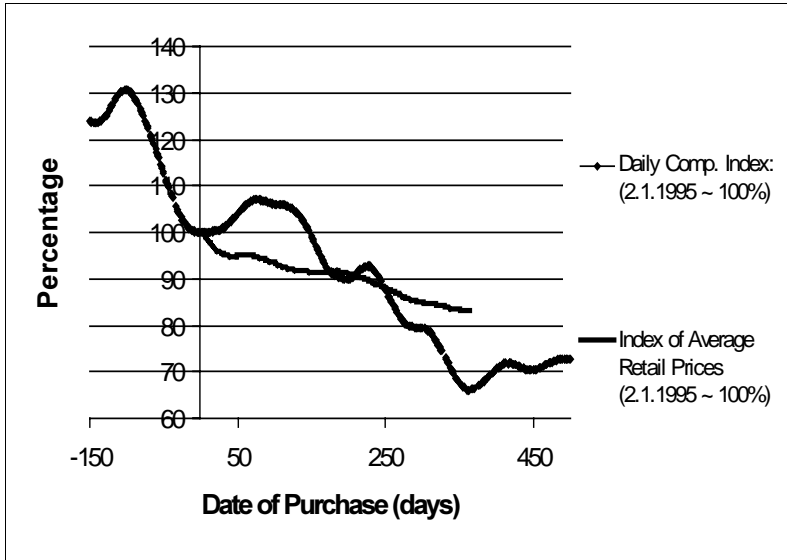
Source: GfK Consumer Panel 1995 (commodity group 12, product ID 24199); own calculations.

The strong decline in retail prices can be related to the decline of green coffee prices that occurred over the course of 1995. In figure 2, we present scaled indices of the average retail price and green of world-market.<sup>8</sup> Note that the price of green coffee is displayed beginning in 1994, almost 150 days before the first retail price observation in our data, and well into 1996. The initial high prices are due to a shortage of green coffee following a frost in Brazil in 1994. This shortage was overcome continuously in 1995. From figure 1, we conclude that retail prices of roasted coffee track green coffee prices and that they are determined by the supply side. Feuerstein (1999) comes to the same conclusion in an econometric analysis of the German coffee market.<sup>9</sup> For our analysis, it is important that the variation of retail coffee prices in the long run is determined exogenously because in this case we can make valid inferences on price adjustment behavior using observed retail prices.

<sup>8</sup> The world market price of green coffee is available on a daily basis; we should like to thank P. Dubois, of the International Coffee Organization (ICO), London, for providing these data.

<sup>9</sup> The study of Feuerstein (1999) provides a detailed analysis of the German coffee market, especially of the role of green coffee prices for price setting in a tight oligopoly.

Figure 2: Green Coffee Prices and Retail Coffee Prices

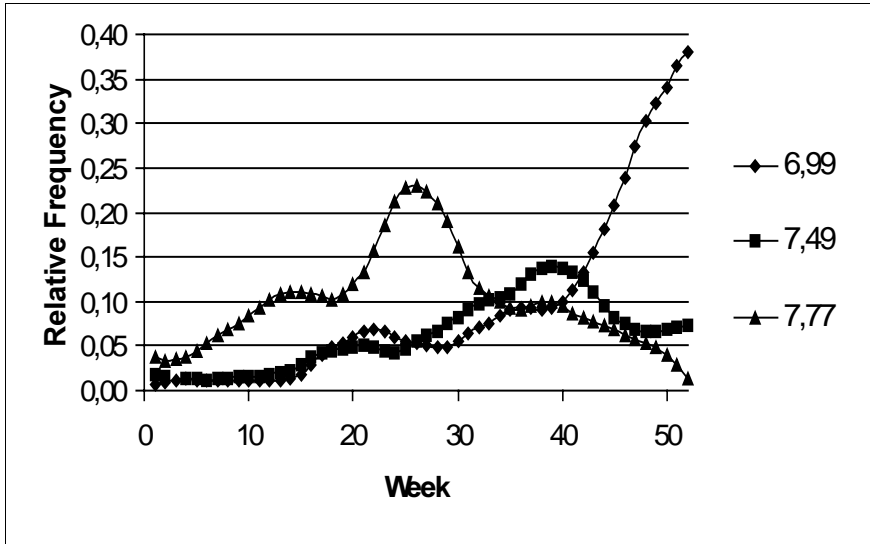


Source: International Coffee Organization, London; GfK Consumer Panel 1995 (commodity group 12, product ID 24199); own calculations.

For the analysis of price adjustment behavior, it is interesting to see how the long run drop of average retail prices is reflected in changes of the *distribution* of retail prices. To this end, we collapse our daily retail price data into an empirical distribution of weekly prices<sup>10</sup> and determine their relative frequency. Using a kernel density procedure, we smooth the relative frequency of the most important (psychological) prices. For convenience of presentation, we divide the coffee market into three segments (lower, middle, and upper segment) that are presented separately (fig. 3, 4, and 5).

<sup>10</sup> Fengler (2000: 61-62) discusses the assumptions under which this procedure is valid.

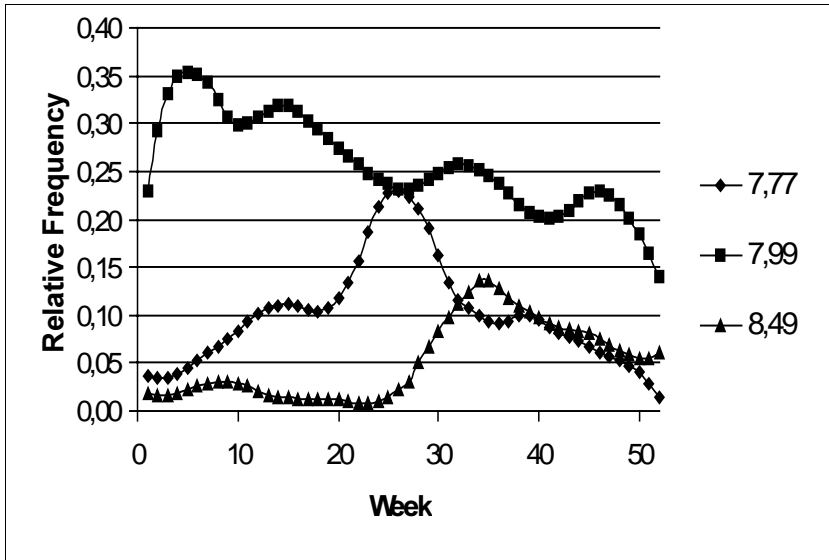
Figure 3: Fractions of Psychological Prices During the Year: Lower Price Segment



Source: GfK Consumer Panel 1995 (commodity group 12, product ID 24199); own calculations.

By inspecting the dynamics of the relative frequencies, it is evident that the slow and smooth decline of the average retail prices masks the sometimes rather abrupt movements of individual prices. These movements do not occur in a balanced fashion over the year, but are limited to a small interval of a couple of weeks during which the relative frequencies of prices change dramatically. Consider first figure 3 for the lower price segment. At the beginning of the year, the smallest prices, 6.99 DM and 7.49 DM, are negligible. With time passing, their relative frequency is rising slowly, until – beginning in the 40<sup>th</sup> week – the price of 6.99 DM, with a fraction of 35% of all quotations, clearly dominates the price segment (and also the whole market; see fig. 4 and 5). This fast upward move is accompanied by declining frequencies of the prices 7.49 DM and 7.77 DM which are the nearest focal prices. Whereas one might expect that during a price decline, the lowest price becomes more significant, the dynamics of the other prices are ambiguous. The rise in significance of the price of 7.77 DM in the first months of the year (see fig 4) is probably a result of the shrinking importance of the 7.99 DM price.

Figure 4: Fractions of Psychological Prices During the Year: Middle Price Segment

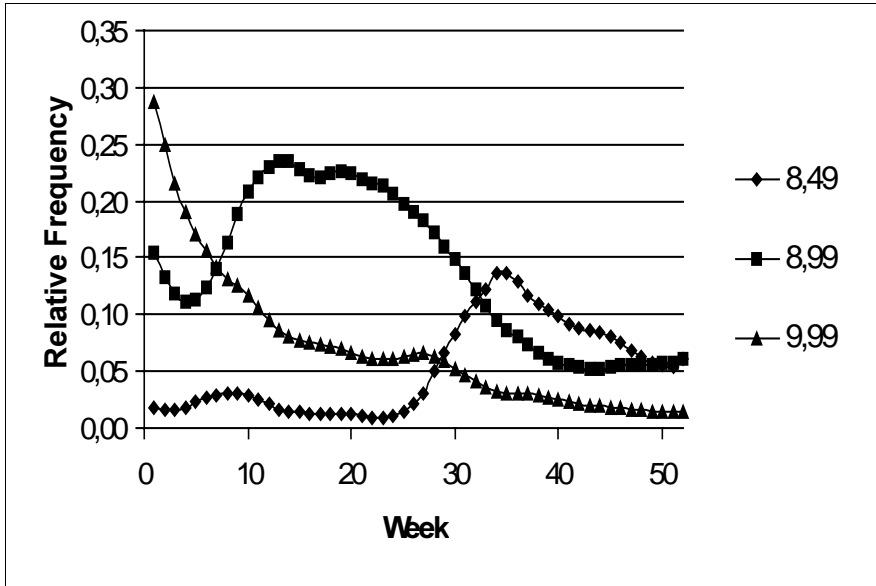


Source: GfK Consumer Panel 1995 (commodity group 12, product ID 24199); own calculations.

Finally, we inspect the upper segment (figure 5). Here, the frequency of the 9.99 DM prices is dropping rather fast, while at the same time the next lower focal price 8.99 DM establishes itself as the dominant price in the upper segment. The declining importance of both prices beginning in week 27 results again in a sharp upward movement in the frequency of 8.49 DM prices which until then had been insignificant.

These empirical observations allow two conclusions: First, psychological prices are very important in German retail markets for fast-moving consumer goods; this replicates the findings of earlier studies. Second, our analysis reveals how complex the dynamics of individual prices is, even if the adjustment of *average* prices to a cost shock looks smooth. Our conclusion from these empirical findings is that retail prices are rigid to an extent which is relevant from an aggregate point of view. We should stress, once again, that assessing the aggregate consequences of price rigidities more adequately would require a structural model of aggregation.

Figure 5: Fractions of Psychological Prices During the Year: Upper Price Segment



Source: GfK Consumer Panel 1995 (commodity group 12, product ID 24199); own calculations.

Finally, let us point out that by the nature of the data, the prices observed refer to actual purchases, meaning that the actual distribution of prices offered by retailers might deviate from the distribution observed in the data set. This will especially be the case when consumers prefer products with psychological pricing points. It would be interesting to investigate this selection problem more formally, but we lack the appropriate data. In any case, our results provide evidence for the importance of psychological prices from a consumer's point of view: The large proportion of these "psychological" prices among all possible prices reveals that consumers prefer these prices even if there were products with other prices available. Moreover, above-mentioned the selection problem might be negligible in practice: Since retail markets are very competitive, products which are purchased only infrequently will not survive on retailers' shelves for long. We are therefore confident that there are hardly any products (and therefore prices) which are not covered by our data set.

## **5. Overview of further empirical approaches to the GfK Consumer Panel**

Based on price data obtained from the GfK Consumer Panel, at least three other topics can be addressed empirically. We discuss these issues in the following sections very briefly. For a more detailed survey of the literature and empirical results on these topics, see Fengler (2000).

### **5.1 The empirical relevance of fixed costs of price adjustment**

In addition to psychological pricing points, many other mechanisms of individual price setting and price adjustment that may lead to price rigidities have been developed in the New-Keynesian and Industrial Organization literatures; Köhler and Winter (1993), Köhler (1996) and Blinder et al. (1998) provide surveys. Some of these approaches can be evaluated using the price data obtained from the GfK Consumer Panel.

One aspect of price adjustment considered in the recent literature is the cost associated with adjusting prices. These costs include actual expenses when printing new price lists or changing price tags, but also fictitious costs like lost reputation among consumers. Mostly, these costs are fixed since they do not depend on the magnitude of the price change. Often, these costs are referred to as “menu costs” (Mankiw, 1985). A nice example for adjustment costs are the costs borne by restaurants when printing new menus, and prices of restaurants are changed only infrequently. In a very detailed study of U.S. supermarket chains, Levy et al. (1996) document the existence of fixed costs and their structure in retail stores.

In theoretical models of price adjustment behavior, one can derive the result that firms which are subject to fixed costs of price adjustment change prices more infrequently, but with bigger jumps resulting in discontinuous price paths. Under certain technical assumptions with respect to the stochastic dynamics of the exogenous variables (for example the nature of the input price process), one can show that pricing behavior can be characterized by certain lower and upper bounds. In these circumstances, actual prices will mostly deviate from their optimal levels.

Inspection of figure 3, 4 and 5 reveals that the discontinuous price adjustment can be observed in the coffee retail market. Therefore, we would conclude that there is some evidence for fixed costs of price adjustment in the price data obtained from the GfK Consumer Panel. A detailed analysis of the relevant lower and upper bounds is, however, limited since – as was already mentioned – we cannot observe individual price changes in the data.



## **5.2 Aggregate consequences of individual price adjustment behaviour**

As noted earlier, a comprehensive assessment of the aggregate consequences of microeconomic price rigidities is possible only within the framework of a structural aggregation model. For example, the lumpy adjustment caused by fixed adjustment costs (which is a microeconomic rigidity) might vanish completely in the aggregate price level. To achieve this remarkable result, the adjustment of individual prices has to be staggered in a certain way.<sup>11</sup>

Comparing the smooth and continuous path of average retail prices (figure 1) with the rigidities documented in individual prices (figures 3 to 5) does not contradict these kinds of models. A more detailed empirical analysis with price data obtained from the GfK Consumer Panel seems possible; however, once again such an analysis would be limited by the fact that no individual time series of price paths are available.

## **5.3 Price dynamics, price dispersion, aggregate price level, and inflation**

The hypothesis that the dynamics and dispersion of individual prices are not independent of changes in the aggregate prices level, i.e., the rate of inflation, is rather old. Mills (1927) found a positive correlation between price changes and the rate of inflation when investigating American whole sale price indices; his findings have been replicated in vast number of subsequent studies.<sup>12</sup> However, we should point out that this empirical regularity has not yet found any satisfying theoretical basis.<sup>13</sup>

The structure of the GfK Consumer Panel suits the purpose of empirical studies in this area perfectly. To our knowledge, there is only one comparable study which uses data with both high frequency and a low aggregation level; these data are from a period of hyperinflation in Argentina (Tommasi, 1991). It is interesting to replicate this study using the GfK data which were recorded in the stable monetary environment of Germany in 1995. Fengler and Winter (2000a) report first results based on an econometric analysis using panel-data methods.

---

<sup>11</sup> See Blanchard (1983, 1987), Caplin and Spulber (1987), Caplin and Leahy (1991), Caballero and Engel (1991), Caballero (1992).

<sup>12</sup> See Vining and Elwertowski (1976), Sheshinski and Weiss (1977, 1983), Parks (1978), Fischer (1981), Danziger (1987), Domberger (1987), Lach and Tsiddon (1992) and, for Germany, Franz (1985) and Gahlen (1988).

<sup>13</sup> Hartman (1991) and Bryan and Cecchetti (1999) maintain that the correlations might possibly be due to methodological problems and therefore could constitute statistical artefacts without any economic content.

## 6. Summary and outlook

In many economic models, it is assumed that prices adjust instantaneously to changes of economic conditions (e.g., to shocks in demand or production costs). Since the existence of price rigidities has been frequently documented, more realistic models require that infrequent and lumpy price adjustment have to be taken into account. There are still many unresolved issues in this area, both theoretically and empirically.

In this paper, we show that the dynamics and dispersion of retail prices can be investigated using price data obtained from the GfK Consumer Panel for 1995. Our results document the importance of psychological pricing points for price setting, confirming results from many earlier studies. A new aspect of our analysis that has not been investigated in the literature is the relevance of psychological prices points for price adjustment and aggregation. We interpret our findings as suggestive evidence for the notion that rigidities are relevant for aggregate dynamics in Germany. However, we also confirm that a structural aggregation theory is necessary for a better understanding of the relevance of micro-level rigidities for aggregate dynamics. In such a more comprehensive model, price data obtained from the GfK Consumer Panel might also prove very helpful in the future.

Among the three other areas of empirical research that could potentially be explored with price data from the GfK Consumer Panel, the analysis of the relationship between individual price dynamics, price dispersion and aggregate inflation proves particularly fruitful. Moreover, the very disaggregated, high-frequency data contained in this data-set are almost unique. In other research areas which require that prices changes (and not only distributions of prices) are observed over time, empirical tests unfortunately suffer from the fact that time series of individual prices can be constructed only under additional strong assumptions.

## Contact

*Dr. Joachim Winter  
Sonderforschungsbereich 504  
Universität Mannheim  
D-68131 Mannheim  
winter@uni-mannheim.de*

---

## References

- Blanchard, O. J., 1983, Price Asynchronization and Price-Level Inertia. In: R. Dornbusch, M. H. Simonsen (eds.): *Inflation, Debt and Indexation*. Cambridge, MA: MIT Press.
- Blanchard, O. J., 1987, Aggregate and Individual Price Adjustment, Brookings Papers on Economic Activity, No. 1, 57-109.
- Blanchard, O. J., 1990, Why Does Money Affect Output? In: B. M. Friedman, F. H. Hahn (eds.): *Handbook of Monetary Economics, Volume II*. Amsterdam: North-Holland.
- Blinder, A. S., E. R. D. Canetti, D. E. Lebow, J. B. Rudd, 1998, *Asking About Prices: A New Approach to Understanding Price Stickiness*. New York: Russell Sage Foundation.
- Bryant, M., S. Cecchetti, 1999, Inflation and the Distribution of Price Changes. In: *Review of Economics and Statistics*, 81, 188-196.
- Caballero, R., 1992, A Fallacy of Composition. In: *American Economic Review*, 82, 1279-1292.
- Caballero, R., E. M. R. A. Engel, 1991, Dynamic (S,s) Economies. In: *Econometrica*, 59, 1659-1686.
- Caplin, A. S., J. Leahy, 1991, State-dependent Pricing and the Dynamics of Money and Output. In: *Quarterly Journal of Economics*, 106, 683-708.
- Caplin, A. S., D. Spulber, 1987, Menu Costs and the Neutrality of Money. In: *Quarterly Journal of Economics*, 102, 703-725.
- Carlton, D. W., 1986, The Rigidity of Prices. In: *American Economic Review*, 76, 637-658.
- Cecchetti, S., 1986, The Frequency of Price Adjustment. A Study of the Newsstand Prices of Magazines. In: *Journal of Econometrics*, 31, 255-274.
- Dahlhäuser, H., 1996, Prominenz der Preise in einem Warenhauskatalog. In: *Zeitschrift für betriebswirtschaftliche Forschung*, 48, 711-737.
- Danziger, L., 1987, Inflation, Fixed Costs of Price Adjustment, and the Measurement of Relative Price Variability. In: *American Economic Review*, 77, 704-713.
- Diller, H., A. Brielmaier, 1996, Die Wirkungen gebrochener und runder Preise: Ergebnisse eines Feldexperiments im Drogeriewarensektor. In: *Zeitschrift für betriebswirtschaftliche Forschung*, 48, 695-710.
- Domberger, S., 1987, Relative Price Variability and Inflation: A Disaggregated Analysis. In: *Journal of Political Economy*, 95, 547-566.
- Fengler, M., 2000, Preisdynamik und Preisdispersion von schnellebigen Konsumgütern. Unpublished diploma thesis, Universität Mannheim.
- Fengler, M., J. Winter, 2000a, Consumer Prices and Aggregate Inflation: Evidence from Germany. Unpublished manuscript, Universität Mannheim.

- Fengler, M., J. Winter, 2000b, Psychological Pricing Points and Economic Rationality. Unpublished manuscript, Universität Mannheim.
- Feuerstein, S., 1999, Do Coffee Roasters Benefit from High Prices of Green Coffee? Unpublished manuscript, Universität Heidelberg.
- Fischer, S., 1981, Relative Shocks, Relative Price Variability, and Inflation. In: *Brookings Papers on Economic Activity*, Nr. 2, 381-431.
- Franz, W., 1985, Nicht-neutrale Effekte der Inflation auf die Preisstruktur: Theoretische Überlegungen und empirische Resultate. In: *Jahrbücher für Nationalökonomie und Statistik*, 200, 41-55.
- Gahlen, B., 1988, Relative Preise und Strukturwandel bei Inflation. In: *Ifo-Studien*, Nr. 1, 1-42.
- Gordon, R., 1990, What is New-Keynesian Economics? In: *Journal of Economic Literature*, 28, 1115-1171.
- Hartman, F., 1991, Relative Price Variability and Inflation. In: *Journal of Money, Credit and Banking*, 23, 185-205
- Huston, J., N. Kamdur, 1996, \$9.99: Can't 'Just Below' Pricing be Reconciled with Rationality? In: *Eastern Economic Journal*, 22, 137-145.
- Kashyap, A., 1995, Sticky Prices: New Evidence from Retail Catalogues. In: *Quarterly Journal of Economics*, 101, 245-274.
- Köhler, A. G., 1996, Nominale Preisrigiditäten auf Gütermärkten: Eine empirische Überprüfung neukeynesianischer Erklärungsansätze (zugl.: CIRET Studien, 51). München: ifo Institut für Wirtschaftsforschung.
- Köhler, A. G., J. Winter, 1993, Nominal Price Rigidities and Business Fluctuations: Theory and Evidence". In: K. H. Oppenländer, G. Poser, G. Nerb (eds.), *The Use of Business Survey Data: New Insights* (also: CIRET Studien, 46). München: ifo Institut für Wirtschaftsforschung.
- Lach, S., D. Tsiddon, 1992, The Behavior of Prices and Inflation: An Empirical Analysis of Disaggregated Price Data. In: *Journal of Political Economy*, 100, 349-389.
- Levy, D., M. Bergen, S. Dutta, R. Venable, 1996, The Magnitude of Menu Costs: Direct Evidence from Large U.S. Supermarket Chains. In: *Quarterly Journal of Economics*, 791-825.
- Mankiw, N. G., 1985, Small Menu Costs and Large Business Cycles: A Macroeconomic Model. In: *Quarterly Journal of Economics*, 100, 529-538.
- Mills, F. C., 1927, *The Behavior of Prices*, New York.
- Monroe, K. B., 1973, Buyers' Subjective Perceptions of Price. In: *Journal of Market Research*, 10, 70-80.

- 
- Müller-Hagedorn, L., S. Zielke, 1998, Das Preissetzungsverhalten von Handelsbetrieben im Zuge der Währungsumstellung auf den Euro. In: *Zeitschrift für betriebswirtschaftliche Forschung*, 50, 946–965.
- Parks, R. W., 1978, Inflation and Relative Price Variability. In: *Journal of Political Economy*, 86, 78-95.
- Rotemberg, J., 1982, Sticky Prices in the USA. In: *Journal of Political Economy*, 90, 1187-1221.
- Sheshinski, E., Y. Weiss, 1977, Inflation and Costs of Price Adjustment. In: *Review of Economic Studies*, 44, 287-303.
- Sheshinski, E., Y. Weiss, 1983, Optimum Pricing Policy under Stochastic Inflation. In: *Review of Economic Studies*, 50, 513-529.
- Slade, M., 1998, Optimal Pricing with Costly Adjustment: Evidence from Retail-Grocery Prices. In: *Review of Economic Studies*, 65, 87-107.
- Slade, M., 1999, Sticky Prices in a Dynamic Oligopoly: An Investigation of (s,S) thresholds. In: *International Journal of Industrial Organization*, 17, 477-511.
- Vining, D., T. Elwertowski, 1976, The Relationship between Relative Price and the General Price Level. In: *American Economic Review*, 66, 699-708.
- Wedel, M., P. S. H. Leeflang, 1998, A Model for the Effects of Psychological Pricing in Gabor-Granger Price Studies. In: *Journal of Economic Psychology*, 19, 237-260.
- Wiswede, G., 1995, *Einführung in die Wirtschaftspsychologie*. München: E. Reinhardt.

# LIVING ARRANGEMENTS AND CONSUMPTION. PLEAD FOR LIFESTYLE ORIENTATED MARKET RESEARCH

*NORBERT F. SCHNEIDER, DORIS ROSENKRANZ AND KERSTIN HARTMANN*

## 1. Thesis

Families, sitting under an apple tree, are eating margarine and are looking forward to get their building society savings agreement. Singles, on the other hand, love chocolate light, sports and candlelight dinner. They are despairing because of non-shining glasses reducing their chances of finding a partner at the marriage market. Reality is as simple as that in advertisement – and partly also in market research, where private living arrangements often are treated undifferentiated. Clichés are surviving although social reality has changed. Family sociology has shown in the last years

- that there is a growing diversity of household types and
- that „the“ family household or „the“ single household never existed as a homogeneous group (see e.g. *SCHNEIDER, ROSENKRANZ, LIMMER 1998*).

Looking at the present situation of private living arrangements in Germany, a “family“ consisting of a couple and its children is still the most frequent household type. Every second person between age 25 and 55 lives in a family of this kind. Besides, other non-conventional household types emerged which is leading to a growing dynamic of living arrangements during the life course: Non-conventional living arrangements are chosen before, after and partly in place of marriage – often temporary but sometimes permanently. Non-marital unions, single households or single parenthood are now common options for the organisation of the individual life, not being stigmatised anymore.

The central question is: Does the change of household structures, demographic structures and living arrangements affect private consumption patterns? Do private living arrange-

ments have their own specific consumption patterns? Which factors are relevant in this context? These questions are to be answered in the following.

Our paper is an attempt to transfer knowledge of family sociology to approaches in market research. We want to show the potentials and limitations of an approach considering living arrangements in market and consumption research. From our point of view the idea of allowing for the effects of living arrangements is not an alternative but a supplement to other approaches, such as lifestyle orientation.

In our opinion, in market research up to now only little attention is given to the influence of living arrangements and the process of family formation on consumption and if this aspect is considered, often obsolete concepts are used. In our opinion it is possible to reach higher efficiency of the instruments in market research, i.e. better market segmentation and better specification of target groups by including characteristics of private living arrangements. Therefore this paper pleads for more interdisciplinary co-operation between marketing research and sociology.

In present research on consumption (see ROSENKRANZ 1998) private living arrangements are referred to only on a very rough and undifferentiated level. Moreover studies dealing with living arrangements are often including:

- obsolete approaches (e.g. the theoretical model of a very rigid family life cycle)
- undifferentiated operationalisations (e.g. description of household structures by using the number of household members) and
- insufficient inclusion of attributes (e.g. transitions in family life).

Therefore it is not possible to get a valid representation of the reality of living arrangements and analyse their influence on consumption sufficiently. Repeatedly reported results showing only little influence of living arrangements on private consumption are not due to their irrelevance but to the insufficient representation of these living arrangements.

Changes of family and household structures and – more general – in the way of living being observed in the FRG and in most of the other western industrialised societies are not taken into account, neither in regard to their expansion nor to their significance. This change is often referred to as *pluralisation of living arrangements* or as *individualisation*. We find three central aspects of these changes:

- The morphology of household structures has changed: Apart from family forms based on marriage other living arrangements like non-marital unions, singles, LAT (Living-apart-together) or voluntary childless couples have emerged. At the moment

more than every third person in younger age groups is living in a non-conventional living arrangement.

- The internal differentiation of living arrangements, having similar external structures, increased, leading to a reduction of the power of morphological attributes (e.g. marital status or number of household members) to tell something about the living conditions of people.
- The dynamic of life courses has changed: Living arrangements are considered less often on long-term perspectives but are more and more linked to a certain phase of life. This leads to new types of modern life course. It is typical for them that more different forms of living arrangements than in the past are following after each other and that more often changes occur between them. As a consequence the life span people are spending in a family has shortened.

In this regard lifestyle orientated consumption research should be based on the following assumptions and theses:

- Household context and the position in the process of family formation as well as changes of living arrangements do have a long lasting effect on consumption behaviour. In this context structural attributes of living arrangements are important in a more complex way than market and consumption research has assumed up to now. For example simple differentiation by marital status doesn't yield to relevant results. There are other criteria having to be recorded more differentiated as well. When looking at the consumption behaviour for example, not only the total sum of the household income is important, also the number of people, who are earning this money have to be taken into consideration.
- As a consequence of their structural attributes, specific living arrangements are leading to specific patterns of consumption. Type and structure of these patterns are formed further by other attributes especially age, income and attitudes.
- Morphological attributes are losing their significance: Equal external structures allow only little conclusions about the reality of a living arrangement. To get a adequate idea of a certain living arrangement, its internal structures and its position in the process of family formation are getting more and more important. For the private demand for technical household equipment is quite different for example, if a non marital union consists of two unmarried or two divorced partners. Furthermore it is important that "classical" structural categories like "singles" or "single parents" are hardly useful because they assume similarity of the living context where heterogeneity is dominating. For example one cannot assume that every person in a one-person



household is a single, i.e. doesn't have a partner. According to latest results, every third person aged 25 to 45, is living alone, but has a partner living in a different household (see SCHNEIDER 1996).

- Life course perspective: Living arrangements not only have to be seen as structural categories but at the same time as phases of life – and this in two aspects: Living arrangements are not static but very dynamic and living arrangements are corresponding with the educational and professional career and the process of family formation. Another example: Non-marital unions differ in duration and in their position in the process of family formation: Young non-marital unions mostly have separate budgets, elderly non-marital unions do often have separate budgets and non-marital unions with the intention to get married and to have children mostly have a common budget (see VASKOVICS, RUPP, HOFMANN 1997).
- Cross-household perspective: More and more living arrangements are existing independent of a household. This is especially true for partnerships ("living apart together") but also for families who are more and more often organised as ego-centered multilocal networks, with relevant intergenerational relationships (e.g. to grandparents).
- In many studies children are only included up to the age of 14 or 18, older children living in and/or outside the household being systematically neglected. Considering the numerous transfer benefits especially from parents and the increasing duration children are living at home with their parents, that procedure excludes a main part of reality.
- Apart from realised purchases it is also important to look at the decision making processes of buying in connexion with living arrangements. There is a lack of analyses which are examining more than decision-making processes of married couples. Do married couples without children have the same decision making-processes as childless non marital unions? How are single parents deciding what to buy?

To sum up, the potentials of a lifestyle orientated consumption research are as follows:

- Possibilities:
  - Better coverage of the diversity and heterogeneity of living arrangements and life courses
  - Differentiated consideration of attributes characterising living arrangements and therefore also the individual reality of life

- Living arrangements and their relevant influence of consumption are seen in the context of societal structures and individual preferences and attributes.
- Potentials for market research:
  - Clearer specification of target groups by few additional and easily measurable attributes
  - Better segmentation of the market consisting of household and living arrangements
  - Improved possibilities for integrated concepts of marketing
  - Improved possibilities to forecast the development of future consumption in many of its spheres
  - Increase of explained variance by the means of a more valid and more differentiated representation of reality.
- This leads to three questions in analysing the GfK data:
  - a. Are there differences between sociologically defined living arrangements in regard to the consumption of selected products?
  - b. What influence do familial transitions (e.g. birth, separation, a new partnership) have on consumption?
  - c. Are the available data of the consumer panel appropriate to answer this questions?

The first question is central for this article. Before going into detail we want to sum up some of our previous empirical results:

- Household income is important for the extent but not for the structure of consumption
- "Classical" determinants like age, number of persons in the household, age of the youngest child and employment status of household members which are usually being analysed do have an independent influence on consumption but its nature and extent depend on other specific attributes of the living arrangements (effect of interaction)
- Two attributes of living arrangements were especially appropriate for explanations:
  - The biographical position, i.e. the age context of living arrangements
  - The position in the life course or in the family formation process, i.e. in relation to the individual succession of different living arrangements.

## 2. The creation of types of living arrangements with the data of the consumption panel

By means of an explorative method it is to be tested to what extent consumption of certain goods can be explained by an effect of private living arrangements. Using the example of a chosen product, it has to be shown that its consumption in different – theoretical and empirical relevant – living arrangements differs significantly. In the following we will report on the formation of living arrangements, the selection of specific products and their application in the analysis.

We choose the data of ‘Panel 6’ of the GfK consumption panel, because it contains more information on household structures than ‘Panel 1’. The available data were only collected once in January 1<sup>st</sup> 1995, so information about changes of the household structure during 1995 is not available. The decision for Panel 6 leads us also to a restricted selection of consumer goods which can be used for the analysis. These consumption data on the other hand, are available for the whole year of 1995.

The data basis consists of 4638 households. To homogenise the household sample in regard to employment status we excluded households where the main income earner is not employed or without any profession, as well as old age pensioners, retirees, pupils, children, housewives, unemployed widows<sup>1</sup> and persons older than 60<sup>2</sup>. Afterwards 2476 households are remaining.

Based on these 2476 households we defined living arrangements relevant in family sociology, using mainly the aspects of partnership and parenthood as indicators. Additionally we considered the degree of institutionalisation of the living arrangements, although only distinctions between „married“ and „not married“ were possible, because of small numbers. Finally only three variables were suited for modelling living arrangements, i.e. size of household, marital status of the head of the household and the number of children under 18 years. The consequence of this insufficient information leads to the fact that the assignment of the households to specific living arrangements often were made by reasons of plausibility. Although these assignments are valid in their results, there is no final security in having classified every single household in the right way. As an example the case of a married couple without children shall be explained. In this case, we assume that in a 2-person-household with no children the statement "the household leading person is

---

1 see var36.

2 Because of unspecific definitions household leading persons as well as heads of the household over 60 years are excluded.

married" can be completed to "the household leading person is married to the second person who lives in the household". Finally, the following living arrangements were defined:

- *Single households*: Persons living in a 1-person-household (n=324)
- *Non-marital unions without children*: 2-person-households where the household leading person is not married and where no children are living (n=134)
- *Married couples without children*: 2-person-households where the household leading person is married and where no children are living (n=442)
- *Married couples with children under 18*: Households where the household leading person is married and where – after subtraction of the number of children under 18 from the total number of persons in the household – a 2-person-household is left (n=968)
- *Non-marital unions with children under 18*: Households where the household leading person is not married and – after subtraction of the number of children under 18 from the total number of persons in the household – a 2-person-household is left (n=44)
- *Single parents with children under 18*: Households where – after subtraction of the number of children under 18 from the total number of persons in the household – one person is left (n=56)

It was possible to classify 1968 households according to these groups. The remaining 508 households are mainly consisting of two types of living arrangements: a) Households where couples with children under 18 and other persons are living (n=196). Even though it can be assumed that in these households couples with children under and over 18 years are living, it is not possible to prove that, so they have been excluded from the analysis. b) There is a similar situation with households where the main income earner is between 45 and 59 years old, the household leading person is married and is living in a 3 or 4-person household (n=267): In this case it is likely that the household consists of a couple with its children over 18 but it is also possible that it is a couple with its older parents. These examples show that no clear assignment is possible because of the restrictions of the data.

Summarising the construction of living arrangements we can state that the available data are only of limited use for our purpose. Important attributes of household structure have already been missing in the concept of the survey. It adds to this that relevant attributes which were collected in Panel 6, e.g. the exact age of the children and the gender of the

household leading person, were not available in the data ZUMA passed on, further restricting the possibilities of analysis. Another problem can be seen in the unclear definition of the variables "household leading person" and "main income earner". We cannot specify who is assigned in which case to which variable and in how many cases the same person is used for both variables. In addition to these difficulties, in most cases, relevant information is available only for one of these concepts. As an example, due to this, we had to use the information on the profession of the main income earner on the one side, and on the other side we had to use the information on the marital status of the household leading person to define different living arrangements. In spite of these difficulties we were able to come up with a differentiated representation of theoretical relevant living arrangements which most of the households could be clearly assigned to with high probability.

One problem that could not be solved in a satisfactory way with the data at hand was the specification of transitions between the different living arrangements. Although data from January 1996 were made available, single members of the household could not be identified, so only few events could be assigned uniquely. Because of the very small number of cases resulting of this approach this idea had to be cancelled.

As a conclusion of this step of the analysis we may note that it is possible to define relevant living arrangements with the panel data but that the differentiation of household structures could not be pursued to the level that would be concerned necessary on theoretical reasons.

With these living arrangements we wanted to explore, on the basis of a selected product, whether there was an independent influence of the living arrangements. As a consumer good with which to test our hypotheses in an exemplary way, we selected crispbread. This choice may be justified on theoretical grounds because of the fact that crispbread as a convenience product covers different aspects of positioning. For example personal scopes like health, fitness and diet, but also durability seem to play an important role for the buying decision. By the dimension of durability we not only mean to indicate the actual durability of that product but also dimensions like immediate and quick possibilities of consumption. These different aspects of positioning should lead to a different consumption of crispbread in the different living arrangements. Furthermore we selected crispbread because it is a very homogeneous product without any restrictions of availability in retail shops or of high prices. After the comparison with other products, leading largely to similar results, an analysis of crispbread is preferable because the results in this case were particularly significant and well interpretable, which adds to the exploratory character of this study.

The consumption of crispbread was aggregated for each household over the whole year 1995 and then we calculated the consumption per-capita of the household. Children under 6 years were weighted with 0.5 because we assume that crispbread is consumed rarely in that age.

### **3. Living arrangements and consumption: An empirical example**

Before the presentation and comparative interpretation of our results, we are going to introduce the statistical method we choose. The analyses should answer the question, if the living arrangements do have an independent influence on the consumption of particular products. In other words: Do the different living arrangements, which have been constructed on the basis of theoretical considerations, yield significant differences in the consumption of crispbread? Considering the measurement scales of the selected variables, we note that the grouping variable is nominal and the independent variable is metric. Furthermore, there is only one dependent variable<sup>3</sup>, so the method of choice is to perform a one-factorial analysis of variance (ANOVA). With the help of the ANOVA we are to find out, if there is a significant difference in the average per-capita-consumption of crispbread between the different living arrangements.

The ANOVA was carried out on different theoretical levels. It starts with two groups in each case, which differ in the status of partnership (yes/no) and parenthood (yes/no). In the following analyses with more than two groups we applied a Tamhane-T2 test after the ANOVA to test for differences between the subgroup-means. This test is based on a t-test and is a comparison in pairs of the means of subgroups. It is suited in our case because the groups to be compared may be of different sizes and it does not require further assumptions as to homoscedasticity between the groups<sup>4</sup>.

This step is followed by a univariate analysis of variance with several factors together in one model, which does not only estimate main effects but additionally takes into account for interaction effects between the factors. We have chosen a factorial design with unequal cell frequencies, where the calculated sums of squares of one factor was corrected for all the other effects.

To test our hypotheses we have applied several steps of analysis. We started with comparing the levels of partnership and parenthood. Next we have extended our analysis to

---

<sup>3</sup> Which for example excludes a discriminance analysis.

<sup>4</sup> The test was always done with a significance level of 0,05.

the classified living arrangements. Finally we added the variables age and occupational status to broaden further the context and validity of our findings.

The first two comparisons show, that persons living in a partnership consume less crispbread than those living alone<sup>5</sup>. This might be caused by the durability aspect we have already discussed, which is of particular importance for small households. As to the aspect of health and diet we cannot provide an interpretation at this point due to the missing information of the age of the consumer, as we assume that the diet aspect is especially important for younger people still being active at the marriage market, while older people are more concerned with their health.

Within the parent-level we compared households with and without children. People in households with children consume significantly<sup>6</sup> less crispbread than those without children. Here again we can only speculate on the reasons. On the one side, one can say that in these households are no problems of durability, as in less time more food is consumed than in those without children, leading to a higher rate of „normal“ bread. An alternative could be, that families with children consume more cooked meals and have less small meals, which leads to a reduced demand for crispbread too.

After these coarse comparisons we combined the parenthood- and partnership-status and additionally took into consideration the degree of institutionalisation of the living arrangements, which lead to the above described groups. As a first result one can say, that there are differences between the subgroups. Graph 1 displays the means as well as the upper and lower values of the confidence intervals. It gives an impression on the significant differences. By the Tamhane-T2-Test we have checked them in more detail. First of all it is remarkable that singles and married couples without children consume substantially more crispbread than all the other groups. The differences between each of these two groups and the other living arrangements are all significant. This finding is not repeating the results of the parenthood analysis, because these would suggest a higher demand of non-marital unions without children.

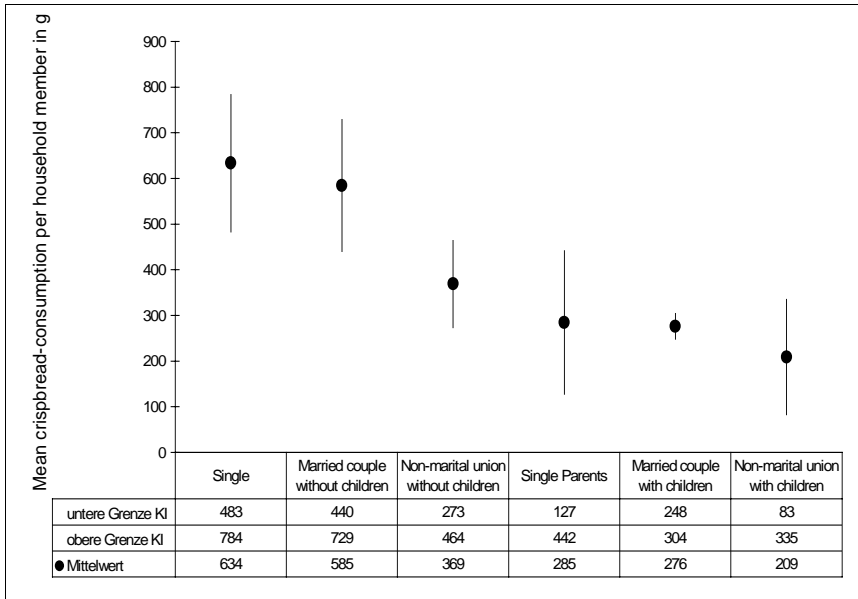
The relatively big confidence-intervals for singles and the married couples without children - compared to those for single parents as well as non-marital unions without children - cannot be explained by a too small number of cases. Obviously these groups have a larger intra-group heterogeneity than for instance married couples with children. This can be taken as an indication for other relevant influences.

---

<sup>5</sup> Level of significance: 0,05.

<sup>6</sup> Level of significance: 0,05.

Graph 1: Mean crispbread demand by living arrangements



Source: ZUMA data of GfK consumer panel 1995, own calculations

The next question to be answered is: Does the effect of living arrangements persist after controlling for other variables and/or do these additional variables have a unique and perhaps an even greater influence on the differences in crispbread consumption? We included occupational status and age into our model, which are essential components of the "Lebensweltkonzept" of *KLEINING* and *PRESTER* (1998). Using the information on the occupation of the main income earner (the person who contributes the largest part to household income), a rough indicator of occupational status was created<sup>7</sup> – distinguishing between lower and higher occupational status. In order to include age, all the cases were divided into two groups: households with a main income earner younger than 40 years and respectively those with a main income earner aged between 40 and 59. To get ade-

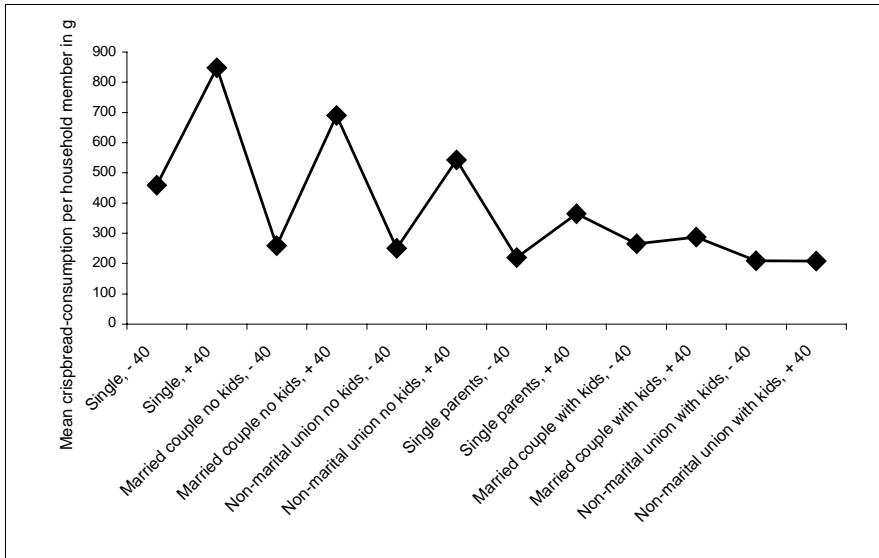
7 Lower occupational status-group: farmer, lower/middle civil servants, instructed employees, semi-skilled workers, skilled workers, apprentices. Upper occupational status-group: Self-employed, civil servants in higher service, qualified and leading employees, qualified skilled workers.



quate group sizes – especially for the combination with living arrangements – we had to decide for such a rough differentiation.

Once more, the influence of occupational status and age was analysed separately at first. As a result no significant difference between the two status groups could be found. This confirms our assumption that crispbread is a product with almost no access-barriers. On the other hand results are different when comparing the age groups. We found significantly less crispbread consumption in households with a main income earner younger than 40 years than in households with older income earners. This might indicate a higher relevance of the health aspects in crispbread consumption than of aspects of diet. At the same time this finding implicitly refers to the position in the family life cycle and therefore stresses the importance of the living arrangement.

Graph 2: Mean crispbread demand by living arrangements and age groups



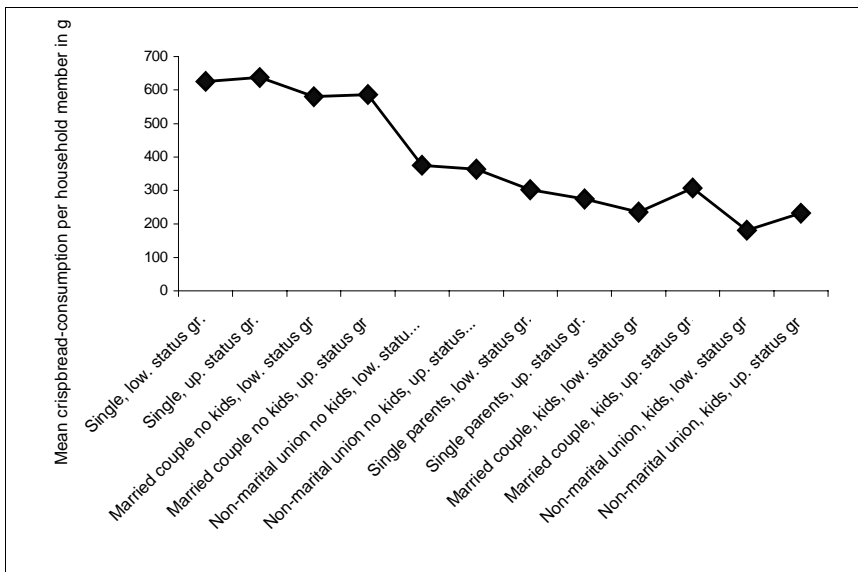
Source: ZUMA data of GfK consumer panel 1995, own calculations

To provide answers to this sort of questions, we now combined both variables with our concept of living arrangements. As a consequence both combinations, living arrangement - status group as well as living arrangement - age group show significant differences. In

Graph 2<sup>8</sup> the influence of age can be seen immediately. In living arrangements without children the demand for crispbread in households with an older main income earner is clearly higher than in households with younger income earners. The effects are disappearing in living arrangements with children<sup>9</sup>. Moreover the result is still valid, that there are differences between living arrangements especially in regard to groups with and without children in the household. So age has an additional effect in special subgroups while the influence of living arrangements does not disappear.

Finally, looking on the influence of occupational-status in combination with living arrangements one gets very different results (see Graph 3): Differences between living arrangements are remaining unchanged, without any further effects of differentiation by addition of occupational status. Only when looking at married couples and non-marital unions with children small but not significant differences can be observed.

Graph 3: Mean crispbread demand by living arrangements and occupational status



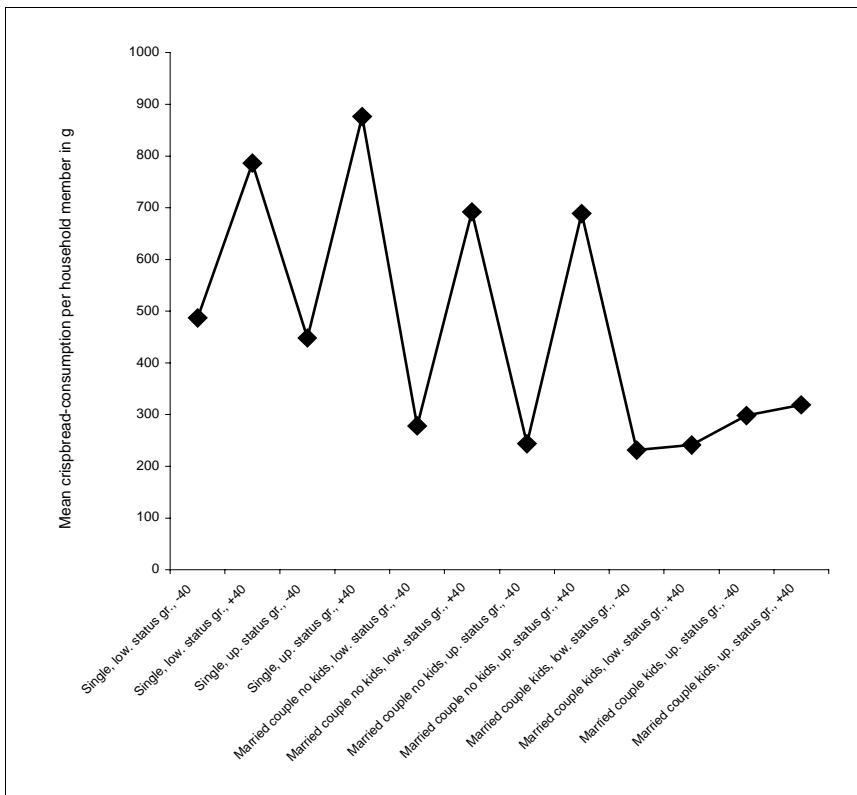
Source: ZUMA data of GfK consumer panel 1995, own calculations

<sup>8</sup> To achieve a higher clarity of the presentation we skipped the borders of the confidence-intervals.

<sup>9</sup> Significance levels are shown in the appendix.

The three largest groups – singles, married couples with and without children – finally have been subdivided by age as well as by status (see Graph 4). According to our former results, it is to be recorded that married couples with children show the lowest level of crispbread consumption of all three groups, without any further observable differentiation by age or status. Another confirmation of our findings presented above is the fact that age but not status has an effect in combination with singles and married couples without children.

Graph 4: Mean crispbread demand by selected living arrangements and age- and status groups



Source: ZUMA data of GfK consumer panel 1995, own calculations

In a last step an univariate analysis of variance with the three factors living arrangements, age and status together was carried out to explain the consumption of crispbread. Although the model explains only a very small part of the variance, it is to be recorded, that only the variable “living arrangements” provides a significant effect in the explanation of variance. Neither age nor status nor the different possible interaction effects have a significant influence on crispbread demand.

It is to be concluded that our hypothesis of a genuine influence of living arrangements on the demand of crispbread can be confirmed. Two considerations may help to explain the findings: The durability of crispbread is relevant because the demand is higher in smaller households with only few members. Secondly, the possibility of substituting cooked meals by foods easy to consume, seems to be relevant for the differences between living arrangements with and without children. With regard to the age effects the influence of the living arrangements is still present, but within some subgroups the crispbread demand was substantially higher in the upper age groups. This leads to the assumption that the aspect of healthiness, which is more relevant for older persons, is more important than the aspect of diet.

All in all this means:

- Living arrangements do have an influence on the demand for crispbread;
- Age does have an effect as a moderating variable;
- Occupational status doesn't have any influence on crispbread consumption.

#### **4. On the capacity of the data for an orientation on living arrangements in market-research**

In connexion with the demands of a orientation on living arrangements in market research the ZUMA-file of the GfK-Consumer-Panel in its present form is only of limited use. On the one hand relevant variables, which have been collected, are missing in the available sample (e.g. gender), on the other hand important variables were not collected at all (e.g. children over 18 in the household). These limitations considerably prevent the specification of differentiated living arrangements.

An especially important restriction in the use of the sample is, that transitions in the family life cycle cannot be depicted at all or only very roughly. Birth, the moving of children out of the parental household, divorce resp. separation of partners and the emergence of new partnerships cannot be represented validly by the present data, although, in our conviction, the consequences are highly relevant for consumption. The same holds

true for the positioning in the family life cycle. While product demand as the variable to be explained is longitudinally recorded, longitudinal information on demographic household structures as the central explaining variable is missing.

Nevertheless our results show, that a differentiation by living arrangements is reflected in differentiated consumption styles. Private living arrangements constitute specific consumption styles – independent of other factors like age, occupational status and income.

## Contact

Norbert F. Schneider and Kerstin Hartmann

FB 12: Institut für Soziologie

Johannes Gutenberg-Universität

Colonel-Kleinmann-Weg 2

D-55099 Mainz

Tel.: 06131/39 22444

E-mail: [norbert.schneider@uni-mainz.de](mailto:norbert.schneider@uni-mainz.de)

Doris Rosenkranz

FH Würzburg

Tel.: 0931/35 111 53

E-mail: [doris.rosenkranz@inetserv.fh-wuerzburg.de](mailto:doris.rosenkranz@inetserv.fh-wuerzburg.de)

## References

Gentry, James W.; Kennedy, Patricia F.; Paul, Catherine und Hill, Ronald Paul, 1995: Family transitions during grief. Discontinuities in household consumption patterns. *Journal of Business Research*, 34, 67-79.

Görtler, Edmund, 2000: Demographische Veränderungen und Konsum. Vom Wandel der Bevölkerungsstruktur zum Demographischen Marketing als neue Methode der Marktforschung. In: Doris Rosenkranz und Norbert F. Schneider (Hg.): *Konsum. Soziologische, ökonomische und psychologische Perspektiven*. Opladen: Leske + Budrich, 305-325.

Görtler, Edmund und Rosenkranz, Doris, 1999: Demographisches Marketing. Der demographische Wandel bestimmt die Marktforschung und das Marketing der Zukunft. In: *Planung & Analyse*. Zeitschrift für Marktforschung und Marketing, 5, 38-42.

Kleining, Gerhard und Prester, Hans-Georg, 1998: Individuelle und Familien-Lebenswelten. Eine neue Marktsegmentation. *Planung & Analyse*. Zeitschrift für Marktforschung und Marketing, 5, 58-63.

Rosenkranz, Doris, 1998: *Konsummuster privater Lebensformen. Analysen zum Verhältnis von familiendemographischem Wandel und privater Nachfrage*. Wiesbaden: Deutscher Universitätsverlag.

Rosenkranz, Doris, 2000: Private Lebensformen und Konsum. Oder: Warum Single nicht gleich Single ist. In: Doris Rosenkranz und Norbert F. Schneider (Hg.): *Konsum. Soziologische, ökonomische und psychologische Perspektiven*. Opladen: Leske + Budrich, 133-148.

Schneider, Norbert F., 1996: Partnerschaften mit getrennten Haushalten in den neuen und alten Bundesländern. In: Bien, Walter (Hg.): *Familie an der Schwelle zum neuen Jahrtausend*. Opladen: Leske + Budrich, 88-97.

Schneider, Norbert F.; Rosenkranz, Doris und Limmer, Ruth, 1998: *Nichtkonventionelle Lebensformen. Entstehung, Entwicklung, Konsequenzen*. Opladen: Leske + Budrich.

Vaskovics, Laszlo A.; Rupp, Marina und Hofmann; Barbara, 1997: *Lebensverläufe in der Moderne. Nichteheliche Lebensgemeinschaften. Eine soziologische Längsschnittstudie*. Opladen: Leske + Budrich.

Wildt, Michael, 1995: *Am Beginn der Konsumgesellschaft*. Hamburg: Ergebnisse Verlag.

**Appendix**

*Table 1: Significant differences of crispbread demand between living arrangements with and without children*

	Single with children	Married couple with children	Non-marital union with children
Single	*	*	*
Married couple without children		*	*

Level of Significance: 0,05, without empty rows and columns

*Table 2: Significant differences of crispbread demand between living arrangements and age*

	Single, older than 40 years	Married couple without children, younger than 40	Married couple without children, older than 40
Married couple without children, younger than 40 years	*		
Married couple without children, older than 40 years		*	
Non-marital union without children, younger than 40 years	*		*
Single, younger than 40 years	*		*
Married couple with children, younger than 40 years	*		*
Married couple with children, older than 40 years	*		*
Non-marital union with children, younger than 40 years	*		*
Non-marital union with children, older than 40 years	*		

Level of Significance: 0,05, without empty rows and columns

Table 3: *Significant differences of crispbread demand between living arrangements by occupational status*

	Single, upper status group	Married couple without children, upper status group
Married couple with children, lower status group	*	*
Married couple with children, upper status group	*	
Non-marital union, lower status group	*	

Level of Significance: 0,05, without empty rows and columns

Table 4: *Significant differences of crispbread demand between selected living arrangements by age and occupational status*

	Single, upper status group, older than 40	Married couple without children, upper status group, younger than 40	Married couple without children, upper status group, older than 40
Married couple without children, upper status group, younger than 40 years	*		
Married couple without children, upper status group, older than 40 years		*	
Married couple with children, lower status group, younger than 40 years	*		*
Married couple with children, lower status group, older than 40 years	*		*
Married couple with children, upper status group, younger than 40 years	*		*
Married couple with children, upper status group, older than 40 years	*		*

Level of Significance: 0,05, without empty rows and columns



# HOUSEHOLD'S INCOME SITUATION AND CONSUMPTION OF ALCOHOLIC BEVERAGES

*GEORGIOS PAPASTEFANOU*

## 1. Introduction

Living conditions of a household's members are highly influenced by the household's financial situation, especially after the available household income has dropped below a significant level.

In poverty research usually identifies a household's the significant income level at or below fifty percent of the mean equivalent net household income. One then assumes that the members of these households are deprived of the standard quality of life in society.

This indicator is easily computed, but there is no theoretical argument supporting the below-50-percent threshold as the significant income level. Net income is just a proxy for the material conditions of living. It would be quite informative, to explicate empirically the living conditions covarying with the income position (Weick 1999).

There are some studies presenting results for covariations of net household income with housing, health, employment and education (Townsend 1979, Ulrich/Binder 1992, Hauser/Hübinger 1993a, 1993b). But there is little information on income effects on the consumption of alcoholic beverages like liquors (Feichtinger 1996), which are a societally defined of the standard of living

Most research on the consumption of alcoholic beverages are focused on the societally and individually negative consequences of abuse of alcoholic beverages. Some few studies provide results on covariations of financial and/or socio-economic status and the demand for alcoholic beverages.

It seems to be, that there is more alcoholic beverages consumption in higher social status households than in lower status strata (Kirschner/Meinlschmidt 1994; Mielck/Helmert 1998, Thiel/Thai 1995, Kußmaul 1995).

But this covariation depends on the type of beverage and its socially defined image. So, less beer is consumed in higher social status groups (Mielck, Helmert 1998), it seems to be that this correlation is only true for men (Kusssmaul 1995).

Contradictory to these results the study of Härtel (1993) found that men of higher educational status overall consume less alcoholic beverages. Except for women with a better educational attainment who tend to consume more alcoholic drinks than those with a lower educational level (after having controlled for age).

In the study of Härtel (1993) consumption of alcoholic beverages was measured generally at the weekend and last working day before interview. Whereas in the study of Kusssmaul (1995) consumption of alcoholic beverages was measured by a seven day nutrition diary.

Overall, these studies give little insight into the covariation of income level and consumption of alcoholic beverages. First, all reports are based on cross sectional correlations. So they cannot provide findings on how a change of household's income does change the demand for alcoholics. Second, highly different methods of data collection are used, three day nutrition diary (Thiel, Thai 1995), seven day nutrition diary Kusssmaul (1995), retrospective data collection of alcoholics consumption at last weekend and the last day before interview (Härtel 1993) and a one spot self report measurement (Hermann-Kunz 1995).

Beside the problem of retrospective measurement of quantitative behavior, the most crucial measurement problem is given by the restrained short referred time interval. As consumption of alcoholics usually depends on a seasonal and socially defined events like silvester or birthdays the data collections might be biased by selective measurement.

Third, only one study differentiated between beer, wine, champagne and liquor as consumed alcoholics. But as contradictory results on the socioeconomic variation of beer consumption shows, each type of alcoholic beverage seems to have different symbolic and social definition.

In sum, to provide results on the effect of household's income position on its consumption of alcoholics, one needs longitudinal time continuous data on the consumption of different categories of alcoholics. This kind of data is provided by the consumer panel data of GfK, including information on the stability or change of a household's income position. Additionally, there is also information available on the socio-demographic structure of the household, which can be used for deeper exploration of the presumable income effect.

In this paper the covariation of level and change of the net household income position with the consumption of liquors, alcoholic beverages with more than 20 % is examined.

First some social psychological aspects of liquor consumption are given then the operationalization of income level and income change is described. The influence of household income position is estimated via linear regression analysis and its moderation by household demographics. After summing up further research options are discussed.

## **2. Social psychological aspects of liquor consumption**

Drinking alcoholic beverages is a normal part of life in western societies. Alcoholics are stimulants whose consumption is tied to social events like birthday parties, graduation parties etc. which organize the meeting of groups of people.

Especially in gatherings which offer opportunities for status demonstration like dinners with friends and relatives liquors are part of the event. By offering high quality and high priced liquors the host uses the possibility of positive self representation as this is signaling a high level life style and taste of the household. Additionally this gesture tells to the guests a high esteem by the host.

Therefore liquors can be seen as a mean of social communication, transporting symbolic meaning (towards others and towards oneself) of participating in a high level quality of life. In a population survey 85 percent of the interviewed confirmed that it is good form to offer alcoholic beverages for guests. (Antons/Schulz, 1990, p. 142).

Beside the social meaning of alcoholic beverages one should not ignore the psychological stimulating aspects. Moderate consumption of alcoholic drinks reduces anxiety, lowers arousal and dissolves inhibitions.

In the context of the psycho-social meanings of alcoholic beverages households with lower income position could experience multiple disadvantages. Because of their financial restrictions they might be not able to afford liquor as a sort of luxury good. Lower income position might be correlated by stressful status decreasing events like unemployment and separation and divorce, thereby favoring the consumption of alcoholics as a kind coping behavior, accompanied with the risk of harmful side-effects.

But one should also have in mind that because of the social meaning of alcoholics an increase of its consumption might be also based in financial improvements of the household. Consuming, purchasing and offering of alcoholic beverages might then help to re-signal participation in an upper standard of living.

### 3. Data and method

The analysis is based on purchase history data as well as on socio-demographic and data of the household, which are provided by the GfK consumer panel waves 1994 and 1995. Alcoholic beverages like appetizers, schnaps (fruit or grain), brandy, cognac, vodka, gin, rum, liquors, and other spirits. In the present study all purchases of these liquors in the first three months of 1995 are analyzed.

The income position of a household is measured by relating its net household income to the overall mean income. On this basis there are four household types defined those in lowest income position, in lower income position, middle income position, upper income position (see table 1).

*Table 1: Types of household by their income position.*

income position	net household income
lowest income position	up to 1499 DM
lower income position	between 1500 and 2999 DM
middle income position	between 3000 bis 4999 DM
upper income position	5000 DM and more

By comparing a household's income position between 1995 and 1994 one gets additional information on the stability of the income level. Logically there are six improvement transitions, six transitions to lower income positions and four no-change combinations of 1994 and 1995 income positions.

The examination of the amount of liquors purchased in the first three months of 1995 for every type of income position transition leads to following results (see table 2).

It seems that the amount of purchased liquors reflects a joint effect of income position and income position change. The higher the income position the more liquors were bought in the period of observation. But the increase of the purchased amount of liquors can be mainly observed as difference between the households with lowest income position compared to all the other households. One can observe this covariation in 1994 as well as in 1995.

Table 2: Amount of liquor (ml) purchased in the first three months of 1995 by household's income position in 1994 and 1995

Household's income position 1995	household's income position 1994				
	lowest income position	lower income position	middle income position	upper income position	overall
lowest income position	722	779	1320	No values	749
lower income position	1063	1410	1834	1830	1376
middle income position	1623	1801	1454	1333	1486
upper income position	2710	1907	1660	1492	1525
overall	855	1454	1541	1485	1384

Additionally there seems to be a positive effect of income position improvement on the amount of purchased liquor beverages, especially when there was a change from lowest to lower income position. But the increase of liquor consumption obviously is also concomitant to a downward change of the household's income position.

To examine these face valid results income position as well as the changes of income position were integrated as dichotomous variables (collapsing those income position groups with similar purchase behavior) into a linear regression model of the amount of purchased liquors.

Additionally the household size was put into the regression model as a factor of its own, to get close to a meaning of income position as an indicator of consumption possibilities. Number of adults and number of children were incorporated as two separate predictors.

Finally the regression model consisted of these variables.

Table 3: Definition of predictors

income position improvement	Household being 1994 in lowest or lower income position and in 1995 in middle or upper income position = 1, all other changes of income position = 0.
income position deterioration	Household being 1994 in middle or upper income position and in 1995 in lowest or lower income position = 1, all other changes of income position = 0
lowest income position	Lowest income position=1, other income position = 0
number of adults	number of persons aged 18 years or more
number of children	number of children under 18

## 4. Results

### 4.1 Effects of income position and its change on the amount of purchased liquors

Table 5 gives the results of the estimated regression of the amount of liquors purchased in the first three months of 1995 on income position and change.

One can find, that a household's income position in the beginning of 1995 covaries with the consumption of liquors in the three-month interval thereafter: lowest income position households significantly buy less liquor beverages. This reflects some degree of income elasticity of liquor consumption.

Looking at the effects of changed income position one gets additional insight: improvement as well as deterioration of income position are related to more consumption of alcoholic beverages.

This simple regression model further shows that the quantity of consumed liquors depends on the number of adults in the household, the more adults the more liquors are consumed. But in households with larger families (more than one child) less liquor is consumed than in small families.

Maybe this reflects the context of liquor consumption as in households with children there are less gatherings of adults which are tied to alcoholic beverage consumption.

Table 5: *Effects of income position and its change on the amount of purchased liquors (b-coefficient) in 1<sup>st</sup> quarter of 1995*

predictors	amount of purchased liquor in 1 <sup>st</sup> quarter 1995
net mean	811
improvement of income position	381
deterioration of income position	403
lowest income position	-455
number of adults in household	331
number of children	-120

\*note: all effects are significant at  $p \leq 0.05$

In sum, the results support the following statements:

- The consumption of luxury products like liquors is reduced in households with low income.
- Improvements of household income level are connected with increased purchases of liquors. This makes sense in the context of the symbolic status meaning of liquor products.
- Deterioration of income level obviously is also followed by more liquor purchases. Maybe this reflects the relaxing function of alcohol and its use as a short-term, impulsive behavior in coping with stressful events.

For further exploration into these interpretations of liquor consumption the income and income change effects are examined for being moderated by the socio-demographic situation of the household (as indicated by the educational status of the household head, employment status of the person who is running the household, children being in the household or not).

## 4.2 Income effects on the quantity of liquor purchases moderated by socio-demographic household structure

### 4.2.3 Educational status of the household head

There are clear differences in the income position dependency of liquor consumption, if the household head has an occupational education compared to those without occupational education (see table 7).

Table 7: *Effects of income position and change on the purchased quantity (ml) of liquor in 1<sup>st</sup> quarter 1995, by educational status of household head (b-coefficient, OLS regression)*

Predictors	household head has no occupational education	household head has occupational education
net mean	738*	804*
Improvement of income position	873*	255
Deterioration of income position	135	506*
Lowest income position	-266	-461*
Number of adults	162*	391*
Number of children	-117	-182*

note: \* means significant with  $p < 0.05$

Income elasticity of consumed quantity of liquor is greater in households with household head having a better occupational education. This might reflect the working of middle class norms of gratification delay and economic-rational organizing of household economics.

The liquor consumption amplifying effect of income improvement is found only in households, whose household head has a lower educational status, not if the household head is better trained. It seems that the symbolic meaning of liquors is only relevant in lower class households. They might try to translate a financial improvement into a social status improvement by consuming more high-status image associated alcoholic beverages, which are assigned on higher image of higher status.

The positive effect of income deterioration on the consumed quantity of liquors can only be found in households where the household head has an occupational education/training.



It seems that only in higher social class households income position losses are seen as serious status declines. So they are in more stressful situation, which pushes more for immediate relaxation and coping.

#### **4.2.3 Employment status of the household running person moderating income effects**

In most two couple families the household is run by the wife (especially if there are children present). So looking at the net income position effects on the consumption of liquor separately for households where the household keeper is fulltime employed versus part-time or not employed, the importance of family as a context of social meaning of alcoholics is getting clearer.

One can assume that households with little occupational activities (employment) by the housekeeper can be understood as households with traditional family orientation. The employment of the housewife indicates a family and marriage system with occupational orientation.

As the results in table 8 show, the type of the family systems seems to be relevant for social and psychological functions of liquor consumption.

If we would assume that the increase of consumed liquor after income position improvement reflects the social status function of alcoholics, one can find in table 8, that only in household with strong family orientation and less job orientation of the housekeeper alcoholics's social function works. As the social status functioning of alcoholics needs social events, their time and energy consuming organization can best be done in households who have the time and self-identity of being "in charge" of this task.

The increase of liquor consumption after income position deterioration can be understood as a kind of "compensating function" of alcohol consumption. This covariation, the compensation by alcohol consumption is observable only in household with a mixed family and job orientation of the household runner. Maybe this is due to the fact that mainly in these households a downgrading of the income position is experienced as stressful because job orientation is an important characteristic of life.

There is also no significant elasticity response, which in this case means no difference between lowest and higher income position households. Wealthier households with family life do not increase their liquor consumption even if they could afford.

Table 8: *Effects of income position and change on purchased quantity of liquors in 1<sup>st</sup> quarter 1995 (in ml), by employment status of housekeeping person.*

model predictors	housekeeping person full-time employed	housekeeping person half-day employed	housekeeping person less than half-day employed	housekeeping person not employed
net mean	213*	423*	806*	1086*
improvement of income position	147	-603	2389*	423*
deterioration of income position	1008*	990*	-135	209
lowes income position	-265	-871	-123	-645*
number of adults	489*	357*	366*	267*
Number of children	310*	116*	-406*	-242*

note: \* means significant with  $p < 0.05$

#### 4.2.3 Having children in the household as moderator of income level and income position change effects on the quantity of liquor consumption

Looking at the income effects on liquor consumption separately for households with versus without children under 18 years, one gets further support for the hypothesis on the crucial meaning of family orientation for the consumption of liquors (see table 9).

The results support the hypothesis, that households with children can be seen as households with family orientation and structure, which makes demand for liquor less strong. Overall these households show a reduced quantity of liquor purchases: further there is no change of liquor consumption after improvement nor after deterioration of the household's net income level.

The social status function as well as the compensation function of alcohol consumption seem to be not working in this family household context.

Table 9: *Effects of income position and change on purchased quantity (ml) of liquors in 1<sup>st</sup> quarter 1995, by household type (b coefficients, OLS regression)*

	household with children (aged less 18 years)	household without children (aged less 18 years)
net mean	627*	773*
improvement of income position	38	582*
deterioration of income position	-4	594*
lowes income position	-163	-470*
number of adults	251	360*
number of children	-68	-

note: \* means significant with  $p < 0.05$

## 6. Summary and conclusions

The main empirical results of this exploratory study are:

- a) households adapt their quantity of liquor consumption to their financial situation: high income restrictions lead to a restriction of liquor consumption;
- b) significant improvements in economic standing lead to short-term increases of alcohol consumption;
- c) significant deteriorations in income level lead also to short-term increases of alcoholic beverages demand.

There are several lines of argumentation supporting the assumption that consumption of alcoholic beverages is not only governed by economic/financial principles, but also by the social and psychological meaning of offering and consuming liquor products. The above results seem to support these hypotheses, they also showed that the sepearate estimation of social status meaning and compensatory meaning of alcoholic beverages is possible.

The exploratory hypothesis were then supported further by looking for moderating effects of the socio-demographic households structure.

The main moderating results are:

- a) if there are children present in the household, liquor consumption is reduced; also the effectiveness of alcoholic beverages consumption is weakened in meaning of social status and compensation .
- b) if the person running household is not or only marginally employed the compensation effect of liquor consumption is reduced, but the social status effect is more effective. The income level restriction is only effective in these households with one money earner.

These results are hints for thinking on the family context as an important context for alcohol consumption. It seems that family foundations of a household's everyday life reduces its consumption of alcohol. In households with more integration into to employment and less family activity spheres there is increased liquor consumption because of less financial constraints and more compensatory needs. Whereas financial restrictions and family integration reduces alcohol purchases, but on the other side pushes alcohol consumption because of its social integration function.

- c) If the household head has a higher educational status (having some formal training) then the income elasticity as well the compensation effect are working. The social status effect of income improvement only works in households with lower educational status.

This result seem to point to social class context of liquor purchases, where not only financial opportunities are the differentiating criteria of class differences but also attitudes and cognitive evaluative regulations of behavior. Middle class culture, possibly indicated by the educational level of the household head, seems to be the context the working of economic rationality. But on the otherside middle class status is more dependent on status consistency therefore reacting more sensible (by short-term increases of alcohol consumption) towards status losses.

The theoretical lines of thinking acknowledge several social, economic and psychological determinants of normal, everyday life consumption of liquors. To analyse, not to say to test the hypotheses, of the theoretical concepts, one needs data representing the ongoing consumption process as well as the process of household stability and change. Especially the measurement of the consumption process of liquors is complex and difficult to operationalize. The best way would be to take the data produced by consumption or purchase process itself. Process produced data are the most adequate data for measuring the process. The present analysis was based on that kind of data, and therefore was successful in precise analysis of different alcohol consumption functions. The results were that clear,

that they are more less straightforwardly interpretable in terms of social, economic and psychic functions. Of course these concepts, basically relied on subjective attitudes and perceptions, were not operationalized in the consumer panel study. This must be done for an adequate hypothesis testing, as well as there is more precise information needed on the social and demographic changes of the household structure over time. Having given the highly costly and effortful process produced data measurement of a household's consumption process as the core of the data collection programm, there would be a big progress in researching the everyday life contexts of alcohol consumption and thereby the societal defined everyday life contexts of possible pathological alcohol consumption.

## Contact

*Georgios Papastefanou*

*Zentrum für Umfragen, Methoden und Analysen (ZUMA) e.V.*

*B 2, 2*

*68159 Mannheim*

*Tel.: 0621 – 1246-278*

*Fax. 0621 – 1246-100*

*papastefanou@zuma-mannheim.de*

## References

Andreß, H., 1999: *Leben in Armut: Analysen der Verhaltensweisen armer Haushalte mit Umfragedaten*. Opladen: Westdt. Verl.

Antons, K./Schulz, W., 1990: *Normales Trinken und Suchtentwicklung*. Göttingen: Verlag für Psychologie.

Arab, L., 1981: *Ernährung und Gesundheit. Eine Untersuchung bei jungen Frauen und Männern in Heidelberg*. Basel: Verlag S. Karger.

Barthels, M., 1991: *Subjektive Theorien über Alkoholismus: Versuch einer verstehend-erklärenden Psychologie des Alkohols*. Münster: Aschendorff.

Becker, G S., 1996: *Accounting for tastes*. Harvard University Press.

Belk, R.W., 1982: *Acquiring, Possessing and Collecting: Fundamental Processes in Consumer Behavior*. In: Bush, R.F., Hunt, S.G. (Eds.): *Marketing Theory: Philosophy of Science Perspectives*. Chicago: American Marketing Association.

Belk, R.W., 1984: *Cultural and Historical Differences in Concepts of Self and their Effects on Attitudes Toward Having and Giving*. In: *Advances in Consumer Research* 11. Provo/UT. Association for Consumer Research.

Bourdieu, P., 1982: *Die feinen Unterschiede*. Frankfurt a.M.: Suhrkamp.

Bundesministerium für Gesundheit [Bonn], 1991: *Repräsentativerhebung 1990 zum Konsum und Mißbrauch von illegalen Drogen, alkoholischen Getränken, Medikamenten und Tabakwaren.*

Cockerham, W.: *Social stratification and health lifestyles in two systems of health care delivery: a comparison of the united states and west germany.* In: *Journal of health and social behavior*, 29: 113-126.

Conrady, R., 1990: *Die Motivation zur Selbstdarstellung und ihre Relevanz für das Konsumentenverhalten. Eine theoretische und empirische Analyse.* Frankfurt/Main, Bern, New York, Paris: Lang.

Dittmar, H., 1992: *The social psychology of material possessions: to have is to be.* Hemel Hempstead: Harvester Wheatsheaf.

Edwards, G., 1997: *Alkohol und Gemeinwohl. Strategien zur Reduzierung des schädlichen Gebrauchs in der Bevölkerung.* Stuttgart: Enke.

Eggen, B., 1998: *Privathaushalte mit Niedrigeinkommen. Familienwissenschaftliche Forschungsstelle im Statistischen Landesamt Baden-Württemberg.* Baden-Baden: Nomos Verl.-Ges.

Eifler, S., 1997: *Einflußfaktoren von Alkoholkonsum: Sozialisation, Self-Control und differenzielles Lernen.* Wiesbaden: Dt. Univ.-Verl.

Grant, M., 1983: *Economics and alcohol: consumption and controls.* London: Croom Helm.

Haseloff, O.W., 1992: Personale und soziale Funktionen des privaten Verbrauchs. In: Rosenberger, G. (Hg.): *Konsum 2000. Veränderungen im Verbraucheralltag.* Frankfurt New York: Campus.

Härtel, U.: *Der Einfluß von Ausbildung und beruflicher Position auf Veränderungen im Zigarettenrauchen und Alkoholkonsum: Ergebnisse der MONICA-Augsburg Kohortenstudie.* In: *Sozial und Präventivmedizin* 38: 133-41

Hermann-Kunz, E., 1995: Ernährung und relative Armut in den neuen Bundesländern. In: *Ernährung in der Armut. Gesundheitliche, soziale und kulturelle Folgen in der BRD.* Berlin : Ed. Sigma.

Hoeltz, J.: *Subjektive Morbidität, Gesundheitsrisiken, Inanspruchnahme von Gesundheitsleistungen.* Gesundheitsberichtserstattung auf Basis des 1. Nationalen Gesundheitssurveys, 2 Bände. München: Infratest Gesundheitsforschung

Kirschner, W. /Meinlschmidt, G.: Gesundheitliche Risikoexposition, Gesundheitsverhalten und subjektive Morbidität von West- und Ostberlinern insgesamt und nach sozialer Schichtzugehörigkeit. In: *Krankheit und soziale Ungleichheit. Ergebnisse der sozialepidemiologischen Forschung in Deutschland:* 253-266

Klein, M., 1992: *Klassifikation von Alkoholikern durch Persönlichkeits- und Suchtmerkmale.* Bonn: Nagel.

- 
- Kluthe, R./Kasper, H., 1998: *Alkoholische Getränke und Ernährungsmedizin*. Stuttgart; New York: Thieme.
- Köhler, B. et.al., 1997: *Poverty and food in welfare societies*. Berlin: Ed. Sigma.
- Kußmaul, B. et.al., 1995: Zusammenhang zwischen Ernährungsverhalten und Bildungsstand: Ergebnisse der Ernährungserhebung 1984/85 des MONICA-Projektes Augsburg. In: *Zeitschrift für Ernährungswissenschaft* Bd.34: 177-182.
- McCracken, G., 1992: Die Geschichte des Konsums: ein Literaturüberblick und Leseführer. In: Rosenberg, G (Hg.): *Konsum 2000. Veränderungen im Verbraucherverhalten*. Frankfurt New York: Campus.
- McGuire, W.J., 1976: Some Internal Psychological Factors Influencing Consumer Choice. *Journal of Consumer Research* 2.
- Müller, W./Marißen, N., 1989: *Leben in Armut*. Hrsg. vom Statist. Amt d. Landeshauptstadt Stuttgart. Stuttgart.
- Nicosia, F. M., 1966: *Consumer decision processes: Marketing and advertising implications*. Englewood Cliffs, NJ: Prentice-Hall
- Pohl, H., 1975: *Problematischer Alkoholkonsum und seine unmittelbaren Kovariablen*. Diss. München.
- Solomon, M., 1983: The role of products as social stimuli: A symbolic interactionism perspective. *Journal of Consumer Research* 10: 319-329
- Thiel/Thai, 1995: *Armut und Ernährung in den neuen Bundesländern. Auswertung von Ernährungsprotokollen*. In: Ernährung in der Armut. Gesundheitliche, soziale und kulturelle Folgen in der BRD. Berlin: Ed. Sigma
- Ward, S., 1974: *Consumer Socialization*. *Journal of Consumer Research* 1.
- Weick, S., 1999: *Relative Einkommensarmut bei Kindern. Untersuchungen zu Lebensbedingungen und Lebensqualität in Deutschland von 1984 bis 1996*. Dissertation.

# CONSUMER PANEL DATA AND RATIONAL CHOICE BASED THEORIES OF MYOPIC HABIT FORMATION. AN EMPIRICAL ANALYSIS

*ROGER BERGER*

## 1. Introduction

No one doubts the existence of habits. We can observe habits daily in the behaviour. Addictions (i.e., particularly strong habits) cause many social problems. In commercial market research a household's last purchase is used as a predictor for the next purchase, assuming some sort of inertia in consumer behaviour.

There are some theories about the formation and persistence of habits. They usually ascribe three main properties to habits:

- Habits are the stronger the longer they already exists. This is the case, e. g., with old people, who often have habits that seem inappropriate today. But the history of successful practices in the past let the habit persist.
- Habits are the stronger the more frequent the behaviour is repeated. We observe this in everyday life. Habits exist very often with behaviour that is repeated frequently, e. g., having breakfast at eight'o'clock, eating the same flavour of yogurt, having a cup of coffee at ten, and so on.
- Habits simplify specific actions, economize on resources, and reduce complexity. For example having found once a sufficiently tasty and low-priced yogurt, we will not search for every new purchase the whole market for yogurt again.<sup>1</sup>

These three aspects of habits correspond to the following three questions that a theory of habit formation should be able to answer to provide a satisfying theoretical model.

---

<sup>1</sup> Carrying on this last point it is often assumed that habits are characterised as unconscious and automatic behaviour, say reflexes. This phenomenon, e. g., is known to every driving learner. While in the beginning, steering, changing gears, and braking takes the whole attention (and sometimes even more), after some practice one is able to do these things automatically (and, e. g., to talk with a passenger simultaneously).



- Under what conditions does a habit develop?
- How does a habit persist?
- Under what conditions is a habit given up?

Especially the first and the last question are of crucial interest. While it is not too difficult to give reasons for the persistence of habits, it is not trivial to determine endogenously how a habit is started, or how it is given up.<sup>2</sup>

Based on these three questions, theories of habit formation roughly can be divided into two classes. One type especially makes use of the first two properties of habits and therefore models the role of time in habit formation. These theories are well suited to answer the first and last question. I call these dynamical theories of habit formation. The other type is based on the third quality presented and therefore considers a habit as a static concept.<sup>3</sup>

In this article I will deal with dynamic theories of habit formation for the following reasons: First, it is obvious to analyse habits as dynamic phenomena. Second, it can be shown that static theories have great problems in explaining habit formation in a theoretically consistent way.<sup>4</sup> Third, consumer panel data are very well suited for the empirical test of dynamic habit formation theories.

Hence, the goal of this analysis is to test empirically the basic conclusions of dynamic habit formation theories.

## **2. Theory: Dynamic models of habit formation**

Dynamic theories of habit formation may be classified into two types. One is grounded in psychological research about behaviour in time.<sup>5</sup> I will concentrate on the other type of habit formation theory, the Rational Choice Models. The latter are based on neoclassical economics. In short their common assumptions can be described as follows: Actors

---

<sup>2</sup> The reader may have noticed that classical behaviouristic learning theories are occupied exactly with the same three questions. This is not accidentally for these theories can be looked upon as theories of habit formation, too. (see below and especially footnote 5).

<sup>3</sup> An example are framing theories as presented by Esser (1990).

<sup>4</sup> See, e.g., Etzrodt (2000).

<sup>5</sup> The best known example of these theories is the so called matching-law (see e.g. Loewenstein and Elster, 1992).

maximize their utility under the given restrictions.<sup>6</sup> To model habit formation according to these principles specific elements are introduced into the utility function.

## 2.1 Rational choice models of myopic habit formation

Several authors present models of rational habit formation. Most of them are very sophisticated models (see e. g. Muellbauer 1988, Orphanides and Zervos 1998, Spinnewyn 1976) and draw on advanced mathematical analysis. I will not introduce these models in detail, because their formal analyses do not add much. Instead we will just present the basic elements of the models.

Dynamic models, and therefore also models of rational habit formation make two main assumptions. The strength of a habit is determined by (a) its frequency and (b) its duration.<sup>7</sup>

In order to define the relation between frequency and duration and their respective effects on habit formation it is assumed that all specific actions (like eating yogurt, spending vacancies in a special place, driving a car etc.) build up a capital stock. Hence this stock becomes bigger the more often the action has been repeated, and the longer the habit persists. In order to take into account the frequency of habit-repetition the habit stock is discounted with the time lag between two consecutive habitual actions. This aspect is based on the following argument: The larger the time interval since the habitual action has occurred for the last time, the smaller is its effect on the present action. In fact the habit stock is modelled analogous to an investment stock that is depreciated, and refreshed by new investments. Formally this can be put as follows:

$$\dot{S}(t) = -\tau S(t) + c(t) \quad \text{with } S(0) = 0 \quad (1)$$

where  $S(t)$  is the habit stock at time  $t$ ,  $c(t)$  is consumption of a habitual good at time  $t$ , and  $\tau$  is the depreciation rate of the habit stock. This habit stock is introduced into a standard economic utility function as follows: If  $U_c$  denotes marginal utility with respect to  $c$  then

$$U_c = f(p_c, c, y, S_c) \quad (2)$$

---

6 I will not point out the theory. Because the crucial points for our analysis can easily be caught by intuition. For an introduction see e. g. Becker (1976).

7 Actually there is no need to explain these assumptions in detail. If we look at the model as an as-if explanation any assumption can be made as long as the resulting hypothesis prove to be empirically correct. In addition the assumptions seem very plausible. Yet if we like these assumption can be grounded on e. g. neurophysiological and psychological theories (see e. g. Leshner (1997)).

This means that the marginal utility is a function of the consumption  $C$ , the relative price  $p_c$ , the income  $y$ , and of the habit stock  $S_c$ .

We neglect a further formal discussion of the utility function and assumptions, but rely exclusively on an intuitive interpretation which will be accessible to most readers. <sup>8</sup>Formal analysis yields, among other things, the following implications: To consume more of a distinct good is always better than to consume less. An increase of the relative price decreases the amount of consumption, while an increase of income will raise consumption. Of course, these statements are standard in economic consumer theory.

The crucial relationship that is to be tested empirically here reads as follows: An increase of the habit stock for a distinct good raises consumption of that good.<sup>9</sup>

In the following we speak of myopic habits because the actor just takes into account his past actions (i.e., consumption acts). He is myopic - not looking into the future - and does not calculate the future effects of his present behaviour. If he would act with taking into account the future effects of current behaviour, we would speak of a fully rational actor and therefore of a fully rational model of habit formation. Then the utility function would contain the future consumption and future price or its expectation of the actor. This extended version of rational habit formation is not analysed here. We restrict our attention to the model to the myopic maximizing actor.

For the interpretation of the model the following aspects are of particular importance:

- The habit stock can be interpreted as an endogenous change of preferences in time.<sup>10</sup>
- The rationality of the model is defined by the maximizing behaviour of the actor, i. e., the household makes choices *as if* he would maximize its consumption. This is the crucial difference to other models of habit formation (especially static models) which assume that habitual actors do not choose or maximize anything but just repeat past behaviour.

---

<sup>8</sup> The problem here is one of dynamical maximization under the given restriction with respect to the consumption. For a broader and exact discussion see, for example, Becker and Murphy (1988)

<sup>9</sup> By that it is not stated that there are no other influences on the actors utility. Of course there are, but first the above stated relations are used routinely in utility theory, second their theoretical effect is clear, undoubted and very well confirmed, third these variables are all easy to measure and included in the used data sample and fourth the model is kept easy to survey in order to determine exactly the effects of habit.

<sup>10</sup> But note that this is not necessary. Becker (see e. g. 1996) does not interpret habits as a change in preferences. Nevertheless this interpretation neither changes anything in the formal treatment nor in the empirical results.

- Therefore the above theoretical model can answer all three questions that were presented in the introduction as desiderata of any habit formation theory. Any actor forms a habit and therefore changes his preferences, as soon as he has chosen the same action more than once. Of course this is the case with almost every action. Therefore habits are an important factor in the formation of preferences. The habit will be maintained as long as the changes in the restrictions, i. e. relative price increases and decreases of income do not outweigh the additional utility by the habit. Therefore a habit will be given up if the increase in relative price of an action and/or the decrease in income is larger than the additional utility from maintaining the action.
- This provides us with a theoretically grounded relation between past and present actions or consumption acts. One may object that there is no need to produce a complicated theoretical model, just to state the more or less trivial assumption that the present action is somehow dependent on the past. This is not correct for the following reasons:

The first counterargument refers to the formulation of "somehow". In contrast to a purely exploratory procedure, which is often pursued in commercial market research, an explicitly formulated theoretical model provides us with several instruments to forecast future behaviour. The use of standard economic utility theory allows us, e. g., to calculate stock elasticities of demand, which are powerful instruments for predicting future market developments.<sup>11</sup> Second, from a scientific point of view, it is preferable to specify precisely how time works on habit formation and to embed this into an explicit action theory, instead of just stating some relation between past and present action.<sup>12</sup> Third, consumer panel research provides us with very rich data. To make use of this amount of information it seems meaningful to use an appropriate theory.

To test the theory we state the following testable hypotheses:

*H<sub>1</sub>: The more often a habitual action is chosen and the longer this habit already persists, i. e., the higher the habit stock S, the higher will be the chosen quantity of consumption c.*

This is our main hypothesis. The following two hypothesis are well confirmed hypotheses of classic economic demand theory. Actually, they are just used to control for these standard variables.

---

<sup>11</sup> We will analyse this procedure further on in the empirical section.

<sup>12</sup> Besides, very often in time series analysis this point is also neglected and theory building is replaced by pure empirical work.

*H<sub>2</sub>: The higher the relative price of consumption  $c$ , the smaller will be the chosen quantity of  $c$ .*

*H<sub>3</sub>: The higher the income  $y$ , the higher will be the chosen quantity of action  $c$ .*

In the next section the hypotheses are put to a test.

### **3. Data, Analysis, Results**

#### **3.1 Data**

In using the available ZUMA data set of the GfK consumer panel data we take those variables that are most exactly measured and provide information on a true interval level. This is the amount of a commodity purchased, its price, the time periods between two consecutive purchases, and the self-estimated income. Therefore, validity and reliability of the data are hardly in question.

#### **3.2 Operationalization**

The theory of habit formation is formulated above in a general form, so that it includes all sorts of consumptions and costs. To test the theory we have to define a specific set of the variables.

The actor is one household or, more precisely, the person in the household which makes the household purchases. However the habit to be analysed actually is one of a household and not of a single actor. If one wants to avoid this assumption one could select just one-person-households, where actor and household coincide. Yet a short test has shown that these estimation do not differ from those using all the households. We will assume that the whole amount of the good will be consumed between two consecutive purchases in time. This means that there is no storage of the good and no giving away. Furthermore we assume that the whole amount of the good is consumed on the day it is purchased. In the absence of further information about the exact consumption pattern this seems to be an adequate approximation.<sup>13</sup>

---

<sup>13</sup> This operationalization of actor and action follows from the assumption that the consumption of a good is the benefiting action. If one wants to avoid this assumptions without any theoretical changes the benefit of the action can be ascribed to the purchase of a good, and not the consumption. For commercial market research this seems even more appropriate, because there the explanandum is the purchase and not the consumption.

As equation (1) shows, it is necessary to determine a depreciation factor  $\tau$  to calculate the habit stock. In the absence of empirical knowledge and of empirical knowledge and theoretical guidelines we have to rely on some plausible guesses to determine the size of  $\tau$ .

Estimations are done with five different factors corresponding to a half-life period of the habit stock of one day, one week, one month, half a year and one year. A half-period of a week means that after one week the habit stock has diminished to the half and so has its habit-generating effect. Common sense tells us that the true value of  $\tau$  probably will lie between a half-life period of a week and half a year. A half-life period of a day or year represent the outer bounds of  $\tau$  which probably will not be crossed.

At last we have to choose distinct products out of the list of all possible goods provided by the panel. Due to reasons of practicability not all possible products can be taken for estimation. Three products are chosen, namely yogurt, roasted coffee, and beer<sup>14</sup>. By choosing these three products we try to cover another property of habits that has not been mentioned yet, but is obvious: Not every good or action forms habit of the same strength. Of course we will expect beer, or the alcohol in it, to generate stronger habits, or even addiction, than yogurt. For coffee, respectively caffeine, we would expect a habit strength in between the habit strength of those two commodities. Note that all three commodities can be expected to be purchased frequently. This aspect is important because the observation period is limited to one year. Because we expect the habit to stem from repeated consumption, it is necessary to observe as much purchases as possible to estimate the habitual effect of it.

### 3.3 Functional specification

To estimate empirically the functional specification must be determined. As it is often the case there are only few theoretical guidelines about the functional specification of the estimation equation. However theory (e.g. Becker and Murphy 1988) tells us that there is a linear relation between the size of the stock and the amount of the good consumed. We will choose a logarithmic specification for price and income for the following reasons. First, this is the function routinely chosen in econometrics and has been proven to be a good estimator in most cases. Secondly, it seems plausible that not absolute values of price and income have the strongest effect on consumption but differences. Third, at least distributions of price of yogurt and coffee are slightly right hand skewed and therefore will better fit the assumptions of OLS-estimation when a logarithmic transformation is used. The same argument holds for income. Moreover, metric income is calculated by

---

14 Corresponding to the product categories number 12 (coffee), 33 (beer) and 78 (yogurt).

replacing ordinal group numbers by the means of the group range. Hence the last group is an open category and the respective mean tend to underestimate the true values and distorts the OLS-regression. By raising the income to the logarithm this distortion is diminished. And fourth, estimating the logarithmic specification will generate directly the respective elasticities of demand that are very useful for interpretation.

This leads to the following estimating equation:

$$c = e^{\alpha + \beta S_c} p_c^\gamma y^\delta \quad (4)$$

or in the linearized transformation

$$\ln(c) = \alpha + \beta S_c + \gamma \ln(p_c) + \delta \ln(y) \quad (5)$$

This last expression can directly be estimated with standard OLS-estimation.

### 3.4 Results

To show the effect of the habit stock  $S_c$  the model is estimated in two steps. In the first step we just estimate a bivariate regression with  $c$  and  $S_c$ , i. e., e following equation

$$\ln(c) = \alpha + \beta S_c \quad (6)$$

As table 1 shows, the main hypothesis (H<sub>1</sub>) is confirmed. The habit stock has a positive effect on the quantity consumed independently of the good and of the discount factor  $\tau$  (i.e., the corresponding half-life period)<sup>15</sup>.

Moreover the coefficients are fairly robust, no matter which depreciation rate is chosen.

Therefore the depreciation rate can be considered as a multiplying factor, which does not change the relative weights in the regression.

Besides, one month seems to be the best approximation for all three commodities for the half-life period of the habit stock.

By estimating equation (4) we take hypothesis (H<sub>2</sub>) and (H<sub>3</sub>) into account and control this relationship for price and income. This yields the following results (see table 2a to 2c).

---

<sup>15</sup> The depreciation factors corresponding to a half-life period of one day and one year are not presented in order to keep the tables easy to survey and because, as expected, they generate the worst results.

Table 1: OLS-regression to logarithmic quantity of consumed good (for beer, coffee and yogurt) on habit stock for depreciation factors corresponding to the half-life periods of habit stock of one week, one month and half a year.

	half-life period	b	std. error	t-value	R <sup>2</sup>	n <sup>16</sup>
<b>Beer</b>	one week	2.241 10 <sup>-5</sup>	3.102 10 <sup>-7</sup>	72.2	0.043	117153
	one month	8.084 10 <sup>-6</sup>	7.962 10 <sup>-8</sup>	101.5	0.077	123601
	half a year	2.002 10 <sup>-6</sup>	2.164 10 <sup>-8</sup>	92.5	0.064	123760
<b>Coffee</b>	one week	1.630 10 <sup>-4</sup>	2.620 10 <sup>-6</sup>	62.2	0.03	127058
	one month	6.461 10 <sup>-5</sup>	8.739 10 <sup>-7</sup>	73.9	0.04	134624
	half a year	1.799 10 <sup>-5</sup>	2.486 10 <sup>-7</sup>	72,4	0.037	134737
<b>Yogurt</b>	one week	1.315 10 <sup>-4</sup>	1.492 10 <sup>-6</sup>	88,2	0.065	112603
	one month	4.482 10 <sup>-5</sup>	4.101 10 <sup>-7</sup>	109,3	0.094	115196
	half a year	1.179 10 <sup>-5</sup>	1.215 10 <sup>-7</sup>	97.1	0.076	115241

Again all three hypothesis can be fully confirmed. Independently of the specification, all variables show always the correct sign. Higher habit stocks still lead to increased demand for the good. The standardised coefficients are slightly decreased by the additional variables but show about the same size. As expected, higher income raises the demand for the good, while higher prices reduce demand. Estimators for price and income are also robust and not affected by changes in the depreciation rate. One month still seems to be the appropriate half-life period of the habit stock.

**16** The number of cases is slightly changing for the different depreciation rates. This stems from the following estimation procedure: All cases that have a habit stock equalling zero are excluded from the estimation. The reason is simple: If we accept the most probable assumption that in each household the product has been consumed already once before the observation period (1995) a stock value of zero is in any case a wrong estimator, because the stock cannot become zero once it was higher. And there is an additional problem. Rounding procedures of the computing program (SPSS) now lead to zero values of the habit stock when this is actually not the case. Imagine for example a household that consumes one bottle of beer on January first 1995 and one bottle on New Year's Eve 1995. Then the habit stock of beer for this specific household is zero for all depreciation rates on January first, because of the lack of past observation to calculate a stock. Assuming a half-life period of a week at New Year's Eve the beer stock will be approximately zero (to be exact  $0.5^{52}$ ). Even for a half-life period of a month the stock will hardly differ from zero ( $S=0.5^{12}$ ), but for a half-life period of a half year the stock will be 0.25.

Taking zero-values to into the estimation does not alter the parameters of the regression except for a smaller explained variance (see also section 3.5 for the problem).



Table 2a: OLS-regression of logarithmic quantity of consumed beer for depreciation factors corresponding to the half-life periods of habit stock of one week, one month and half a year on the habit stock, logarithmic price and the logarithmic income.

Beer	half-life period	std. b	std. error	t-value	adj. R <sup>2</sup>	N
habit stock S		0.118	3.880 10 <sup>-7</sup>	48.0	0.094	60720
price p	one week	-0.195	0.014	-35.0	0.094	60720
income y		0.195	0.009	51.9	0.094	60720
habit stock S		0.243	9.870 10 <sup>-8</sup>	65.1	0.119	63984
price p	one month	-0.118	0.0136	-31.5	0.119	63984
income y		0.202	0.008	54.2	0.119	63984
habit stock S		0.217	2.616 10 <sup>-8</sup>	57.9	0,107	64077
price p	half a year	-0.125	0.014	-33.4	0,107	64077
income y		0.204	0.008	54.5	0,107	64077

Table 2b: OLS-regression of logarithmic quantity of consumed coffee for depreciation factors corresponding to the half-life periods of habit stock of one week, one month and half a year on the habit stock, logarithmic price and the logarithmic income.

Coffee	half-life period	std. b	std error	t-value	adj. R <sup>2</sup>	N
habit stock S		0.152	3.664 10 <sup>-6</sup>	40.6	0,105	64713
price p	one week	-0.273	0.010	73.4	0,105	64713
income y		0.054	0.004	14.6	0,105	64713
habit stock S		0.160	1.260 10 <sup>-6</sup>	44.1	0.111	68581
price p	one month	-0.270	0.010	-74.5	0.111	68581
income y		0.053	0.004	14.7	0.111	68581
habit stock S		0.137	3.644 10 <sup>-7</sup>	37.1	0.103	68637
price p	half a year	-0.260	0.010	-70.8	0.103	68637
income y		0.056	0.004	15.5	0.103	68637

Table 2c: OLS-regression of logarithmic quantity of consumed yogurt for depreciation factors corresponding to the half-life periods of habit stock of one week, one month and half a year on the habit stock, logarithmic price and the logarithmic income.<sup>17</sup>

Yogurt	half-life period	std. b	std. error	t-value	adj. R <sup>2</sup>	N
habit stock S	one week	0.255	1.471 10 <sup>-6</sup>	89.7	0,091	112603
price p		-	0.006	-57.0	0,091	112603
		0.162				
habit stock S	one month	0.303	4.049 10 <sup>-7</sup>	109.3	0.117	115196
price p		-	0.005	-55.1	0.117	115196
		0.152				
habit stock S	half a year	0.271	1.200 10 <sup>-7</sup>	96.8	0,099	115241
price p		-	0.005	-54.6	0,099	115241
		0.153				

Just looking at the best estimation, we see that for explaining the chosen quantity of  $C$  the habit stock statistically is at least as important as is price. Nevertheless we should not stress this fact to much. The reason is the small variance of price especially for beer and yogurt. For these two commodities prices are about the same for a wide range of different brands. By raising the price to logarithm we even reduces range. This could be a reason for the relatively small proportion of explained variance by beer and yogurt prices. This view is confirmed by the estimation of the demand for coffee. There, the price range is much bigger and so a greater proportion of variance is explained by price.<sup>18</sup>

Yet standardised b is a difficult to interpret measure for the effect of habits on demand. But we can calculate the respective elasticities of demand for all three explaining variables. Because we used the double-logarithmic functional specification for price and income the respective elasticities of demand just correspond to the unstandardised b coefficients. For the habit stock we applied a logarithmic-linear functional specification

<sup>17</sup> The variable "income" is not available for all households in the consumer panel. And it is totally missing yogurt and partly for beer and coffee. Hence for yogurt the estimation had to be done without the variable. Due to missing income data the number of cases for the analysis of beer and coffee consumption is greatly reduced.

<sup>18</sup> A test with alcoholic beverages ("spirits") for which price vary in a wide range also confirms this explanation. Though no one would doubt the strongly addictive character of alcoholic beverages, most of the demand is explained not by the habit stock but by the price of spirits.

so that the elasticity is not constant for every size of the stock. As a consequence elasticity increases linear with habit stock and so does the reaction of the actor to changes in the stock. <sup>19</sup>

For matters of simplicity the results are just shown for the half-life period of the habit stock of one month (see table 3a to 3c).

*Table 3a: Elasticities of beer demand for habit stock, price and income (half life period of one month).*

<b>Beer</b>	mean	median	Mode
habit stock S	0.206	0.144	0.064
price p	-0.428	-0.428	-0.428
income y	0.453	0.453	0.453

*Table 3b: Elasticities of coffee demand for habit stock, price and income (half life period of one month).*

<b>Coffee</b>	mean	median	Mode
habit stock S	0.100	0.056	0.028
price p	-0.74	-0.74	-0.74
income y	0.055	0.055	0.055

*Table 3c: Elasticities of yogurt demand for habit stock, price and income (half life period of one month).*

<b>Yogurt</b>	mean	median	Mode
Habit stock S	0.218	0.155	0.013
Price p	-0.297	-0.297	-0.297

<sup>19</sup> The elasticity is defined as follows:  $\mathcal{E} = \beta S$  (see e. g. Braun, et al 2001).

For the chosen functional specification the respective elasticities of demand for price and income are constants. For all three commodities price and income elasticity of demand is below unity. In standard economic terms this is called inelastic demand and the goods are necessary goods. Nevertheless, again one should be careful with overinterpreting this results. First, as stated above, the variance of the price for beer and yogurt is very small. Therefore the reaction of demand on price variety is confined to a small range which leads to underestimation of price elasticity. Second, we did not spend to much effort on finding the true functional specification. Though we have good reason to choose a double logarithmic, this is not necessarily the true specification. Further analysis eventually would lead to different functional specifications and hence to different elasticities. The same argument is true for income elasticity.

The opposite situation is given for the stock demand elasticity of demand. There, variance is very high, and we have theoretical reasons for the chosen logarithmic-linear functional specification. Hence we can conclude the following: The habit stock elasticity of demand for beer is a linear increasing function with an extremely flat slope of  $6.423 \cdot 10^{-6}$  (coffee:  $5.555 \cdot 10^{-5}$ , yogurt:  $4.427 \cdot 10^{-5}$ ).<sup>20</sup>

The habit stock elasticities of demand of all there goods are fairly small and lie distinctly below price elasticity of demand. This can be interpreted as follows: The habits of consuming beer, coffee and yogurt seem to be just slightly addictive. A one percent increase of the stock<sup>21</sup> generates an increase of demand in the next period of 0.084% (coffee) to 0.16% (yogurt). For heavy consumers this rate increases linearly in their respective stock. This means that frequent consumption for a long time leads to stronger habits. But as we would expect it, habits turn out to be fairly persistent, too. Splitting the stock into two halves after one month reduces demand just about 8% for yogurt (beer: 6.8%, coffee: 4.2%). We could therefore conclude that yogurt is the least addictive good, while coffee (in contrary to what we expected) is the most addictive of the three goods.<sup>22</sup>

---

<sup>20</sup> Again this numeric results are almost the same if we use another half-life period of stock.

<sup>21</sup> Translating this to the measures of everyday life yields the following results: For an average consumer (measured at the median level of the habit stocks) the consumption of a bottle of beer, (cup of coffee, a cup of yogurt), leads to an increase of demand in the next period of about 0.3% (0.14%, 0.8%). Note that the "ranking" of the three commodities is different from ranking of the elasticities. This happens because elasticities are unit free measures, while the above numbers reflect the unit of measurement of the respective commodity (i.e., centiliter for beer and gram for coffee and yogurt).

<sup>22</sup> All these statements are actually just true for the so called steady state. This state is reached when the system is in the equilibrium again after the disturbance of one or more parameter. This means that the actor will exactly consume the amount of a commodity that has been depreciated. Hence his habit stock will not change anymore. This steady state can - by definition - just be reached asymptotically.

### 3.5 A note on autocorrelation

One problem with estimating time-dependent equations is the appearance of autocorrelation. This leads to distorted estimators which are no longer efficient. Especially standard errors would be too small and hence, e. g., the adjusted determination coefficient too large. One may wonder if autocorrelation characterises the above estimations, especially when keeping in mind the very high t-values. To answer this question we first have to determine the structure of the model. Though we take time into account, it is not a time series model. Remember that information about time is used just for the calculation of the habit stock. Hence we use a true cross-section model where each consumption  $C$  at any point in time  $t$  is explained by its corresponding habit stock  $S_c$ . Note the difference between our estimation model and the often used so-called autoregressive model which regresses consumption  $C_t$  on the preceding consumption  $C_{t-1}$  (say:  $C_t = \alpha + \beta C_{t-1}$ ) without further theoretical modelling of the relation between time and chosen action. Such models suffer usually from a high degree of autocorrelation (and hence of the above described distortion in standard errors).<sup>23</sup>

A test (Durbin-Watson d statistic) shows a rather high degree of autocorrelation in our model. This means that we can not be sure about the efficiency of the equation estimators. However, what could cause autocorrelation in the model? To answer this question we have to reanalyse the structure of the predicting variables: Recall that all cases with a habit stock value of zero were excluded. The latter because this values are certainly bad predictors for any habit. The same argument holds for the respective consecutive habit stocks. Habit stocks with a short past are worse estimators than those with a long past. Assuming that habit stocks with a short past will be on average smaller than those with a long past, this will generate a bias between the size of the habit stock and the respective residuals will become smaller on average, the bigger the habit stock is.<sup>24</sup> However, an even higher degree of autocorrelation is discovered in the price residuals. This may stem from the rather low variance in prices. If we assume in addition also some habit of choosing the same prices, this may explain also the halo effect in prices (and generate autocorrelation in the respective residuals).

Fortunately, we don't have to stress these points too much. The tremendous number of cases allows us to test the validity of the estimations with a simple procedure. For that

---

<sup>23</sup> Note that this autocorrelation (or serial correlation, how it is often called then) is NOT indicated by the common measure for autocorrelation Durbin-Watson d statistic (see e. g. Gujarati 1995). Instead the Durbin-Watson h test should be used. This is difficult because this test is not done by most statistical programs.

<sup>24</sup> I owe this idea to Martin Abraham.

purpose we select just the last case of every actor - respectively household - in the sample. There the habit stock will be the best predictor available for the respective consumption because it has the longest past. This procedure reduces the number of cases to the number of households in the sample. But note that we do not lose information about consumption history of the households because that is taken into account by the habit stocks.

This procedure yields the following results (see table 4).

*Table 4: OLS-regression of logarithmic quantity of consumed good and price (for beer, coffee and yogurt) on habit stock, for depreciation factors corresponding to the half-life periods of habit stock of one month, just using the last case of every household.*

		std. b	std. error	t-value
<b>Beer</b>	habit stock S	0.293	$6.912 \cdot 10^{-7}$	18.7
	price p	-0.097	0.062	-6.0
	income y	0.18	0.034	11,3
adj. $R^2 = 0.145$		Durbin-Watson d = 2.004		n = 3429
<b>Coffee</b>	habit stock S	0.173	$6.319 \cdot 10^{-6}$	11.6
	price p	-0.238	0.043	-16.0
	income y	0.080	0.018	5.3
adj. $R^2 = 0.105$		Durbin-Watson d = 1.997		n = 4124
<b>Yogurt</b>	habit stock S	0.342	$3.272 \cdot 10^{-6}$	22.7
	price p	-0.13	0.030	-8.6
	adj. $R^2 = 0.136$		Durbin-Watson d = 2.075	

The Durbin-Watson d statistic shows practically no autocorrelation for these estimations. The signs of the parameters are unchanged and have still the correct directions. The b-parameters are about the same. In fact, the habit stock becomes even a little bit more important in these models. This is not astonishing if we remind us of the fact that these predictors are the best ones available. We can conclude that estimating with the full sample of observed consumption cases does not generate overestimation of model parameters or an artefact. Autocorrelation is not a problem of the model estimation - the results are valid and reliable.

## 4 Conclusion

The empirical test of myopic models of rational habit formation with consumer panel data shows that this theory seems to be an appropriate model for habit formation and persistence. Its main idea that frequency and duration of a habit determine its strength confirmed. Thereby this theory provides us with a theoretically founded and empirically tested relation between time and habit formation. In addition controlling for prices and income, the robustness of the model shows up. This means that the theory suggests appropriate answers to the question for the formation, persistence and breaking of habits. Additionally it allows us to calculate elasticities of the habit stock on demand. This provides us with a practical instrument for concrete estimation of the strength of habits.

It would be interesting to improve the empirical analysis in several ways. The observation period could be prolonged providing us with more cases that could be supposed to be near the theoretical steady state. For commercial applications it would be necessary to define more subtle habits. For example we did not distinguish between different flavours of yogurt, or different brands of beer. Undoubtedly this would lead to more sophisticated results which could be the base for rational and empirically grounded marketing decisions.

## Contact

Roger Berger  
Institute of Sociology  
University of Munich  
Konradstr. 6  
80801 München  
Tel. +49 (0)89 - 2180-3256  
Fax. +49 (0)89 - 2180-2922  
[roger.berger@soziologie.uni-muenchen.de](mailto:roger.berger@soziologie.uni-muenchen.de)

## References

- Becker, Gary S., 1976, *The Economic Approach to Human Behavior*. Chigago: The University Press of Chigago
- Becker, Gary S, Kevin M. Murphy, 1988, A Theory of Rational Addiction. In: *Journal of Political Economy*, 96 (4), 675-700
- Becker, Gary S., 1996, *Accounting for Tastes*. Cambridge: Harvard University Press

Braun, Norman, Bruno Nydegger-Lory, Roger Berger, Claudia Zahner, 2001, *Illegale Märkte für Heroin und Kokain*. Bern: Haupt Verlag

Esser, Hartmut, 1990, "Habits", "Frames" und "Rational Choice. In: *Zeitschrift für Soziologie*, 19 (4), 231-247

Etzrodt, Christian, 2000, Alfred Schütz - Ökonom und/oder Soziologe? Eine Kritik an Hartmut Essers Interpretation der Theorie von Alfred Schütz und an seiner "Definition der Situation". In: *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 52, 761-782

Gujarati; Damodar N., 1995, *Basic Econometrics (Third Edition)*. New York: McGraw-Hill

Leshner, Alan I., 1997, Addiction Is a Brain Disease, and It Matters. In: *Science*, 278, 45-47

Loewenstein, George, Jon Elster (ed.), 1992, *Choice over Time*. New York: Russell Sage

Muellbauer, John, 1988, Habits, Rationality and Myopia in the Life Cycle Consumption Function. In: *Annales d'économie et de statistique*, 9, 47- 70

Orphanides, Athanasios, David Zervos, 1998, Myopia And Addictive Behaviour. In: *The Economic Journal*, 108, 75-91

Spinnewyn, Frans, 1981, Rational Habit Formation. In: *European Economic Review*, 15, 91-109

References



---

# CONTINUITY OR DISCONTINUITY IN THE DECISION OF THE APPLICATION OF INCOME OF PRIVATE HOUSEHOLDS?

*UWE FACHINGER*

## 1. Introduction

### 1.1 Some preliminary remarks

Why are we interested in the continuity or discontinuity of income decisions due to the application of income of private households? The following preliminary remarks will hint at some relevant economic aspects.

In micro-economic theory it is generally assumed, that the behaviour of people is risk averse. To explain the distribution of earnings or wealth, this assumption is partly eased and different attitudes towards risk are admitted. One result of the softened assumption is, that on average households with a higher propensity in risk taking gain higher earnings. It is argued, that this causes the special shape of the distribution of earnings or wealth. In other words, this means, that more households with a higher risk taking propensity are to be found in higher earnings classes.

As the opportunity costs of a wrong decision for a household with higher earnings are *ceteris paribus* lower, higher earnings could lead to lower risk aversion in buying products. So risky decisions are *ceteris paribus* more likely in higher income classes<sup>1</sup>. In this context it has to be mentioned, that the consequence of *ceteris paribus* high opportunity costs of gathering information about the product<sup>2</sup> could also lead to brand loyalty.

To trace the risk aversity of private households, many ways could be taken, for example:

---

1 In the purchasing behaviour of households with lower income, opportunity costs of wrong decisions could lead for example to brand loyalty.

2 This depends on the characteristics of the special good, whether it is a neoclassical homogeneous good, for which these costs are very low, or not.

- direct consultation of people, asking about their risk behaviour as is done in the applied data base<sup>3</sup>,
- experiments by putting people in specific situations to examine their reactions<sup>4</sup>, or
- using surveys and identifying adequate indicators.

In economics the identification and analysis of the risk aversion of people is mainly based on experiments. But this is only one way to do so. Due to the existing longitudinal data sets, which are necessary for such analysis, the third mentioned possibility should more often be taken into account.

Another aspect belongs to the explanation of income distributions and empirical tests of the life-cycle theory. The shape of income distributions is always stable over time, but that does not mean, that the household income is constant over time. On the contrary, it could be shown in longitudinal income analyses, that a lot of income mobility exists<sup>5</sup>. Risk averters would prefer – given the same mean income – an income path with lower erratic fluctuations. Therefore, utility maximising would *ceteris paribus* lead to life-time income profiles with relative slight fluctuations. Hence it follows, that the life-cycle theory has a problem to deal with the observed income mobility. But income could be the wrong indicator: income decisions could be better planned out, because it is possible to smooth the fluctuations of actual income through dissolving savings. Another aspect is, that a lot of income decisions over short time periods are not free. The amount of spending money is fix e. g. for rents, credit redemption's or energy (exc. fuel). Therefore, the time-profiles of spending money may be better suitable as an indicator for testing the life-cycle theory.

A third aspect is characterized with the term "habit formation". Habit formation evolves over time and takes place with the ageing of the individuals. So, if habit formation is completed, the behaviour of people may be more continuous (and predictable) as in the case, where habit formation comes into being and therefore erratic behaviour occurs more often. As the latter may be more the case in younger ages, and it may be possible, that the behaviour of the elderly is more continuous.

---

<sup>3</sup> But this reflects mainly the opinion of the people about themselves. For a theory, which relies on the behavioural and normative beliefs on people, and its application see Ajzen and Fishbein, 1980.

<sup>4</sup> See for a short description and overview e. g. Lunt, 1995.

<sup>5</sup> See for example Atkinson et al., 1988, Atkinson et al., 1992, Fabig, 1999, Fachinger, 1991, Meier, 1983, Müller and Frick, 1997, Trede, 1997, and Zurbrügg, 1990.

## 1.2 Scope of the paper

The main purpose is to test, whether it is possible to show differences in risk taking, as it is reflected in the continuity of the purchasing behaviour on the basis of goods for daily requirements. The underlying idea is, to identify risk averters through their daily behaviour. The data on household expenditures should give the possibility of identifying different behaviour of risk taking on the basis of income decisions, especially those decisions concerning products of daily requirements<sup>6</sup>.

The paper is structured as follows. First, some comments are given on the used method. As Papastefanou did already describe the data set in general, in the following, comments are only given on special issues of the data set. After the short description of the data and method, the results are presented in two sections. The first one contains a description of purchasing behaviour on the basis of the product-limit estimation. Within this section, the continuity res. discontinuity of the purchasing behaviour is identified. In the following section, results of the proportional hazards model, which is used only as an explanatory tool, are presented. The paper concludes with a short summary and outlook.

## 2. Some comments on the method

Income decisions of households are processes over time and the purchase of a good is a time discrete event, which can occur at any point of time. To examine such changes over time it is necessary to use a method, which enables the handling of the dynamic of the process. Statistical methods for analysing these kinds of processes are summarized under the terms event history analysis or survival analysis. The event history analysis res. survival analysis is an appropriate statistical method especially for processes where at every given time a discrete change can occur.

In the event history analysis there are two main variables for the description of a process and which have to be operationalized: the state and the event<sup>7</sup>:

- State means in the context mentioned here, that a special product in the household exists.

---

<sup>6</sup> This approach follows a line, as it was stated for example by Griliches, 1985, 200: “We should be using the newly available data sets to help us find out what is actually going on in the economy and in the sectors that we are analysing, without trying to force our puny models on them. The real challenge is to try to stay open, to learn from the data, but also, at the same time, not drown in the individual detail. We have to keep looking for the forest among all the trees.”

<sup>7</sup> The basic terminology is stated e. g. in Blossfeld and Rohwer, 1995: 33 pp.

– Event means the change of the state, that is buying a new product.

The period between two successive events is called episode or spell. The elapsed time until an event occurs, or the duration of an episode, is the random variable  $T$ , which is the core variable of any event history analysis. For the description and explanation of the variable  $T$ , three terms are of special interest:

1. The distribution of the continuous random variable gives information about the level of the probability of a change up to a specific point of time.
2. The survival function  $S(t) = P(T \geq t)$  informs about the level of probability, that up to a specific point of time, no event occurs.
3. The hazard rate or hazard function is the instantaneous probability of having an event at time  $t$ , given that an event has not occurred before  $t$ .

$$\lambda(t) = \lim_{\substack{\Delta t \rightarrow 0 \\ \Delta t > 0}} \frac{1}{\Delta t} \cdot P(t \leq T < t + \Delta t | T \geq t)$$

One main problem arises, when the whole process over time can not be measured. When it is unknown

- how long the period of the state before the observation period has been – this is called left censoring –,
- when the last state in the observation period will end– the so called right censoring –, special assumptions have to be made<sup>8</sup>. If one can assume, that the process did not depend on the history, i. e. the prehistory has no impact on the process, left censoring would not be serious (Diekmann and Mitter, 1984: 23). In the case of right censoring, the information will not be dropped, instead it will be used for the description and the estimation of the process. Because to know, that the last state is longer than the remaining time of the observation period, means we know at least a little bit about the process and this information should be used.

---

<sup>8</sup> For a detailed description see Blossfeld and Rohwer, 1995: 34 pp., or Kalbfleisch and Prentice, 1980: 39 pp.

### 3. Definitions and operationalization

As mentioned earlier, the dataset is suitable to analyse a process over time even so it makes special requirements. As usual, you have to deal with the data at hand. The main problem is, how to identify continuous res. discontinuous behaviour within a given data set. What are the indicators or proxy variables and how are they to be operated?

The event history analysis offers several options to model the process, which is of interest here. The „extreme positions“ are characterised in short:

1. the process of buying could be modelled as a multistate-model with reversible events (Diekmann and Mitter, 1984: 37 and 177 pp.)<sup>9</sup>. That means that several states are existing and, moreover, a state can be achieved several times in the observation period.
2. the process of buying as modelled as a two state model with absorbable finite state. That means only one, i. e. the first, event is of interest.

Some problems belonging to these definitions arise. In the following, short comments on the main problems are given.

#### *Problem one: Selection of the group of goods / commodity group.*

It is necessary to use commodity groups, in which the goods are substitutes and not complementary goods.

Suitable are mainly goods, which are only used once and of which not several goods with different objectives do exist at the same time in the household, as for example:

- WGS 35: Wine: ordinary wine for the „daily“ consumption and higher class wine for special occasions.
- WGS 39: Shoe polish and means for maintenance leather: the different colours of shoe polish,
- WGS 50: cream: cream and sour cream
- WGS 89: Rice: full grain rice and rice pudding.

For these groups, buying another good of the same group does not necessarily indicate discontinuity of household decisions.

---

<sup>9</sup> It is also known as a multistate-multi-episode model; Blossfeld and Rohwer, 1995: 34.

So the following goods are qualified for the analysis:

- WGS 12: roasted pure coffee,
- WGS 21: universal detergent,
- WGS 44: sherry / port,
- WGS 75: filter paper only for coffee or only for tea,
- WGS 83: yeast,
- WGS 99: toilette paper.

Furthermore the product should be short-lived and may be multiply purchased in a year. So the group WGS 44 (sherry/port) is not suitable for the purpose of the analysis, because it is possible, that households buy sherry or port only once a year.

Also, the information of the product group must be available for the members of Panel 6. So the WGS 75 and WGS 99 can not be used.

Concerning the opportunity costs of a wrong decision it could be stated, that all products are of low price. Therefore, the opportunity costs of wrong decisions are low, even for households with low income and these costs are meaningless in the context of the analysis.

#### *Problem two: changing environmental conditions*

It has to be considered, that products res. product groups are chosen, whose purchase is not induced through (due to) changing environmental conditions. For example, purchasing a dishwasher reduces the amount and art of washingup liquid. The maintenance of the floor (ground care means ) is another example: to move house, with the older one having a carpeted floor and the new one having a parquet, changes the purchase of detergent.

#### *Problem three: seasonal commodity goods*

For seasonal goods and goods, for which the decision of purchase is dependent of the weather (for example ice cream), the buying decision is not continuous. On the given information in the data set, it is not possible to decide, whether the purchase depends only on household internals or is induced through other factors. Therefore, these goods must also be excluded.

#### *Problem four: Selection of the households*

As it is important to use covariables, which are constant over time, to identify those changes in the consumer behaviour of households that are not due to environmental changes, those households have been selected that showed no changes in the following variables between 1994 and 1996.

- V1: federal states,
- V2: number of inhabitants of the town,
- V5: professional life of the housekeeping person,
- V6: occupational group of the main earner,
- V8: Number of people living in the household,
- V9: housing condition, tenancy, distribution of property,
- V10: Rooms in the house,
- V12: secondary / college education of the main earner,
- V13: marital status of the housekeeping person,
- V36: Occupation of the main earner,

*Problem five: gifts*

It has to be considered that the acquired products would be consummated in the household and not used as gifts or something like that. In particular the variable v7: personal use, must have the status “yes”.

*Problem seven: relative price changes*

Changes in the purchasing behaviour could be induced by a lot of factors, for example changes in the environment as mentioned, but also changes in the relative prices. For the latter it is not possible to control such effects, because only a few goods are included in the data set.

*Problem eight: changes in income*

Another problem is the budget restraint. The decision of purchasing especially in lower income classes may only depend on the relative price of a good: a household would always buy the cheapest product. Because the data set contains only information on the average net income in 1995, the income mobility can not be measured. Therefore it is not possible to identify, whether purchasing a good is due to income changes.

*Problem nine: external factors*

For identifying risk behaviour the decision to buy a product should not depend on additional factors such as the cost of getting more information about the character of the good. Fortunately, the groups of goods which are analysed in the following, are cheap. Therefore, such opportunity costs are very low in respect to the net income of the household, even if the household belongs to the lowest income class.

## 4. Empirical analysis

### 4.1 Some comments on the methods used

Also the definition of the underlying process is given, there are three main categories of statistical techniques which could be used. Non-parametric, semi-parametric and parametric methods. In the analysis, only the first two methods come to application.

The most unrestricted methods are the non-parametric, in which only descriptive measures are used. To estimate the hazard and the survival function, two approaches exist: the life table estimator and the Kaplan-Meier estimator<sup>10</sup>.

The semi-parametric method, that is the so called proportional hazards model, is more restrictive<sup>11</sup>. But the baseline hazard is an unspecified function. Therefore, the proportional hazards model is a „robust“ model in respect to parametric models. But one necessary condition to work with this model is the homogeneity of the underlying population.

To measure the amount and the direction of the influence of  $n$  exogenous variables  $(x_1, \dots, x_n)$ , the proportional hazards model is used.

$$\lambda(t) = \lambda_0(t) \cdot e^{\beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \dots + \beta_n \cdot x_n}$$

and

$\lambda_0(t)$  = baseline hazard rate, which is unknown and not specified,

$x_1, \dots, x_n$  = the exogenous variables, and

$\beta_1, \dots, \beta_n$  = the parameter of the exogenous variables  $x_i$  ( $i = 1, \dots, n$ ).

The use of the model requires the assumption, that the influence of the exogenous variables is monotone over time, i. e. the exogenous variables are time-independent. The most restrictive is the parametric method. Using parametric methods requires a completely specified functional form of the underlying process. To use such methods it is necessary to make specific assumptions about the shape of the survival function. That means also that the groups must be homogenous with the same underlying distribution.

<sup>10</sup> This procedure is also often referred to as the product-limit estimator; see for example Kalbfleisch and Prentice, 1980: 12 pp.

<sup>11</sup> These models also called Cox models; see for a general description e. g. Lawless, 1982: 275 p, and 343 pp., or Klein and Moeschberger, 1997: 229 pp.



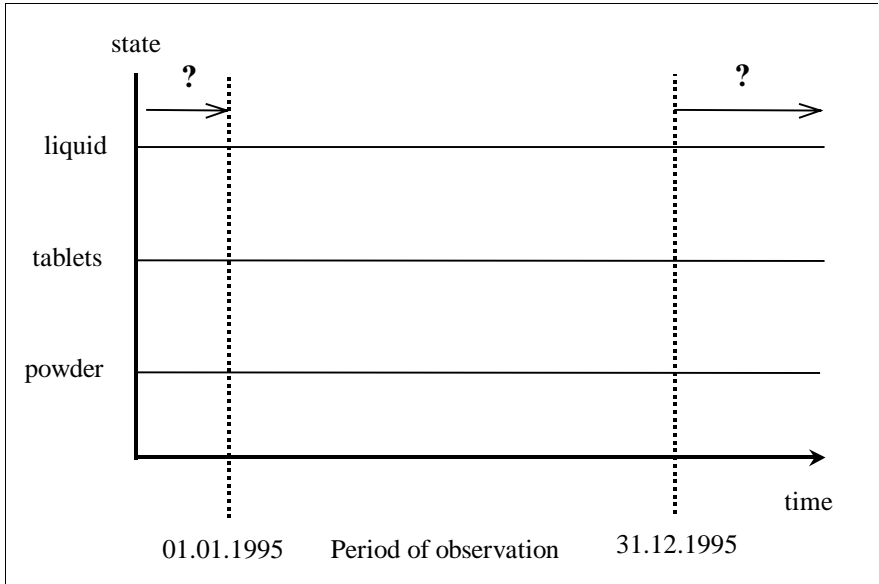
The approach is exploratory and the analysis is separated into two parts. First the hazard rate res. survival functions are computed. Second a semi-parametric model is carried out, to identify special variables which may influence the hazard rates or the survival function.

To gain a first insight in the process, a two state model with absorbed finite state is analysed. For further analysis, the terms „state“ and „event“ have to be substantiate. Whereas „event“ is easily defined as a change of the present state, there is more than one way to define „state“. The predefinition is essential for what continuity res. discontinuity means. For example: a state could be determined with the existents of a product of the same brand and the identical amount the household always buys. In this case, an event occurs by buying a good of another brand or another amount of the same good. But this seems too differentiate for a first insight of the time process of purchasing goods. To exclude to a larger part coincidence and to be conservative, a main aspect of the analysis is therefore, that the purchase of a good is not always an event. Only the purchase of a good of another class or category is stated as an event or as a state change. An event is therefore not a new acquisition of the same product. It is the change from a group of goods with the same characteristics to a group with other characteristics such as buying tablets or liquid detergent instead of washing powder.

What does this mean? To clarify the proceeding, the following figures enclosed two discontinuously and one continuously decision in a schematic manner. A specific peculiarity of the product „detergent“ is chosen as initial state. As state is mentioned the existence of the good in a specific characteristic in the household for consumption purposes. The question marks in the figures denote, that neither the state before the time of observation started is known nor the state after the end of the observation period. The former case is known as left censoring, the latter as „right censoring“.

In survival analysis especially left censoring can cause major problems. But in the present analysis, left censoring is not of larger concern, as it can be stated, that in former time periods, the attitudes of the household which are investigated here are mainly the same and have not altered. Figure 1 shows the underlying structure. The observation period started at January 1, 1995, and ended at December 31, 1995, i. e. the observation period covers a whole year. No information is available before and after 1995, so the process is left and right censored.

Figure 1: The underlying structure



In the following Figure 2 the change between the states „powder“ and „tablets“ is pictured. This is a household, which income decision is mentioned as discontinuous in respect to the specific product.

The continuity in decision is described in Figure 3, where no event occurs in the observation period. That denotes, that this household has chosen the same characteristic of goods in each purchase. But this holds true only for the observation period. It applies not for the time span before and after the observation period. Such case without any event is named as censored. These are households which have continuous income decision in the chosen approach here.

Figure 2: Left censored data with an event

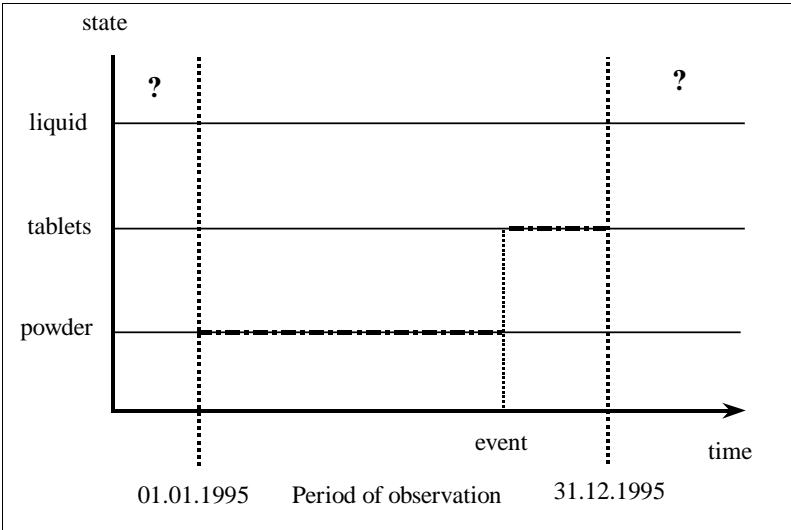
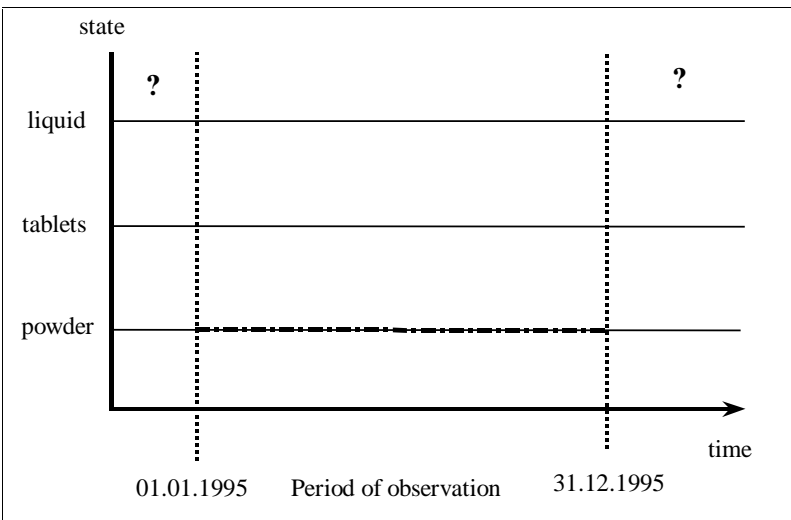


Figure 3: Left and right censored data



In comparing the survivor function or hazard rates for specific groups it is possible to identify differences in the continuity of the purchasing behaviour of households.

To gain more information about the reasons, which variables are of relevance to explain different behaviour, the Cox-Regression is used. As exogenous variables only objective variables are chosen. All subjective variables are neglected, because it is not the main goal of the analysis to test, whether the own view of ones behaviour is identical with the observable behaviour<sup>12</sup>.

All variables which could be mentioned as „objective“ are taken into account as exogenous with one restriction: for variables with a high correlation, for example age of the housekeeping person and the age of the main earner, it has to be decided, which one to use. In this case, the variable age of the main earner are omitted, because of the products which would be purchased. The purchase of low price goods are rather decided by the housekeeping person<sup>13</sup>.

The household panel 6 is used for the estimation, because only in this panel information about the income of the household and the number of persons, living in the household are known and these variables are notable as very important for income decisions.

## 4.2 Results

### Descriptive analysis

As already mentioned, only three WGS-groups could be analysed due to the chosen approach: roasted pure coffee, universal detergent, and yeast. The analysis is done separately for each commodity group. In the introduction it is mentioned that the differences in income decisions could depend mainly on two factors: the household income and the age<sup>14</sup> of the housekeeping person. These variables could be employed as proxies to represent household differences.

---

<sup>12</sup> As this is another question of research, this will be done in a later state of the analysis after identifying specific groups with different behaviour in spending money. For a critical discussion about this point of view in gaining some information on the behaviour of people see for example Ajzen and Fishbein, 1980. They neglect the relevance of external (objective) variables in explaining and predicting the behaviour of people. But, for example, to demonstrate the utility of a theory, it is necessary to measure the behaviour itself.

<sup>13</sup> For the behaviour of customers and the variables which – depending on the product price – influence it, see for an overview e. g.: Kroeber-Riel, 1990.

<sup>14</sup> It has to be mentioned, that age is only a proxy variable for experience, risk aversity, earnings capacity in sense of human capital theory, or for physical capability and it can not be considered as an

Therefore, to gain information, whether the behaviour of households is different, two subgroups are created in respect to the given household income classes and to the classes of age of the housekeeping person.

The WGS 12: roasted pure coffee

In the good group WGS 12: roasted pure coffee, two characteristics are examined to try to identify household behaviour. The WGS group is

- separated into three kinds of coffee: (1) grind in the factory, (2) grind in the shop and (3) not grind, and
- discriminated into three sorts: (1) with caffeine, (2) mild res. caffeine reduced and (3) caffeine free.

In the category kinds of coffee, the dominant kind is the first peculiarity: grind in factory, with 85 % of all households (3,682), 11.5 % are in the initial state of grind in the shop and only 3.1 % are purchasing not grind coffee. The households are more even distributed over the characteristics of the category “sorts”: From 4,326 households 48.8 % are belonging to with caffeine, 36.5 % to caffeine reduced and 14.7 % to caffeine free. The results of the Kaplan-Meier estimation are stated in Table 1.

Table 1: Results of the Kaplan-Meier estimation for Subgroups of WGS 12

Subgroups of WGS 12	Mean survival time in days	Number of events	Censored data in %
<b>Kinds of coffee</b>			
(1) grind in the factory	184	12,764	82.5
(2) grind in the shop	24	10,322	14.2
(3) not grind	21	70,251	3.7
<b>Sorts of coffee</b>			
(1) with caffeine	53	36,414	50.1
(2) caffeine reduced	36	26,972	37.0
(3) caffeine free	24	63,521	12.9

Source: Own calculations.

explanatory variable for itself. Age should be viewed “...as an index of change, just as inches are an index of person’s height, ...”; Hayslip and Panek, 1989: 10.

For the category “kinds of coffee”, there are no remarkable age or income effects in the purchasing behaviour due to category 3. Only a few censored cases – about 3.7 % – occurred, what means that this state is only temporarily adopted: the mean survival time is about 21 days. Virtual the same holds for the category 2: the mean survival time is 24. Only the censored events are higher with around 14.1 %, so there is a little bit more continuity in the household behaviour due to the purchase of coffee which is grind in shops. Purchasing coffee, which is ground in the factory, is the most continuous case with 82.5 censored events and a mean survival time of 184 days.

Overall, in the category “sorts of coffee”, the results are the same regarding to the household classification. No income effects exist and an age effect in the purchasing behaviour appears only in the category 1: with caffeine. Here, the older the people, the lower is the mean survival time, but around 50.0 % of the events are censored, as shown in Table 2.

*Table 2: Results of the Kaplan-Meier estimation for the subgroup “with caffeine” of WGS 12*

Age classes	Mean survival time in days	95 % Confidence Intervall	Number of events	Censored data in %
20 to 24 years	100	(69 ; 131)	75	59.0
25 to 29 years	85	(76 ; 94)	897	55.6
30 to 34 years	67	(62 ; 72)	2,258	53.1
35 to 39 years	63	(59 ; 68)	2,869	53.4
40 to 44 years	52	(48 ; 57)	3,536	49.3
45 to 49 years	46	(43 ; 49)	3,654	47.1
50 to 54 years	55	(50 ; 60)	2,922	54.7
55 to 59 years	53	(49 ; 57)	4,012	49.3
60 to 64 years	49	(45 ; 52)	4,410	50.7
65 to 69 years	47	(43 ; 50)	4,269	41.6
70 or older	43	(41 ; 46)	6,057	49.9

Source: Own calculations.

Whether there are differences in the survivor function of the subgroups could also be tested by means of test statistics<sup>15</sup>. The SPSS program package – which is used for the analysis – yield the Log Rank statistic, the Breslow-statistic, and the Tarone-Ware statistic, which are sensible in different areas of the process. The null hypothesis is that all survival curves are the same.

<sup>15</sup> For a description of the tests for two or more samples see e. g. Klein and Moeschberger, 1997: 191.

All three statistics delivered results which are statistically highly significant<sup>16</sup>, so the null hypothesis has to be rejected: the survival functions for the age classes are different. This means, that parametric analysis is not appropriate for these data.

The WGS 21: universal detergent

The purchase behaviour for the group “universal detergent” is far more differentiated. The group is split up into six subgroups, as mentioned in Table 3. Also shown in Table 3 is the mean survival time, the number of events and the percentage of censored data for each subgroup. In totally, 21,042 purchases are done.

Table 3: Results of the Kaplan Meier estimation for Subgroups of universal detergent

Subgroups of WGS 21	Mean survival time in days	Number of events	Censored data in %
(1) detergent construction systems	56	20,053	4.7
(2) traditional powder	63	18,272	13.2
(3) liquid detergent	58	19,635	6.7
(4) tablets	54	21,042	0.0
(5) concentrate powder	99	11,397	45.9
(6) super concentrate powder	78	14,811	29.6

Source: Own calculations.

The Table 3 indicates clear differences in the continuity of the purchasing behaviour of households between the subgroups. The mean survival time for the subgroups (1), (3) and (4) is the lowest with 54 to 58 days and for (5) and (6) are the highest. In these two groups, the percentage of censored data is, in respect to the other sub-categories, high with 29,6 % res. 45,9 % cases without an event in 1995.

The question now is, whether the differences are due to the age of the housekeeping person or to the household income. For this reason, the households are divided into subgroups to these two variables.

For all subgroups of universal detergent, the following results obtained:

- in respect to the household income: the higher the income, the lower the mean survival time,

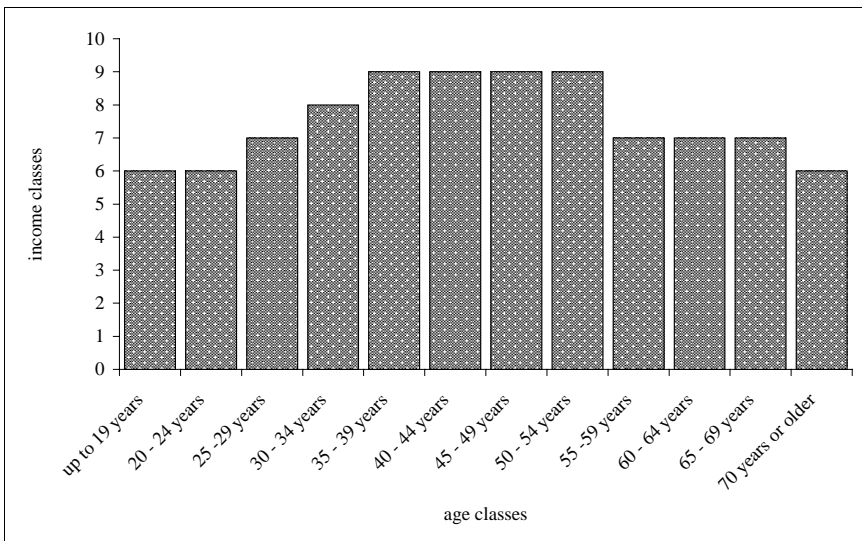
---

<sup>16</sup> The test statistics are 652.44 (Log Rank test), 549.17 (Breslow test), and 632.75 (Tarone-Ware-test), with 10 degrees of freedom.

- for the classification in age groups, no linear relationship occurs, but often an u-profile with higher mean survival time in the younger and older age classes.

The age profile is remarkable, but for the interpretation it has to be taken into consideration, that it could be caused by the income differences between the age classes. Households with young and old housekeeping persons have lower income than the middle-aged as shown in Figure 4, therefore it could be deduced, the u-shape profile is caused by the distribution of household income, which shows – as is well know from numerous analyses – an invers-U-shape profile.

Figure 4: The median household income in age classes



Note: 6 : 2.000 DM to 2.499 DM

7 : 2.500 DM to 2.999 DM

8 : 3.000 DM to 3.499 DM

9 : 3.500 DM to 3.999 DM

Source: Own calculations.

To gain an insight in the dependency of the purchase behaviour and age, the income has to be held constant. Unfortunately, the numbers of households are insufficient to make analysis in this direction. The numbers of households per cell are high enough (over 50



households) only for the middle income classes – between 2,000 DM and 4,000 DM – and for age groups older than 25 years. Even though the analysis contains hints, that the u-shape profiles may exist in special income classes, there is no simple pattern overall.

So only the differentiation between income classes yielded differences in the purchasing behaviour for this group of goods as stated as a hypothesis.

**The WGS 83: yeast**

The group 83, yeast, is differentiated into two subgroups: blowing yeast and dry yeast. For blowing yeast, the percentage of censored data for a change to dry yeast is about 92 %, and around 40.7 % for a change from dry to blowing yeast, as shown in Table 4. That means, that the continuity in buying blowing yeast is very high<sup>17</sup>.

*Table 4: Results of the Kaplan Meier estimation for Subgroups of yeast*

Subgroups of WGS 83	Mean survival time in days	Number of events	Censored data in %
(1) blowing yeast	237	606	92.0
(2) dry yeast	71	4,512	40.7

Source: Own calculations.

For the twelve age groups only about 50 events occur in the average, whereas the total number of cases is 7,606. The mean survival time for a household in the state “blowing yeast” is 237 days, whereas the mean time of being in the state “dry yeast” is only 71 days.

Although the differences in the survivor functions and hazard rates are statistically significant, no special pattern exists. Therefore, no systematic differences in respect to the behaviour of buying dry yeast could be detected.

**5. Summary**

All things considered, it must be stated, that only weak hints are detected on the continuity of the purchasing behaviour in respect to differences between households belonging to their income or to the age of the housekeeping person. An age effect appears only for one category of “sort of coffee”: the older the housekeeping person, the lower the mean survival time of the state “with caffeine”, and the numbers of censored events is slightly

---

<sup>17</sup> It seems, as it is necessary to differentiate the group blowing yeast further into subgroups of special brands or something like that. Unfortunately, this is not possible with the data set at hand.

lower. An income effect occurs in the WGS 21 group as the higher the income-class, the lower the survival time.

It seems as if the differences due to continuous res. discontinuous behaviour belongs to the specific subgroups of the goods analysed because the differences in percentage of censored data for the subgroups in each commodity group are large. For example, 82.5 % of the households always purchase coffee, grind in the factory, whereas only 3.7 % data are censored in the subgroup “not grind”.

But this analysis could not yield any information about the reaction of households due to changes in the exogenous variables income or (the proxy variable) age. Even more, it is not possible to consider, whether the events – the discontinuity – are caused through price changes. Some hints due to this aspects could be gained using the method of semi-parametric analysis<sup>18</sup>.

### **Semi-parametric analysis**

The purpose of the semi-parametric analysis is not to test a theory about purchasing behaviour. The aim is rather to use the method as an explanatory approach and try to identify some variables, which could induce the differences in the survivor function.

As the previous pure descriptive analysis shows, the household income and the age of the housekeeping person should not be neglected as exogenous variables, although there are only weak hints on the effect. Besides the other objective variables in panel 6, the variable price per unit was used to deal with the price elasticity of the behaviour. The age of the housekeeping person and the age of the main earner are highly correlated (the Spearman correlation coefficient is 0.946). Therefore, only the first variable is used. Furthermore, the variable V1: federal states, is merged in only two data: West and East Germany. There are two variables on the occupation of the principal, so one has to be omitted and the variable occupation of the main earner is used. Concerning the variables number of people in the household and number of children, only the first one is chosen. The remaining variables are

- price per unit,
- regions: West res. East Germany,
- V3: age of the housekeeping person
- V5: professional life of the housekeeping person,

---

<sup>18</sup> For the theoretical relevance of the relationship between the actual evolution of price and the demand over time in demand theory see for example Hildenbrand, 1994, chapter 1.

- V7: net income of the household1
- V8: number of people living in the household,
- V12: secondary / college education of the main earner,
- V13: marital status of the housekeeping person,
- V36: Occupation of the main earner.

In the following, I will only point out the main results of the analysis. First the results of the Cox-regression for the group roasted pure coffee are presented in Table 5. In Table 5, the so called  $\alpha_i$ -coefficients (or  $\alpha_i$ -effects) are presented. Their interpretation is easier than that of the original parameters  $\beta_i$ : they inform directly over the direction and the magnitude of the influence of the exogenous variable. As the function of the hazard rate is

$$\lambda(t) = \lambda_0(t) \cdot e^{\beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \dots + \beta_n \cdot x_n}, \quad i = 1, \dots, 9,$$

the  $\alpha_i$ -effects are calculated as:  $\alpha_i = e^{\beta_i}$ .

Therefore, the  $\alpha_i$  are multipliers, which are always positive. If  $\alpha_i > 1$ , the effect of the variable on the rate is positive, for  $\alpha_i < 1$  the effect is negative and for  $\alpha_i = 1$ , no effect exists at all. The  $\alpha_i$  inform also about the magnitude of the effect on the hazard rate: with a change of the independent variable  $x_i$  of one unit ( $\Delta x_i = 1$ ), the rate alters about  $100 \cdot (1 - \alpha_i)$ -percent. A value of  $\alpha_i = 0.98$  means, that the effect of the exogenous variable is negative and the rate will be reduced by 2 percent, if the  $x_i$  increases by one unit with all other variables are constant.

First, the commodity group roasted pure coffee, divided in two specific groups, is considered. In the Table 5, the values of the  $\alpha_i$ -coefficients are presented, which are statistically significant at 0.05 level.

As Table 5 shows, not all variables are statistically significant. The effect of age and income is always positive, which means that positive changes in age ore income raise the hazard rate. The strength of the effect for both variables is not very high: between 2 % and 5 %. The direction of the effect for the variable education is always negative, but the effect with a percentage between around 1 % and 4 % is not very strong. The  $\alpha_i$ -effects of the variables price per unit and regions are the highest for most subgroups, but the direction is positive in the first subgroup (1) and negative in the others (2) and (3). For the other independent variables the effects are not clear: sometimes they are insignificant, sometimes the effect is positive or negative. Table 6 presents the results for the group W21: universal detergent.

Table 5:  $\alpha_i$ -coefficients for subgroups of WGS 12: roasted pure coffee

Variables	Kinds of coffee		
	(1) grind in the factory	(2) grind in the shop	(3) not grind
Price	4.204	0.731	0.924
Regions	0.227	1.210	1.104
Age	1.052	1.031	1.031
Professional life	n.s.	1.012	1.013
Net income	1.039	1.023	1.025
Number of people	1.014	1.028	1.029
Education	0.956	0.985	0.980
Marital status	n.s.	n.s.	n.s.
Occupation	0.997	n.s.	n.s.
Variables	Sorts of coffee		
	(1) with caffeine	(2) caffeine reduced	(3) caffeine free
Price	1.331	0.959	0.883
Regions	0.561	1.380	1.165
Age	1.035	1.036	1.028
Professional life	1.046	n.s.	n.s.
Net income	1.030	1.021	1.025
Number of people	n.s.	1.041	1.030
Education	0.963	0.984	0.988
Marital status	n.s.	0.980	n.s.
Occupation	n.s.	1.001	0.999

Note: n.s.: not statistically significant at 0.05 level. Source: Own calculations.

Table 6: Results of the Cox-regression for Subgroups of WGS 21: universal detergents

Variables	Subgroups of universal detergent					
	(1)	(2)	(3)	(4)	(5)	(6)
Price	0.846	1.869	0.985	1.225	3.765	0.415
Regions	1.162	n.s.	1.127	1.114	1.191	1.103
Age	0.981	0.961	0.974	0.973	0.991	0.968
Professional life	1.039	1.041	1.042	1.040	1.040	1.034
Net income	1.042	1.049	1.037	1.043	1.047	1.042
Number of people	1.039	1.040	1.041	1.041	1.041	1.043
Education	0.978	0.987	0.981	0.982	0.976	0.987
Marital status	1.079	1.075	1.078	1.071	n.s.	1.068
Occupation	n.s.	0.998	0.998	0.999	n.s.	n.s.

Note: (1) detergent construction systems; (2) traditional powder; (3) liquid detergent; (4) tablets, (5) concentrate powder; (6) super concentrate powder;

n.s.:not statistically significant at 0.05 level.

Source: Own calculations.

In this group, five variables provide the same results for all subgroups due to the direction of the effect: professional life, net income and the number of people affect the hazard rate positive, that holds also for the variables regions and marital status for most subgroups. That means, that, for example, the raise in net income or in the number of people living in the household, increase the hazard rate. The influence of the variables age and education on the hazard rate is in the opposite direction. In other words, households, in which the housekeeping person is older or has a higher education, are more continuous in purchasing, independent of the special subgroup. The effects of all the variable occupation is not significant in three of six subgroups. The price variable is a little bit suspicious, because the  $\alpha_i$ -effect is sometimes very high but the effect is not constant in one direction: for three subgroups the direction is negative and for the other three groups it is positive. Table 7 summarises the results for the group yeast.

Table 7: Results of the Cox-regression for Subgroups of WGS 83: yeast

Variables	Subgroups of yeast	
	Blowing yeast	Dry yeast
Price	1.069	1.024
Regions	0.954	0.829
Age	0.949	n.s.
Professional life	1.201	1.162
Net income	1.058	1.018
Number of people	n.s.	1.052
Education	n.s.	1.022
Marital status	0.860	n.s.
Occupation	n.s.	1.005

Note: n.s.: not statistically significant at 0.05 level. Source: Own calculations.

The group yeast is the only one, in which the effect of price change on the hazard rate is consistent for all subgroups: it is always positive but not the effect is not very high with 6.9 % res. 2.4 %. There is also a positive linkage between the hazard rate and the variables net income and professional life, with the latter has the strongest effect on the hazard rate with 20.1 % for blowing yeast and 16.2 % for dry yeast. Another strong  $\alpha_i$ -effect occurs for the variable region: for dry yeast, households in West-Germany are far more discontinuous than their counterparts in East-Germany, for those households the effect is 17.1 % lower. Another strong negative effect has to be stated for the variable marital status for blowing yeast with 14 %, whereas for dry yeast this independent variable is not statistically significant.

To summarize the results of the semi-parametric analysis, it has to be stated, that no simple pattern occurs even within the specific groups. The continuity res. discontinuity is not

always caused by the same variables and, moreover, the effects of the exogenous variables on the hazard rate between the subgroups of one commodity group are sometimes divergent. On the other hand, it could be shown, that it is possible to describe the purchasing behaviour – as it is operationalized here – with method of the semi-parametric analysis and to shed some light on the direction and amount of the influence of independent (objective) variables.

The results obtained in the analysis are not compatible with the micro-economic theory on the price elasticity of the goods, on the whole. The analysis delivers hints, that the reaction of households due to changes in the price of goods is not always high or even in the same direction for all subgroups within a commodity group. This is in contrast to the assumption, which is often made in consumption analysis and which may be correct for the group itself. It is also obvious, that the purchasing behaviour is far from being as simple as is often stated in economic theory. The analysis indicates that the influence of factors for which age is used as proxy variable or of income is not linear or even works in the same direction. In specific commodity groups, the effect of the chosen exogenous variables is often different for the subgroups. Therefore no empirical evidence could be obtained on the hypothesis on risk aversity res. continuity of purchasing behaviour, as it is operationalized in the analysis: the younger the people or the higher the income, the lower is the continuity.

## **6. Summary**

The main purpose of the paper was to analyse the purchasing behaviour of households due to aspects of continuity res. discontinuity. To describe the process, methods of survival analysis are applied. The process of purchasing was modelled as a two state model with absorbed finite state. Even this is a quite crude approach to model the process, it delivers a lot of information about purchasing behaviour.

In the descriptive analysis based on the product-limit estimation, large differences occur in respect to the subgroups concerning the continuity of the purchasing behaviour. But only a few hints are detected in regard to differences in the purchasing behaviour due to the household net income or the age of the housekeeping person. Therefore it seems as if these exogenous variables are not as important in explaining differences in the purchasing behaviour of households as often stated.

To throw some more light on the relevance of exogenous (objective) variables, a semi-parametric analysis was additionally carried out. To measure the direction and the amount of the influence of the exogenous variables, the proportional-hazards model was used. The

results of the analysis are for some parts encouraging: even in the relative coarse operationalisation of the discontinuity in the purchasing behaviour of households, it could be shown that, with the method used, it is possible to describe the direction and amount of the influence of the independent variables. This section of the analysis also delivers results, which are not compatible to the micro-economic demand theory concerning the price elasticity of the goods: The reaction of the households due to price changes are not coherent even within the same commodity group.

The data set is unique. It offers a lot more possibilities to analyse the purchasing behaviour than could be offered in this paper. So a lot is left to be done especially due to longitudinal analysis on the purchasing behaviour of households.

## Contact

PD Dr. Uwe Fachinger  
Zentrum für Sozialpolitik  
Wirtschaftswissenschaftliche Abteilung  
Universität Bremen  
Parkallee 39  
28209 Bremen

## Bibliography

- Ajzen, Icek, Martin Fishbein, 1980, *Understanding Attitudes and Prediction Social Behavior*. Englewood Cliffs, N. J.: Prentice-Hall.
- Atkinson, Anthony Barnes, Francois Bourguignon, C. Morrisson, 1988, *Earnings Mobility*. In: *European Economic Review*, 32, 619-632.
- Atkinson, Anthony Barnes, Francois Bourguignon, C. Morrisson, 1992, *Empirical Studies of Earnings Mobility. Fundamentals of Pure and Applied Economics 52*. Chur u. a. O.: Harwood Academic Publishers.
- Blossfeld, Hans-Peter, Götz Rohwer, 1995, *Techniques of Event History Modeling. New Approaches to Causal Analysis*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Diekmann, Andreas, Peter Mitter, 1984, *Methoden zur Analyse von Zeitverläufen - Anwendung stochastischer Prozesse bei der Untersuchung von Ereignisdaten*. Stuttgart: Teubner.
- Fabig, Holger, 1999, *Einkommensmobilität im internationalen Vergleich. Eine empirische Mobilitätsanalyse mit Panel-Daten*. Wirtschaftswissenschaft. Frankfurt – New York: Campus.

- Fachinger, Uwe, 1991, *Lohnmobilität in der Bundesrepublik Deutschland. Eine Untersuchung auf der Basis von prozeßproduzierten Längsschnittsdaten der gesetzlichen Rentenversicherung*. Volkswirtschaftliche Schriften, Heft 409. Berlin: Duncker & Humblot.
- Griliches, Zvi, 1985, *Data and Econometricians – The Uneasy Alliance*. In: American Economic Review, 75, 196-200.
- Hayslip, Jr., Bert, Paul E. Panek, 1989, *Adult Development and Aging*. New York u. a. O: Harper & Row.
- Hildenbrand, Werner, 1994, *Market Demand: Theory and Empirical Evidence*. D. M. Kreps and T. J. Sargent, Frontiers of Economic Research. Princeton, New Jersey: Princeton University Press.
- Kalbfleisch, John D., Ross L. Prentice, 1980, *The Statistical Analysis of Failure Time Data*. New York – Chichester – Brisbane – Toronto – Singapore: John Wiley & Sons.
- Klein, John P., Melvin L. Moeschberger, 1997, *Survival Analysis. Technics for Censored and Truncated Data*. K. Dietz, M. Gail, K. Krickeberg and B. Singer, Statistics for Biology and Health. New York: Springer.
- Kroeber-Riel, Werner, 1990, *Konsumentenverhalten*. München: Vahlen.
- Lawless, Jerald F., 1982, *Statistical Models and Methods for Lifetime Data*. New York – Chichester – Brisbane – Toronto – Singapore: John Wiley & Sons.
- Lunt, Peter, 1995, *Psychological Approaches to Consumption. Varieties of Research – Past, Present and Future*. In: D. Miller, Acknowledging Consumption. A Review of New Studies. London – New York: Routledge,
- Meier, Christoph, 1983, *Lebenszyklus und ökonomische Ungleichheit: Eine Analyse der Einkommens- und Vermögensdynamik anhand von Längsschnittsdaten*. Freiburger Studien zur Verteilung von Einkommen und Vermögen 3. Freiburg, Schweiz: Universitätsverlag.
- Müller, Klaus, Joachim Frick, 1997, *Die Äquivalenzeinkommensmobilität in den neuen und alten Bundesländern 1990 bis 1994*. In: S. Hradil and E. Pankoke, Aufstieg für alle? Beiträge zu den Berichten zum sozialen und politischen Wandel in Ostdeutschland, Beiträge zum Bericht 2 "Ungleichheit und Sozialpolitik". Opladen: Leske + Budrich, 103-154.
- Trede, Mark, 1997, *Statistische Messung der Einkommensmobilität*. Göttingen: Vandenhoeck & Ruprecht.
- Zurbrügg, Fritz, 1990, *Einkommensentwicklung im Lebenszyklus. Eine Längsschnittuntersuchung mit Daten der staatlichen Rentenversicherung der Schweiz*. Bern – Stuttgart: Haupt.



# NONPARAMETRIC MODELING OF BUYING BEHAVIOR IN FAST MOVING CONSUMER GOODS MARKETS

*YASEMIN BOZTUĞ AND LUTZ HILDEBRANDT*

## 1. Introduction

Much empirical research of consumer behavior has been done using discrete choice modeling. In general, model selection in this area is a decision between complexity of the model and the simplicity of estimation (Ben-Akiva et al., 1997). The well-known multinomial logit model (MNL) (Guadagni and Little, 1983), which is often used on the basis of scanner panel data, especially shows this dilemma. It is easy to estimate but has many restrictive assumptions. Relaxing assumptions however leads to estimation problems. Much research has been done to improve the parametric MNL model (e.g. Kannan and Wright, 1991; Chintagunta, 1992; Fader et al., 1992; Chintagunta, 1993; Gupta and Chintagunta, 1994; Erdem and Keane, 1996; Erdem, 1996; Papatla, 1996). The estimation of nonparametric and semiparametric variants of the MNL model may offer useful alternatives to circumvent its intrinsic constraints. These models encompass a large latitude for modeling, and are based on statistical theory that allows for relatively simple estimation.

This article starts with a short explanation of the logit model in Section 2, where the assumptions are also discussed. The theoretical aspects of nonparametric density estimation with respect to discrete choice models are investigated in Section 3. Accordingly the case of kernel density estimation for binary data is shown. Also, we document the semiparametric approach with a typical data structure of mixed binary and continuous explanatory variables. In addition, we will demonstrate the advantages and benefits of pursuing nonparametric and semiparametric methods. Section 4 presents an application of a semiparametric method to a real panel data set, estimated by two different algorithms.

## 2. Logit models in marketing

### 2.1 General description of a logit model

The multinomial logit model (MNL) captures the individual choice behavior between several alternatives. Here, the theorem of utility maximization for the consumers is assumed (Ben-Akiva and Lerman, 1985). That means that a consumer  $n$  chooses the alternative which maximizes his utility  $U_n$ . The choice set  $C_n$  is comprised of the alternatives  $i = 1, \dots, I$ . The probability  $Pr_n(i)$  for individual  $n$  to choose the alternative  $i$  is described by

$$Pr_n(i) = Pr(U_{in} \geq U_{jn}, \forall j \in C_n, i \neq j), \quad (1)$$

with  $U_{in}$  and  $U_{jn}$  the utilities of the alternatives  $i$  and  $j$ , where  $U_{jn}$  specifies all alternative utilities to  $U_{in}$ . Usually, the utility function can be separated additively into a systematic part ( $V_{in}$ ) and a random part ( $\varepsilon_{in}$ ) with  $U_{in} = V_{in} + \varepsilon_{in}$ . In the MNL, the systematic utility function is assumed to be linear in the parameters ( $V_{in} = \beta^T x_{in}$ , where  $\beta$  is a parameter vector to be estimated and  $x_{in}$  a characteristic of alternatives in the opinion of the individual  $n$ ). With these assumptions, the choice probability has the form

$$\begin{aligned} Pr_n(i) &= Pr(U_{in} \geq U_{jn}, \forall j \in C_n, j \neq i) \\ &= Pr(V_{in} + \varepsilon_{in} \geq V_{jn} + \varepsilon_{jn}, \forall j \in C_n, j \neq i) \\ &= Pr(\varepsilon_{jn} - \varepsilon_{in} \leq V_{in} - V_{jn}, \forall j \in C_n, j \neq i). \end{aligned} \quad (2)$$

A further supposition of the MNL model is that the differences of the errors  $\varepsilon_{jn} - \varepsilon_{in}$  from equation (2) are i.i.d. logistic distributed. Following McFadden (1974) and using the information about the distribution of the errors, equation (2) can also be written as

$$\begin{aligned} Pr_n(i) &= Pr(U_{in} \geq U_{jn}, \forall j \in C_n, i \neq j) \\ &= \frac{\exp(V_{in})}{\sum_{j \in C_n} \exp(V_{jn})} \\ &= \frac{\exp(\beta^T x_{in})}{\sum_{j \in C_n} \exp(\beta^T x_{jn})}. \end{aligned} \quad (3)$$

A maximum-likelihood estimator can obtain the parameter values contained in  $\beta$ . Here, the sign and also the absolute values of the elements of  $\beta$  are of interest. The sign informs us of the direction of connection, e.g. if the sign for a specific explanatory variable of a brand is positive, than a decrease of this variable means also a decrease in the probability

of choice for that brand and the same holds for negative signs in the other direction. If all variables are scaled the same, the absolute value of  $\beta$  can be interpreted regarding the strength of the connection between the explanatory and the dependent variables.

## **2.2 Assumptions of the logit model**

The logit model follows assumptions that restrict the interpretation of the estimation results, and also the application of the model is limited. The main criticism of the logit model is related to the IIA assumption (independence of irrelevant alternatives), which implies that the relative utility of one alternative to a second one is independent of the existence of a third one.

The second weak point of the logit model is the assumption about the logistic distribution of the differences of its error terms. It is not obvious, why the differences should follow this distribution, because there are several possibilities for modeling the error terms.

The third weakness relates to the assumption of linear formulation of the utility function. Due to this assumption, a lot of possibilities to describe the utility are excluded. Also, there is no economic justification for the linearity.

To address the first two weak points of the logit model often probit models are used. They use an i.i.d. normal distribution assumption for the differences of the error terms. But for specifying a probit model for the multidimensional case it is difficult to estimate because higher order integrals must be solved. An alternative is to test the appropriateness of the logit model (Bartels et al., 2000). Another solution to deal with the strong assumptions of the logit model is a semi- or nonparametric formulation of the model, which will be presented in the next section.

## **3. Non- and semiparametric discrete choice models**

A nonparametric or semiparametric discrete choice modeling approach allows a nonparametric systematic utility component and/or a distribution-free random component. With this type of modeling one has much more flexibility to specify a choice process. To classify the different characteristics of model types (parametric, non- and semiparametric), see Table 1.

Table 1: Different model types for parametric and nonparametric choice models.

Model type	Systematic utility function	Random component	Possible resulting method
Parametric	parametric	parametric	MNL, probit, etc.
semiparametric I	parametric	distribution free	various
semiparametric II	nonparametric	parametric	Generalized Additive Model (GAM)
nonparametric	nonparametric	distribution free	Nonparametric Density Estimation (NDE)

Parametric models could be e.g. the MNL, or a probit model (Manrai, 1995). The type of “semiparametric I” is described by a parametric utility function, but a distribution free random term. These kind of models are discussed by a wide group of researchers, e.g. Horowitz et al. (1994); Horowitz and Härdle (1996); Matzkin (1991); Chintagunta and Honore (1996). Because of some disadvantages for practical use (see below) this approach will not be described here. The “semiparametric II” type includes a nonparametric utility function, but a parametric random term. The Generalized Additive Models (GAM), introduced by Hastie and Tibshirani (1986), could be used to specify these models. The Nonparametric Density Estimation (NDE) is one possibility to model a pure nonparametric approach. Abe (1995) introduced this approach in the marketing context. The NDE approach will be discussed in the next section in more detail.

### 3.1 Nonparametric density estimation

The availability of scanner panel data sets has made the application of the nonparametric methods in marketing a feasible alternative to the existing parametric models. Nonparametric methods are based only on very few assumptions, so that they have a lot more structural freedom than the models in the parametric model class, e.g. the MNL. On the other hand, they need sufficient large data sets to produce a good fit of the data. One popular nonparametric method is nonparametric density estimation.

#### 3.1.1 General formulation of the density estimation

The choice decision formulated with a nonparametric method is usually described as a conditional expectation, where the condition is the actual marketing-mix situation at the purchase time. So the conditional expectation  $E[y|x]$  is needed, with  $x$  the marketing-mix condition and  $y$  the choice decision. Usually, the choice is coded binary, so that the following identity holds  $E[y|x] \equiv P(y|x)$  with the assumption of  $0 \leq E[y|x] \leq 1$  (e.g.

Hosmer and Lemeshow, 1989). Also the restriction that the sum over the choice probabilities is 1 for given covariates  $x$  is needed. Using Bayes' theorem it follows that

$$E[y | x] = \frac{Pr(y)f(x | y)}{f(x)} \tag{6}$$

under the assumption of  $y$  binary and  $f(\cdot)$  a probability density function. For the estimation of  $E[y | x]$  through  $E[y | x]$ , the estimator can be partitioned into several different components. One is the estimator  $\hat{Pr}(y)$  of  $Pr(y)$ , which can simply be described as

$$\hat{Pr}(y) = \frac{\text{number of choices made for the inspected brand}}{\text{number of all possible choices for all brands}}.$$

The components remaining to complete the expectation of equation (6) are the densities functions  $f(x)$  and  $f(x|y)$ . A kernel density estimator can estimate them. The density estimation  $\hat{f}_h$  at  $x$  for the density function  $f$  of the  $n$  continuous multivariate i.i.d. random variables  $X_1, \dots, X_n$  with  $d$  dimensions can be described as

$$\hat{f}_h(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) = \frac{1}{nh^d} \sum_{i=1}^n K(x, h) \tag{7}$$

with  $K(\cdot)$  a kernel function, which gives the smoothing instruction for the inner distance from  $(\cdot)$  as described in equation (7). The distance measure is usually chosen Euclidian for continuous variables. The kernel function also has to fulfil the following conditions (e.g., Silverman, 1986)

- $\int_{R^d} K(x) dx = 1$
- $K(\cdot)$  is symmetric and positive
- $K(\cdot)$  is  $l$  times continuously differentiable

The parameter  $h$  is called the bandwidth of the kernel. It determines with the window width  $h$ , in other words which observations are included in the calculation. It is also a measure for the balance between the bias and the variance of the estimation. There are different methods to determine the value for  $h$  as well as different expressions for the kernel function that will not be discussed further here. For a detailed explanation the reader is referred to Härdle (1991). It is shown in several statistical research projects that

the choice of the bandwidth  $h$  is important, whereas the choice of the kernel  $K(\cdot)$  influences the estimates only slightly (Härdle, 1991; Silverman, 1986; Fan and Marron, 1992).

The conditional expectation can also be seen as a kernel regression on a 0-1 binary response variable. Therefore, the conditional expectation with the estimation of the densities as in equation (7) can be expressed as

$$\hat{E}[y | x] = \frac{\sum_i y_i K\left(\frac{x - X_i}{h}\right)}{\sum_i K\left(\frac{x - X_i}{h}\right)} = \frac{\sum_i y_i K(x, X, h)}{\sum_i K(x, X, h)}, \quad (8)$$

with  $y$  a multivariate binary variable with  $y_i$  a vector of  $J$  elements with

$$y_{ij} = \begin{cases} 1 & \text{if alternative } j \text{ is chosen} \\ 0 & \text{otherwise} \end{cases}$$

Up to now a general description of a kernel density estimation has been given. In the marketing context some particularities exist, e.g. the fact that the explanatory variables have a discrete or binary character.

### 3.1.2 Extension of the kernel density estimation to marketing data

Usually kernel density estimation is applied for continuous, and most of the time multivariate, data. In the application of this nonparametric method in marketing we have the special situation of mixed (binary and continuous) explanatory variables. We therefore need a suitable modification of the estimator for this case. The use of kernel density estimation for binary data has been well known for a long time (e.g., Aitchison and Aitken, 1976; Silverman, 1986). The idea is to build a kernel similar to the spherical normal kernel. In the  $k$ -dimensional binary space  $B^k = \{0,1\}^k$ , the needed kernel for binary variables is defined as

$$K(x, X, b) = b^{k-d(x,X)} (1-b)^{d(x,X)} \quad (9)$$

where  $x$  is the variable at which the density should be estimated,  $X$  the explanatory variable and  $b$  the smoothing parameter with  $\frac{1}{2} \leq b \leq 1$ . The distance function  $d(x, X)$  is defined as a measure of the disagreements between  $x$  and  $X$  with

$$d(x, X) = (x - X)^T (x - X). \tag{10}$$

The analyst should determine the value of the smoothing parameter  $b$ . It describes the weight that a binary observation is given that lies not completely at the point where the estimate is wanted. A value close to 1 gives those not totally equal observations little importance and a value close to  $\frac{1}{2}$  produces a uniform weighting of right and wrong specified observations. The value close to 1 produces the effect of a small, but existing, weighting of values “close” to the one wanted. In a classic continuous kernel approach, the value of  $b$  would equal 1, which leads to a very strict separation of “right” and “wrong”. This is not wanted here for the binary variables.

Now we have a kernel density estimator for strict continuous (equation (7)) or strict binary (equation (9)) data, and so both expressions must be combined to a mixed non-parametric density estimator. This model is given in the form of

$$K(x, X, b, h) = b^{k_1 - d_1(x, X)} \bullet (1 - b)^{d_1(x, X)} \bullet \frac{1}{h^{d_2}} \bullet \frac{15}{16} \left\{ 1 - \left[ \frac{1}{h} d_2(x, X) \right]^2 \right\} \bullet I \left\{ \left| \frac{d_2(x, X)}{h} \right| < 1 \right\} \tag{11}$$

where  $d_1(x, X)$  is the distance function for the binary data as defined before in equation (10), and  $d_2(x, X)$  is the Euclidian distance of the continuous variables. The value of  $k_1$  describes the number of binary components and  $k_2$  the number of the continuous components in the explanatory variables. For the continuous components of the observation the Quartic kernel was chosen (e.g., Härdle, 1991), which is a common choice for a kernel.

Now the conditional expectation can be written with respect to the mixed kernel of equation (11) and due to equation (8) as

$$\hat{E}[y | x] = \frac{\sum_i y_i K(x, X, b, h)}{\sum_i K(x, X, b, h)}$$

to model the brand choice decision. The conditional expectation is modeled for each brand separately, and from there the choice probabilities dependent on the explanatory variables can be obtained. No assumption like the “utility maximization” for the MNL is made. Moreover, it is not necessary to make an assumption about the distribution of the random component as in the MNL. These two points represent the main differences between nonparametric density estimation and parametric models (e.g., the MNL). The nonparametric approach affords a wide field of applications but often some parameters are wanted to calculate market shares or to make predictions. With a pure nonparametric model it is not possible to give these parameter values. This leads us to another possible brand choice model formulation, a semiparametric approach.

### 3.2 The semiparametric approach

Brand choice models can also be specified and estimated in a semiparametric way. This type of model is closer to the MNL model than the purely nonparametric methods. Semiparametric models specify one part of the MNL in the common parametric way (in our approach the error term distribution), but the other component (the utility function) is formulated by a nonparametric method. The resulting semiparametric model is called the “Generalized Additive Model” (GAM) (Hastie and Tibshirani, 1986; 1987; 1990). In modeling brand choice, Abe (1997) introduced the method to the choice modeling research area. He worked with direct additive components in the specification of the utility function. Our model has a more general formulation with nonparametric modeling of the explanatory variables.



### 3.2.1 The general formulation of a GAM

The general form of a GAM can be written as

$$E[y | x] = G\left(\sum_{k=1}^K f_k(x_k)\right). \quad (12)$$

Here,  $G(\cdot)$  is a logistic link function with

$$G(x) = \frac{1}{1 + \exp(-x)}, \quad (13)$$

which implies a parametric description of the error term. The terms of  $f_k$  describe one-dimensional nonparametric functions, which must be estimated, and  $x_k$  are the explanatory variables. In this formulation, the GAM is close to the classic MNL approach with an additive, but nonparametric, utility function  $f_k$  and with the same error term distribution as in the MNL, a logistic i.i.d. one for the error term differences. There exist many alternatives to model the influence of the explanatory variables in a nonparametric way. In the statistical context, modeling by the GAM approach is often used, because it gives the benefit of a large model class including a proved theory.

### 3.2.2 Application of the GAM theory to brand choice models

The general formulation of the GAM has the limitation of being only formulated for continuous explanatory variables. But the usual data sets in the choice-modeling context include discrete, e.g. binary, explanatory variables as well as continuous ones.

For this kind of data structure a new approach is needed that allows binary variables in the model formulation. This extended approach (e.g., Hastie and Tibshirani, 1990) has the form of

$$E[y | x] = G\left(\sum_{k=1}^K f_k(x_k) + \beta^T x_l\right). \quad (14)$$

Here,  $\beta$  describes the parameter vector of the linear part of the model, which must be estimated. The terms  $f_k$  and  $x_k$  have the same meaning as described before. All binary explanatory variables must be included in  $x_l$ . Continuous variables can be modeled parametrically by inclusion in  $x_l$ , or nonparametrically by inclusion in  $x_k$ . The use of the linear formulation of some explanatory variables (in  $x_l$ ) has the advantage of using much less computational time for estimation, because the nonparametric functions  $f_k$  needs most of the estimating time. The extended GAM supplies a good starting point to model choice

behavior in a semiparametric approach. Modeling the nonparametric functions  $f_k$  and the parameter  $\beta$  supplies two possibilities to act with the resulting estimates. One way of using the results is to calculate the conditional expectation  $E[y|x]$  as described in equation (14). The other possibility is to get ideas of a functional form of the underlying explanatory variables  $x_k$  described by nonparametric functions  $f_k$ . Plotting the explanatory variable versus its nonparametric representation should provide some clues. This functional form supplies one possibility of how the explanatory variable could be included in the classic MNL approach. These two possibilities of using the estimates from the extended GAM open up a wide field of research opportunities.

### 3.2.3 Estimation algorithm for the GAM

Estimation of the GAM can be performed by the backfitting algorithm, introduced by several authors (e.g., Friedman and Stuetzle, 1981; Hastie and Tibshirani, 1986; Buja et al., 1989; Hastie and Tibshirani, 1990). This method works with variance decomposition of the additive part of the model as described in equation (14). It can also be described as a projection from a Hilbert space into a lower dimensional subspace. The backfitting algorithm deals with the assumption of an additive form of the “true” underlying model and splits the whole variance into components. But if the underlying model is misspecified (e.g. by omitting variables or interactions), the estimates for the nonparametric functions  $f_k$  can only be interpreted with care, because you do not know exactly which other additional parts are in the estimated  $f_k$ . Here you deal with a bias which cannot be attached to the separate explanatory variables. Another problem of this method consists in the lack of statistical measures, e.g. standard errors or other asymptotic properties. Also the exact behavior of the algorithm is unknown, even the convergence to the correct nonparametric functions is not established (Hastie and Tibshirani, 1990, pp. 117-118). The advantage of this estimation method is that it is an established algorithm, which is implemented in many common statistical software packages (e.g., S-Plus, R, XploRe etc.). And if the underlying data structure is specified correctly, the method gives usually quickly correct estimates for  $f_k$ .

There exists a second method for estimating a GAM, the marginal integration estimator (Chen et al., 1995; Linton and Nielsen, 1995; Linton and Härdle, 1996; Sperlich et al., 1997; Nielsen and Linton, 1998). This method estimates the marginal influence of the additive components specified in equation (14). It works on integration over a product kernel, usually the Nadaraya-Watson estimator. Because this method belongs to the well-known field of kernel estimation, all common statistical measures, e.g., bias, variance, confidence intervals and other asymptotic properties are available. Also, the estimated functional form of  $f_k$  is correct, even if the model is to a certain degree misspecified for the underlying data set. But if strong interaction effects exist in the data, the estimates for

the  $f_k$  of the marginal effects are not sufficient for interpreting the results. The method is not well known and due to this, it is only implemented in one statistical software package, XploRe (Härdle et al., 2000).

If a GAM is estimated by both methods, and the estimation results for the functional forms of the explanatory variables differ greatly, then this is usually due to interaction terms, which are not taken into consideration by the backfitting algorithm. If different results are observed usually the underlying model structure is misspecified. Here, more precise analysis is needed. First approaches to estimating interaction terms via the marginal integration estimator are made by Sperlich et al. (1998).

#### 4. Application of the model to panel data of consumer brand choice

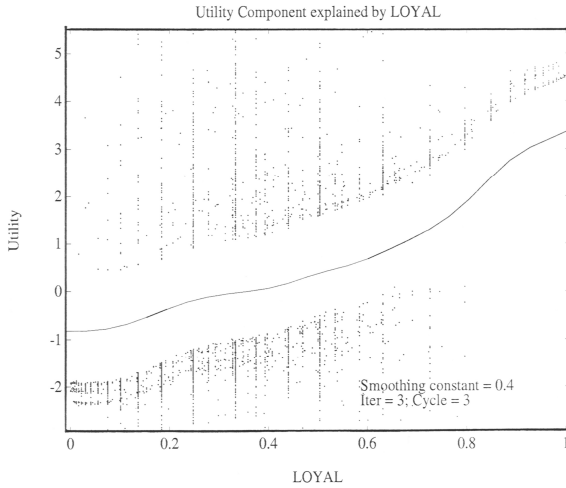
In this part of the paper we apply the two estimation algorithms for a GAM to a real data set in a consumer research setting. GfK, Germany, provided the data. They contain panel purchase records at one store over a period of 104 weeks. Also included are price and binary promotion indicator variables (feature and display) for each brand. We create a subset of the data by extracting purchases of panelists who have bought only one of the three leading brands. This results in a database with 2651 purchases made by 964 households.

The aim of the analysis is first to discover what kind of shape the response function of the marketing instrument price has, controlling for other impact variables (loyalty), instruments (promotion) and the activities of the competing brands using backfitting. The structure of the model is comparable to a model of van Heerde et al. (1998) of the shape of the price response with store level data. Second, we use marginal integration estimation to estimate these effects again, but allowing for interaction between the key variables. If the results are different we are able to conclude that the model based only on direct (order 1) effects may be misspecified.

For the model specification, we use two continuous explanatory variables, *PRICE* and *LOYALTY*, and one binary explanatory variable, *PROMOTION*. *LOYALTY* is defined comparable to Guadagni and Little (1983), a weighted geometric sum over past purchases, to capture household heterogeneity through the purchase history. *PROMOTION* is defined to be 0 if neither feature nor display occurred and 1 otherwise. This is defined this way due to high correlation between the two promotional activities feature and display.

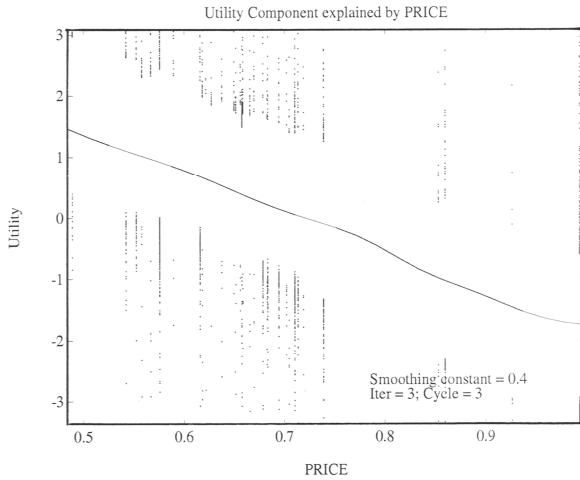
First, we estimated this semiparametric model with backfitting. As shown in Figures 1a and 1b, utility increases with *LOYALTY* in a slightly nonlinear fashion and decreases linearly with *PRICE*.

Figure 1a: Estimation results of the continuous variables made by backfitting



Source: ZUMA data of GfK Consumer Panel 1995, own calculations

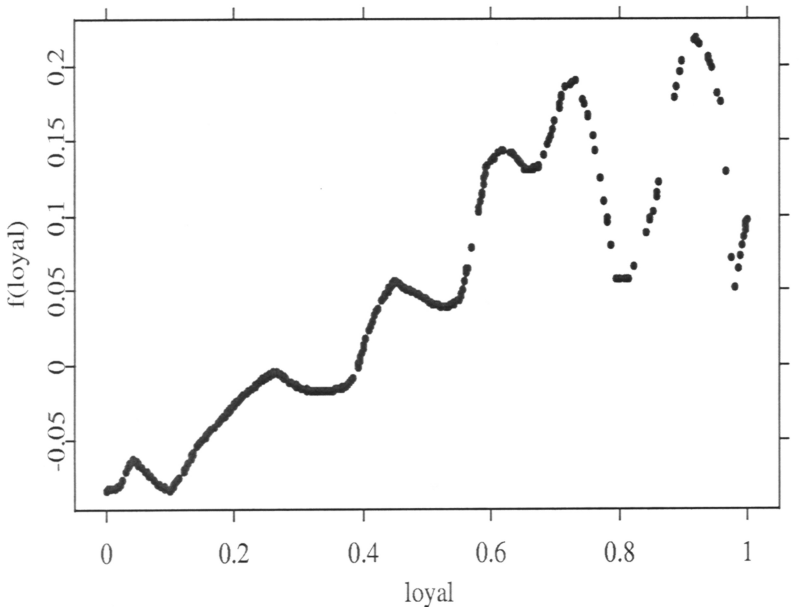
Figure 1b: Estimation results of the continuous variables made by backfitting



Source: ZUMA data of GfK Consumer Panel 1995, own calculations

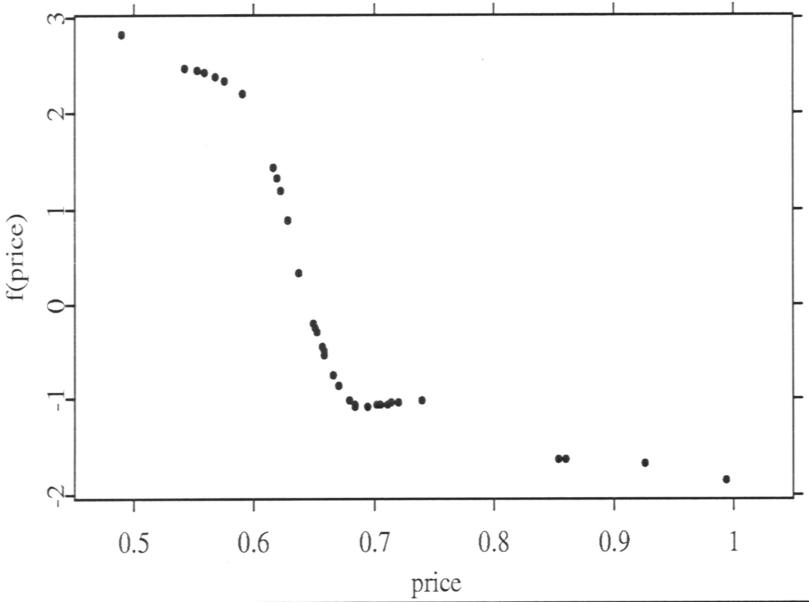
The model was also estimated with marginal integration. The estimation of this small model took a much larger computation time (50x). The results are shown in Figures 2a and 2b. The functional form of *LOYALTY* is again described in a linear increasing way (the slopes at the right side are due to bandwidth effects). But the estimated functional form of *PRICE* indicates a polynomial form of order 3, which is far from a linear form. The form could also be interpreted as a “loss and gain” function, which is often assumed for a price process. This estimation result is different from the one estimated by the backfitting algorithm which leads to the assumption that the underlying data structure may have an interaction effect of *PRICE* with other variables in the model. Therefore the estimation of the model by the backfitting procedure might produce results, which in this case are not valid and may be misleading if the interaction is ignored.

Figure 2a: Estimation results of the continuous variables made by marginal integration



Source: ZUMA data of GfK Consumer Panel 1995, own calculations

Figure 2b: Estimation results of the continuous variables made by marginal integration



Source: ZUMA data of GfK Consumer Panel 1995, own calculations

We have to keep in mind that the marginal integration estimation indicated an additional effect due to the different shapes of the functions of the price. To get a valid model a search for and integration of an additional effect might be necessary.

## 5. Summary

From the statistical point of view a nonparametric formulation of a brand choice model (NDE) is a powerful alternative to the logit model. But in the marketing context, researchers in general want to have parameter values to make predictions or to estimate market shares. This leads to a semiparametric model (GAM) formulation with two possible ways of using the results. One is to perform estimation of choice probabilities, but there one is confronted with the same problem as in the nonparametric approach, because no parameters are estimated for the nonparametric part of the model. The second possibility of a semiparametric model formulation overcomes this problem. In addition, with the estimation results a modified parametric model formulation can be estimated. This

also gives the possibility to work with the parameter values to estimate market shares or make predictions. Especially for this use of modeling, the underlying data structure should be detected correctly. Therefore, two different estimation algorithms for a GAM were presented and the application of the semiparametric model to a real data set was reported. The estimations were made by the two common algorithms, backfitting and marginal integration, and are compared to each other. An interaction effect in the variable price in the data set was discovered, which leads to the need of additional studies of the data set.

## Contact

Yasemin Boztuğ and Lutz Hildebrandt  
Humboldt University of Berlin  
Institute of Marketing; SFB 373  
Spandauer Str. 1  
D-10178 Berlin  
boztug@wiwi.hu-berlin.de

## References

- Abe, Makoto, 1995, A Nonparametric Density Estimation Method for Brand Choice using Scanner Data. In: *Marketing Science*, 14 (3), 300-325.
- Abe, Makoto, 1997, A Generalized Additive Model for Discrete Choice Data. Technical report, The University of Illinois at Chicago.
- Aitchison, J., C. G. G. Aitken, 1976, Multivariate binary discrimination by the kernel method. In: *Biometrika*, 63, 413-420.
- Bartels, Knut, Yasemin Boztuğ, Marlene Müller, 2000, Testing the multinomial logit model. In: Wolfgang Gaul, Reinhold Decker (Hg.), *Classification and Information Processing at the Turn of the Millennium*, Heidelberg: Springer.
- Ben-Akiva, Moshe, Steven R. Lerman, 1985, *Discrete Choice Analysis*. Cambridge: The MIT Press.
- Ben-Akiva, Moshe, Daniel McFadden, Makoto Abe, Ulf Böckenholt, Denis Bolduc, Dinesh Gopinath, Takayuki Morikawa, Venkatram Ramawamy, David Revelt, Dan Steinberg, 1997, Modeling Methods for Discrete Choice Analysis. In: *Marketing Letters*, 3 (3), 273-286.
- Buja, Andreas, Trevor Hastie, Robert Tibshirani, 1989, Linear Smoothers and Additive Models. In: *The Annals of Statistics*, 17 (2), 453-555 (with discussion).

Chen, R., Wolfgang Härdle, Oliver B. Linton, E. Severance-Lossin, 1995, Nonparametric Estimation of Additive Separable Regression Models. Discussion Paper, Sonderforschungsbereich 373.

Chintagunta, Pradeep K., 1992, Estimating a multinomial probit model of brand choice using the method of simulated moments. In: *Marketing Science*, 3 (3), 386-407.

Chintagunta, Pradeep K., 1993, Investigating Purchase Incidence, Brand Choice and Purchase Quantity Decisions of Households. In: *Marketing Science*, 12 (2), 184-208.

Chintagunta, Pradeep K., Bo E. Honore, 1996, Investigating the effect of marketing variables and unobserved heterogeneity in a multinomial probit model. In: *International Journal of Research in Marketing*, 13, 1-15.

Erdem, Tülin, 1996, A Dynamic Analysis of Market Structure Based on Panel Data. In: *Marketing Science*, 15 (4), 359-378.

Erdem, Tülin, Michael P. Keane, 1996, Decision-Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Market. In: *Marketing Science*, 15 (1), 1-20.

Fader, Peter S., James M. Lattin, John D. C. Little, 1992, Estimating Nonlinear Parameters in the Multinomial Logit Model. In: *Marketing Science*, 11(4), 372-385.

Fan, Jianqing, James S. Marron, 1992, Best Possible Constant for Bandwidth Selection. In: *The Annals of Statistics*, 29 (4), 2057-2070.

Friedman, Jerome H., Werner Stuetzle, 1981, Projection pursuit regression. In: *Journal of the American Statistical Association*, 76, 817-823.

Guadagni, Peter M., John D. C. Little, 1983, A Logit Model of Brand Choice Calibrated on Scanner Data. In: *Marketing Science*, 2, 203-238.

Gupta, Sachin, Pradeep K. Chintagunta, 1994, On Using Demographic Variables to Determine Segment Membership in Logit Mixture Models. In: *Journal of Marketing Research*, 31, 128-136.

Härdle, Wolfgang, 1991, *Smoothing Techniques With Implementation in S*. New York: Springer.

Härdle, Wolfgang, Sigbert Klinke, Marlene Müller, 2000, *XploRe Learning Guide*. Berlin: Springer.

Hastie, Trevor J., Robert J. Tibshirani, 1986, Generalized Additive Models. In: *Statistical Science*, 1 (3), 297-318.

Hastie, Trevor J., Robert J. Tibshirani, 1987, Generalized Additive Models: Some Applications. In: *Journal of the American Statistical Association*, 82 (398), 371-386.

Hastie, Trevor J., Robert J. Tibshirani, 1990, *Generalized Additive Models*. London: Chapman and Hall.



- Heerde, Harald J. van, Peter S. H. Leeflang, Dick R. Wittink, 1998, Semiparametric analysis to estimate the deal effect curve. Working Paper, University of Groningen.
- Horowitz, Joel L., Denis Bolduc, Suresh Divakar, John Geweke, Füsün Gönül, Vassilis Hajivassiliou, Frank S. Koppelman, Michael P. Keane, Rosa L. Matzkin, Peter Rossi, Paul Ruud, 1994, Advances in Random Utility Models. In: *Marketing Letters*, 5, 311-322.
- Horowitz, Joel L., Wolfgang Härdle, 1996, Direct Semiparametric Estimation of Single-Index Models with Discrete Covariates. In: *Journal of the American Statistical Association*, 91 (436), 1632-1640.
- Hosmer, David W., Stanley Lemeshow, 1989, *Applied Logistic Regression*. New York: Wiley.
- Kannan, P. K., Gordon P. Wright, 1991, Modeling and Testing Structured Markets: A Nested Logit Approach. In: *Marketing Science*, 10 (1), 58-82.
- Linton, Oliver B., Wolfgang Härdle, 1996, Estimation of additive regression models with known links. In: *Biometrika*, 83 (3), 529-540.
- Linton, Oliver B., Jens Perch Nielsen, 1995, A kernel method of estimating structured nonparametric regression based on marginal integration. In: *Biometrika*, 82 (1), 93-100.
- Manrai, Ajay K., 1995, Mathematical models of brand choice behavior. In: *European Journal of Operational Research*, 82, 1-17.
- Matzkin, Rosa L., 1991, Semiparametric Estimation of Monotone and Concave Utility Functions for Polychotomous Choice Models. In: *Econometrica*, 59 (5), 1315-1327.
- McFadden, Daniel, 1974, Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Hg.): *Frontiers in Econometrics*, New York, 105-142.
- Nielsen, Jens Perch, Oliver B. Linton, 1998, An optimization interpretation of integration and back-fitting estimators for separable nonparametric models. In: *Journal of the Royal Statistical Society, Series B*, 60 (1), 217-222.
- Papatla, Purushottam, 1996, A Multiplicative Fixed-effects Model of Consumer Choice. In: *Marketing Science*, 15 (3), 243-261.
- Silverman, B. W., 1986, *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.
- Sperlich, Stefan, Oliver B. Linton, Wolfgang Härdle, 1997, A Simulation Comparison between Integration and Backfitting Methods of Estimation Separable Nonparametric Regression Models. Discussion Paper, Sonderforschungsbereich 373.
- Sperlich, Stefan, Dag Tjøstheim, Lijian Yang, 1998, Nonparametric Estimation and Testing of Interaction in Additive Models. Discussion Paper, Sonderforschungsbereich 373.

# THE ZUMA DATA FILE VERSION OF THE GfK CONSUMERSCAN HOUSEHOLD PANEL

*GEORGIOS PAPASTEFANOU*

## 1. Data structure

The ZUMA scientific use file of consumer panel data is based on the GfK ConsumerScan household panel. In this consumer panel GfK is running two independent samples with about 12000 household. They are designed as quota samples, quoted by federal state, administrative district, size of town and household size, with marginal distributions adapted to Mikrozensus distributions. The data, which are delivered by ZUMA as scientific use files are subsamples, containing those households, which in 1995 reported continuously their purchase diaries. These two subsamples comprise 9064 households.

The ZUMA version file consists of two main types of data: a) data on each household's product purchases in 1995 and b) data on the socio-economic and demographic structure of the households at the beginning of 1994, 1995 and 1996, as well as data on attitudes towards nutrition, environment and consumption of the person running the household in october 1994 and october 1995. These two types of data can be merged by household identification number.

### *Household purchases in 1995*

Core of consumer panel data collection is the history of a household's purchases all over the time with precise information on the purchases's point of time. Usually, household budget diaries were used for measurement. Households voluntarily use a booklet, in which they note all the purchases plus some characteristics.<sup>1</sup> The ZUMA scientific use file consumer data are bases on this type of paper-and-pencil household budet diaries.

---

1 Because of technological and product diversity the paper and pencil approach is replaced by home screening method of data collection. The participating households do have a manual screening device, by which they can collect the information of EAN code. The data then automatically are stored in a base

---

Methodologically, the collected data are event history data, which are not gathered retrospectively, but in a kind of measurement by process, in continuous time. By this, the data are organized as an asymmetrical data set, in which the purchased product is unit of observation. So, for each household there may be an unequal number of purchase events. Their household assignment is achieved by an added household identification number, being constant for each purchase records, as long it does belong to the same household.

The GfK ConsumerScan household panel is focused on fast moving consumer goods, especially on packed nutrition products and products of household convenience. For each purchased product there are available several variables:

- *Date of purchase*
- *weekday of purchase*
- *product category (measured in total purchases in 81 product categories, see table 1)*
- *product subcategory (like flavor, taste etc.)*
- *type of retailer*
- *number of purchased products*
- *type of price (standard vs. special price)*
- *total quantity purchased*
- *total amount of expenses*
- *waiting time since last purchase*
- *sort id of purchase (the sort id indicates the sequence number of purchase since the beginning of 1995)*
- *specific characteristics of the product, like type of package etc.*

---

station and are transferred later by telephone to the GfK data bank of the consumer panel. The home screening method was installed at GfK in 1998.

Table 1: Product categories of the GfK ConsumerScan Household Panel as being available in the ZUMA file version

No of product category	Sample 1	Sample 2	Label of product category	No of purchased products in 1995
0	x	X	Window/carpet/toilet-cleaner	25.059
1	x	x	Only sample 1: tomato puree, only 6: pastasauces, both: Ketchup, spicesauces	47.651
2	x	x	Only sample 1: mayonese, tartar sauce, both: dressing	27.527
3	x	x	Special detergents	22.304
4	x	x	Detergents for dishes	34.556
5	x	x	Detergents for the household	20.940
6	x	x	Only 1: denture detergents both: toothbrushes/mouthwash	26.818
7		x	Tinned vegetables	63.275
8	x		Milk	204.339
10	x	x	Water softener	29.542
11	x	x	Dental care products	55.497
12	x	x	Pure coffee (roasted)	143.194
13	x	x	Pure coffee-instand	31.576
14		x	Salt	13.100
16		x	Poultry	33.907
17	x	x	Frozen food	230.841
18	x	x	Tea	47.426
19	x	x	Cocoa	17.680
20	x	x	Spirituos beverages	59.129
21	x	x	General detergents	41.225
22		x	Fats	233.124
23	x		Mustard / horseradish	15.055
24		x	Prefabricated cake	23.931
25	x		Soap/washing lotion	13.105
26	x	x	Sparkling wine	30.127
27	x	x	Cider	3.229
28	x	x	Ground care	4.536
29	x	x	Bathing additives	33.206
30	x	x	Finished potato products	39.077
31	x		Pudding/dessert	33.399
32	x		Household tissues	12.265

Table 1 (continued): Product categories of the GfK ConsumerScan Panels as being available in the ZUMA file version

33	x	x	Beer	131.245
34	x	x	Vermouth / appetizer	2.767
35	x	x	Wine/ mulled wine	27.614
36	x	x	Alcoholfree uncarbonated beverages (fruit juice)	187.349
37		x	Winned cabbage	48.502
38		x	Crispbread	17.367
39	x		Shoe- and leather care	3.584
40	x	x	Salt for the dishwasher	4.703
42	x	x	Coffee stuff (like Caro)	3.564
43		x	Delicatess salads	41.375
44	x	x	Sherry/Port wine	3.343
45	x		Pudding powder	14.454
46	x	x	Alcoholfree carbonated beverages (no mineral water)	155.447
47	x		Cream cheese / soft cheese	144.046
48	x	x	Detergents for curtain	6.364
50	x		Cream	74.009
51	x	x	Milk for coffee	111.247
52	x		Air fragrant	2.683
53	x	x	Beverages with wine	5.292
54		x	Gingerbread etc.	19.778
55	x		Napkins	4.432
56	x		Lights, matches etc.	2.730
57		x	Vinegar	10.925
59	x		Food care	1.113
64		x	Sweet pastries	59.473
66	x	x	Animal food/ cat litter (not prefabricated: only 6)	123.133
68		x	Rusk	7.355
71		x	Dry ready-to-serve meals	9.369
72		x	Products for plants	3.598
73	x		Curd/ junkets	94.575
74		x	Salty pastries	19.174
75	x		Filter paper	12.648

*Table 1 (continued): Product categories of the GfK ConsumerScan Panels as being available in the ZUMA file version*

76		x	Instant soups	6.062
77		x	Canned meals	11.244
78	x		Yoghurt	119.270
79		x	Flour	25.886
80	x		Cleaning sponge	9.417
81	x		Cheese	96.766
82	x	x	Detergents for the bath	6.278
83		x	Baking powder / yeast	19.084
84	x	x	Mineral water	174.470
85	x		Sandwich paper	18.756
86	x	x	Icecream	59.108
87		x	Ready-to-bake blends	6.875
89		x	Rice	16.158
90	x		Cereals (Conflakes)	33.256
91		x	Pasta	53.201
93		x	Ready-made menus	4.641
94		x	Roasted nuts etc.	58.008
99	x		Toilet paper	3.245

***Socio-economic and demographic situation of the household and attitudes of the person running the household***

The socio-economic and demographic situation of the household is measured by a paper and pencil survey at the beginning of each year respectively. The following household characteristics are available in the ZUMA dataset:

- federal state of residence
- size of community
- age of the household leading person
- number of children (up to the age of 6, under the age of 14, under the age of 18)
- occupational status of the household leading person
- current occupation of the main income earner
- occupation of the main income earner
- former occupation of the main income earner
- educational status of the main income earner
- size of house
- equipment of the household (washing machine, dishwasher, microwave, VCR)

- garden ownership/-use
- size of garden
- pets (cat, dog, budgie)

The socio-economic and demographic data in the ZUMA dataset are available for three waves of the ConsumerScan panel, namely of 1994, 1995 and 1996. The files of the different years can be merged by a household identification number, which makes possible analyzing individual household structure changes.

### *Attitudes of the person running the household*

Surveys on attitudes of specific persons in the household are inserted into the consumer panel.

Assuming, that persons, who are running the household represent some function of gatekeepers, they are the reference persons for the attitudinal survey. The ZUMA file version contains two attitude surveys of Octobre 1994 and Octobre 1995. These data are part of the variable list in the household structure file and can – like them – be merged with the product purchase history data.

The attitude measures cover the following items:

Attitudes towards nutrition: Items on medical health, natural, joy of discovering, german products, convenience orientation, slimness orientation, plain fare, full grain nutrition, superior savour, freshness orientation, pro branded goods, pro vitamins/minerals, un-critical style of nutrition

Attitudes towards aspects of daily life: Items on tendency to innovate, traditional living, experience-orientation, nostalgia, mistrust towards new products, quality-orientation, convenience-orientated cooking

Attitudes towards environment: Items on ecological awareness, environmental behaviour, environment and mobility, state and industry

Price consciousness of the household leading person: This variable contains a composite index which is based on ratings of three item pairs.

## **2. How to combine purchase history data and household structural data**

As we made clear above, socioeconomic and attitudinal data are organized in one type of file with a rectangular data matrix. This means, that for each observation unit there is on

data record. This type of data we will call household structure data (including the attitudinal data of a reference person). The purchase history data are asymmetrical data files.

The purchase history data of 1995 are organized as separate files for each product category. Be aware, that these files do contain all the purchase of households of both samples, and that some of the socioeconomic demographic or attitudinal data might not be available for both samples.

To combine purchase history data with household structure data, one has to transform the asymmetrical purchase history file into a rectangular data matrix. This can be done by a preselection or by household specific aggregation of purchase variables.

In case of preselection a specific purchase in the order of all purchase will be selected by using the sort id. Thereby it is possible, to condense the asymmetric purchase history into a rectangular household data matrix with the first or the second or the third etc. purchase. It is not possible to choose by this preselection the last purchase done in 1995, because there is not known the sort id of the last purchase in each household.

For household specific aggregation of purchase variables in SPSS AGGREGATE procedure there are different functions. The main logic is to indicate the household context by a break variable (which in our case is represented by the household identification number), which is on each purchase record, but which changes if the household assignment changes. According this household context the purchase variables can be aggregated into one value by using different functions mean, standard deviation, minimum value, maximum value, sum of values, number of purchases for each household (in SPSS terminology these are the number of cases for each value of the break variable). Besides these central tendencies of the household specific distribution of the purchase variables one can also get statistics for parts of the household specific distributions. To get this one has to define a threshold value or an interval and then to compute the percentage or the fraction of purchases being below resp. above the threshold or inside resp. outside the interval.

At last it might be interesting, that by the AGGREGATE procedure one can also run a selection of the first or the last purchase record and its variables.

By aggregating the household's purchase history into one central tendency value one gets a new file containing the household id number plus the aggregated values of the purchase variables. This is a rectangular data matrix which can be merged with the data matrix on household structure and attitudes measures. So, covariances of central behavioral tendencies of the household with its socio-economic structure and the attitudes of its reference person can be analyzed.



## **Einverständniserklärung**

Der Datenempfänger verpflichtet sich durch Unterschrift, nach Erhalt der ZUMA-Daten des GfK-Verbraucherpanels von 1995 zur Einhaltung folgender Bedingungen:

- 1 Der Datenempfänger und alle Personen, denen die Daten zugänglich gemacht werden, verpflichten sich, keine De-Anonymisierungsversuche zu unternehmen und die Daten nicht an andere Personen - außer den Mitarbeitern des Datenempfängers - weiterzugeben oder sie ihnen zugänglich zu machen. Dies gilt auch für modifizierte Daten.
- 2 Die Daten dürfen nur in der eigenen wissenschaftlichen Forschung und der universitären Lehre des Datenempfängers eingesetzt werden. Eine Nutzung für gewerbliche Zwecke ist nicht gestattet.
- 3 Für die Überlassung des Nutzungsrechts an den Daten wird ein Preis von Euro 20.- inkl. MWST\* vereinbart. Der Betrag wird mit meinem Einverständnis von meinem Konto abgebucht.
- 4 In den Forschungsarbeiten, in denen Ergebnisse der ZUMA-Daten des GfK-Verbraucherpanels von 1995 eingehen, wird folgender Quellenverweis aufgenommen:

*Die in diesem Beitrag benutzten Daten entstammen dem Datensatz, den die GfK, Nürnberg als Unterstichprobe des ConsumerScan Haushaltspanels von 1995 ZUMA zur Verfügung gestellt hat. Dieser ZUMA-Datensatz enthält alle Haushalte, für die im Jahr 1995 durchgehend Kaufdaten gesammelt worden sind. Für eine nähere Beschreibung der Verbraucherpaneldaten siehe: Papastefanou, Georgios: The ZUMA scientific use file of the GfK ConsumerScan Household Panel, in: Papastefanou, Georgios / Schmidt, Peter / Börsch-Supan, Axel / Lüdtke, Hartmut / Oltersdorf, Ulrich (eds.), 2001, Social and Economic Analyses of Consumer Panel Data, Zentrum für Umfragen, Meinungen und Analysen (ZUMA), Mannheim.*

---

\* Dieser Preis ist ein Subskriptionspreis. Er gilt in Verbindung mit Erwerb des ZUMA –Spezial Nr. 7 und ist gültig bis 31.12.2001.

Mit meiner Unterschrift verpflichte ich mich die umseitig aufgeführten Bedingungen einzuhalten.

Adresse des Datenempfängers

Ort, Datum

Unterschrift

*Einzugsermächtigung*

Ich erteile hiermit dem Zentrum für Umfragen, Methoden und Analysen (ZUMA) die Berechtigung, den Betrag von Euro 20.- von meinem Konto Nr. .... bei Bankinstitut ..... abzubuchen.

Ort, Datum

Unterschrift

An  
ZUMA -Abt. Einkommen und Verbrauch  
Postfach 122155  
68072 Mannheim

---

## Terms of Agreement

By signing this document the recipient of data complies to the following rules concerning usage of the ZUMA scientific use file of the 1995 GfK ConsumerScan Household Panel Data (in short: ZUMA/GfK Data):

1. The recipient of the ZUMA/GfK Data, as well as all other persons to whom these data are made accessible, are committed not to undertake any attempts to de-anonymize these data. The data must not be made accessible to anyone with the exception of collaborators of the data recipient. This applies also to modified versions of these data.
2. The ZUMA/GfK Data must not be used for purposes other than the data recipient's personal scientific research and academic teaching. It is not allowed to use the data for commercial purposes.
3. The data recipient will pay a fee of Euro 20.-\* for obtaining the ZUMA/GfK Data and being granted the right to use them for scientific purposes as indicated above.
4. All publications reporting findings derived from the ZUMA/GfK Data must refer to the data source by including the following annotation:

*The data used for this analysis are part of a subsample of the 1995 GfK ConsumerScan Household panel data and were made accessible by ZUMA. The ZUMA data set includes all households having continuously reported product purchases during the entire year 1995. For a description of this data set cf. Papastefanou, Georgios, 2001, The ZUMA scientific use file of the GfK ConsumerScan Household Panel, in: Papastefanou, Georgios / Schmidt, Peter / Börsch-Supan, Axel / Lüdtke, Hartmut / Oltersdorf, Ulrich (eds.), 2001, Social and Economic Analyses of Consumer Panel Data, Zentrum für Umfragen, Meinungen und Analysen (ZUMA), Mannheim.*

Name and address of data recipient

Date and signature of data recipient

---

\* This subscription price is effective until December 31, 2001, and only in combination with the purchase of this ZUMA Special Volume No. 7.

## LIST OF CONTRIBUTORS

### **Berger, Roger**

Institut für Soziologie  
Universität München  
Konradstr. 6  
D-80801 München  
roger.berger@soziologie.uni-muenchen.de

### **Boztuğ, Yasemin**

Institut für Marketing; SFB 373  
Humboldt Universität zu Berlin  
Spandauer Str. 1  
D-10178 Berlin  
boztug@wiwi.hu-berlin.de

### **Fachinger, Uwe**

Zentrum für Sozialpolitik  
Universität Bremen  
Parkallee 39  
D-28209 Bremen  
ufach@zes.uni-bremen.de

### **Fengler, Matthias**

Institut Statistik und Ökonometrie  
Humboldt Universität zu Berlin  
Spandauer Str. 1  
D-10178 Berlin  
fengler@wiwi.hu-berlin.de

### **Hartmann, Kerstin**

Institut für Soziologie  
Johannes Gutenberg-Universität  
Colonel-Kleinmann-Weg 2  
D-55099 Mainz  
kerstin.hartmann@uni-mainz.de

**Hildebrandt, Lutz**

Institut für Marketing; SFB 373  
Humboldt Universität zu Berlin  
Spandauer Str. 1  
D-10178 Berlin  
hildebr@wiwi.hu-berlin.de

**Lüdtke, Hartmut**

Institut für Soziologie  
Philipps-Universität Marburg  
Ketzerbach 11  
D-35032 Marburg  
luedtkeH@mail.uni-marburg.de

**Ohr, Dieter**

Institut für Angewandte Sozialforschung  
Universität zu Köln  
Greinstarße 2  
D-50939 Köln  
dieter.ohr@uni-koeln.de

**Papastefanou, Georgios**

Abteilung Einkommen und Verbrauch  
Zentrum für Umfragen, Methoden und Analysen (ZUMA)  
Postfach 12 21 55  
D-68072 Mannheim  
papastefanou@zuma-mannheim.de

**Prester, Hans-Georg**

Operations Consumer Panel  
GfK AG, Panel Services  
Nordwestring 101  
D-90309 Nürnberg  
hans-georg.prester@gfk.de

**Rosenkranz, Doris**

Fachbereich Sozialwesen - Pflegemanagement  
Fachhochschule Würzburg  
Münzstraße 12  
D-97070 Würzburg  
Doris.Rosenkranz@soziale-arbeit.fh-wuerzburg.de

**Schneider, Jörg**

Institut für Soziologie  
Philipps-Universität Marburg  
Ketzlerbach 11  
D-35032 Marburg  
schneidj@mail.uni-marburg.de

**Schneider, Norbert F.**

Institut für Soziologie  
Johannes Gutenberg-Universität  
Colonel-Kleinmann-Weg 2  
D-55099 Mainz  
Norbert.Schneider@uni-mainz.de

**Thøgersen, John**

Department of Marketing  
Aarhus School of Business  
Haslegaardsvej 10  
DK-8210 Aarhus V  
John.Thogersen@mar.hha.dk

**Winter, Joachim**

Sonderforschungsbereich 504  
Universität Mannheim  
L13, 15  
D-68131 Mannheim  
winter@uni-mannheim.de