Statistics Norway
Research Department

Yun Li

Documents

An Analysis of the Demand for Selected Durables in China

Statistics Norway
Statistisk sentralbyrå

Documents 99/13 • Statistics Norway, June 1999

Yun Li

An Analysis of the Demand for Selected Durables in China

Abstract:
Traditional consumer theory assuming infinite divisibility of consumer goods is not appropriate for the analysis of the demand of durable goods. In addition, traditional approaches to consumer demand modeling ignores the problem associated with product heterogeneity when important product characteristics are latent. The point of departure in the present study is a particular framework developed in Dagsvik (1996a,b) and Dagsvik et al. (1998). In this approach the consumer is assumed to make his choice form a discrete set of product variants, and the resulting choice probabilities are derived from behavior arguments. The empirical application is based on household consumption data abstracted from the State Statistical Bureau's Urban Household Survey (UHS) of Sichuan and Liaoning provinces during the 1988-1990 period.

Address: Yun Li, Statistics Norway, Research Department, P.O.Box 8131 Dep., N-0033 Oslo, Norway.
E-mail: yun.lie@ssb.no.

Contents

1. Introduction 3

2. A general framework of qualitative choice models 4

3. The modeling framework for the demand of durable goods 7
3.1. Heterogeneous products 7
3.2. Microeconomic utility theory 8
3.3. Problems with conventional discrete choice modelling when the choice set of variants are not observed 10
3.4. A general model for analyzing the demand of durable goods 11

4. Model with particular functional form assumptions 15
4.1. Alternative I 15
4.2. An alternative assumption about the $\lambda(p)$-function 20

5. Description of the data and aspects of the economic development in China 21
5.1. A brief sketch of the urbane economic development in Sichuan and Liaoning 21
5.2. The data set 23

6. Estimation strategies and empirical results 25
6.1. Estimation; Method I 25
6.2. Estimation; Method II 30

7. Estimation when market prices are available 34

8. Elasticities 34
8.1. Price elasticity 34
8.2. Income elasticity 36

9. Conclusions 37

References 38

Appendix A 39

List of tables

Table 5.2.1. Summary statistics for the whole sample during 1988-1990. 24
Table 5.2.2. Summary statistics for the households that made a purchase during 1988-1990. 25
Table 6.1.1. c-value (the minimal expenditure per year) 27
Table 6.1.2. Estimation results for refrigerator; method I. 29
Table 6.1.3. Estimation results for color TV; method I 29
Table 6.2.1. Annual estimates for refrigerator when ML method is applied 30
Table 6.2.2. Annual estimates for color TV when ML method is applied 30
Table 6.2.3. Estimation results for refrigerator; method II 33
Table 6.2.4. Estimation results for color TV; method II 33
Table 8.2.1. Income elasticities for refrigerator 37

1. Introduction

The traditional theory of individual choice behavior, as it usually is presented in textbook of consumer theory, presupposes that the goods offered in the market are infinitely divisible. However, this theory is not suitable for the analysis of consumption of durable goods, as one can easily realize. The recent theory of discrete choice can be used to model these kinds of choice settings, and to provide the corresponding econometric methodology for empirical analyses. Since some variables are unobservable to the econometrician (and possibly also to the individual agents themselves), the observations from a sample of agents' discrete choices are, in the theory of discrete choice, viewed as outcomes generated by a stochastic model. The purpose of the theory of discrete choice is to provide a structure of the choice probabilities (the probabilities of choosing the respective alternatives) that can be justified from behavioral arguments, that is to express the choice probabilities as functions of the distribution of agents' preferences and the choice constraints.

Products, such as durables, are often differentiated with respect to quality- and location attributes. Also prices may vary with respect to these attributes. This paper, based on a particular approach that was proposed by Dagsvik (see Dagsvik (1996) and Dagsvik et al.(1998)), develops a framework for analyzing consumer demand for differentiated products taking into account of quality- and location attributes when some of these attributes are unobservable to the analyst.

The approach proposed by Dagsvik uses ideas from the literature on discrete choice to express the probability of choosing a variant with given price and quality as a function of the distribution of preferences, market prices and quality attributes. In the presence of latent quality/location attributes and unobserved heterogeneity in preferences, the distribution of unit prices, i.e. expenditure per unit of the commodity, may differ from the distribution of market prices due to the selection effect that arises from consumers having preferences over the variants. By mean of the model developed below it is possible to derive a convenient expression for the distribution of unit prices as a function of the distribution of market prices and parameters relating to the preferences.

After some key assumptions about the utility functional form and the distribution of the random utilities have been made, we obtain a convenient parametric expression of the distribution of unit values and the probability of buying. We then apply this model to analyze the demand of two durable goods, refrigerator and color TV, in China during the period of 1988-1990. Though in each year there are about 1150 households participating in the survey, the households that did not already own one piece of the durable goods are only about half of the sample, and the number of the households that did buy a durable goods in those respective years are only about one hundred. Two alternative estimation methods will be used to estimate some of the unknown parameters of the model. However, not all of them are identifiable due to the fact that the distribution of market prices is unknown. We also show how to derive price elasticities and income elasticities that perhaps are the most valuable information from a policy point of view. Unfortunately, only income elasticities can be calculated in this study, which is also a consequence of limited data information. But the results are interesting and show that income elasticities of poor families are much higher than those of rich families, which seems very reasonable.

SAS- and TSP programs are used in estimation. The data set consists of individual household data obtained from the State Statistical Bureau's Urban Household Survey (UHS) of Sichuan and Liaoning provinces during the period of 1988-1990. It is a cross sectional one, and we only focus on the consumption of the durables.

The paper is organized as follows. In Section 2 the general choice setting and modeling framework are briefly presented. A basic model of analyzing the demand of durable goods is introduced in Section 3. In section 4, assumptions about functional forms are introduced. An empirical application on microdata from China is in Section 5 and Section 6, where we describe the data and report different estimation results based on several alternative methods. Section 7, some extensions of the model are discussed. The final section contains concluding remarks.

2. A general framework of qualitative choice models

Qualitative choice models (or discrete choice models) are concerned with choice behavior when the set of alternatives is discrete. Such situations arise in a variety of contexts in such areas as transportation, energy, telecommunication, housing, and labor supply, to name a few. Here, the commodity can not be treated as infinitely divisible, as assumed in the textbooks of consumer theory.

The problem confronted by discrete choice theories is how to formulate models of choice from a finite and exhaustive set of mutually exclusive alternatives. The set of alternatives may be "structurally" discrete or only "observationally" discrete. The set of feasible transportation alternatives is an example of a structurally categorical setting while different levels of labor supply such as "part time" and "full time" employment may be interpreted as observationally discrete since the underlying set of feasible alternatives, "hours of work", may be viewed as a continuum. Thus, in practice, the distinction between discrete and continuous dependent variables is not always evident and hence does not automatically serve to guide the researcher in determining whether to use qualitative choice methods. In many situations it may be just a matter of taste, or a strategy decision by the researcher.

Usually, qualitative choice theories lead to formulations of the probability that a decisionmaker will choose a particular alternative from a set of alternatives. The models often differ in the functional form that relates the representation of the choice alternatives to the choice probabilities.

Denote the decisionmaker in a qualitative choice situation by i and the set of alternatives he faces by B, and let z index the alternatives. Let $A_i(z)$ denotes the corresponding response variable, where $A_i(z)=1$ if alternative z is chosen by individual i and zero otherwise. Then $EA_i(z)=P(A_i(z)=1)\cdot 1+P(A_i(z)=0)\cdot 0 \equiv \tilde{P}_i(z)$, where $\tilde{P}_i(z)$ is the probability that individual i shall choose alternative z. We can therefore write

$$A_i(z)=\tilde{P}_i(z)+\mu_i(z), \tag{2.1}$$

where $\mu_i(z)$ is a random variable with zero mean. The "systematic" term $\tilde{P}_i(z)$ can be specified as a function of explanatory variables that consist of attributes associated with the alternatives and characteristics of the individual. Note that $\tilde{P}_i(z)$ depends also on attributes of other alternatives than alternative z. The specific qualitative choice models, such as logit and probit, are obtained by specifying the particular structure of the choice probabilities.

We shall next elaborate upon this general description with concepts from the standard microeconomic theory of utility maximization. Utility theory provides a context for motivating and deriving various specifications of the function relating observed data to the choice probabilities.

The derivation of qualitative choice models from utility theory is based on a precise distinction between the behavior of the decisionmaker and the information of the researcher. Consider first the decisionmaker. The decisionmaker is assumed to have preferences over the alternatives represented by a utility function, $U_i(z)$. The utility of one alternative depends on the attributes of the alternative and the characteristics of the decisionmaker. He will choose, of course, the alternative from which he derives the greatest utility. That is, the decisionmaker will choose alternative z from the choice set B if and only if $U_i(z)>U_i(k)$ for all $k \neq z$ and $k \in B$. This completes the specification of how the decisionmaker behaves. Note that the decisionmakers' choice is assumed to be deterministic: he chooses the alternative with the highest utility.

To specify the choice probabilities, we focus on the researcher. Suppose that a researcher is interested in predicting the decisionmaker's choice. Since the researcher does not observe all the relevant factors and does not know the utility function exactly, the utility of the decisionmaker is partitioned into two parts, one that depends only on factors observed by the researcher and whose form is assumed known by the researcher up to a vector of parameters, to be estimated. Let the component known by the researcher labeled $V_i(z)$, and the other that represents all factors and aspects of utility that are unknown by the researcher be labeled $\varepsilon_i(z)$. The last component $\varepsilon_i(z)$ is assumed to be a random variable. That is,

(2.2) $$U_i(z)=V_i(z)+\varepsilon_i(z) \quad z\in B .$$

The possible sources of uncertainty that contribute to the randomness of the utility function may be: unobservable attributes of the alternatives, unobservable characteristics of the decisionmakers, measurement errors, functional misspecification and bounded rationality of the decisionmakers.

We are now capable to introduce choice probabilities. The probability that person i chooses alternative z, $\widetilde{P}_i(z)$, can be expressed formally as

(2.3) $$\widetilde{P}_i(z)=P\left(U_i(z)=\max_{x\in B} U_i(x)\right)=P\left(U_i(z)>U_i(k),\ \mathit{for\ all}\ k\ \mathit{in}\ B,\ k\neq z\right).$$

By substitution of (2.2), we obtain

(2.4) $$\widetilde{P}_i(z)=P\left(V_i(z)+\varepsilon_i(z)>V_i(k)+\varepsilon_i(k),\ \mathit{for\ all}\ k\ \mathit{in}\ B,\ k\neq z\right).$$

Rearranging, we get

(2.5) $$\widetilde{P}_i(z)=P\left(\varepsilon_i(k)-\varepsilon_i(z)<V_i(z)-V_i(k),\ \mathit{for\ all}\ k\ \mathit{in}\ B,\ k\neq z\right).$$

Since all the $\varepsilon_i(z),\ z \in B$, are random variables, the differences between $\varepsilon_i(z)$ and $\varepsilon_i(k),\ k \neq z$, are also random variables. The researcher can observe $V_i(z),\ z \in B$, so he can calculate all the differences between $V_i(z) - V_i(k),\ k \neq z$. Consequently, the right hand of (2.5) is simply a joint cumulative distribution, namely, the probability that each random variable $\varepsilon_i(k) - \varepsilon_i(z)$ is below $V_i(z) - V_i(k)$, respectively, for all k in the choice set B, $k \neq z$.

By assuming a specific distribution function of the random variable $\varepsilon_i(z)$, the researcher can derive the joint distribution of $\varepsilon_i(k) - \varepsilon_i(z)$ for all k in B, $k \neq z$, and calculate the probability that the decisionmaker will choose alternative z as a function of $V_i(z) - V_i(k)$ for all k in B, k≠z. This function is the one we mentioned above which relates the representation of the alternatives in the choice set B to the choice probabilities.

All qualitative choice models are obtained by specifying some distribution for the unknown component of utility and deriving functions for the choice probabilities. Different qualitative choice models are obtained by specifying different distributions for the εs, giving rise to different functional forms for the choice probabilities.

3. The modeling framework for the demand of durable goods

3.1. Heterogeneous products

Our aim is to estimate a model for the probability of buying a variant of a differentiated product with a given price as a function of the whole price distribution and household characteristics. The model developed in this paper can be applied to many cases, but we are only going to use it to analyze the demand of two durable goods, namely, refrigerator and color TV, in China during the period between 1988 and 1990. Since the products are differentiated, qualitative choice models stand out as the suitable models to do such an analysis.

Take the refrigerator as an example. The Chinese consumers who do not already owned one piece of it face the choice of not buying or buying one variant of the refrigerator each year.

The major difference between the approach discussed in this paper and the conventional discrete choice method is that we will try to take into account of the problem of product heterogeneity.

Products consumed are often differentiated with respect to quality attributes rather than homogeneous as usually assumed in the textbook theory of consumer demand. Also prices may vary with respect to quality attributes as well as with respect to geographical location of the stores. Therefore quality- and location attributes of the goods will matter in addition to prices when the consumer tries to work out the utility he can obtain from the purchase. But these attributes are often unobservable, and thus often are ignored or treated rather superficially in standard empirical demand models. By taking these latent attributes into account we will have to deal with difficult issues in modelling and estimation arising from the fact that a wide variety of differentiated products are offered in a market with heterogeneous consumers and producers.

In fact, consumers face an almost infinite set of different variants/locations characterized by price and quality attributes. Since the variants and stores are unobservable we may, without loss of generality, treat stores and variants symmetrically in the formalism.

3.2. Microeconomic utility theory

One distinction between durable goods and ordinary goods is that a durable good can last for quite a long period and it depreciates gradually. Thereby, the actual price related to the consumption of a durable good is different from the purchase price. The relevant price in this contest is the user cost, which takes into account the depreciation and interest rates. See for example A. Deaton and J. Muellbauer (1980). The user cost is the price that should be taken into the utility function. If we assume that prices, the interest rate and the depreciation rate change only little over time, then the user cost will be proportional to the purchase price. We therefore can use the original purchase prices in our analysis, as it will only have a scalar effect on those parameters.

Since durable goods depreciate only gradually, and those two durable goods were available to Chinese families not so long before the period we are examining, we can just ignore those families that had already owned one piece of them. It is natural to assume that no family will buy one more right after the purchase. Since the durable goods considered in this paper are quite expensive to the Chinese households during that period, those assumptions are not at all restrictive. Apart from durable goods, consumers also demand other goods of which the quantities will be denoted by a vector **q**. Therefore the utility obtained from all these consumptions can be formulated as

(3.2.1) $$W_i(q,z)=\Phi_i(q)+\eta_i(z),$$

where z denotes the variant bought by the consumer, $W_i(\boldsymbol{q}, z)$ is total utility, $\Phi_i(\boldsymbol{q})$ and $\eta_i(z)$ are utilities obtained from purchasing of other goods and durable goods, respectively. We have assumed implicitly in Eq. (3.2.1) that utility generated from consumption of durable goods does not relate to those from other consumptions, which leads naturally to the additively separable structure of utility function.

As usual, consumers are assumed to be pricetakers, which means that for any commodity the consumption of anyone is only a small fraction of all and therefore will not have impact on the price. Then the task of everyone will be to choose how much to buy for each item, given the prices and the consumer's income, such that the total utility can be maximized. Indirect utility function is derived therefore by maximizing utility function (3.2.1) subjected to the prices and the income level of the person/family. Conditional on the purchase of product variant z the budget constraint can be written as

(3.2.2) $$\boldsymbol{q}'\boldsymbol{r} = y - P(z),$$

where **q** is the quantity vector, **r** is the price vector, y is the income, and P(z) is the price of the variant bought by the consumer. Due to the additive structure of the direct utility function, maximization of $W_i(\boldsymbol{q}, z)$ with respect to the vector **q** subject to (3.2.2) is equivalent to maximization of the first term on right side of (3.2.1) with respect to the quantity vector. Thus, the indirect function can be partitioned into two additive terms, one relating to the consumption of ordinary goods, and the other representing the utility obtained from consumption of the durable goods. Let $V_i(\boldsymbol{r}, P(z), y)$ be the total indirect utility. Then

(3.2.3) $$V_i(\boldsymbol{r}, P(z), y) = V_i^*(m, y - P(z)) + \eta_i(z),$$

where m is a price index generalized from the price vector **r,** and $V_i^*(m, y - P(z))$ is the maximal utility that can be obtained from consumption of goods other than the durable one conditioned on that variant z is chosen. From utility theory, one knows that the indirect utility function is homogeneous of degree zero with respect to prices and income, which in other words can be described as that the total utility would not be altered if all the prices and income had changed proportionally. Thus (3.2.3) can be rewritten as

(3.2.4) $$V_i(\mathbf{r}, P(z), y) = V_i^*\left(1, \frac{y-P(z)}{m}\right) + \eta_i(z).$$

Let

(3.2.5) $$u(y-P(z)) = V_i^*\left(1, \frac{y-P(z)}{m}\right),$$

where we for notational simplicity have suppressed m in the notation on the left hand side.

3.3. Problems with conventional discrete choice modelling when the choice set of variants are not observed

Consider now for a moment a standard type of discrete choice setup. Suppose the utility from buying a durable for consumer i has the form

(3.3.1) $$U_{i1} = \eta_{i1} + u(y - \hat{P}_i) + \varepsilon_{i1},$$

where $\hat{P}_i$ denotes unit price, the expenditure on the durable, η_{i1} captures the consumer's utility generated from the consumption of the durable, ε_{i1} is a random term, and U_{i1} is his total utility.

Similarly, the utility of not buying is

(3.3.2) $$U_{i0} = u(y) + \varepsilon_{i0}.$$

A utility-maximizing consumer will choose to buy if $U_{i1} > U_{i0}$, that is

(3.3.3) $$A_i = \begin{cases} 1 & if \quad U_{i1} > U_{i0} \\ 0 & if \quad U_{i1} < U_{i0} \end{cases}.$$

The problem with this formulation is that it has been implicitly assumed that the consumer must choose between not buying and buying a variant with price equaling to $\hat{P}_i$. It ignores the fact that a consumer actually faces a whole set of alternatives, that is, he could have chosen a variant with price other than $\hat{P}_i$. Therefore, the model that will be introduced in next section is more suitable to the situation.

3.4. A general model for analyzing the demand of durable goods

We begin with writing down a general form of utility function for consumer i, that is,

(3.4.1) $$U_i(z) = \eta_i(z) + u(y - P(z)),$$

where for notational simplicity we have replaced $V_i(\boldsymbol{r}, P(z), y)$ by $U_i(z)$.

It can be easily understood that the enjoyment of owning, or more precisely, of using e.g. a refrigerator depends on its quality. Therefore, from the point of view of an econometrician, $\eta_i(z)$ is a random variable that captures the consumer's subjective evaluation of the "quality" of variant z and all the other randomness in the utility function.

Without loss of generality we can write

(3.4.2) $$\eta_i(z) = log\, T(z) + \varepsilon_i(z),$$

where $log\, T(z)$ represents the mean evaluation of the quality of variant z (across consumers) and $\varepsilon_i(z)$ is the error term representing the randomness in the utility function which as mentioned in section 2 is mainly due to the unobservable attributes of the alternatives, the unobservable characteristics of the consumers, measurement errors, functional misspecification and bounded rationality of the consumers. In other words, $\varepsilon_i(z)$ captures the unobservable heterogeneity of preferences. Note that $log\, T(z)$ depends not only on the variant z but also on which group the household belongs to, e.g. the type of the household. For the simplicity of notation, we skip an extra footnote indicating the group belonging at present. However, the dependence will be expressed explicitly in the formulas later on.

Accordingly, the utility of not buying a refrigerator has the general form

(3.4.3) $$U_i(0) = u(y) + \varepsilon_i(0)$$

where $\varepsilon_i(0)$ is a random variable.

We will further assume that $\varepsilon_i(z)$, $z = 1,2,\ldots, i = 1,2,\ldots,$ are identically and independently distributed with the extreme value distribution

$$P(\varepsilon_i(z) \leq x) = exp(-e^{-x}). \tag{3.4.4}$$

One of the possible justifications for this assumption is that it is consistent with the notion of "Independence from Irrelevant Alternatives" that we shall discuss further below.

It follows from (3.4.4) that the probability of choosing a variant z, $\tilde{P}(z)$, from the choice set can be expressed as

$$\tilde{P}(z) = P\left(U_i(z) = \max_{k \in B} U_i(k)\right) = \frac{e^{v(z)}}{e^{v_0} + \sum_{k \in B} e^{v(k)}}, \tag{3.4.5}$$

where

$$v(z) = u(y - P(z)) + log\, T(z), \tag{3.4.6}$$

$$v_0 = u(y) , \tag{3.4.7}$$

and B is the set of available variants and stores in the market. A proof of this result can be found for example in Ben-Akiva and Lerman(1985).

We shall next briefly review an important property of these probabilities. Consider the ratio of the choice probabilities for two alternatives, z and k. We obtain that

$$\frac{\tilde{P}(z)}{\tilde{P}(k)} = \frac{e^{v(z)}}{e^{v(k)}}. \tag{3.4.8}$$

Note that the ratio of these two probabilities does not depend on any alternatives other than z and k. The ratio therefore can be said to be independent from "irrelevant" alternatives, that is, alternatives other than those for which the ratio is calculated. Therefore the logit probabilities exhibit the "Independence from Irrelevant Alternatives property" (IIA property for short). For more details about this IIA property, see for example Ben-Akiva and Lerman(1985).

As mentioned above, a consumer faces a choice set consisting of different variants characterized by price and quality attributes. Let *B(p,t)* be the subset of B consisting of the set of variants in the market with $P(z)=p$ and $T(z)=t$, and let *b(p,t)* be the number of variants *within B(p,t)*. Then, it follows from (3.4.5) that the probability of choosing a variant z with $P(z)=p$ and $T(z)=t$ for consumer i given his income, denoted by $\varphi(p,t|y)$, can be expressed as

(3.4.9)
$$\varphi(p,t\mid y)\equiv P\left(\max_{z\in B(p,t)} U_i(z)=\max\left(U_i(0),\max_{w,r}\max_{z\in B(w,r)} U_i(z)\right)\right)=\sum_{z\in B(p,t)}\tilde{P}(z)$$
$$=\frac{\sum_{z\in B(p,t)} e^{v(z)}}{e^{v_0}+\sum_{k\in B} e^{v(k)}}=\frac{b(p,t)te^{u(y-p)}}{e^{u(y)}+\sum_k\sum_r b(k,r)re^{u(y-k)}}$$

Now let

(3.4.10)
$$b=\sum_w\sum_r b(w,r),$$

and

(3.4.11)
$$g(p,t)=\frac{b(p,t)}{b}.$$

Note that now g(p,t) becomes a density function and is interpreted as the fraction of variants in the market that have price p and mean quality t. Thus, from the point of view of an econometrician we may interpret g(p,t) as the probability that a variant with price p and mean quality t exists in the market. Using these new terms, Eq. (3.4.9) can be rewritten as

(3.4.12)
$$\varphi(p,t \mid y) = \frac{g(p,t)te^{u(y-p)}}{e^{\gamma+u(y)} + \sum_w \sum_r g(w,r)re^{u(y-w)}},$$

where

(3.4.13)
$$\gamma = -\log b.$$

The discrete setting applied above is somewhat unsatisfactory for several reasons. First, the wide variety in product quality, location and service of the stores makes it difficult to classify variants and stores into a few groups. Furthermore, also as a result of the first reason, the distribution of market prices may vary nearly continuously across variants and stores. It is therefore desirable to extend the discrete setting as to allow for continuous distributions of prices and quality attributes. Thus, we may modify Eq. (3.4.12) to a continuous version as

(3.4.14)
$$\varphi(p,t \mid y) = \frac{g(p,t)te^{u(y-p)}}{e^{\gamma+u(y)} + \int\int g(w,r)re^{u(y-w)}dwdr}.$$

Second, only information about prices of the purchased variants is available and quality is an attribute which can not be observed by the econometrician. Thus, we are more interested in knowing the probability that consumer i given his income shall choose a variant with a specific price, and it will be denoted as $\varphi(p|y)$. From (3.4.14) it follows immediately that

(3.4.15)
$$\varphi(p \mid y) \equiv \int \varphi(p,t \mid y)dt = \frac{\int g(p,t)te^{u(y-p)}dt}{e^{\gamma+u(y)} + \int\int g(w,r)re^{u(y-w)}dwdr}.$$

Let

(3.4.16)
$$\lambda(p) \equiv E(T(z) \mid P(z) = p) = \int \frac{tg(p,t)}{g(p)}dt,$$

where

(3.4.17) $$g(p) = \int g(p,t)dt .$$

We can interpret $\lambda(p)$ as the conditional mean of T(z) across variants given the price level. In other words, the function $\lambda(p)$ is in fact a quality index given the price level p. By inserting Eq. (3.4.13) into (3.4.12), we can rewrite the probability for consumer i to choose a variant z with $P(z) = p$ as

(3.4.18) $$\varphi(p \mid y) = \frac{\lambda(p)e^{u(y-p)}g(p)}{e^{\gamma+u(y)} + \int \lambda(w)e^{u(y-w)}g(w)dw} .$$

Similarly, the probability of not buying for consumer i can be expressed as

(3.4.19) $$\varphi(0 \mid y) \equiv 1 - \int \varphi(p \mid y)dp = \frac{e^{\gamma+u(y)}}{e^{\gamma+u(y)} + \int \lambda(w)e^{u(y-w)}g(w)dw} .$$

Eq. (3.4.18) and (3.4.19) are general forms that are not immediately applicable in empirical analysis, since the functional forms of the utility function, the conditional quality index and the distribution of the market prices are all unknown to the econometrician. Therefore we need to make some assumptions about those functional forms, and we will do it in the next section.

4. Model with particular functional form assumptions

4.1. Alternative I

In this section we shall demonstrate that the functional form of probabilities can be simplified dramatically and becomes convenient for empirical analysis by assuming a special form of utility function, distribution of market prices and the quality index function. We assume first the utility function of household i has the form

(4.1.1) $$u(y) = \begin{cases} \beta(y-c), & if \quad y > c \\ -\infty, & otherwise \end{cases} ,$$

where c denotes the subsistence expenditure level or in other words the minimal living expenditure, y is the income of the household, β>0 is an unknown parameter. We see that, the utility a consumer obtains from owning a durable good depends negatively on the price and positively on the mean quality given his (or the family's) annual income, which is in accordance with intuition. Below you shall see the convenience by assuming such a particular form of utility function.

Given those assumptions about the utility function and eq.(3.4.15), the probability of buying a variant with price p given the income can be derived as

(4.1.2) $$\varphi(p \mid y) = \frac{\lambda(p)e^{-\beta p} g(p)}{e^{\gamma} + \int_0^{y-c} \lambda(w)e^{-\beta w} g(w)dw}.$$

Similarly, the probability of not buying given the income can be expressed as

(4.1.3) $$\varphi(0 \mid y) = \frac{e^{\gamma}}{e^{\gamma} + \int_0^{\infty} \lambda(w)e^{-\beta w} g(w)dw}.$$

Before we can proceed to the estimation stage, there are several further assumptions needed to be made. First, we need to make assumption about the functional form of $\lambda(p)$. We will assume for simplicity that

(4.1.4) $$\lambda(p) = n(x)e^{\theta p},$$

where x denotes the group to which the household belongs, $n(x) > 0$ and $\theta \geq 0$ are constants. From a theoretical point of view it is a weakness that we have to make this ad hoc assumption. In such a functional form the mean quality of the variants will increase with the price level, though not in proportion. It is this property and the relative simple function forms that make this assumption attractive. The unknown parameter θ determines how the consumer adjusts his evaluation about the quality when a variation in the price is observed. A high value of θ means that an increase in price will cause a substantial increase in the evaluation of the quality. θ depends generally on the type of the durable.

Secondly, we also need to specify the distribution of market prices, g(p). We assume that market prices are Gamma distributed, apart from a scale parameter. The reason behind this assumption lies in the convenient properties of the Gamma distribution. The importance of these properties will be unfolded later. Let $f(p;s,v)$ denote the scaled Gamma density function with parameters s and ν, then it leads to

(4.1.5) $$g(p) = f(p;s,v) = \frac{s^{v} p^{v-1} e^{-sp}}{\Gamma(v)},$$

where v>0 and s>0 are parameters. From (4.1.5) it follows that

(4.1.6) $$E(P(z)) = \frac{v}{s}$$

and

(4.1.7) $$var(P(z)) = \frac{v}{s^2}.$$

If we insert (4.1.5) and (4.1.4) into (4.1.2), we obtain

(4.1.8) $$\varphi(p \mid y) = \frac{n(x)\left[e^{-(s-h)p}(s-h)^{\nu} p^{\nu-1}/\Gamma(\nu)\right]\left(\frac{s}{s-h}\right)^{\nu}}{e^{\gamma} + n(x)\int_{0}^{y-c}\left[e^{-(s-h)w}(s-h)^{\nu} w^{\nu-1}/\Gamma(\nu)\right]\left(\frac{s}{s-h}\right)^{\nu} dw} = \frac{f(p;s-h,\nu)}{e^{\gamma^{*}(x)} + F(y-c;s-h,\nu)},$$

where

(4.1.9) $$\gamma^{*}(x) = \gamma - \log n(x) - v \log\left(\frac{s}{s-h}\right),$$

$h = \theta - \beta$, and $F(x;s-h,v)$ is the scaled cumulative Gamma distribution function with parameters (*s-h*) and ν. Similarly, (4.1.10) $$\varphi(0 \mid y) = \frac{1}{1 + e^{-\gamma^{*}(x)} F(y - c; s - h, v)}.$$

In Eq. (4.1.8) and (4.1.10) , the probabilities of buying and not buying given the income have been expressed as functions with unknown parameters of the utility function, the quality index function and distribution of market prices. Theoretically, we can apply those to equations directly in estimations program and estimate all the parameters which are identifiable simultaneously. Unfortunately, there are some technical problems of programming which prevent us from doing so. Therefore, we need to derive the following equations, which will make it possible to divide the estimation process into two steps to circumvent those technical problems.

The probability that consumer i, given his income y, shall buy a variant with price p provided that he makes a purchase will be denoted by $\hat{g}(p \mid y)$, and it is given by

(4.1.11) $$\hat{g}(p \mid y) \equiv \frac{\varphi(p \mid y)}{1 - \varphi(0 \mid y)} = \frac{f(p; s - h, v)}{F(y - c; s - h, v)}.$$

Let Y_i denote the actual income of the household i. From (4.1.11) it follows that the unconditional probability of buying a variant with price p given that a purchase takes place, equals

(4.1.12) $$\hat{g}(p) \equiv E\hat{g}(p \mid Y_i) = f(p; s - h, v)K$$

where K is a constant which can be expressed as

(4.1.13) $$K = E\left(\frac{1}{\int_0^{Y_i - c} f(w; s - h, v) dw} \right) = E\left(\frac{1}{F(Y_i - c; s - h, v)} \right).$$

In equation (4.1.13) and (4.1.12) the expectations are evaluated with respect to Y_i. Note that $\hat{g}(p)$ is the (unconditional) probability density function of unit prices. In other words it represents the fraction of households that had bought a durable with price p among those households that had made the purchase. Since both $\hat{g}(p)$ and $f(p; s - h, v)$ are probability density functions, the integral of the probabilities over all the possible value of p is necessarily equal to one. Thus, the constant term K must therefore be equal to one. Consequently, we obtain

$$\hat{g}(p) = f(p; s-h, v). \tag{4.1.14}$$

In words, the distribution of unit prices is equal to a scaled Gamma distribution with parameters (s-h) and v. Notice that the distribution of unit prices depends in a simple way on the distribution of market prices through s and v. This is one advantage that ensues from the assuming that the market prices are Gamma distributed. These two distributions are different due to the selection effect that arises from consumers having preferences over the variants. We are now able to estimate parameters (s-h) and v by using the data of unit prices.

If we insert equation for γ, Eq.(3.4.13), into (4.1.9), we obtain

$$\gamma^*(x) = -log\, b - log\, n(x) - v\, log\left(\frac{s}{s-h}\right). \tag{4.1.15}$$

As mentioned above, $n(x)$ is a constant whose value depends on certain characteristics of the household. We will now assume that

$$\gamma^*(x) = \delta + \tau x_i \tag{4.1.16}$$

where δ and τ are parameters and x_i is a variable which characterizes the household. The probability of not buying given the family income can now be rewritten as

$$\varphi_i(0 \mid Y_i) = \frac{exp(\delta + \tau x_i)}{exp(\delta + \tau x_i) + F_i}, \tag{4.1.17}$$

and correspondingly, the probability of buying a variant with price p can be rewritten as

$$\varphi_i(p \mid Y_i) = \frac{f(p; s-h, v)}{exp(\delta + \tau x_i) + F_i} \tag{4.1.18}$$

where

(4.1.19) $$F_i = F(Y_i - c; s - h, v).$$

We can now estimate parameters δ and τ by applying Eq. (4.1.17) and (4.1.18). This completes the derivation of the probability of buying a variant with price p (unit price). This achievement is reached by some restrictive assumptions about the utility function, the quality index, the distribution of market prices, etc. We however realize that we are unable to identify s, θ and β separately in this setting. In section 7 it will be discussed how one could identify and estimate additional parameters if the distribution of market prices was available.

4.2. An alternative assumption about the $\lambda(p)$-function

We shall now consider the implication from postulating that the $\lambda(p)$ -function has another form, namely

(4.2.1) $$\lambda(p) = n(x)p^{\kappa},$$

where x denotes the group to which the household belongs, $n(x)$ and κ are parameters.

Provided that all the other assumptions are unchanged, the probabilities of buying and not buying given the income can in this case be expressed as

(4.2.2) $$\varphi(p \mid y) = \frac{f(p; s+\beta, v+\kappa)}{e^{-\gamma'(x)} + F(y-c; s+\beta, v+\kappa)},$$

and

(4.2.3) $$\varphi(0 \mid y) = \frac{1}{1 + e^{-\gamma'(x)} F(y-c; s+\beta, v+\kappa)},$$

where

(4.2.4) $$\gamma''(x) = -\log b - \log n(x) - v \log s + (\kappa + v)\log(\beta + s) - \log \Gamma(v+\kappa) + \log \Gamma(v).$$

Thus, unit prices will also be distributed with respect to a Gamma density function, though with different parameters, that is,

$$\hat{g}(p) = f(p; s+\beta, \nu+\kappa). \tag{4.2.5}$$

The process of deriving (4.2.5) from (4.2.3) and (4.2.4) is completely analogous to the one described in section 4.1. In this case we are only able to identify $(s+\beta)$ and $(\nu+\kappa)$, which means none of the parameters in utility function, quality index function and distribution function of market prices can be identified without data of market prices. Although the same kinds of estimation methods as those that will be described in details in section 6 can be used, the meaning of the estimators will not be the same.

5. Description of the data and aspects of the economic development in China

5.1. A brief sketch of the urbane economic development in Sichuan and Liaoning

China's economy has since liberation in 1949 been dominated by a high degree of central planning. However, a series of market-oriented economic reforms have been introduced since the late 1970s in order to increase productivity and the living conditions of Chinese households. The economic reforms were aimed at the rural economic system at first, and the significant urban economic reforms were actually introduced at the end of 1984. Important aspects of the reforms were to decentralize decisions to the local government level or even to the firm level and allow firms to retain a larger fraction of profits and to make use of performance-linked bonus payment. The economic reforms resulted in a considerable increase in productivity and output and on average in the level of living. One distinctive mark of the consumption pattern during the 1980s is the rapid acquisition of durables. Before 1979 most durable goods available in other countries were 'unaffordable, unavailable, or even unknown' for Chinese households. The rush buying, which expressed people's desire for more comfortable living, was met by the growing consumption goods industries.

Although the urban reform of 1984 concerns all provinces, the south eastern coastal provinces seem to have been in the forefront of the reform process. This is probably due to the establishment of special economic zones and the introduction of an urban price reform in a few eastern coastal provinces. The broadness and complexity of the reform package of 1984 makes it, however, hard to distinguish between the other provinces with respect to how they have succeeded in implementing the reforms. We shall here only look at two representative provinces, Sichuan and Liaoning province.

Sichuan is by population the largest province in China with more than 100 million inhabitants, nearly 10 percent of the total population, and situated landlocked in the southwest central part of China on the Yangtze river. There were 26 cities in Sichuan by 1988, of which two very large cities with population above 1 million people, no large cities, seven medium-sized and 17 small cities. The number of non-agriculture residents in urban areas totaled only 7.4% of the province population. The relatively low degree of urbanization is mainly due to the isolated location and the self-supported economics in the old days which benefited from the rich natural sources. Sichuan had very little industry before 1949, mostly producing basic consumer goods. During the First Five Year Plan (1953-1957), development of the industry production in this region was given high priority according to the regional development policy at that time aimed at reducing income disparities between provinces. The military industry was developed rapidly after the liberation, since Sichuan was considered as the strategic rear base for the defense of China. It may explain why urban Sichuan on average experienced a considerable increase in level of living between 1965 and 1975. However, Sichuan has not been in the front-line of the urban economic reforms and export-oriented development strategy that were initiated in 1984.

Liaoning province is situated in the northeast of China facing the Bohai Sea and the northern part of the Yellow Sea. Liaoning had by the end of 1990 a population of 39.7 million or 3.4 per cent of China's population. The urban population of Liaoning counted in 1988 of 50.3% of the total population. Hence, the degree of urbanization is vastly different in the two provinces. Liaoning is undoubtedly China's most important base for heavy industry. Founded by the Russians and the Japanese under the semicolonial period in the first half of this century, Liaoning's heavy industry experienced another major booming period under the First Five-Year plan strongly supported by Soviet aid. Due to the abundant natural resources of the region, the well-established basic industrial facilities, heavy industry was still on top of the agenda during the economic reforms in the 1980s. Most industrial enterprises in Liaoning suffer from the decay of 'aging': the technologies, as well as the facilities and the products are too old to meet the needs of a market with ever harder competition from more recent production facilities in other regions and from other countries. However, impressive efforts have been made to upgrade the traditional industry and the old enterprises in the last few years.

For more details about the economic development in those two provinces we refer to Olav Bjerkholt and Yu Zhu (1993).

5.2. The data set

This study is based on individual household data abstracted from the State Statistical Bureau's Urban Household Survey (UHS) of Sichuan and Liaoning provinces during the 1988-1990 period. The data is a three years cross sectional one. We, here, focus only on the annual expenditure on consumer durables. The two items, color TV and refrigerator, are chosen to give a representation of the expenditure on the durables every year.

Every year, about 550-600 households participated in the survey in each of the provinces. Purchasing of the two items reached its peak in 1988 and declined afterwards. However, the average purchasing price for each item arose to its top in 1989 and went to the downside afterwards. We get the impression, by comparing the annual expenditure on the durables between these two provinces, that the households in Liaoning province bought goods with higher price tags on average and more family made the purchasing decisions. It is probably due to the higher income level in Liaoning province during that period, which consequently led to more attention on the quality. Price is usually treated as an indicator for quality. Households that had made the purchase have unsurprisingly higher income level on average.

Almost no family bought two pieces of color TV or refrigerator within one year. Due to the character of cross sectional data, we are unable to obtain the information about the household's purchasing history over the whole period. Although Liaoning province lies in north-east China, the area which is one of the coldest in China, and Sichuan province in contrast locates in China's south-west part where it is relatively hot, there is no evidence indicating that families in Liaoning bought fewer refrigerators. This can also be viewed as an indicator of families' being wealthier in Liaoning province.

Table 5.1 gives some general information about the average income level, price level, the ownership of the households, etc. The summary statistics for the households that made a purchase during the period are given in table 5.2.

Table 5.2.1. Summary statistics for the whole sample during 1988-1990.

Province	Year	No. of obs.	Mean income	Mean no. of persons per house-hold	No. of households that already owned one piece		No. of households that bought one piece during the year though already owned one	
					Refrigerator	Color TV	Refrigerator	Color TV
Liaoning	1988	600	4600 (1644)	3.43 (0.93)	141	233	3	7
	1989	600	5057 (2070)	3.40 (0.87)	214	306	9	3
	1990	597	5682 (1888)	3.32 (0.83)	312	386	6	6
Sichuan	1988	550	4186 (1414)	3.39 (0.92)	172	223	0	3
	1989	550	4704 (1684)	3.28 (0.96)	255	309	0	3
	1990	550	5329 (1990)	3.24 (0.98)	333	358	4	1

*) Standard deviations are given in the parentheses.

Table 5.2.2. Summary statistics of the households that had made the purchase during 1988-1990.

	Province	Year	No. of observation	Mean income	Mean price (unit value)	Mean no. of persons per household
Refrigerator	Liaoning	1988	66	5101 (1603)	2120 (710)	3.55 (0.98)
		1989	75	6074 (1854)	2273 (671)	3.62 (1.00)
		1990	47	6010 (1647)	1865 (416)	3.24 (0.59)
	Sichuan	1988	65	4909 (1459)	1570 (552)	3.57 (0.77)
		1989	51	5240 (1680)	1763 (547)	3.50 (0.89)
		1990	43	5839 (1960)	1349 (426)	3.32 (0.90)
Color TV	Liaoning	1988	83	5075 (1614)	2384 (773)	3.51 (0.94)
		1989	46	5944 (2227)	2849 (389)	3.40 (0.70)
		1990	36	6266 (2083)	2780 (700)	3.40 (0.77)
	Sichuan	1988	52	4798 (1311)	2165 (553)	3.27 (0.79)
		1989	44	5640 (1650)	2685 (416)	3.38 (0.90)
		1990	31	6304 (1863)	2392 (421)	3.62 (0.93)

*) Standard deviations are given in the parentheses.

6. Estimation strategies and empirical results

6.1. Estimation; Method I

The first approach to estimation is the most intuitive one. It can be divided into two steps. Since we have proved that the distribution of unit prices is a scaled Gamma density function with parameters (s-h) and v provided the unknown market prices are distributed according to a scaled Gamma density function with parameters s and v, we are able to use the following method.

I will give a short repetition of the properties belonging to a Gamma function from section 4 for convenience. Recall that we assumed that unit prices have a scaled Gamma density function with parameters (s-h)>0 and v>0, that is,

(6.1.1)
$$\hat{g}(p) = f(p; s-h, v) = \frac{(s-h)^v p^{v-1} e^{-(s-h)p}}{\Gamma(v)}.$$

From (6.1.1) it follows readily that

(6.1.2)
$$E(\hat{P}_i) = \frac{v}{(s-h)}$$

and

(6.1.3)
$$var(\hat{P}_i) = \frac{v}{(s-h)^2},$$

where $\hat{P}_i$ denotes unit price, $E(\hat{P}_i)$ and $var(\hat{P}_i)$ are expectation and variance of the unit prices.

As the data of unit prices is available, we are able to get consistent estimators for (s-h) and v <u>for each year</u> by using the empirical counterparts of the expectation and variance of the unit prices. It can be written down as

(6.1.4)
$$\overline{(s_t - h_t)} = \frac{Est(E(\hat{P}_{it}))}{Est(var(\hat{P}_{it}))}$$

and

(6.1.5)
$$\overline{v_t} = \overline{(s_t - h_t)} \cdot Est(E(\hat{P}_{it})),$$

where $\overline{v_t}$ and $\overline{(s_t - h_t)}$ are the estimators, $Est(E(\hat{P}_{it}))$ and $Est(var(\hat{P}_{it}))$ are the empirical expectation and variance of the distribution of unit prices. Footnote t denotes the year as we estimate parameters annually. Unfortunately, we cannot get a separate estimator for s, which makes it impossible to identify the unknown parameters in the utility functions and in the λ-function, and which also makes it impossible to estimate the distribution of market prices. But we are now capable of calculating the cumulative densities of the distribution of unit prices, which means one term in the equation (4.1.17) and (4.1.18), $F(Y_i - c; s - h, v)$, the cumulative probability of buying an variant with price less than (Y_i-c), can be estimated first before we go further to the estimation of the other unknown parameters. To assess the value of c, we are using three years' average of the minimal expenditures with respect to the type of the households. In the data set, households are classified into five types:

Type	Definition
1	Single
2	Couples without children
3	Couples with one child
4	Couples with more than one child
5	All others

Thus, for each type an average minimal living expenditure is calculated. Note that we assume here that c is constant through years. Since three years are a rather short period, this assumption seems not to be very restrictive. The value of c is displayed in table 6.1.1.

Table 6.1.1. c-value (the minimal expenditure per year)*

	Type 1	Type 2	Type 3	Type 4	Type 5
c	560	908	1008	1849	1381

*) Unit of the minimal living expenditure is Renminbi.

In the next step, maximum likelihood method is used to estimate the remaining parameters specifying the probability of not buying. From section 4, we know the probability of not buying can be expressed as

(6.1.6)
$$\varphi_{it}(0 \mid Y_{it}) = \frac{1}{1 + F_{it}\, exp(-\delta_t - \tau_t x_{it})},$$

where Y_{it} is the income, x_{it} is the number of persons per household, δ_t and τ_t are two parameters needed to be estimated, and $F_{it} = F(Y_{it} - c; s_t - h_t, v_t)$. F_{it} has been calculated in the first step by using the estimates for $(s_t - h_t)$ and v_t.

Our task now is to maximize the joint probability of observing such decisions of the households, that is,

(6.1.7)
$$\prod_{i \in S_0} \varphi_{it}(0 \mid Y_{it}) \cdot \prod_{i \in S_1} (1 - \varphi_{it}(0 \mid Y_{it})),$$

where S_0 is the set of household that did not buy a refrigerator that year and S_1 is the set of household that did buy.

The loglikelihood function can therefore be written as

(6.1.8)
$$log\, L = \sum_{i=1}^{N} A_{it}\, log(1 - \varphi_{it}(0 \mid Y_{it})) + \sum_{i=1}^{N} (1 - A_{it}) log(\varphi_{it}(0 \mid Y_{it})),$$

where $A_{it} = 1$ if the household made a purchase and 0 otherwise.

By the maximum likelihood principle the unknown parameters δ_t and τ_t are estimated by maximizing logL with respect to them, which is obtained by setting all the first order conditions of the loglikelihood function with respect to the parameters to zero.

Note again that since durable goods will only depreciate gradually and since those two durables that we are analyzing here are recently available to Chinese families, the families which had already owned one piece will not be taken into account.

We have applied the theory and the estimation method discussed so far on refrigerator and color TV, and the estimation results are displayed in Table 6.1.2 and Table 6.1.3.

Table 6.1.2. Estimation results for refrigerator; method I*

	1988		1989		1990	
	Estimate	t-statistic	Estimate	t-statistic	Estimate	t-statistic
v_t	7.134		11.577		10.814	
$(s_t - h_t)$	0.0039		0.0056		0.0067	
δ_t	1.88 (0.37)	5.03	2.17 (0.38)	5.63	1.46 (0.53)	2.73
τ_t	-0.088 (0.10)	-0.88	-0.224 (0.10)	-2.15	0.025 (0.16)	0.16

*) Standard deviations are given in the parentheses.

Table 6.1.3. Estimation results for color TV; method I*

	1988		1989		1990	
	Estimate	t-statistic	Estimate	t-statistic	Estimate	t-statistic
v_t	10.71		45.87		17.85	
$(s_t - h_t)$	0.0047		0.0166		0.0068	
δ_t	0.245 (0.30)	0.61	0.53 (0.42)	1.25	1.75 (0.50)	3.52
τ_t	0.23 (0.11)	1.96	0.18 (0.12)	1.44	-0.13 (0.14)	-0.92

*) Standard deviations are given in the parentheses.

Notice that the estimates for τ_t are insignificant for both cases. It means that the size of the household does not cast much light on explaining the difference in choices among households.

One shortcoming of this estimation method is that the standard deviations of the estimates of $(s_t - h_t)$ and v_t can not be obtained straightforwardly. Though it is possible to get them, but only by a great deal of calculation. It would be desirable to have the standard deviation so that we could have knowledge about how precise these estimates are.

In the second stage of the estimation, the results from the first stage, namely, the values of F_i, are treated as one of the input variables as if they are non-stochastic. But the truth is that they actually are random since they are calculated by using the estimates of (s-h) and v which are of course stochastic. As a result, the standard deviations and t-values of the estimates of δ and τ may be biased.

6.2. Estimation; Method II

The estimation process can be divided into two steps as well as in method I. First, estimates of $(s_t - h_t)$ and v_t will also be obtained by the unit price data set at hand, though maximum likelihood method will be used this time instead. The estimates will be those that maximize the joint probability of observing the obtained set of unit prices. The loglikelihood function can be written as

$$
\begin{aligned}
log\, L_t &= \sum_{i \in S_1} log\, f\left(\hat{P}_{it}; (s_t - h_t), v_t\right) \\
&= \sum_{i \in S_1} \left\{v_t\, log((s_t - h_t)) + (v_t - 1) log\, \hat{P}_{it} - (s_t - h_t)\hat{P}_{it} - log(\Gamma(v_t))\right\}
\end{aligned}
\qquad (6.2.1)
$$

where footscript t denotes the different years, i.e. $t = 1$ represents 1988, etc. One particular feature, that makes the ML method in the first step more attractive than the one used in method I, is that we are able to get standard error of estimates readily.

The estimation results are displayed in table 6.2.1 and table 6.2.2.

Table 6.2.1. Annual estimates for refrigerator when ML method is applied

Parameter	Estimate	Standard error	t-statistic
v_1	7.52	0.88	8.59
v_2	7.70	0.90	8.60
v_3	9.59	1.46	6.58
$s_1 - h_1$	0.0041	0.00047	8.58
$s_2 - h_2$	0.0041	0.00048	8.55
$s_3 - h_3$	0.0059	0.00102	5.82

Table 6.2.2. Annual estimates for color TV when ML method is applied

Parameter	Estimate	Standard error	t-statistic
v_1	12.68	0.98	12.87
v_3	19.01	2.61	7.28
$s_1 - h_1$	0.0055	0.00039	13.97
$s_3 - h_3$	0.0073	0.00098	7.41

Note that in table 6.2.2, there are no estimates for 1989. This is because we were unable to estimate the parameters by the ML method in this case. We suspect that there may be serious measurement errors in the data set of 1989. From those two tables, we also find that for both durables estimates of ν do not change significantly over time since the respective 95 percent confidence intervals overlap. It may indicate that market prices change proportionally cross years.

Had all the market prices of refrigerator changed in proportion over years, the true value of ν should be constant over time. This statement can be proved easily as follows. Some repetition from section 4 will be made for convenience. As the market prices are Gamma distributed, the expectation and the variance of the prices can be expressed as functions of parameters existing in the density function, that is,

(6.2.2) $$E(P(z)) = \frac{v}{s}$$

and

(6.2.3) $$var(P(z)) = \frac{v}{s^2}.$$

If prices change proportionally, namely, $P_{new}(z) = k \cdot P_{old}(z)$, then

(6.2.4) $$E(kP(z)) = k \cdot E(P(z))$$

and

(6.2.5) $$var(kP(z)) = k^2 \cdot var(P(z)).$$

Thus,

(6.2.6) $$\frac{(E(P(z)))^2}{var(P(z))} = \frac{(E(kP(z)))^2}{var(kP(z))} = v,$$

which means that ν does not change while prices increase proportionally.

Since one can see from table 6.2.1 and table 6.2.2 that estimates of v do not change significantly over time, the assumption can be considered as an rather reasonable one, and we shall next estimate a common ν for every year during the period. The loglikelihood function can be written as

$$
(6.2.7) \qquad \begin{aligned} \log L &= \sum_{t=1}^{3}\sum_{i\in S_1} \log f\left(\hat{P}_{it};(s_t-h_t),v\right) \\ &= \sum_{t=1}^{3}\sum_{i\in S_1}\left\{v\log((s_t-h_t))+(v-1)\log\hat{P}_{it}-(s_t-h_t)\hat{P}_{it}-\log(\Gamma(v))\right\} \end{aligned}
$$

We then proceed to calculate $F_{it}=F(Y_{it}-c_t;(s_t-h_t),v)$ for all the household that did not own a refrigerator or a color TV in the beginning of the year, which is due to the same reason mentioned in method I.

Step two, our task is the same as in step two in method I as to work out estimates of the unknown parameters in the probability of buying or not buying. But we are going to introduce one more assumption. Since three years are only a short period, we may assume that the isolated influence on the preference from the characteristics of the household i.e. the effect of the number of persons of the household will change very little. In other words, we assume that τ_t, the parameter in front of the variable representing the size of households, is constant over time. We will see if it may improve the estimation results, since the annual estimates of τ_t are insignificant by using method I. Thus, the maximum likelihood function can be written down as

$$
(6.2.8) \qquad \log L = \sum_{t=1}^{3}\sum_{i=1}^{N}\left[A_{it}\log(1-\varphi_{it}(0\mid Y_{it}))+(1-A_{it})\log(\varphi_{it}(0\mid Y_{it}))\right],
$$

where

$$
(6.2.9) \qquad \varphi_{it}(0\mid Y_{it}) = \frac{1}{1+F_{it}\exp(-\delta_t-\tau x_{it})}.
$$

Note that parameters ν and τ are constant over years and parameters (s-h) and δ vary over time. The other symbols have same meanings as described above in method I.

Estimation results for refrigerator and color TV are displayed in table 6.2.3 and 6.2.4.

Table 6.2.3. Estimation results for refrigerator; method II

Parameter	Estimate	Standard error	t-statistic
v	8.72	0.64	13.71
$s_1 - h_1$	0.0047	0.00036	13.27
$s_2 - h_2$	0.0042	0.00033	12.69
$s_3 - h_3$	0.0054	0.00048	11.11
τ	-0.1205	0.0353	-3.41
δ_1	1.8547	0.2072	8.95
δ_2	1.6812	0.2092	8.04
δ_3	1.8069	0.1669	10.83

From table 6.2.3 we see that household size has a significant positive effect on the probability of buying a refrigerator. Also the δ_t parameters seem to be nearly constant over time. This is also true for the parameters $s_t - h_t$.

Table 6.2.4. Estimation results for color TV; method II

Parameter	Estimate	Standard error	t-statistic
v	17.79	0.88	20.06
$s_1 - h_1$	0.0077	0.00036	21.53
$s_2 - h_2$	0.0064	0.00044	14.74
$s_3 - h_3$	0.0068	0.00038	17.79
τ	0.1514	0.0837	1.81
δ_1	0.4930	0.3007	1.64
δ_2	0.6175	0.3077	2.01
δ_3	0.8791	0.3224	2.73

From table 6.2.4, we realize that household size does not seem to affect the probability of buying a color TV significantly. The parameters $s_t - h_t$ and δ_t seem not to change significantly over time.

This method still cannot deal with the problem mentioned in section 6.1 that stems from the two-stage estimation process. The values of $F(Y_{it} - c_t;(s-h)_t, v)$ are taken into step two as though they are deterministic. We know, however, they are stochastic since they depend on estimates of (s-h) and v which are stochastic. Had we been capable of applying more advanced programming tools, the problem could be solved. We might estimate all the variables simultaneously in one stage, which could be done by using the following loglikelihood function

(6.2.10)
$$\log L = \sum_{t=1}^{3}\sum_{i=1}^{N}\left\{A_{it}\log\left(\frac{f(\hat{P}_{it};s_t-h_t,v)}{\exp(\delta_t+\tau x_{it})+F(Y_{it}-c_t;s_t-h_t,v)}\right)\right\}$$
$$+\sum_{t=1}^{3}\sum_{i=1}^{N}\left\{(1-A_{it})\log\left(\frac{1}{1+F(Y_{it}-c_t;s_t-h_t,v)\exp(-\delta_t-\tau x_{it})}\right)\right\}$$

Unfortunately the available programming tool did not allow us to do this.

7. Estimation when market prices are available

If data of market prices were available, we could estimate the parameter s, which can be done by applying maximum likelihood method as described in section 6.1. The loglikelihood function will be then

(7.1)
$$\log L = \sum_{z=1}^{M}\log f\big(P(z);s,v\big),$$

where $P(z)$, , $z=1,2,M$, denote observations of market prices. Here, these parameters are estimated annually. Knowing parameter s and (s-h) opens to us a possibility of comparing changes in market prices and in purchasing prices. In other words, it makes us able to explore the underlying relationship between market prices and unit prices.

8. Elasticities

8.1. Price elasticity

It is of interest to know how the probability of buying would be altered by a marginal change in the market prices. We shall now discuss how this can be done.

We assume that market prices change proportionally, therefore v will be constant over time. We also assume that changes in prices of the durable will have no impact on the general price index m. Recall from section 3, we let $u(y-P(z))$ represents $V_i^*\left(1,\frac{y-P(z)}{m}\right)$ for notational simplicity. In section 4 we have specified $u(y-P(z))$ as $\beta(y-P(z))$. Thus, parameter β is in fact equal to a genuine parameter β^* divided by the price index. Parameter θ in that $\lambda(\cdot)$ function can also be formulated as a genuine parameter θ^* divided by the price index. Intuitively, the conditional quality index would not be altered had all prices including prices of the durable and prices of the other commodities changed in proportion. Thereby, if changes in prices of the durable have no influence on the price index, parameter h which is equal to θ-β will not be influenced by the changes in market prices of the durable.

Since parameters ν and h would not be altered by changes in market prices. Then the only parameter that will be influenced by changes in market prices is s. We realize that changing $\frac{1}{s}$ while keeping ν fixed corresponds to changing prices while the relative variance of the price distribution is kept fixed. The price elasticity can be derived as

(8.1.1)
$$\begin{aligned} El_{EP(z)}(1-\varphi_i(0|y)) &= \frac{\partial\, log(1-\varphi_i(0|y))}{\partial\, log(EP(z))} \\ &= \frac{\varphi_i(0|y)EP(z)}{F(y-c;s-h,\nu)}\frac{\partial F(y-c;s-h,\nu)}{\partial(EP(z))} - \varphi_i(0|y)EP(z)\frac{\partial \gamma^*(x)}{\partial(EP(z))} \end{aligned}$$

where

(8.1.2)
$$\frac{\partial F(y-c;s-h,\nu)}{\partial(EP(z))} = \frac{\partial F}{\partial s}\cdot\frac{ds}{dEP(z)} = -f((y-c)(s-h);1.\nu)\cdot(y-c)\cdot\frac{s^2}{\nu},$$

and

(8.1.3)
$$\frac{\partial \gamma^*(x)}{\partial(EP(z))} = \frac{\partial \gamma^*(x)}{\partial s}\cdot\frac{ds}{dEP(z)} = -\frac{hs}{s-h}.$$

By inserting (8.1.3) and (8.1.2) into (8.1.1), we can write the form of price elasticity as

(8.1.4) $$El_{EP(z)}\left(1-\varphi_i(0|y)\right) = -\varphi_i(0|y)\left(\frac{f\left((y-c)(s-h);1,v\right)\cdot(y-c)\cdot s}{F\left((y-c)(s-h);1,v\right)} - \frac{vh}{s-h}\right).$$

Notice that in deriving (8.1.2), we have used

(8.1.4) $$F(y-c;s-h,v) = F\left((s-h)(y-c);1,v\right).$$

A proof of this is presented in appendix A. It is a pity that we do not have the data of market prices, which would have made it possible to calculate price elasticities.

8.2. Income elasticity

Often one is also interested in knowing how the probability of buying would be altered by a marginal increase in the income of the family. These income elasticities can be calculated even if the data of market prices is not available. Differentiating the probability of buying with respect to income y leads to

(8.2.1)
$$\frac{\partial\log\left(1-\varphi_i(0|y)\right)}{\partial\log y} = \frac{y\varphi_i(0|y)\cdot f(y-c;s-h,v)}{F(y-c;s-h,v)} = \frac{y\varphi_i(0|y)\cdot(s-h)\cdot f\left((y-c)(s-h);1,v\right)}{F\left((y-c)(s-h);1,v\right)},$$

We can calculate the income elasticities by using those estimates we obtained in section 6.2. For instance, in analyzing the demand for refrigerator, we have estimates of ν, (s-h), τ and δ for every year during the period 1988-1990. Let us now calculate income elasticities basing on estimates obtained from the demand of refrigerator in 1988. I will for the sake of illumination only choose three representative values of income for each family type, and all of those values are hypothetical. The results are in table 8.2.1.

Table 8.2.1. Income elasticities for refrigerator

Y	type	$1-\varphi_i(0\|y)$	Elasticity
8000	1	0.1501	0
8000	2	0.1661	0
8000	3	0.1834	0
8000	4	0.2022	0.0000401
100000	5	0.2223	0
3415	1	0.1434	0.3904
3572	2	0.1546	0.6171
4047	3	0.1785	0.2884
4488	4	0.1880	0.7862
5413	5	02220	0.0260
1000	1	0.00016	14.7058
1500	2	0.0011	14.8731
2000	3	0.0184	8.7054
2500	4	0.0025	21.6435
2500	5	0.0379	8.5160

Those values of income were chosen because they represent a high, an average and a low level of income for each group. It can be seen from the table that the income elasticities for those families with high income level are almost 0, which indicates that a little increase in income has no influence on those families' decisions. But a little extra means something for those families belonging to low income class, income elasticities are much higher. The results seem quite compatible with the reality.

9. Conclusions

In this paper we have discussed a particular econometric approach to analyze the demand of durable goods. A major advantage of this approach compared to conventional discrete choice models is that we have taken account of the fact that products are differentiated with respect to unobserved quality attribute. We have demonstrated that under particular functional form assumptions the empirical model becomes convenient to estimate. We have shown that the model can be estimated by using two versions of a two-step procedure.

We have applied this approach to analyze consumption of color TV and refrigerator in China during the period 1988-1990. Income elasticities have been calculated, the results of which are rather reasonable in the sense that poor families have large income elasticities while rich families' income elasticities are almost zero. Price elasticities can not be calculated in this case due to the lack of information of the distribution of market prices.

References

Ben-Akiva, M., and S. Lerman (1985): *Discrete Choice Analysis: Theory and Application to Predict Travel Demand*, Cambridge, MA: MIT Press.

Bjerkholt, O. and Y. Zhu(1993): Living conditions of Urban Chinese Households around 1990. Discussion Papers, No. 87, Statistics Norway.

Dagsvik, J.K. (1996a): Consumer Demand with Unobservable Product Attributes. Part I: Theory. Discussion Papers, No.166, Statistics Norway.

Dagsvik, J.K., Y. Zhu and R. Aaberge(1998): A Framework for Empirical Modelling of Consumer Demand with Latent Quality Attributes. Discussion Papers, No.229, Statistics Norway.

Deaton, A. and Muellbauer, J.(1980): *Economics and consumer behavior*, Cambridge.

Derivation of Eq.(7.2.8)

In deriving price elasticity, we have applied a property of Gamma distribution as granted. We here shall give a proof of that. Recall that the density function of a scaled Gamma distribution with parameter ν and s has the form

(1) $$f(p;s,\nu)=\frac{s^{\nu}p^{\nu-1}e^{-sp}}{\Gamma(\nu)}.$$

Now let

(2) $$Z=sP,$$

then

(3) $$\Pr(Z\le z)=\Pr(sP\le z)=\Pr\left(P\le\frac{z}{s}\right)=F\left(\frac{z}{s};s,\nu\right),$$

which implies that variable Z has a density function as

(4) $$f\left(\frac{z}{s};s,\nu\right)\cdot\frac{1}{s}=\frac{z^{\nu-1}e^{-z}}{\Gamma(\nu)}=f(z;1,\nu),$$

where $f(z;1,\nu)$ is a standard Gamma function with parameter ν>0. Until now, we have shown that variable Z is actually distributed with respect to a standard Gamma density function.

Since unit prices $\hat{P}_i$ have a Gamma distribution function with parameters (s-h) and ν, and

(5) $$\Pr\left(\hat{P}_i\le y-c\right)=\Pr\left(\hat{P}_i\cdot(s-h)\le(y-c)(s-h)\right),$$

we have proved that

(6) $$F(y-c;s-h,\nu)=F((y-c)(s-h);1,\nu).$$

Recent publications in the series Documents

97/14 E. Berg, E. Canon and Y. Smeers: Modelling Strategic Investment in the European Natural Gas Market

97/15 A. Bråten: Data Editing with Artificial Neural Networks

98/1 A. Laihonen, I. Thomsen, E. Vassenden and B. Laberg: Final Report from the Development Project in the EEA: Reducing Costs of Censuses through use of Administrative Registers

98/2 F. Brunvoll: A Review of the Report "Environment Statistics in China"

98/3: S. Holtskog: Residential Consumption of Bioenergy in China. A Literature Study

98/4 B.K. Wold: Supply Response in a Gender-Perspective, The Case of Structural Adjustments in Zambia. Tecnical Appendices

98/5 J. Epland: Towards a register-based income statistics. The construction of the Norwegian Income Register

98/6 R. Chodhury: The Selection Model of Saudi Arabia. Revised Version 1998

98/7 A.B. Dahle, J. Thomasen and H.K. Østereng (eds.): The Mirror Statistics Exercise between the Nordic Countries 1995

98/8 H. Berby: A Demonstration Data Base for Business Register Management. A data base covering Statistical Units according to the Regulation of the European Union and Units of Administrative Registers

98/9 R. Kjeldstad: Single Parents in the Norwegian Labour Market. A changing Scene?

98/10 H. Brüngger and S. Longva: International Principles Governing Official Statistics at the National Level: are they Relevant for the Statistical Work of International Organisations as well?

98/11 H.V. Sæbø and S. Longva: Guidelines for Statistical Metadata on the Internet

98/12 M. Rønsen: Fertility and Public Policies - Evidence from Norway and Finland

98/13 A. Bråten and T. L. Andersen: The Consumer Price Index of Mozambique. An analysis of current methodology – proposals for a new one. A short-term mission 16 April - 7 May 1998

98/14 S. Holtskog: Energy Use and Emmissions to Air in China: A Comparative Literature Study

98/15 J.K. Dagsvik: Probabilistic Models for Qualitative Choice Behavior: An introduction

98/16 H.M. Edvardsen: Norwegian Regional Accounts 1993: Results and methods

98/17 S. Glomsrød: Integrated Environmental-Economic Model of China: A paper for initial discussion

98/18 H.V. Sæbø and L. Rogstad: Dissemination of Statistics on Maps

98/19 N. Keilman and P.D. Quang: Predictive Intervals for Age-Specific Fertility

98/20 K.A. Brekke (Coauthor on appendix: Jon Gjerde): Hicksian Income from Stochastic Resource Rents

98/21 K.A.Brekke and Jon Gjerde: Optimal Environmental Preservation with Stochastic Environmental Benefits and Irreversible Extraction

99/1 E. Holmøy, B. Strøm and T. Åvitsland: Empirical characteristics of a static version of the MSG-6 model

99/2 K. Rypdal and B. Tornsjø: Testing the NOSE Manual for Industrial Discharges to Water in Norway

99/3 K. Rypdal: Nomenclature for Solvent Production and Use

99/4 K. Rypdal and B. Tornsjø: Construction of Environmental Pressure Information System (EPIS) for the Norwegian Offshore Oil and Gas Production

99/5 M. Søberg: Experimental Economics and the US Tradable SO_2 Permit Scheme: A Discussion of Parallelism

99/6 J. Epland: Longitudinal non-response: Evidence from the Norwegian Income Panel

99/7 W. Yixuan and W. Taoyuan: The Energy Account in China: A Technical Documentation

99/8 T.L. Andersen and R. Johannessen: The Consumer Price Index of Mozambique: A short term mission 29 November – 19 December 1998

99/9 L-C. Zhang: SMAREST: A Survey of SMall ARea ESTimation

99/10 L-C. Zhang: Some Norwegian Experience with Small Area Estimation

99/11 H. Snorrason, O. Ljones and B.K. Wold: Mid-Term Review: Twinning Arrangement 1997-2000, Palestinian Central Bureau of Statistics and Statistics Norway, April 1999

99/13 Y. Li: An Analysis of the Demand for Selected Durables in China

Documents

PORTO BETALT
VED
INNLEVERINGA
P.P.
NORGE/NOREG

Tillatelse nr.
159 000/502

B

Returadresse:
Statistisk sentralbyrå
Postboks 8131 Dep.
N-0033 Oslo

Statistics Norway
P.O.B. 8131 Dep.
N-0033 Oslo

Tel: +47-22 86 45 00
Fax: +47-22 86 49 73

ISSN 0805-9411

Statistisk sentralbyrå
Statistics Norway