NHH

Norges Handelshøyskole

Norwegian School of Economics and Business Administration

Producer Assessment of Demand and Equilibrium in Differentiated Markets

A Monograph

By

Nils E. Joachim Høegh-Krohn

A dissertation submitted for the degree of doctor oeconomiae

Producer Assessment of Demand and Equilibrium in Differentiated Markets

A monograph on how producers who utilize a limited and specific set of information, anticipate consumer reactions, and how this assessment influences the market equilibrium.

Nils E. Joachim Høegh-Krohn
Norwegian School of Economics and Business Administration

November, 1994

To my Mother and Father.

Contents

List of variables

$\Re$	The real numbers.
W	The population of consumers.
B	The population of producers; population of players.
N	The number of consumers in W.
I	The number of products in the whole economy.
I-1	The number of producers in B; the number of differentiated products.
X	The choice space of W.
x	A consumer point in, or an element of X.
x^i	The consumer point of X containing ψ^i.
x_I	The general product.
x_I^i	The general product associated with x^i.
Ψ	The set of all possible elements of the differentiated product.
ψ^i	Alternative *i* of Ψ.
S	The strategy space of B.
C'	The price-income space.
C^{i}'	The set of all possible prices of the i-th alternative of ψ.
C	The price space.
c	A price vector in C.
c_i	Price of product *i*.
$x^i(c)$	A consumer point for a given *c*.
m	Individual income.
M	A budget set for members of W.
L	A budget plane for members of W.
R	A binary relation on X.
Υ	The power set of X.
T	The set of all subsets of X.
x^0	The greatest element of M.

C^	The set of prices making every *x* of M obtainable to some member of W.
D(c)	The demand for a given *c*.
u(x)	The utility of *x*.
Q_i	The information set of producer *i*.
Γ	A σ-algebra.
p	A probability measure; probability distribution of W over L(c).
(X,Γ,p)	A probability space over X.
P	The set of possible probability distributions.
p*	The ultimate chosen subject probability distribution of a member of B
p^i	The subjective probability distribution of producer *i* over L(c).
p_i	The probability of observing the choice of x^i.
ω	A microstate, i.e. a distribution indicating individual position.
ω^0	The real microstate.
Ω	The space of microstates.
z	A frequency distribution of the members of W.
z_i	The frequency indicating the number of consumers in individual state *i*.
$Z(\Omega)$	The space of frequency distributions; macrostates.
$p^h(\omega)$	The h-th alternative probability of observing microstate ω.
$A(\omega)$	The activity level in microstate ω.
$c(\omega)$	The cost level in microstate ω.
ξ	A monotone increasing function over N.
E(x)	The expectation of *x*.
K_i	The costs of producer *i*.
κ_i	Fixed costs of producer *i*.
ν_i	Variable costs of producer *i*.
x_i^s	Production quantity of producer *i*.
s^g	A strategy vector for a specific game *g*.
G	Number of possible strategies open to the players.
π	Profit; payoff.

s^*	A fixed point; equilibrium point.
R_i	Producer i's best reply function.
s_i	Producer i's strategy.
s_{-i}	The vector of strategies selected by all other members of B.
x_i^*	Number of items of x^i transacted in the market.
n_i	Number of members of W demanding x^i.

Preface

The purpose of this dissertation has been to address how firms in a differtiated and imperfect market behave faced by a given, but limited, set of information. In particular we analyse this problem in a discrete choice context, i.e. in a situation where the consumers are allowed to choose only one brand and one item of each brand at a time.

The specific problem facing a firm in such a situation is how to assess the demand from consumers. Since the producer has limited information the demand is uncertain. Hence it has to be assessed, i.e. the expected demand has to be calculated. In doing so, every producer will have to decide on a probability distribution over the number of brands available in the market for the industry in question.

Assuming rational behaviour at the hands of the producers, the probability distribution will have to be set so that it accounts for all the available information, but nothing else. Hence the probability distribution which will be used by a producer will be a distribution which is unbiased with respect to the available information, and at the same time which is in accordance with rational economic behaviour.

Assuming that the information set is equal for all firms involved, we first show that there exists a unique probability distribution, which will be used by every producer. This distribution is a specific and explicit exponential distribution over the available information. Consequently an expected demand relation can be constructed for every firm which specifically indicates how a wide range of common information concerning prices, marketing effort, information gained by a market survey, quality etc. will influence the expected demand, to the extent that this information is seen to be relevant to the problem under consideration. The specific parameters for each of these variables can be given a precise interpretation based on how the assessment problem of the producers is constructed in particular. Hence a theoretical foundation describing how the demand function should be specified, in this particular context, is put forward.

Since a direct theoretical specification of expected demand is provided, such a demand function can easily be submitted to empirical measurement. However, it should be pointed out that this expected demand is demand as viewed by the producers, and as such only an assessment of demand, not an expression of the demand that will materialize.

Secondly we prove that an expected equilibrium for such a differentiated market can be established, assuming price to be the only strategic variable controlled by firms. This expected equilibrium is a part of the total assessment done by firms in preparation for their offer in the market. Based on this assessed equilibrium the firms present their offer in the market, i.e. they offer a given quantity at a given price. If the firms have made a perfect assessment, the equilibrium now reached will be equal to the expected equilibrium. However, due to incomplete information, the actual equilibrium which occurs when the offers are presented in the market will in principle differ form the actual equilibrium since actual and expected demand usually will be different. However, such historical discrepancies constitutes themselves information which can be utilized by producers in the next period, and subsequently such information can be comprised by the forwarded model.

In conclusion a relatively comprehensive and axiomatic model is provided of how producers behave faced with incomplete but specific information in a differentiated and discrete market. However, it should be emphasized that the model presented does not contain a meticulous analysis of interactions between producers. The aim is to model the basic behaviour of producers when faced with uncertainty and incomplete information. Hence the focus of this analysis is somewhat different from the bulk of the literature on product differentiation, which is mainly directed towards the problem of equilibrium, and producer interaction.

This dissertation was initiated out of my studies of the efficiency principle in the context of discrete choice. The efficiency principle states that any population behaviour is efficient if and only if the probability of observing a specific distribution, which represents a lower overall cost, is greater than the probability of observing a more costly distribution of the population. This principle places a specific constraint on any acceptable probability distribution.

To present an axiomatic model of producer behaviour in a differentiated context being my objective, I soon recognized that the problem facing the producer had structural features in common with the solution to the problem of efficient population behaviour. But the problems were not equal. However, the crucial point was that producers faced with limited but specific information would have to make an assessment of how likely it was to observe the choice of a particular brand. Furthermore, this assessment, which had to be a probability distribution, should not be biased with respect to the information used in the assessment. Hence the problem could in principle be solved by fitting a probability distribution to the available information in the only unbiased way there is, i.e. by maximizing Shannon's (1948) measure of uncertainty. The idea which was derived from the analysis done on the efficiency principle, was to insert the available market information as constraints to this maximization problem. Thus an explicit and unique solution could be deduced. This solution would have a specific interpretation, since the objective function, as well as all the constraints, were clearly defined. Furthermore, this solution was the only possible unbiased solution to the clearly defined assessment problem which faced the producers. Consequently the producers, if acting rationally, would use the derived probability distribution as a measure of consumer reactions to the available information.

Hence, what is of particular interest in this dissertation is the acknowledgement that: Producers will have to measure consumer reactions to a particular set of information, that this measurement is uncertain, and that this measurement can in principle only be carried out as described, if producers and consumers act rationally.

The work and comments of Professor Sven Erlander and of Lecturer Jan Lundgren of the University of Linköping have been valuable. The comments of and conversations with Professor Agnar Sandmo at the Norwegian School of Economics have brought me encouragement and insight. I am greatly indebted to my father Nils Høegh-Krohn for his comments and encouragement. I should also thank Professor Robert Grubbström at the University of Linköping for making his time available to discuss my work; and finally I will express my indebtedness to my friend and mentor, Professor Kurt Jörnsten at the Norwegian School of Economics for guidance, discussions, and inspiration.

The inspiration and support given by my spouse Vibeke, and my Mother Eva should be mentioned when acknowledgement is to be granted.

I also wish to express my indebtedness to the Norwegian School of Management and to the Norwegian School of Economics for financial support and general encouragement.

Bergen, November 11, 1994

Nils E. Joachim Høegh-Krohn

Chapter 1. Introduction.

1.1 The problem.

In this treatise I shall address and treat thoroughly the following problem in connection with economic behaviour in differentiated markets:

Consider two different populations of economic agents. Let one population named B consist of the producers of a particular product named ψ which is seen as similar by the members of another population W. A member of B is by definition someone who produces a particular brand of ψ. A member of W is by definition someone who agrees that every different brand of ψ is quite similar and who seeks the consumption of a brand of ψ, and who buys only one brand and one item of that brand at a time. The different brands of ψ are not in fact entirely similar, they are differentiated. The members of W behave according to the general accepted axioms of choice. The members of B seek to maximize their profit. The only knowledge the members of B have about the behaviour, i.e. the preferences and choices of the members of W, is that they behave in accordance with these axioms. This is common knowledge. Nothing else is known, except the prices set by the members of B and a set of incomplete market information, which are also common knowledge. The only decision variable which the members of B control is the price of their own brand. The problem which we eventually address is the characterization of the equilibrium in this market.

We will first clarify some terms. The term differentiated product is based on the notion that some firms or producers and their products can be distinguished as a group from all other firms and products in the economy, by the consumers. Such a group of firms or producers is called an industry. The consumers view a particular group of products within an industry to be strong if not perfect substitutes for one another, and weak substitutes for any other product in the economy. The term differentiated products is thus defined on the basis of the subjective notions of the consumers. No producer can by himself completely decide which industry he is in. A producer may opt for a particular industry but end up in another, depending on the subjective views of the population of consumers.

A particular sub-population of consumers W which demand the brands of a particular industry B constitute the market for this industry's brands.

Such a market is discrete when the members of the sub-population W is only interested in demanding one and only one alternative of all the alternative brands offered in the industry, and only one item of such a brand at a time. Thus the term discrete is also based on the subjective notions of the consumers. Only those consumers that view a particular industry B as a differentiated and discrete collection of alternatives will be members of W. All other consumers are excluded.

The problem stated above has not been extensively treated and solved as a particular problem of interest, i.e. as a general, although restricted problem analysing how producers view the consumers, and how this view affect the behaviour of producers in a differentiated and discrete market.

However, several topics related to this problem have been treated in the existing literature - some quite extensively. The particular problem of strategic producer behaviour in a differentiated market where the demand for each brand is assumed to be known to the producers prior to their own decisions, have been treated most extensively. The classical works in this field are those of Hotelling (1929) and Chamberlin (1933) analysing the interaction between producers assuming consumers to be a completely predictable entity. This tradition has been passed on by the bulk of later literature, to the effect that most analyses of differentiated markets are based on a population of consumers endowed with whatever characteristics are required to make the analysis of producer or firm behaviour interesting. Usually this result in presumptions about a representative consumer, where predictable consumer behaviour is represented by a given demand relation. Thus these models are based on the assumption that demand is known in advance to the producers, and that it is continuous. These assumptions stated, the analysts proceed to discuss the most intricate and detailed problems of firm or producer strategic behaviour, i.e. the interaction between firms.

Consequently there are not many analyses dealing with the simpler yet interesting, problem as far as strategic producer behaviour is concerned, of how producers' view of consumer behaviour affect producer behaviour per se. It should be evident that from the point of view of the producers, the demand from consumers is not known, nor continuous, in most cases. In principle the future demand from the consumers may be known by the consumers themselves, although this is not obvious, but it is certainly not revealed to the producers before they have made their offer to the market, by at least setting their initial prices, and determining initial quantity available for transaction. It is not obvious whether this is recognized or not by most analysts in this field of study, but the fact is that usually the demand from consumers is assumed to be known by the producers, and to be continuous. Hence the problem of how producers view the behaviour of consumers and, how this view affect the behaviour of producers, should be regarded a fundamental problem which, if solved - easily or not, should have implications for several more well known problems in this field of study. This includes the problem of strategic producer behaviour, which is afforded so much attention in the literature on market differentiation.

The only work worth mentioning dealing with the problem of producer assessment of demand, is Anderson, de Palma & Thisse (1992), who also sum up all the earlier attempts of incorporating this problem into the analysis of differentiated markets. Anderson et al (1992) do not attempt to solve the problem of how rational consumer behaviour affect producers. Their emphasis are on strategic firm behaviour in discrete and differentiated markets. However, the discreteness of demand adds a new dimension to the problem of strategic behaviour in differentiated markets, compared to the more usual approach of known continuous demand relationships for every firm. When discrete choice is introduced the consumer side of the strategic problem is given increased weight. Consequently Anderson et al (1992) have to add assumptions about how firms view such situations, to solve their particular problem. The pertinent question is of course what assumptions do Anderson et al (1992) introduce to cope with this added problem of the strategic analysis. The solution is the same as has been applied to pure demand analyses of discrete demand, e.g. Börsch-Supan (1985), McFadden (1974, 1975, 1976), Rust (1985), and Train (1986), to mention a few. Accordingly they solve the problem by introducing additional assumptions about consumer behaviour - not by introducing assumptions about producer behaviour, which would be the reasonable course to take. They quite simply assume that consumers have random preferences, and that the distribution of these preferences is known to all agents involved. Both these assumptions are of course unreasonable. We will show that they are also unnecessary. Despite the obvious misconception embedded in the assumption of random preferences this approach is in fact quite common when attempting to solve such problems. In an econometric context of discrete choice this approach may have some precedence, to the extent that the behaviour can be viewed as random from the outside. It is, nevertheless, based on a misconception. The greater misconception, however, is the assumption that the distribution of these random preferences is known to every economic agent, including the producers. Anderson et al (1992) do not state explicitly that the producers possess this information, but this assumption follows indirectly from their representation of demand in their altogether too complex model.

Based on these assumptions of convenience Anderson et al (1992) derive a so called logit model, based on a certain assumed distribution of the preferences, to represent demand from the consumers. This problem having been solved they then turn to strategic considerations at the hand of the producers.

The problem of using such a model as is advocated by Anderson et al (1992), can be attributed to the underlying reason for the assumption of random preferences: The preferences are of course not random, but the modellers, that is Anderson et al, do not know the preferences of the consumer, and therefore have to perceive these as random. Thus they are not facing up to the problem of unknown consumer preferences irretrievable connected with the inclusion of actual economic agents included in the model, but instead deal with it as a problem of the modellers. Since the randomness of the consumer's preferences is not believed to be an actual problem of real life for any relevant economic agents, the distribution of this artificial randomness also have to be decided upon by the modellers themselves and specified as an additional assumption in their model. The assumption of a particular preference distribution is consequently the fundamental weakness of the random utility models in dealing with analysis of differentiated markets.

We will approach the problem of how producers should view the behaviour of consumers, and how this view affects producer behaviour, directly. First we recognize that demand from consumers is not observable before the producers have presented their offers in the market. When the offers are presented then the consumers respond with action. Consumers may of course express their wishes and desires in advance of the offers from the producers, but there is no way that the producers can trust these expressions to be matched by realizations when the offers are presented later on. Thus the producers will in any case have to make a subjective assessment of the demand they will face for different price levels. This assessment on the hands of the producers will have to take into consideration not only the behaviour of the consumers, but also the interaction with all the other producers in the industry. Hence the offer a single member of industry B will present to the market is the price, and demand he expects to receive in equilibrium. However, only when the producer has presented his price and the corresponding quantity that he expects to be consumed at this price will he be able to observe the real demand, which may very well differ considerably from his expectations.

This being so it should be evident that the producers will have to assess demand. Thus the problem arises at the hand of the producers, and it is the behaviour of this group that has to be elaborated further, not the behaviour of the consumers. That this is so is due to the fact that the demand from consumers is deterministic when the price is known to the consumers, but from the point of view of the producers the demand is unknown prior to their own decisions and hence has to be measured.

Accordingly we will only assume normal rational consumer behaviour, within the context of discrete and differentiated markets. No other assumption is made concerning consumers. In particular we do not assume that consumers have random preferences, and that the distribution of these preferences is known to the producers or to anybody else. To the general accepted assumptions about consumer behaviour we add some easily accepted assumptions of firm behaviour, which in summary say that firms seek to maximize their profits, that demand is measured by probabilities, and that these probabilities are based only on the information available to the producers. On the basis of these assumptions we arrive at the following solution to our problem:

Confronting a discrete and differentiated market the members of B will have to assess the demand from the members of W. Due to the lack of information as to the preferences of the members of W over the available alternatives, this assessment is made using the concept of probability. Assuming that all producers know how all other producers set their prices, we next show that all producers will use a particular class of probability distributions which are unbiased, and incorporate all the information available. For certain assumptions this class of distributions are similar to the distribution that is derived using the logit approach. However, the interpretation of the results are quite different since the fundamental assumptions are different. When the logit is based on random utility, and a known distribution of preferences, our model is based on a subjective derivation by economic agents of an assessment of how other economic agents will behave. Thus our model have a meaningful economic interpretation, where the logit solution at the best has an unclear interpretation as to the fundamental notions concerned.

Our class of distributions is then used by the members of B to calculate the expected demand, and profit, and eventually to derive the expected market equilibrium, which in turn is used to determine the producers' offer to the market.

Since the assessment used by the producers is uncertain by nature, and since the producers use this assessment to set their price and quantity, the actual observed equilibrium may differ from the expected equilibrium. This is due to the fact that the real demand from consumers is unknown, and at given prices this demand may turn out to be different from the demand expected, and the quantity offered, by each producer.

In consequence of what has been stated above, the idea behind this treatise can be outlined as follows: Whereas Anderson et al (1992) choose to introduce a random element into their analysis in order to solve a technical problem, our aim is to model human economic behaviour. In doing so we place our emphasis on the challenges facing the different agents, depending upon the different roles of these agents. In particular we recognize the role played by the members of the population of producers in a particular market. Adhering to the same underlying principles of economic behaviour as do the consumers, the producers confront a certain lack of information concerning the preferences of the consumers. This predicament has to be considered a fact of life. A central theme of our treatise is to answer the question of how an economic agent, playing the role of a producer, solves the problem of assessing demand from the consumers in face of his lack of information. Instead of solving this real problem by introducing a particular distribution of preferences, presumably adhered to by all the members of the population of producers, we place the emphasis on deducing how the producers will behave in such a situation given the limited information they possess, and the implications of this behaviour for equilibrium in the market.

Let us spell out our view in greater detail. A market is identified by a group of agents demanding a sort of product, i.e. consumers, and a group of suppliers supplying the product, i.e. producers. We are studying differentiated markets where the product under consideration is a group of brands that are perceived as similar, but non-identical, and where each producer produces his particular brand of the product. Characterizing the equilibrium in this market being our objective we have to describe both consumer and producer behaviour. Since both consumers and producers are human economic agents, both are assumed to behave according to the basic principles of human economic behaviour, i.e. the general accepted axioms of choice. But as a consequence of the different roles that the agents play as consumers and producers the situation differs between them, and consequently they preform different actions in order to gain their objectives.

In concerning ourselves with differentiated markets we assume that each consumer chooses only one alternative or brand at a time, and that he or she then only chooses one item of each alternative. The income which is not spent on the alternative chosen by the consumer is used to consume the general product, which represents all other products in the economy except the alternatives included in the set of differentiated products under consideration. Saving is also included in the general product. Hence, in a later period, which in theory may be a few seconds later, the consumer can make use of his savings to demand another item of the same alternative or an item of another alternative. Thus the assumption that the consumers only buy one item of one alternative at a time is no real constraint on our model's generality. But, it is a necessary assumption in order to derive our model.

Characterizing the behaviour of the producers, our next assumption is that the producers assume that the consumers behave according to the axioms of choice. That is - the producers know how consumers behave. But they do not know the preferences of the consumers, and consequently which brand each consumer will eventually choose. Consequently each producer will have to make an assessment of the demand he is facing given the price of his brand, price being the only strategic variable he controls.

The question therefore is how this assessment is to be done. We assume that the producers measure the choices of the consumers using probabilities. Thus a producer measures the probability that his particular brand will be chosen, given the set of prices of all brands in question - which is assumed to be generally known. In this way a distribution of choice probabilities is formed, which consists of a probability of choice for every brand in the differentiated product set under consideration. This probability distribution, being formed by each producer to measure the choices of the consumers, will be shaped so that it is unbiased, and so that it satisfies the axioms of choice, known to every producer to guide the consumers in making their choices, and the assumption that each consumer chooses only one alternative and one item of this alternative at a time - since this is all the information that a producer possesses.

A simple but important theorem, characterizing the probability distribution in question, can now be deduced from the assumptions stated above: Assuming that the price of one alternative i is not greater than that of an other alternative j, the probability distribution p* will be such that the probability of a consumer choosing alternative i is not less than the probability that the same consumer will choose j. It should be noted that this is a result deduced from the assumptions laid down, which the probability distribution of choice for a single consumer has to satisfy to be in accordance with the assumptions set forth and in which the producers believe. But, at this stage we also know that in principle this result p* in not necessarily the only distribution which satisfies the behavioural assumptions we have stated above.

Next we recognize that a corresponding distribution is achieved when analysing the distribution of particles in a thermodynamic system, assuming that the particles behave in accordance with the first and second laws of thermodynamics. This distribution problem of thermodynamics, as a problem of logics, is in principle similar to the assessment problem confronting the producers, which is, and this should be underlined, the problem of deciding upon their own subjective probability distributions of consumer demand, i.e. how the consumers are to be distributed among the alternatives in question. Subsequently we observe that our result p* represents the only unbiased distribution with respect in both the case of producer assessment in a differentiated market, and in the thermodynamic case - given the state of information. Thus the resulting p* is the only probability distribution an economic agent in the role of a producer will use in calculating the expected demand from a population of consumers. This conclusion follows from the assumption of rational producer behaviour, which implies that every producer will use the probability distribution that takes into account all the relevant information, and nothing more.

At this stage we have arrived upon a probabilistic model. But it is not a probabilistic model of behaviour. It is a model that duly respects the fact that the choices of the consumers are deterministic. The probabilistic component consists in the fact that we recognize the necessity of the producer to measure the demand from the consumer, and that he is left with no other choice but to measure this through the use of probabilities. Thus the model has probabilistic elements which are assumed to be present in the problems of the real world, and which are not merely technical assumptions of convenience. Hence the use of probabilities is an element characterizing the behaviour of the producers, and not that of the consumers.

Furthermore, we do not assume a particular underlying probability distribution of consumer preferences. On the contrary we shall show that there is a unique form of the probability distribution that satisfies the results spelled out above, and thus satisfies the above stated assumptions about consumer behaviour.

The choice probability distribution p* is of a particular exponential form and is used by the producers to calculate their expected demand and thus their expected income and costs, and therefore, to decide upon which price they will offer to the market, considering the prices of their competitors. We then turn to the question of the existence of market equilibrium, upon which we conclude that the existence of a price equilibrium can be assured for our particular model.

Finally, in chapter 8, we extend our information set to also included in principle, a wide range of firm specific and industry related information such as marketing effort, quality, and market surveys. The equilibrium solution which is obtained for the basic model, will also apply to the enlarged model. Consequently a model which account for a wide range of information in addition to uncertainty is provided.

One last question remains. Some analyst may argue that in the aggregate, and at the limit, the expected demand function in our model may be perceived as a deterministic demand relation. This is wrong and the argument is built on a misconception.It is a fact that the demand from consumers is unknown to the producers. Hence they will have to make an assessment of this demand which is not yet to be observed. Since this demand is uncertain to the producers this uncertainty has to be incorporated into the model - and it is not obvious that a model incorporating uncertainty will have the same properties that a model avoiding to cope with this uncertainty.

However, in principle the producers of differentiated products calculate the expected demand using one or another probability distribution of consumer choice. Therefore, when some analysts use deterministic demand relations and do not mention choice probabilities at all, the same analysts are implicitly assuming an aggregate demand situation, where the concepts set forth in this treatise are the unrevealed underlying concepts. Our claim is thus that the way of analysing the particular problem that concerns us in this treatise is revealed through the idea of producer assessment, as it is set forth below.

In summary, we specify a model of price equilibrium in differentiated markets, based only on reasonable descriptions of consumer and producer behaviour, without assuming anything like random utility behaviour, or that demand can be represented by a logit model. The focus of this analysis is on the modelling of producer behaviour. It is this view of how producers measure demand from consumers in a differentiated market that brings about our results, i.e. a consistent model of price equilibrium in differentiated markets. Hence, our basic axiom or hypothesis is that producers actually calculate expected demand assuming that consumers behave according to the axioms of consumer choice. If this hypothesis is true, then it follows that the distribution of choice probabilities can be represented by a particular exponential form over prices.

1.2 Structure.

The structure of this treatise is laid out so that in chapter 2 the fundamental concepts of consumer behaviour in a discrete and differentiated context are presented. The concepts introduced in this chapter represent the basic notions on which we elaborate in the succeeding chapters. In chapter 3 we state our problem, and the notion of probability is introduced. In chapter 4 previous models such as logit are presented and discussed. In chapter 5 we proceed to discuss the general problem of assessment of probabilities from uncertain information. Here we consider problems both in physics and in information theory. In chapter 6 we set forth how producers will behave faced with the fact that the actions of the consumers are unknown; and finally in chapter 7 we derive the expected and the real market equilibrium for our problem. Eventually in chapter 8, we present an extension of the basic model, in addition to some comments on econometrics. In this chapter we also sum up our results.

1.3 On notation.

Since this is a work of economics, mathematical, logical and other technical terms and expressions which are not assumed universally known by the average economist, will be explained in the text or in footnotes.

I shall now introduce my terminology and state the meaning of these terms. In this treatise we are going to raise questions and give answers within a limited field of study. One of the simplest yet most effective ways of answering a logical question is through the use of a **definition**. A definition is an agreement, by all parties concerned, as to the meaning of a particular term. Thus, the definitions set forth in this treatise is expected to be generally accepted, although some might be of a more controversial character. When an answer to a question of interest can not be given as a definition we have to give it as a statement. A **proposition** is a true statement of interest that we are trying to prove. Some more important statements of interest are referred to as **theorems**. A **proof** is a convincing argument of the truthfulness of a proposition or a theorem. In executing a proof we may use supporting statements that are each proved separately, to support the proof of the main statement. These supporting statements are referred to as **lemmas**. In proving a statement we usually have to build on statements that are assumed to be true, but that can not be proved to be true. These unproved statements are referred to as **axioms**.

Some notation. The sign $\wedge$ is *logical and*; $\vee$ is *logical or*; a *set* is a collection of elements and denoted x = (....); $\times$ is the *Cartesian product* of two or more sets; $x \subset y$ means that the set x is *contained in* the set y; $x \in y$ means that the set x is an *element of* the set y; $\neg$ stands for the *logical not*; $\forall$ is the logical quantifier *for every*.

Chapter 2. Fundamental concepts of economic behaviour.

We concern ourselves in this chapter with the fundamental concepts of consumer and producer behaviour in a differentiated discrete market. Based on these fundamental concepts of behaviour, we will analyse in the following chapters the interaction between these groups of agents in this particular market. Thus this chapter sets forth the fundamental concepts describing the economic behaviour of both consumers and producers. The economic behaviour of both consumers and producers is assumed to be common knowledge to both groups, and anything which can be deduced by logic from the concepts introduced in this chapter is of course also common knowledge. The consequences of these concepts, i.e. the specific choices of the consumers, are however not assumed to be known, since the concepts only say how preferences are formed and not which preferences the consumers have. But, as we will show, the choices of the producers can be deduced from these concepts.

This theory concerns those markets which can be described as differentiated and discrete. The term differentiated means that the products among which the consumers may choose are similar but not equal in quality. By discrete we mean that the consumers can choose only one item of one alternative at the time.

2.1 The primitive concepts.

We start by assuming the existence of two different groups of individuals called the population of consumers and the population of producers, denoted W and B, respectively. A specific individual can of course be a member of both groups. The groups are separated by their purposes. The population of consumers have as their common purpose to consume, whereas the purpose of the population of producers is to produce.

There exists many different populations of consumers and of producers and a particular individual can be a member of several such populations. But, there exists no population of consumers, however, without a corresponding population of producers. Thus every population of consumers is distinguished from other populations of consumers by its relationship with a particular and unique population of producers, and vice versa. Two such connected populations of consumers and producers seen as a whole is called a market. Thus the population of consumers in a differentiated and discrete market, and the corresponding population of producers, together form the market of differentiated and discrete products.

A market is described through a description of the behaviour of the individuals of the populations, i.e. the behaviour of the populations, and the interaction between the different populations. Such a description is called a theory of the particular market in question. In the following we will set forth a specific theory of differentiated and discrete markets.

Two particular populations defining a market are distinguished from all other populations by the concepts which connect them, directly or indirectly. Thus a theory of a particular market is characterized by a certain set of concepts. But, as nothing comes from nothing, therefore some of the concepts characterizing our theory have to be primitive[1], in terms of which all other concepts are defined. Thus, we start out to set forth a set of primitive concepts which represent our chosen starting point. To start with, these concepts may be perceived as purely formal, but they are chosen on the basis of an impression or a general idea of what confronts the consumers and the

[1] A primitive concept is an initial and general idea or notion of a theory that can not be defined by or deduced from any other notions. It is a basic element of a theory without which the theory is meaningless.

producers when they are to make their choices. For all practical purposes the reader may choose to consider these concepts as chosen by the author as a "reasonable" set of basic concepts.

To complete the theory we shall present axioms which are statements about properties of the primitive concepts. From these axioms other properties which we shall name propositions or theorems are deduced. Thus the primitive concepts, the axioms, the propositions, and theorems together represents our theory. That is, a theory about the properties of the primitive concepts. If the primitive concepts are interpreted as economic concepts, as they are in our case, we have an economic theory. Thus by interpreting the primitive concepts as concepts describing the populations of consumers and producers forming together a non-specific differentiated and discrete market, we have a general theory of differentiated and discrete markets.

The primitive concepts are:

PC1: A population W.

PC2: A population B.

PC3: An abstract set X.

PC4: An abstract set C.

We will in the following first give a preliminary interpretation of these concepts, and then in later sections elaborate the properties of the concepts further.

Since we are to set forth an economic theory we will interpret the primitive concepts in economic terms.

PC1, the population W, is the population of consumers making choices from the choice space X. In the following we will describe how the members of W are behaving in making such choices.

PC2, the population B, is the population of producers, providing some of the elements of X, and choosing a price form the set C.

PC3, the abstract set[2] X, is called the choice space[3], which is the set of all elements that might possibly be elements of a choice set, denoted M. A choice set M is the set of elements from which a member of W has to choose one and only one element ,or point. Thus, a choice set M is a subset of X. On X we will also assume the existence of a binary relation R on X, which is a preference relation, which orders the elements of X in more preferred and less preferred.

PC4, the abstract set C is called the price-income space, which we interpret as the space of all prices that may be chosen by the producers of population B for the commodities that in combination can constitute the elements of X, and the incomes that may occur for the consumers in a given population W.

---

[2] Abstract set. On the term *set*: A set or a class is a group of things or objects, i.e. a collection of distinct numbers, objects, etc. that is treated as an entity in its own right, and with identity dependent only upon its own members. In other words, a collection of anything is a set. On the term of *abstraction*: Abstraction is the process by which allegedly we form concepts on the basis of experience or of other concepts, i.e. an abstract set is a set formed on a basis of a particular concept. If we have experienced the concept of human choice, then a choice set is an abstract set formed on the concept of choice. Thus the difference between a general set and an abstract set is that a general set may be a collection of anything, e.g. the set (5, the sun) is a general set, when an abstract set is a collection of something that is selected through the use of a particular concept.

[3] Choice space. On the term *space*: A space is a set of elements or members, called points, which are structured by using a set of axioms which the points have to satisfy. The elements of a choice space are those elements that are elements of a choice set. Thus the choice space contains as elements all the alternatives that can be chosen.

2.2 The choice space.

We have interpreted X as the choice space of the population W. As a space containing points or elements which are bundles of commodities, it has to have specific properties concerning the relationship between the points in question. A reasonable assumption is that the points of choice space X are related by a distance function or a metric describing the exact relationship between any two points of X. We therefore assume X to be contained in the metric space[4] of the Euclidian *I*-space[5]. Thus, we state:

S1 Axiom.

$X \subset \Re^{I}_{+}$, where $\Re$, $\Re_{+}$, and $\Re_{++}$ denote the Real numbers, the nonnegative reals, and positive Reals, respectively.

The assumption of axiom S1 that the choice space is contained in the Euclidian space of reals is not especially reasonable since we usually observe that consumers buy either 1, 2 or more items of a product, not, say, 0.456 of a car. But the assumption is stated as a matter of analytical convenience.

---

[4] A space where for any two points x and y there is associated a real number d(x,y) called the distance, such that (i) d/x,y) > 0, if $x \neq y$; (ii) d(x,y) = d(y,x); (iii) d(x,y) $\leq$ d(x,z) + d(z,y), for any $z \in X$. A function with these properties is called a distance function or a metric.

[5] Euclidian space or Cartesian space: The basic axiom the points of an Euclidian space have to satisfy is that the points are vectors. On the term *vector*: An element that can be located by a single n-tuple of coordinates. Thus the Euclidian space can be said to be an abstraction of the three-dimensional space of daily experience, i.e. the n-fold cartesian product of real fields, i.e. a space of n-vectors. In our particular problem every element is a point in the Euclidian *I*-space.

S2 Definition.

A choice set $M \subset X$, containing as elements all possible choices for a individual member of W.

From definition S2 follows that all possible choice sets for every member of W are subsets of X. Now we denote the set containing all the subsets of X that can possibly be choice sets for all individuals of W concerned, as T, i.e. $M \in T$, $T \subseteq X$.

Each choice set M consists of a number of elements or points. One set may contain only a few elements, when an other contains several.

S3 Axiom.

Every choice set $M \in T$ consists of consumer points $x=[\psi, x_I] \in \Re^I$, where $\psi=[x_1,...,x_{I-1}] \in \Psi$, where $\Psi=\{[x_1=1,x_2=0,...,x_{I-1}=0], [x_1=0,x_2=1,x_3=0,...,x_{I-1}=0],\cdots, [x_1=0,...,x_{i-1}=0,x_i=1,x_{i+1}=0,...,x_{I-1}=0], \cdots, [x_1=0,...,x_{I-2}=0, x_{I-1}=1]\} \subset \Re^{I-1}$, $x_I \in X_I \subset \Re$.[6]

In axiom S3 the consumer points of the choice sets can be interpreted as consisting of two commodities, a differentiated commodity denoted by ψ which is differentiated into *I-1* different alternatives of which the set Ψ is the complete collection of the *I-1* possible choices; and a general good x_I consisting of a blend of a all other commodities in the economy that a consumer may possible choose, of which the set X_I is the complete collection of all possible blends of this general commodity.

Our problem is to model the market of differentiated and discrete products. The

[6] A set is denoted by $(\cdots)$ or $\{\cdots\}$. A point, an element or a vector is denoted $[\cdots]$. A set is a collection of points, elements, vectors. A point or a vector is an ordered n-tuple of coordinates.

set Ψ contains all the possible elements that can be chosen by any consumer of the consumer population concerning the differentiated product. Since only one item of this product can be chosen by each consumer the set Ψ consists only of the unit and zero values of each of the *I-1* alternatives of this product. It should be noted that there is no such restriction is placed on the general good which can be consumed in several combinations, and in various quantities.

X_I is the set of all possible quantities of the general good x_I. It is now clear that the choice space consists of the Cartesian product[7] of Ψ and X_I:

[E1] $X = \Psi \times X_I \subset \Re^I$.

The elements or the consumer points of the choice space X are every combination of the elements of Ψ and the elements of X_I. If the set X_I contains V different elements, i.e. quantities, then since the set Ψ contains *I-1* elements the total number of consumer points or elements in X is $V \cdot (I-1)$.

In modelling a differentiated and discrete market, the focus is on how consumers choose from the set Ψ and how the producers who produce the alternatives of this set, determine their prices. We will return to how producers behave in later sections, in the following sections our focus will be on how consumers choose, i.e. on how the choice sets M are formed and how consumers choose a single element $x=[\psi,x_I]$ from these sets.

[7] Cartesian product or cross product: The set of ordered n-tuples the elements of which are respectively members of any sequence of the given sets. In our case of $\Psi \times X_I$, this means that any combination of the the members of these sets will be an element of the Cartesian products of these sets.

In the following sections we will describe the economic behaviour of the members of the consumer population W.

2.3 On consumers' ability to choose.

By consumer behaviour we mean how a member of W makes his choice from the choice space X. We are thus to describe how members of the consumer population W choose an element, that is a consumer point or vector $x=[\psi,x_I]$, $\psi\in\Psi, x_I\in X_I$, from their choice sets $M\subset X$.

Observe two such elements $x,y\in X$. It is reasonable to assume that a member of W have a preference for one of the elements or is indifferent between them. That is, the member of W has a perception of which of the elements he would like to receive if he had to choose among them, or he would be indifferent if he received one instead of the other. Thus it is reasonable to assume that every member of W is able to compare every two elements $x,y\in X$ and prefer one to the other or be indifferent between them. Such behaviour represents in fact a binary relation between every couple of elements in X:

S4 Definition.

If for every couple of elements $\{x,y\}\in X$, a certain statement about these elements, in the given order, can only be true or false, this statement establishes a binary relation on X. If this relation is denoted by R, we have xRy if the statement is true and $x\neg Ry$ if false.

There are many binary relations, but we are only interested in ordering relations which we define to have one of the following properties:

S5 Definition.

If $x,y,z \in X$ and R is a ordering relation on X, then R is said to be

a. transitive if $xRy \wedge yRz \Rightarrow xRz$,

b. complete if $\forall x,y \in X$: $xRy \vee yRx$,

c. a-symmetric if $xRy \Rightarrow y \neg Rx$.

That the ordering relation R is transitive: If *x* is ordered before *y*, and *y* before *z* then *x* should be ordered before *z*.

That R is complete means that there exists an ordering: For two elements *x* and *y*, either *x* is ordered before *y* or y before *x*. A situation where the elements of X is not ordered can not exist.

That R is a-symmetric means that a specific order implies that the reverse order can not be true.

Three different ordering relations of interest have at least one of these properties:

S6 Definition.

The ordering relation R means "at least as good as".

R(B) means "better than", and

R(E) means "equivalent to".

The relation R is transitive. The strict relation R(B) is transitive and a-symmetric. The relation R(E) is transitive. Neither of the relations are obviously complete. We assume that it is a property of the population of consumers W that the consumers have to choose an element from their choice set. They can not choose to choose nothing. Thus we assume that:

S7 Axiom.

All consumers of the consumer population W have an ordering relation on X that is a complete preordering, i.e. that it is transitive and complete. This requires

$xRy \wedge yRz \Rightarrow xRz,$

$\forall x,y \in X: xRy \vee yRx.$

From this assumption the following property can be deduced:

S8 Proposition.

$\forall x,y \in X: xR(B)y \vee xR(E)y \vee yR(B)x.$

Proof: Weddepohl (1970). It follows from the definition of R and R(E) that $xR(E)y \Leftrightarrow xRy \wedge yRx$, and from the definition of R(B) that $xR(B)y \Leftrightarrow xRy \wedge y\neg Rx$. Since it follows from axiom S7 that the ordering relation is complete and transitive the relations R(E) and R(B) have to exist, and proposition S8 is true. QED.

Proposition S8 says that if a consumer ranks the alternatives open to him in a transitive and complete way, then there exists an ordering relation that is such that either x is preferred to y, or y is preferred to x, or the consumer is indifferent between x and y. This is reasonable behaviour.

Proposition S8 implies that every member of W has a most preferred element or a group of most preferred elements among the elements open to him for choice from X. If the choice set $M \subset X$ denotes the set of elements of X that is open for choice to a member of W, then among the elements of M there has to be an element that is most preferred or a group of elements that is most preferred, since a binary relation which is defined on X is also defined on any subset of X. Thus if X is a choice space, and M is a choice set, certain points of M can eventually be considered as "best elements" to the consumer since it follows from proposition S8 that the consumer will prefer some points over other points in the choice set.

S9 Definition.

If R is a binary relation on a set X and if R is transitive and $M \subset X$, then $x^0 \in M$ is said to be a greatest element of M, if $\forall y \in M$: $x^0 R y$.

Then we may only have that $x^0 R(E) y$ or $x^0 R(B) y$. It is excluded that y and x^0 are not comparable, and that $y R(B) x^0$.

Hence a choice set M, always contains a greatest element, or a preferred choice, i.e. a "best element". That is, a consumer will always be able to say that there is a most preferred element or a group of elements that are most preferred to all the other elements, even though the consumer is indifferent between the elements within this group.

So far we have concluded that with every member of W is associated a choice set. With every choice set is associated its maximal elements, i.e. the elements preferred by the member to all other elements included in M. Thus we may define:

S10 Definition.

A set Υ is called a power set of a set X, if

$M \in \Upsilon \Leftrightarrow M \subset \Upsilon$.

Thus, the power set is the set of all subsets of X. If X is preordered, with every element of Υ, that is, with every subset of X, can be associated its maximal elements. We define a correspondence H: $\Upsilon \rightarrow X$.

S11 Definition.

If X is a set, preordered by a relation R, and Υ is its power set, we have for every $M \in \Upsilon$: $H(T) = \{x \in M \mid y \in M \Rightarrow xRy \vee y \neg Rx\}$.

If R is a complete preordering, then by proposition S8, H(M) is the set of greatest elements of $M \in \Upsilon$. That is the set of elements or the element that is preferred over all other elements of the choice set M of the consumer.

2.4 On how consumers make choices.

We have to this point assumed that R is transitive and complete. These assumptions say something about the members of W's ability to make a choice, in its own right, from a choice set. They do not say everything about how these choices are made. We are therefore to proceed by completely describing how the members of the population W make choices, but not which choices they make.

Assume a member of W is confronted by the problem of choosing between two elements $x,y \in X$ consisting of the same commodities, but in different quantities, i.e. $x \geq y$. If this commodity is desired in its own right by the consumer, then it is reasonable to assume that the consumer will prefer more to less of it:

S12 Axiom.

$x \geq y \Rightarrow xRy,$

$x > y \Rightarrow xR(B)y.$

Axiom S12 is called the monotonicity axiom since it assumes the existence of a ordering relation R on X, such that it orders a point *x* as better than an other point *y* if this first point *x* represents more of something that is wanted than the other point does. Together with axiom S7 which specifies that all consumers have a binary relation that is a complete preordering, this axiom defines what has been termed rational behaviour. Thus, we are assuming that the consumers are rational, which is reasonable.

S13 Axiom.

$\forall x^0 \in X$: $\{x \mid xRx^0\}$ is convex[8].

To assume that the ordering relation R on X is such that it ranks the elements of the choice space into a convex set of more or gradually less preferred choices, implies

[8] Convex set: A set $A \subset \Re^l$ is called convex, if $x \in A \wedge y \in A \wedge \lambda \in [0,1] \Rightarrow \lambda x+(1-\lambda)y \in A$.

that all points on a line segment connecting two points that are preferred to a third, are also preferred to this point, or equivalently, all points on a line segment are preferred to the worst of the two points it connects, which is reasonable.

2.5 A universal ordering.

We shall now introduce the concept of a utility function. Conceptually a utility function may be viewed as an expression of the mechanism that orders the elements $x \in X$. It is obvious from the previous discussion that the binary relationship R only says something about the relationship between to elements $x,y \in X$. Thus the relationship R can not be used directly to say something about the order of, say, the set of elements $X_K = (x_1, x_2, x_3, x_4, x_5)$. But, it is obvious that indirectly and with the help of some elaboration, a complete order can be established. The more extensive a set is the more elaboration is needed to establish the order of the set for a given consumer. It would be almost impossible to use the relationship R directly to establish the order for a given consumer of the universal set X. Since the complete order can be established by R, then an order-preserving function can be derived from the properties of R which completely orders the elements of X directly. The utility function is such a function. Thus the utility function u(x) for all $x \in X$ says something about the preference relationship between all the elements of X, when R only says something about the binary relationship between the elements of X.

S14 Definition.

If X is a set, completely ordered by a binary function R, then a mapping[9] $u: X \rightarrow \Re$ is said to be an order-preserving function, if only if

$u(x) > u(y) \Leftrightarrow xR(B)y,$

$u(x) = u(y) \Leftrightarrow xR(E)y,$

where $x,y \in X$.

Note that the choice space X now can be assumed to be completely ordered by the ordering relation R which have the properties set forth in axioms S6, S11, and S12. This means among other things, that the choice sets of X, T, are sets where the elements *x* are ranked according to the preferences of the consumers of the population W. That is, it follows that the choice sets can be viewed as ordered sets of elements which contain a best element that will be chosen by each consumer. Since the mapping *u* maps this ordering of elements from a point in the space $\Re^I$, to a number in $\Re$ in such a way that the rank among the elements of each choice set is contained, this mapping is called a utility function which associates with every point in X a real number, so that a point which is preferred to another gets a higher value, while equivalent points get the same value. We call the value u(x) a utility. In a subset, or choice set, $M \subset X$, the best element has the highest utility. From one particular utility function many others can be derived. Hence, a utility function is not uniquely determined. Thus, a utility function only indicates order and it does not measure intensities, that is, it constitutes an "ordinal scale".

We have so far used the axioms of choice to define an order-preserving function.

[9] A mapping *u* of a set X into a set $\Re$ ($u: X \rightarrow \Re$) is a law that connects with every $x \in X$ at least one element $u(x) \in \Re$.

The existence of such a function for all consumers concerned can be proved if some technical assumptions are made[10]:

S15 Theorem.

There exists a quasi-concave and continuous function

$u: X \rightarrow \Re$

where $u(x) \geq u(y)$, if xRy and $u(x) > u(y)$, if $xR(B)y$,

$x > y \Rightarrow u(x) > u(y)$,

$x \geq y \Rightarrow u(x) \geq u(y)$, if $x,y \in X$.

Theorem S15 is not proved here. For proof see among others, Weddepohl (1970).

Since the binary relationship R does not say anything directly about the universal preference relationship between the elements of X, only something about the binary preference relationship, it can not be used directly to decide which elements of a consumer's choice set M that is the greatest element or the preferred choice.

---

[10] These technical assumptions are:

The axiom of continuous preferences: $xRy \wedge yRz \Rightarrow \exists\alpha: 0 \leq \alpha \leq 1 \wedge yR(E)\alpha x+(1-\alpha)z$.This axiom says there exists a relation R on X which have the property that it is always possible to form a combination of x and z such that the consumer is indifferent between this combination and y. It is reasonable to assume that the consumers behave according to this axiom. If they did not, we would observe that if $xRy \wedge yRz$, than there would be no possible combination of a choice x, that is at least as good as y, and of z, which is not better than y, that are at least as good as y. Since y is not better than x, and not worse than z, this would seem strange.

The axiom of weak satiation: $xR(E)x+t \wedge t \geq 0 \Rightarrow \forall\varepsilon > 0, \exists\lambda > 0:[y \in B_\varepsilon(x) \Rightarrow y+\lambda tR(E)y+(\lambda+1)t]$.The axiom states that if addition of *t* to x is not appreciated, then after addition of some *t* to any alternative y, a new addition of *t* is not appreciated either. It can be said that the consumer is satiated with *t*, given x.

However, the utility function is a universal representation of the preference relationship between all the elements of the choice space, and more specifically, between the elements of a given consumer's choice set M. The greatest element of a choice set will, therefore, be given by the utility function simply by finding the element that has the highest utility for a given choice set M. This element is the preferred choice. Thus the term utility maximizing behaviour.

2.6 Price-income space.

A consumer can not pick whatever element he chooses from his choice set. Every element $x=[\psi,x_I]\in M\subset X$ consists of *I* coordinates or products where the members of the population of producers B produce the *I-1* alternatives of the differentiated product ψ. The general product x_I is produced by a different population of producers which do not interact with the population of consumers W. Consequently the price of the general product is given and fixed for the members of W. But the population of consumers W interact with the population of producers of ψ. We will return to how the members of the population of producers B set their prices, but we reveal that they are constrained to choose a price from the set of prices C. Since each member of the population W has a given income included in the set C and since every choice implies a price, every consumer can only choose that alternative or element which is obtainable given his income and the set of prices he faces.

Since the choices of the consumers is dependent upon the prices set by the producers and since the producers set their prices after assessing the preference positions of the products in the population W, the choice of product by the consumer and the choice of price of the producer is an intertwined process which we shall describe and analyse.

We are now going to introduce prices and income, and therewith constrain the choices open to the consumers of the population W.

Let C^{i}' be the set of all possible prices of the *i-th* alternative of the commodity ψ or the price of commodity x_I. Let the set C^{I+1}' be the set of all possible disposable incomes. Then $c_i' \in C^{i}$' and $m \in C^{I+1}$', and then

$$[E2]\ C' = \prod_{i=1}^{I+1} C^{i}{}'$$

is the set of all possible combinations of prices and incomes. Prices and incomes are always non-negative, hence $C' \subset \Re_{+}^{I+1}$. Now every point $c' \in C'$ is an I+1-vector $c'=[c_1',\ldots,c_I',m]$.

S16 Axiom.

Every member of the population of consumers W has the same income $m \in C$.

This assumption is not especially reasonable. But, it is essential for the technical results, i.e. we would hardly be able to arrive upon an equilibrium solution for our problem if this assumption is not to be stated. We will return later to why this is so, but indicate that this is due to simplicity. By assuming that all consumers have the same level of income we avoid the complicating problem of income distribution and the implications of such a distribution for our analysis. Removing the problem of income distribution is not believed to have any grave implications for our results in principle, but it simplifies the technical procedure in arriving upon a unique solution. This assumption should thus be viewed as an assumption of technical convenience.

Using axiom S16 another price-concept is derived: $c=[c_1,\ldots,c_I]\in C$, where

[E3] $c_i = \frac{c_i'}{m}$, $i=1,\ldots,I$.

This means that the points of C' are absolute prices, whereas the elements of C are "relative" prices expressed in disposable income. The income-component is here always set equal to 1 for all $i=1,\ldots,I-1$, and therefore omitted, and $c_i = [0,1]$ for all $i=1,\ldots,I$, i.e. all possible relative or normalized prices are between 0 and 1.

S17 Definition.
$C = \{[c_1,\ldots,c_I] \mid \exists c'\in C': c_i=\frac{c'_i}{m},\ i=1,\ldots,I\}$.

C is now called a price space.

It should be noted that X contains only elements which a consumer may eventually choose. The points of C represent price and income situations, which eventually may occur, seen from the perspective of the consumer. In fact C represents the relative prices that are chosen by the producers, given fixed incomes.

We now distinguish the following sets of elements of X and the sets of elements of the price space C:

S18 Definition.

If $c \in C$,

$L(c)=\{x \in X \mid c^T x=1\}$,

$M(c)=\{x \in X \mid c^T x \leq 1\}$, where c^T is the transpose of *c*.

The set M(c) contains all points or elements that can be bought at the relative price vector *c*, since their value does not exceed the amount of 1. Therefore, the sets M(c) will be named budget sets. The sets L(c) of elements that cost exactly 1 will be named budget planes. Clearly $L(c) \subset M(c)$ for any $c \in C$. If $M(c) \in T$ then this budget set is a choice set. Hence, M(c) represents a possible choice situation for the consumer. Thus, we introduce

S19 Definition.

$C^\wedge = \{c \in C \mid M(c) \in T\}$.

Definition S19 gives the set of prices that makes all consumer points or elements obtainable to one or more consumers in a given population W. Since the set C contains as elements all possible prices that may occur, this means that there may be prices for which no one can afford to buy one or more of the elements of the choice space. Thus, C^ gives the set of prices that makes all points obtainable to someone.

From definition S19 it follows that

[E4] $c \in C^\wedge \Rightarrow M(c) \in T$, which means that if c is an element of C^, then the subset M(c) is an element in the set of choice sets. In other words, the set C^ is the set of price vectors $c \in \Re^I$ for which all the elements of X are obtainable to some members of the population W. Thus if the realized price vector is an element of C^ then every element of X is eligible for choice to some member of W. But it does not mean that every element of X will be chosen.

S20 Axiom.

$M(c) \in T \Rightarrow c \in C^\wedge$.

Axiom S20 says that if the subset M(c) is a choice set, then c has to be a vector of prices that makes the elements of M(c) obtainable to some consumer. This is reasonable, since a budget set M(c) is a choice set of elements that are obtainable for some members of W. Thus c has to be an element of C^.

By combining [E4] and axiom S20 we get:

[E5] $C^\wedge = \{c \in C \mid H(M(c)) \neq \varnothing\}$,

which says that C^ is the set of price vectors for which the choice sets M(c) for all c, have a greatest element. Thus, the set C^ is the set of price vectors that secure the existence of a set of eligible elements since a greatest element is an eligible element. It is only among a set of eligible elements that a greatest element can be found.

The existence of a set of eligible elements is established. It follows from the above axioms of choice that the consumers of the population W will choose the greatest element of such a set open to them. We therefore turn to the question of how this greatest element is identified. This is the question of demand.

2.7 The demand function.

A demand function is an expression of which element a consumer will choose given the prices and the income he faces. Since we are operating through disposable income prices C this reduces to the problem of which element the consumer will choose given the disposable income prices.

It follows from theorem S15 that a utility function for a member of W is a universal ordering of this member's choice set M(c), where the most preferred element of the member's choice set has the greatest utility value. It follows from [E5], i.e. axiom S20 and definition S19, that for $c \in C^\wedge$, there exists at least one greatest element in M(c). Thus, if M(c), $c \in C^\wedge$, is the choice set of a member of W, and D(c) denotes the demanded element or group of elements for $c \in C^\wedge$, then the demanded element is that element of M(c) for which the utility value is the highest:

S21 Theorem.

$\forall c \in C^\wedge$: $D(c) = x^0$, where $u(x^0)=u^0= \max\limits_{x \in M(c)} u(x)$.

Proof: Assume $c \in C^\wedge$, then from [E5] it follows that $H(M(c)) \neq \varnothing$, thus there exists a greatest element. From theorem S15 if follows that any greatest element or group of greatest elements has the highest utility value. And from definitions S9 and S14 it follows that the element with the highest utility is the most preferred element in M(c). QED.

We have now completed the description of how members of the consumer population W choose. But it follows from the previous description that the choice is both dependent upon which sets are obtainable and which sets are eligible. Thus the consumer will choose the obtainable and eligible element from his choice set. Which elements are eligible can not be influenced by any action by members of the producer population B. But these members of B set the prices and thus decide which elements are obtainable by the members of W. The choices of the consumers can therefore be influenced by the actions and choices of the members of B.

On the other hand, the preferences of the members of W influence the choices of the members of B. The preferences of the members of the population W is not known to the members of B and have thus to be assessed. Based on this assessment the members of the population B will set their prices $c \in C^\wedge$ and consequently the members of W will make their choices, i.e. the elements $x^0 = D(c)$ which will be chosen for each member of W. For each member i=1,...,I-1 of B the question is how many members of W will choose a consumer point $x^0 = [\psi^0, x_I^0]$ containing $x_i=1$, $x_i \in \Psi$. Define the general commodity x_I to include all products except the differentiated product ψ, included savings, i.e the choice set is a budget plane L(c):

S22 Axiom.

Every choice set is a budget plane, $L(c) \in T$.

As a consequence of axiom S22, since the income is assumed to be equal for all members of W, the quantity of the general commodity x_I will be equal to all members who choose a particular alternative i=1,...,I-1, and every member of W will face the same common choice set L(c) for a given price vector $c \in C^\wedge$. Denote an element of X that contain $x_i = 1$ by x^i. Let x_I^i denote any value of the general good for an element $x=x^i$. We state:

S23 Proposition.
$\forall c \in C^\wedge \Rightarrow L(c)^T = (x^1(c),...,x^i(c),...,x^{I-1}(c))$, where $x^i(c) \in X$ for all i=1,...,I-1, for all members of W.

Proof: In virtue of axioms S3 and S22, $c \in C^\wedge \Rightarrow x_I^i = x_I^i(c)$ for all i = 1,...,I-1, and for a member of W. But, since every member of W has the same income due to axiom S16, then $x_I^i(c)$ $\forall c \in C^\wedge$ has to be the same for every member of W. Thus, every choice set is the same to every member of W, and this common choice set L(c) for a given $c \in C^\wedge$ contains *I-1* elements. QED.

The implication of proposition S23 is that every member of population W will have to choose among the same choice set $L(c) \subset X$, for a given price vector $c \in C^\wedge$. Thus there exists as many choice sets as there are different possible price vectors.

We may now view the choice situation of every member of W as choosing an alternative $x^i = (\psi^i, x_I^i)$, where $\psi^i=(x_1=0,...,x_i=1,...,x_{I-1}=0) \in \Psi$, among *I-1* alternatives included in the choice set L(c), $c \in C^\wedge$.

Chapter 3. Outline of the problem of producer assessment[11].

In chapter 2 we described how consumers make their choices within the context of discrete and differentiated markets. We note that there is no uncertainty at the hands of the consumers. However, the underlying assumption of our model is that the consumers make their choices after the producers have decided upon their offer to the market, i.e. their prices, and the quantities they will offer at these prices. Hence the sequence of actions in our model is such that all producers first and simultaneously set prices and quantities, and then the consumers decide which brand they want to consume, i.e. every consumer decide which brand $i=1,\dots,I-1$ is $x^i \in D(c)$.

It is obvious that the actual demand $D(c)$ will not be observed by the producers prior to their own decisions. Hence consumer demand is by nature uncertain from the point of view of producers. It is this uncertainty which is assumed away in most of the literature on producer behaviour in differentiated markets. We are to focus on this uncertainty, and will show how producers, if rational, will cope with it. Finally we will give some indications of how this uncertainty affects the equilibrium in a differentiated market.

The actual question facing the producers in population B is how the members of W are distributed over the elements $x^i \in L(c)$, for all $i=1,\dots,I-1$, for a given price vector $c \in C^\wedge$, and for whatever information is available. This problem of the distribution of the members of W facing the producer population B is to be the subject of our attention in the following chapters. In the next chapter we will present and discuss the most common approach to this problem: The logit model, which is derived assuming that the distribution of choices is known for the population W as a

[11] A part of this chapter leans to some extent on Jaynes (1957).

whole. In the successive chapters an alternative approach is presented where the distribution is derived using only market information.

The fundamental problem facing the members of population B is to measure the choice of the members of W from the set X. Since the members of B do not know the preferences of the members of W, they do not know their choices either. Thus the members of B is confronted with lack of information, or, in an other word, uncertainty, in making their own decisions.

The members of B therefore have to assess in one way or another the choices of the consumers of W. A useful concept in making such an assessment is the concept of likelihood or probability, i.e. how likely the occurrence of an event is, expressed in fractions. Thus the degree of information is reflected in the probability distribution conceived by the decision maker, i.e. among the members of B. The greater the extent of information available to the decision maker, the more concentrated the distribution.

The central question which we are to answer later is: How is the decision maker to specify the relevant probability distribution? This problem of specification is, in cases where little or no information is available, as old as the theory of probability. Laplace's "Principle of Insufficient Reason" constitute an attempt to supply a criterion of choice in which one said that two events are to be assigned equal probabilities if there is "no reason to think otherwise". However, except in cases where there is an evident element of symmetry that clearly renders the events "equally possible", this assumption may appear just as arbitrary as any other criterion. Consider the following example submitted by Borch (1968): View the effect of the outcome of a political election, say, in Britain on the exchange rate. The crucial factor is if Labour gets a majority or not. There can be three outcomes in the election: Labour wins, the

Conservatives win, or the Socialdemocrats win. In this case Laplace's principle assigns a probability of 1/3 to each event. However, if the problem is structured as to the relevant events "Labour gets a majority" or "Labour does not get a majority", Laplace's principle would assign a probability of 1/2 to the event that Labour wins the election. However, this contradicts the fact that we first assigned a probability of 1/3 to the same event using the same principle.

Therefore, since Laplace, this way of formulating such problems has been largely abandoned, mainly because the lack of any constructive principle which would give the decision maker reason for preferring one probability distribution over another in cases where both agree equally with the available information.

For further discussion of this problem, one must recognize the fact that probability theory has developed into two different "schools" as regards the fundamental notions. The "objective" school of thought regards the probability of an event as an objective property of that event, always capable in principle of empirical measurement by observation of frequency ratios in a random experiment. In calculating a probability distribution the objectivist believes that he is making predictions which are in principle verifiable in every detail. The ultimate test of a good objective probability distribution is : Does it correctly represent the observable fluctuations of the variable of interest, i.e. x?

The "subjective" school of thought regards probabilities as expressions of human ignorance. Thus the probability of an event is merely a formal expression of our expectation that the event will or did occur, based on whatever information is available. In the eyes of the subjectivist, the sole purpose of probability theory is to help us in forming reasonable conclusions in cases where there is not enough information available to lead to certain conclusions. Thus detailed verification is not expected. The test of a good subjective probability distribution is: Does it correctly represent our state of knowledge as to the value of x?

It should be evident that the assessment problem facing the producers arise as a consequence of the producers ignorance as to the actions of consumers. This is the reason why producers have to make assessments. Hence any probability distribution used by any producer has to be viewed as this particular producers subjective assessment of the uncertainty he faces.

It should be recognized that the theories of subjective and objective probability are mathematically identical. But the concepts themselves cannot be united. In analysing human economic behaviour, and the involvement of probability in the modelling of the assessment made by a group of economic agents of the steps taken by an other group of economic agents, we have to adopt the subjective point of view. The subjective point of view is certainly one of the fundamental ideas behind the model presented in this treatise. Our idea is that since the producers of B face uncertainty or lack of information they have to assess the probability distribution of choice based on the information available to them about the preferences of the members of W and the prices they themselves decide upon. It should be evident that the subjective thought is fundamental to our analysis.

This having been said it should be noted that producers are not expected to set their probability distributions randomly. Assuming rational behaviour one expects to observe that the producers analyse the situation, and, based on the available information, decide which distribution to use. It is this analysis we are to model in the subsequent chapters.

Since any probability distribution is assumed to be based entirely on the information available to the producers as to how consumers will react, we are to discuss and state what information it is reasonable to assume that the producers will possess.

Let Q_i denote the information set of the i-th member of B. The question we will have to ask is: What relevant information is it reasonable to assume that the producers will possess? Surely any member of B know how his fellow producers determine their offers in the market. We assume this procedure is the same for every producer, and that every producer knows this. Furthermore it is reasonable to assume that the producers know how the consumers make their choices, but of course not which choice they make, i.e. the principles guiding the consumers in making their choices are included in the information set. Usually the producers will have access to historical information of some kind. The question remains how relevant such information is. The only historical information which can be said to have direct relevance to the assessment problem facing the members of a particular industry is the historical market information for the industry. This information gives a sort of feedback to the industry members and ought therefore to be considered relevant. Among other possible historical information is information regarding the whole economy, such as inflation, and expected growth rates. However, how relevant such information is for a particular industry is difficult to say, and this sort of information is also usually embedded in the historical market information in some way.

Furthermore, such firm specific information as marketing effort and brand quality, in addition to the information provided by market surveys, should be considered relevant. As we will show in chapter 8, such information can easily be comprised by our model. However, this sort of information is not included in our basic model, which we are now to derive, since we will in the following chapters focus on how producers assess consumer demand and market equilibrium. When these issues are concluded, we will in section 8.1 extend the information set to account for a broad range of relevant information, and show that our basic results concering producer assessment and market equilibrium also applies to the extended information model.

S24 Axiom.

The information set $Q_i = Q$ for i=1,...,I-1, and contains:

1. The principles of producer behaviour which will be stated below.
2. The principles of consumer behaviour which are stated above.
3. Market information for the previous periods the industry has been in operation.

The market information for previous periods is represented by the transactions and the prices that materialized in different periods. The transactions of the last foregoing period *t-1* is expressed $n_{i,t-1}$, for all i=1,...,I-1, where $\sum_i n_{i,t-1} = N$, where *N* is the total number of consumers in population W. The market share of firm *i* in period *t-1* is $d_{i,t-1} = n_{i,t-1}/N$. The price information for the previous period *t-1* is expressed by $c_{i,t-1}$ for all i=1,...,I-1.

The market information can be represented by an expected industry price level for the present period:

$$[E6]\ \bar{c} = \sum_i d_{i,t-1} \cdot c_{i,t-1}.$$

Now the question is how this information is to be used by the members of B to assess or measure the demand from the members of W. Specifically this is the problem of measuring the choices of the members of W from the choice space X. However, only measurable events can be measured. Hence we will have to assure the existence of a measurable space of events on X. We observe that the choices made by the consumers can be classified as events. A measurable space (X,Γ) is now defined on X, where Γ is called a σ-algebra and is the set of all measurable subsets of X. A measurable subset is a set of X which is capable of being measured in some way. The measure we apply is probability. Let $p(\Gamma \mid Q)$ denote a probability distribution over the measurable space. The subjective probability distribution of any member of B is defined:

S25 Definition.

$p: (X,\Gamma) \rightarrow p(\Gamma \mid Q)$, such that $p(X) = 1$.

Thus the measurable space (X,Γ) is the space of subsets or events capable of being assigned a probability. The space of all possible probability distributions $p(\Gamma)$ over (X,Γ) is called the probability space, and is denoted (X,Γ,p).

We note that axiom S3 directs that the consumers can choose only one brand and one item of this brand. Consequently the measurable subsets of Γ have only one element, i.e. one of the alternatives $x^i \in L(c)$, $i=1,\dots,I-1$, $c \in C^\wedge$. Every producer recognizes that $L(c) \subseteq (X,\Gamma)$ for a given $c \in C^\wedge$. The producers observe that for a given price vector $c \in C^\wedge$ the probability distribution which they seek is the distribution of the N members of W over the L(c), i.e. $p \in \Re^{I-1}$, $p=[p_1,\dots,p_i,\dots,p_{I-1}]$ where $p_i =$

$p(x^i \in D(c) \mid Q,c)\ \forall\ i=1,...,I-1$. Accordingly we observe that the probability space $(X,\Gamma,p) \in \Re^{I-1}$ is the set of all probability distributions that can be assigned to all the possible choice planes for every $c \in C^\wedge$.

As we postulate that this procedure is being observed by all producers we conclude that the assessment problem of the members of B is to decide upon which of the probability distributions that are included in (X,Γ,p) they are to use in measuring demand from the consumers. The answer to this question is in principle that the members of B should use the probability distributionwhich conforms best with the information set, and which does not take in to account any other information than the information that is included in Q. Subsequently the question is: Which element of (X,Γ,p) do conform best with the information included in Q? This problem is to be the objective of the discussion in the next two chapters.

Finally it should be recognized that the questions stated above coincide with the question stated in connection with the test of subjective probability. Hence every producer will have to pick out the probability distribution that he believes best represents his ignorance as to how the members of W will choose. To answer this question we have to describe how producers, if rational, would behave in a discrete and differentiated context. An analysis of how producers behave in such a context, involving the problem of assessment of consumer behaviour and strategic considerations, is the theme of chapters 6 and 7.

It should be noted that from definition S25 a kind of "expected" price level can be calculated, using the prices that will be set, and the probabilities assessed by the producers, i.e. $\dot{c} = \sum_i p_i \cdot c_i$, where p_i is the probability that an arbitrary consumer will choose brand *i*. However, $\dot{c}$ and $\bar{c}$ are not in principle identical since $\bar{c}$ is a bit of the information that is used to derive p_i and hence c_i for all $i=1,...,I-1$, and since $\bar{c}$ is a maximum accepted price level and hence not necessarily equal to $\dot{c}$ in any case.

Chapter 4. Previous models and logit.

4.1 Introduction.

The producer's problem of assessing demand in a differentiated and discrete context, i.e. the question of which element of (X,Γ,p) that conform best with Q, is a problem which has not been solved nor focused in the litrature. However, some attention have been given to the isolated problem of how consumer demand in a discrete choice context is to be represented from the analyst's point of view.

When analysing consumer demand in isolation in a discrete choice context several analysts have turned to the so called random utility model. The most prominent is perhaps McFadden (1974, 1975) using the logit model, but also Berkovec & Rust (1985), Bos (1970), Börsch-Supan (1987), Cameron (1984), Li (1977), Train (1980,1986), and others, use this approach.

These models deal only with how demand as such is to be represented by analyst in a discrete choice context. However, one may identify the analyst's problem with the producer's assessment problem. Then the random utility model may be assumed to be applied by producers in assessing the demand from consumers in a differentiated and discrete context.

The most prominent contribution dealing with the whole problem of interaction between consumers and producers in a differentiated and discrete market is made by Anderson, de Palma & Thisse (1992). In short, Anderson et al assume the consumers to behave according to the so called random utility maximizing behaviour, and the demand is consequently represented by an expected demand function using logit to describe the underlying probability distribution. A population of producers are introduced and an equilibrium is derived. Anderson et al (1992) do not explicitly assume that the producers believe the consumers to behave according to the random utility model. However, in our context this is the essential consequence.

In the follwing sections we will briefly discuss this approach to the assessment problem and in particular analyse how the random utility model conform with the information in Q. We will show that such a model will not be used by the producers since it assumes the existence of additional information not included in Q. A model based on such fundamental notions as random utility have also to be regarded as controversial.

We will first of all give a short presentation of the random utility approach, and what assumptions that have to be made to arrive at a probability distribtuion over L(c) that can be used by the producers. Some of the properties of this distribution is pointed out, and finally we show that the producers acting rationally, will not use this approach to their assessment problem since it assumes that the producers have knowledge of how the preferences of the consumers are distributed. Such information is not included in the inforamtion set Q.

4.2 Random choice.

The random utility theory is based on the above referred axioms of consumer choice, with an additional assumption of importance. This additional assumption is that there is a random element in the utility function representing the consumers preferences. The motivation for introducing this random element seems to be the analyst's problem of uncovering the factors influencing the choice of a given consumer.

The utility of a choice of $x^i \in L(c)$, $c \in C^\wedge$, to a member *n* of population W is:

[E7] $u_{inc} = u(x^i \in L(c)), \forall i=1,\ldots,I\text{-}1, \forall n \in W, \forall c \in C^\wedge$.

A consumer *n* is maximizing his utility if and only if he chooses alternative *i* , that is $x^i = (\psi^i, x_I^i)$, if and only if $u_{inc} > u_{inc}, \forall j=1,\ldots,I\text{-}1, j \neq i$, for a given $c \in C^\wedge$.

We will now introduce a random element into the utility index [E7]. This introduction of a random element is due to different reasons. It is argued that from the analyst's point of view the choice of any consumer seems to be random, when dealing with discrete choice from the set X. Since a consumer only chooses one alternative from a set of alternatives a small change in price, income, preferences or attributes of an alternative may cause a fundamental change in the demand from the consumer in population W, in contrast to what we expect to find in the continuous case. Thus we may experience unforeseen and discrete jumps in demand which seem to be generated by random preferences. An other argument forwarded by some researchers is that the preferences of the consumers are truly random. That is, the consumers do not know themselves which alternative they prefer until the moment they actually make the choice.

The idea is that the utility index [E7] consists of a deterministic component and a random component:

[E8] $u_{in} = v_{in} + \varepsilon_{in}, \forall i=1,\ldots,I\text{-}1, \forall n \in W$, for a given $c \in C^\wedge$,

where ε_{in} is a random variable.

[E8] says that the utility function of every member of W is stochastic.

Next we introduce the notion of a representative consumer by assuming that the deterministic component, v_{in}, is equal for all $n \in W$ for a given $c \in C^\wedge$:

[E9] $u_{in} = v_i + \varepsilon_{in}, \forall i=1,\ldots,I-1, \forall n \in W$, for a given $c \in C^\wedge$.

In [E9] v_i is a value that is equal to the population of consumers as a whole, and ε_{in} is an individual stochastic utility value.

It now becomes impossible to determine with certainty which alternative a given consumer will prefer since the consumer's utility index is stochastic. The only measure that can be forwarded is the probability that a consumer will choose a particular alternative. The concept of probability is already defined. The choice probability, that is the probability that a member of W will choose a particular alternative, is denoted by $p_i(c)$, $i=1,\ldots,I-1$, for a given $c \in C^\wedge$. Within the context of the theory of random utility the choice probability may be interpreted as the probability that $u_{in}>u_{jn}$ for all $i=1,\ldots,I-1$, $j \neq i$, for a given $c \in C^\wedge$, that is, the marginal utility of alternative *i* is greater than any other alternative:

[E10] $p_{in}(c) = \text{prob}(u_{in}>u_{jn} \text{ for all } i=1,\ldots,I-1, j \neq i, c \in C^\wedge)$.

From [E9] and [E10] we derive

[E11] $p_{in}(c) = \text{prob}(v_i - v_j > \varepsilon_{in} - \varepsilon_{jn} \text{ for all } i=1,\ldots,I-1, j \neq i, c \in C^\wedge)$.

Since ε_{in}, ε_{jn}, for all i=1,...,I-1, j≠i, are stochastic values, the difference is also stochastic.

Since all members of W are guided by the same principles when their preferences are formed it is not unreasonable to assume that the distribution of the stochastic component is the same for every member of W:

[E12] $F(\varepsilon_n) = F(\varepsilon)\ \forall n \in W$.

$F(\varepsilon_n)$ denotes the distribution of the stochastic component for the n-th member of W.

From [E11] and [E12] we derive the probability that a consumer will choose brand *i* if the price vector is $c \in C^\wedge$:

[E13] $p_i(c) = \int_{\varepsilon_i=-\infty}^{\infty} (\varepsilon_j < \varepsilon_i + v_i - v_j,\ j=1,...,I-1,\ j\neq i,\ c\in C^\wedge)dF(\varepsilon),\ \forall x^i \in L(c),$

$\forall c \in C^\wedge$.

If $F(\varepsilon)$ is known then the choice probabilities p_i for all i=1,...,I-1, can be computed.

4.3 The logit model.

[E13] is the general expression derived using stochastic utility maximization. A meaningful interpretation and computation is denied us, however, if we do not know the functional form of $F(\varepsilon)$, the general distribution of the stochastic utility component.

The most frequently assumed distribution for the stochastic component is the extreme value distribution and the normal distribution. Of these the extreme value distribution is the most used according to Train (1986). If the stochastic component is assumed to be distributed by the extreme value distribution then the resulting model for the choice probabilities expressed by [E13] is the logit model.

The extreme value distribution is defined as follows:

S26 Definition.

The extreme value density function for every ε_i, $i=1,...,I-1$, is given by $\exp\{\varepsilon_i\}\cdot\exp[-\exp\{\varepsilon_i\}]$, where $\exp[-\exp\{\varepsilon_i\}]$ is the cumulative distribution.

If we now assume that the preferences of the consumers are distributed by the extreme value distribution:

[E14] $F(\varepsilon) = \exp\{\varepsilon_i\}\cdot\exp[-\exp\{\varepsilon_i\}]$, for $i=1,...,I-1$.

If [E14] is inserted into [E13] this will result in the logit model:

Assume that ε_i takes on a particular value, φ. It then follows from [E11] that the probability that alternative *i* is chosen is equal to the probability that $\varepsilon_j<\varphi+v_i-v_j$ for all $j=1,...,I-1$, $j\neq i$. The probability that $\varepsilon_i=\varphi$, and at the same time $\varepsilon_j<\varphi+v_i-v_j$ for all $j=1,...,I-1$, $j\neq i$, is the density of ε_i evaluated at φ multiplied by the cumulative distribution of every ε_j with exception of ε_i which is evaluated at $\varphi+v_i-v_j$. If definition S26 is applied we may write $p(\varepsilon_i=\varphi \wedge \varepsilon_j<\varphi+v_i-v_j\ \forall j=1,...,I-1,\ j\neq i)$

$= \exp(-\varphi)\exp(-\exp(-\varphi))\prod_{j,j\neq i} \exp(-\exp(-\varphi-v_i+v_j))$. But, since $v_i-v_i=0$, we write

$$[E15]\ p(\varepsilon_i=\varphi \wedge \varepsilon_j<\varphi+v_i-v_j\ \forall j=1,...,I-1,\ j\neq i) = \exp(-\varphi)\prod_{j}\exp(-\exp(-\varphi-v_i+v_j)).$$

We know that the stochastic component ε_i does not have to be equal to φ. ε_i can take on any value for which it is defined. The right-hand side of the expression in [E13] may thus be expressed as the sum of all values of φ:

$$[E16]\ p_i = \int_{\varphi=-\infty}^{\infty} \exp(-\varphi)\prod_{j,j\neq i} \exp(-\exp(-\varphi-v_i+v_j))d\varphi.$$

An evaluation of the integral gives

$$[E17]\ p_i = \int_{\varphi=-\infty}^{\infty} \exp(-\varphi)\{-\exp(-\varphi)\sum_{j} \exp(-\varphi-v_i+v_j)\}d\varphi.$$

We now substitute $\exp(-\varphi)$ with ϕ, then $-\exp(-\varphi)d\varphi = d\phi$ and $d\varphi = -(d\phi/\phi)$. These expressions are now inserted into [E17]:

$$[E18]\ p_i = \frac{\exp(v_i)}{\sum_{j} \exp(v_j)},\ \text{for all } i=1,...,I-1.$$

Thus the logit model is derived based on especially two crucial assumptions, which is not to be found in traditional analysis of economic behaviour:

1.The utility function of a member of W has a stochastic element.

2. The stochastic utility element is extreme value distributed.

The assumption regarding the existence of a stochastic utility element is to some extent discussed above. It is mainly due either to a practical way of analysing the particular problem of discrete choice, or to the belief that the utility function is in fact stochastic to some extent.

The assumption that the assumed stochastic component is extreme value distributed is mainly due to the fact that this assumption generates choice probabilities that are easily interpreted and computed. How reasonable this last assumption is we are to elaborate on to some extent below.

An implicit assumption that follows from the assumed extreme value distributed stochastic component, is that the stochastic utility component has a mean value of zero. The deterministic component v_i is thus characterized as the representative utility or the expected utility.

The logit model of choice probability has three important characteristics:

The first to be mentioned is that every choice probability has a value between 0 and 1: $0 \leq p_i \leq 1$, for all i=1,...,I-1. This characteristic implies that if an alternative is not strongly sought after, then the expected utility will approach $-\infty$, and the choice probability will subsequently go towards zero.

The second characteristic of the model is that the probabilities add up to one.

The third characteristic says that the relationship between the choice probability of alternative $x^i \in L(c)$, and the expected utility of alternative $x^i \in L(c)$ is simoid, i.e. S-curved. This implies that if the expected utility is small compared to the other alternatives, then the choice probability of alternative $x^i \in L(c)$ will respond only in a very limited way to an increase in the expected utility of this alternative. The interpretation of this is that if an alternative is not at all preferred, than a marginal change in its status does not increase its likelihood of being chosen. This seems reasonable.

The three above mentioned characteristics should be considered reasonable for an economic model. But, the model has a fourth characteristic which also can and should be discussed. This characteristic we label independence from irrelevant alternatives.

S27 Definition.

When the relationship between two choice probabilities p_i and p_j, for all $i,j=1,...,I-1$, are dependent upon the alternatives alone, then these probabilities are independent from irrelevant alternatives.

That the logit model of choice probabilities are independent of irrelevant alternatives is easily shown. Study the relationship between p_i and p_k:

$$[E19]\ \frac{p_i}{p_k} = \frac{\exp(v_i)/\sum_j \exp(v_j)}{\exp(v_k)/\sum_j \exp(v_j)} = \exp(v_i - v_k),\ \text{for all } i,k=1,...,I-1.$$

We observe that the relationship between these two probabilities does not depend on other alternatives then $x^i, x^k \in L(c)$. This implies that the relationship between these choice probabilities is constant and independent of the other elements of L(c).

The Independent of Irrelevant Alternatives characteristic could be considered reasonable in some situations, and obviously unreasonable in others. This can be illustrated by an example constructed by Train (1986):

Consider that a traveller has to choose between using his own car or riding a red coloured bus. Assume that his deterministic utility component is equal for using the car or taking the red bus. The choice probabilities for using the car or taking the red bus are computed to $p_{oc} = p_{rb} = 0.5$ using the logit model. This seems reasonable.

We now introduce a third alternative, a blue bus. Assume that the traveller's deterministic utility components for the red and the blue bus are equal. This implies that $p_{rb}/p_{bb} = 1$. We already know that $p_{rb}/p_{oc} = 1$. The only probability that satisfies these IIA conditions are $p_{oc} = p_{rb} = p_{bb} = 1/3$. Obviously we would expect the probability of choosing his own car to be the same before and after we introduce a substitute to the red bus. That is, we would expect $p_{oc} = 0.5$ and $p_{rb} = p_{bb} = 0.25$. Thus the Independent of Irrelevant Alternatives characteristic of the logit model underestimates the choice probability of the own car alternative.

4.4 Random utility and market equilibrium.

The problem analysed in this treatise is the modelling of the interaction between consumers and producers in a differentiated market of discrete choice. This problem has been addressed by some other authors, primarily Anderson, Palma & Thisse (1992). Anderson et al approach the problem using the logit model to represent consumer behaviour. Thus they assume that the utility indexes of the consumers are stochastic, and that the distribution of the stochastic element is extreme valued. This may be arguable, but one thing is certain: It is not a general principle that the tastes of the consumers are extreme value distributed. The use of the extreme value distribution is obviously a choice, not a result derived from general accepted axioms. Hence this assumption cannot be regarded as a fact an included in the information set Q of the producers. This at least excludes the logit model as a starting point in the discussion of the actual problem.

The starting point should be fundamental assumptions about consumers' and producers' behaviour, and from these assumptions the microeconomic structure of the problem under study should be derived. We will show in this treatise that a model that characterizes how the consumers view themselves and their opposite numbers, the producers, and how the producers view the consumers and their own behaviour, can be derived from general accepted assumptions about consumer and producer behaviour, such as utility maximizing consumer behaviour and profit maximizing producer behaviour.

Thus, there is no need for disputable assumptions such as extreme value distributed tastes, or random utility consumer behaviour. However, it should be mentioned that technically the model derived in this treatise is in some ways similar to the model derived using the extreme value distribution and random utility.

Chapter 5. Assessment of behaviour: A discussion.

5.1 Introduction.

As stated in chapter 3, the problem facing the members of the population of producers is to assess the distribution of the members of the population of consumers over the *I-1* different alternatives open to them. The only information in the possession of the producers is the information revealed above in the analysis of consumer choice. This is known by every member of the producer population B. This information, however, only reveals how the members of the consumer population W will choose in general, i.e. the principles guiding their choices. It says nothing about the specific choices made by the consumers at given prices. Therefore, these choices have to be assessed by the producers. It is obvious that the assessment made by the producers has to be in accordance with the principles of consumer behaviour which the producers know, i.e. the axioms of choice and their implications.

The assessment problem of the producers can now be stated more precisely:

Assume that the price vector $c \in C^{\wedge}$ is presented to the members of W by the members of B. Then the individual member of population B has to decide on which distribution of the members of population W over the alternatives $i=1,\ldots,I-1$ he believes will occur as a result of the introduction of this particular price vector.

It should be recognized that the previous passage contains two important points: Firstly it is assumed that a distribution will occur over the members of W. This follows from the axioms of choice. Second, each member of population B has to make a decision on what he believes will be the resulting distribution. This belief has to be in accordance with the axioms of choice. Thus any distribution which satisfies the axioms of choice could be used by some member of population B. On the other hand, if there exists a unique an unbiased distribution which is implied by these axioms that distribution will be used by every member of the producer population B.

In general the problem facing the producer is a problem of allocating the consumers to a finite number of states, or over a finite number of alternatives for the given information, e.g. for $c \in C^{\wedge}$. This problem could be solved by 1) either assuming that the producer knows more than we in fact are allowed to assume, i.e. that the producer knows more than what is assumed to be included in the information set Q, or 2) by deriving a model of producer assessment based on the assumption that the producers have only the information included in Q. In the first case it is usually assumed that the distribution of preferences of the population W is known. This was the case in the solution presented in the previous chapter. Of course if the producers know both how many of the members of W that rank x^i for all $i=1,...,I-1$ as their greatest element, and they know the obtainable sets of the population W, the assessment of the choices of the consumers is simple.

Nevertheless, we recognize that it is not reasonable to assume that the producers know the distribution of the preferences of the members of W. This assumption should therefore not be introduced. How then is it reasonable to model the behaviour of the producers, i.e. how will the producers react if faced with this limited information set Q?

In principle the problem facing the producers is to allocate a finite number of subjects over a finite number of states in accordance with a given set of principles. This is a fundamental problem which arises in several different contexts. In general it is a technical problem of logic and thus we may turn to how this problem has been solved in other contexts. If the logical structure of a problem in another context is the same as the logical structure of the problem facing the members of population B, then it can be argued that in principle the problems are identical. They differ only in the interpretation of the elements included in the problems.

One such problem can be identified in one of the branches of physics, i.e. thermodynamics.

5.2 The distribution of physical particles: Carnot's problem.

Carnot (1824) was not confronted by a population of consumers which should be distributed over a set of elements. He studied the efficiency of steam engines and observed that heat, for no obvious reason, moved from a warm substance to an initially colder one. In a nutshell this is what thermodynamics[12] is about: The movement of particles within a physical system. A physical system[13] in general consists of elements. A given number of water molecules in a closed container constitute a relevant example. The elements of any system are characterized as particles[14]. Any physical system will always be in a particular state at any given point in time. The state of the system is a function of the state of the individual particles. Let us presume that the structure of the problem of the producers of population B is

[12] Thermodynamics: The study of the interrelation between heat, work and internal energy.

[13] Physical system: Any identifiable collection of physical substance which can be distinguished from anything else through a defined surface in such a way that a change in everything else do not change the system and its contents.

[14] A particle: Molecule,atom, electron or foton.

similar to that facing Carnot in analysing the thermodynamic conditions of a physical system. The thermodynamical system correspond to the population W and the particles to the individual members of W. The set of possible particle states correspond to the set of individual choices X.

Return to Carnot's (1824) original problem of heat diffusion in a physical system. Let a property of the system be denoted entropy. Classical thermodynamic entropy is defined as a measure of the unavailable energy of a thermodynamic system in a given state. In Carnot's problem initially the physical substance first was warm only in one part, and gradually the heat diffused throughout the entire system or substance. In the beginning the available energy of this system in the beginning was the energy represented by the warmth of the part that was heated. Then the system had low entropy. At a later stage when the whole system had reached an even temperature, when the state of the system had changed, some of the previous available energy had been transformed into unavailable energy, and therefore the entropy was higher. It should be noted that the first law of thermodynamics says that the energy of a system is constant. Thus a system's energy is only transformed from available, represented by low entropy, to unavailable form, represented by high entropy.

The second law of thermodynamics says that the entropy of any system will move towards its maximum, i.e. that the thermodynamical equilibrium is reached when there is no available energy left in a system. Within the frame of Carnot's problem this implies that there exists an entropy value for any state the system or substance is in. Carnot described his system as consisting of, say, four particles x, y, z and w, of which three particles, say, x, y and z were particles of high velocity, i.e.

particles generating heat, and one particle was a particle of low velocity. Let us also specify two possible particle states, which in a sense spatial states, where particles are either in one part of the system denoted *i* or in the other part, denoted *j*. A system microstate A1 may now exist for which particles *x*, *y* and *z* are in particle state *i*, and particle *w* in state *j*. In this microstate one part of the substance or system would be perceived as hot since all the high velocity particles are ordered in one part of the system. But now the second law of thermodynamics says that the particles to a greater and greater extent will be distributed among the different states. Assume that driven by the second law the system changes from microstate A1 to microstate B1 where particles *x* and *y* remain in state *i*, when particles *z* and *w* are in state *j* . Now a heat diffusion has taken place and some of the available energy as been transformed to unavailable energy; the entropy value of state B1 is higher then the value of state A1. This phenomenon can be viewed as a transformation from a higher order to a lower order and eventually into disorder or chaos. In state A1 the system is in comparative order where all the high velocity particles are in the same state, whereas in state B1 the high and low velocity particles are mixed together in the same state. Thus entropy can be viewed as a measure of order.

That view is taken by a branch of thermodynamics: Statistical mechanics. In statistical mechanics the entropy of a system as a measure of order, has to satisfy two conditions:

1. The disorder of a microstate of the system can be measured ordinally in relation to the disorder that exists in the corresponding macrostate.

2. The disorder of a macrostate is proportional with the number of

corresponding microstates.

In his fundamental work on statistical mechanics Boltzman (1871) states that two arrangements of particles are in two different microstates, if and only if the particles are rearranged from one state to another. In comparison with microstate A1 above, a microstate A2 where *y*, *z* and *w* are in state *i* and *x* is in *j* , will be a different microstate, but it represents the same macrostate since there are three particles in state *i* and one particle in state *j*. If we restrict the space of individual states or events X to contain only two possible states, i.e. *i* and *j*, then there exists four possible microstates, A1, A2, A3, and A4, for a given macrostate ALPHA.

Due to the principles, marked by 1 and 2, above the ordinal measure of the disorder in the different microstates and the corresponding macrostate is the number 4. In general this can be measured in the following way: If there are *I-1* different states and N particles then the measure of disorder of a given macrostate is given by:

[E20] $T = \frac{N!}{z_1!z_2!...z_{I-1}!}$,

where $N = \sum_i z_i$, and z_i denotes the frequency or the number of particles in state *i* for all i=1,...,I-1.

In our particular example the disorder was 4. The highest level of disorder in this particular system is 6 , i.e. a situation where there is two particles in each state, e.g. macrostate B above. The lowest level of disorder for this system is when all the particles are in the same state. Then T=1.

It follows from principles 1 and 2 above that the measure of disorder has to be derived from the value *T*. Boltzman thus stated entropy as

[E21] $S = k \cdot \ln T$,

where k is Boltzman's constant[15]. He then stated:

S28 Theorem.

$S = k \cdot N \cdot H$,

where $H = -\sum_i p_i \ln p_i$, $p_i = z_i/N$.

Proof: Boltzman (1871). Let $z_i \to \infty$, for all $i=1,\ldots,I-1$. Then [E20] $\Rightarrow \ln T = N \cdot \ln N - N - \sum_i z_i \cdot \ln z_i + \sum_i z_i \Leftrightarrow -\sum_i z_i \ln (z_i/N)$. Thus $N \cdot H$ is equal to $\ln T$ and $S = k \cdot N \cdot H$. QED.

Theorem S28 says that if T or ln T is a measure of disorder, then the expression

[E22] $S = -\sum_i p_i \ln p_i$

is also a measure of disorder. If the event space for the population of particles is X and the σ-algebra is Γ, then the space (X,Γ,p) is a probability space for a physical

[15] Boltzman's constant $k = 1.38 \cdot 10^{-16}$ is a fundamental physical constant.

thermodynamic system where the number of particles is N and the number of possible states is *I-1*, since $p(X) = \sum_i p_i = N/N = 1$. Thus $p_i \in (X,\Gamma,p)$ is the probability that a particle will be in state *i*; and the probability distribution *p* describes the macrostate of the system since every p_i gives the proportion of particles that are in particle state *i*. It follows from the second law of thermodynamics that the probability distribution p* that maximizes [E22] given the constraints implied by the physical system under consideration will constitute the equilibrium distribution of the particles of this system:

$$\max_p - \sum_i p_i \ln p_i$$

[E23]

subject to the boundaries and the conditions of the system.

We recognize that Carnot's and Boltzman's problems are similar to that of the producers in population B. The problem of Carnot and Boltzman is to distribute a given population of particles into different states. Applying the first and second laws of thermodynamics this is done by expression [E23]. Thus there exists a solution to the problem if additional assumptions, which are reasonable are made, such as the first and second law of thermodynamics. The question which now raises is if there exists a solution to our corresponding problem of economics, like the one we found for the physical thermodynamic problem. It is emphasised that we are not looking for an analogous solution, but we recognize that the solution of the physical problem is a solution derived from a problem that is similar to our economic problem and that the solution is derived as a consequence of a description of the behaviour of the particles. Thus we have:

1. The physical problem of heat diffusion: A population of N particles, which are to be distributed among the *I-1* possible states of the state space X, in accordance with specific laws of behaviour, i.e. the first and second laws of thermodynamics, and given the conditions of the system, i.e. the boundaries and other facts that are known about the system and that defines the population contained in the system. Based on this information a solution to the distribution problem is derived.

2. The economic problem of assessing the choice of a population of consumers: A population of N consumers, which are to be distributed among the *I-1* possible alternatives of the choice space X, in accordance with specific laws of behaviour, i.e. the axioms of choice, and given the conditions of the problem or the system, i.e. the given price vector $c \in C^{\wedge}$. Based on this information, is it possible to derive a solution without stating any assumptions that is not credible?

The answer to the latter question is affirmative, and the solution to our specific problem is in fact quite similar to the solution obtained for the physical problem. This is not due to an analogous application of the solution of the physical problem to the economic problem. We merely have a coincidence in methodology. But the physical problem and its solution should be recognized as a motivator in finding the solution to our problem. The reason why the solutions are similar is the similarity of the logical problems. Thus one may be tempted to say that both problems are the same general problem where the elements of the problem have different interpretations, i.e. as particles, states, physical laws, and boundaries and conditions of physical systems; and as consumers, consumer points, laws of behaviour, and market information. Strictly speaking, however, this is not the case, and the respective solutions are deduced independently of the other problem.

We will in section 5.5 and chapter 6 return to how our economic distribution problem is solved. We recall that our problem is how a member of B will assess the probability distribution of a member of W over L(c) based on the producers knowledge about the individual behaviour of the members of W. Before we solve this problem, we will present another but closely related problem, i.e. the problem of how to distribute the population of W not assuming a particular individual behaviour, but a particular population behaviour. As will be seen the solutions to these different problems are quite similar, and as such it is relevant to present some of the ideas incorporated in this problem of population behaviour, but the crucial difference remains: Our problem is a solution to a particular problem confronting an economic agent of our model where the solution is based on assumptions of *individual* behaviour, when the problem of distribution based on assumptions of *population* behaviour has to be viewed more as a problem of interest in situations where information about individual behaviour is not relevant or available, e.g. in empirical analyses.

But due to the similarities and the relationship between the problems we find it appropriate to present the problem in connection with our context.

5.3 The efficiency principle.

Erlander & Smith (1991) reveal that the efficiency principle was first introduced by Zipf (1949) and later formally expressed by Smith (1978), and has been developed further especially by Erlander (1985) and Erlander & Smith (1990).

Assume the existence of a population of individuals. This population can be in several different states. The state of each individual is described as the individual state, e.g. a particle state or the choice of an individual member of W. The sum of all individual states gives the state of the population as a whole, and are described as the population state. The population may be in a given number of states. Each possible state represents a particular level of activity, e.g. that all the individuals have chosen one alternative. Each population state also represents a particular cost level, e.g. the total population expenditure. Consider two population states A and B. Assume that the activity level of both A and B is the same. Then assume that the cost level of A is lower than the cost level of B. The principle of cost efficiency can be defined as:

> The system or the population is cost efficient if and only if the probability p(A) that state A occurs is at least as great as the probability p(B) that state B occurs.

The most cost efficient population *state* is defined as the state for which the cost level for the population as a whole is the lowest. Cost efficient population *behaviour* is defined so that the probability that the population is in a cost efficient state is at least as great as the probability that the population is not.

The distribution problem at hand can in general be described as follows: The population of consumers W can be in different individual states $x^i \in X$, $i=1,...,I-1$, depending on the members' choices, where X is the individual state space. On the individual state space X a probability space (X,Γ,p) is defined. Notably in the context of consumer choice among discrete alternatives, the subsets of Γ will be elements of the sets L(c), which are the choice sets that contain as elements those choices that are obtainable to some consumers. Those elements of X that are not elements of any L(c), can not be elements of the collection of measurable subsets of (X, Γ), since those elements that can not be chosen can not be measured either. In the general context the elements of the measurable space (X,Γ) can be perceived as the set of possible individual states that at least one individual member of the population can be in, and that is in fact also the definition of any L(c), given a certain cost level $c \in C^\wedge$. Now *p* is defined as a measure on the measurable space $(X,\Gamma)=T$. Now there exists a set $P \subseteq (X,\Gamma,p) \in \Re^{I-1}$ of possible probability distributions. Since the measurable space (X,Γ) is identical to the choice set $L(c) \in \Re^{I-1}$ in the event of a particular price vector $c \in C^\wedge$, the probability distributions are elements of $\Re^{I-1}$.

As a consequence of the distribution of individual members of W over the state space a population microstate will occur where the individiual state of each member is indicated. The set of possible microstates is denoted Ω, and a specific microstate is denoted ω. If the population has N members, then $\omega=[\omega^1,...,\omega^n,...,\omega^N]$ may represent the microstate of the population where ω^n indicates the individual state of the n-th member of W. The microstate being the result of a given price vector $c \in C^\wedge$ the microstate can be denoted $\omega(c)=[\omega^n(c)]$ where $\omega^n(c)=D_n(c)$, i.e. the greatest element among L(c) for the n-th member of W.

Let every population microstate imply a macrostate or frequency denoted $z(\omega)=[z_1(\omega),...,z_i(\omega),...,z_{I-1}(\omega)]$, where $z_i(\omega)$ is the number of individuals in state *i*, or consumers of population W who have chosen element $x^i=D(c) \in L(c)$ as their greatest element L(c). The macrostate space implied by Ω is denoted $Z(\Omega)$.

A resulting microstate represents a definite solution to the allocation problem at hand. Thus the question is which microstate will occur given a specific price vector $c \in C^\wedge$. The answer to this question can not be found a priori. However, one can ask how probable it is that one will observe a specific microstate. A population state $\omega \in \Omega$, where $\Omega \subset X$ is the set of population states, is not an element of (X,Γ), but a subset of elements of (X,Γ), i.e. for a given $c \in C^\wedge$ ω is a subset of elements of L(c). Thus a population state can not be assigned a probability direct, but the probability of observing a particular population state can be derived form the probabilities of observing jointly the different events of individual states which make up the population state, such that:

$$[E24]\ p(\omega) = \prod_{i=1}^{I-1} p_i^{z_i(\omega)}, \text{ for all } p \in P,$$

where $p_i = p[x^i=D(c)]$ denotes the probablity that an individual will be in state i given the use of probability distribution p and given the price vector $c \in C^\wedge$, and assuming independence between p_i, p_j for all $i,j=1,\ldots,I-1$, $i \neq j$.

Now the more precise definition of the efficiency principle given by Erlander & Smith (1990) states:

S29 Definition.

For any given activity matrix[16] A and cost matrix *c* the probability distribution $p \in \Re_+^{I-1}$ is cost efficient with respect to A if and only if for all comparable population states[17] $\omega, \omega' \in \Omega \subset X$ it is true that

$$A(\omega) = A(\omega'),\ c(\omega) \leq c(\omega') \Rightarrow p(\omega) \geq p(\omega'),$$

where $A(\omega)$ is the resulting activity level, and $c(\omega)$ is the resulting cost or expenditure level for a given microstate $\omega \in \Omega$.

The population states $\omega, \omega' \in \Omega \subset X$ are subsets of X. Thus the principle states that for given prices or costs, and for a given activity level, it would be more likely to find a distribution of individual states that implies lower total costs for the population as a whole, than a distribution that implies higher total costs. This definition seems reasonable. The definition is based on the general assumption that any human population will have a tendency towards displaying economically efficient behaviour, i.e. using as few resources as possible.

It follows from [E24] that there are several possible probability distributions of which only one, denoted p*, can be the correct one based on the available information, i.e. a given price vector. What the efficiency principle therefore says is

[16] The activity matrix gives the activity values/attributes for the problem at hand. If the problem is a trip dsitribution problem, the activity martix gives how each trip is to be valued due to whatever activity is assumed to exist in the problem.

[17] Comparable population states: Two population states ω, $\omega' \in \Omega$ are comparable if they are of the same size, i.e. the number of individuals in each of the states are the same, when the population states are generated by exactly the same populations.

that if the population behaviour is to be efficient, then for given activity and cost matrices the correct probability distribution $p^*\subset(X,\Gamma,p)$ describing this behaviour, of all possible probability distributions $P\subset(X,\Gamma,p)$, should be such that for

$$[E25]\ Az(\omega) = Az(\omega'),\ cz(\omega) \leq cz(\omega') \Rightarrow \prod_{i=1}^{I-1} p_i^{*\,z_i(\omega)} \geq \prod_{i=1}^{I-1} p_i^{*\,z_i(\omega')}.$$

In other words: The only possible true probability distributions, if the population behaviour is cost efficient, are the distributions that satisfy [E25].

We note that the distributions of individual states $z(\omega)\in Z(\Omega)$, where $Z(\Omega)$ is the set of distributions that are implied by Ω, are not actually known distributions but possible distributions of individual states. If it was known which $z(\omega)$ that was true it would be meaningless to assess probabilities.

5.4 The representation theorem for the efficiency principle.

In the previous section we laid out which behavioural conditions a population of individuals has to satisfy in order to be labelled cost efficient. Yet we have not provided any precise description of which probability distributions $p\in P$ satisfy definition S29. We will set forth in this section which class of distributions satisfy the principle of efficient population behaviour.

The general problem can be outlined in greater detail, and it should be observed that it is quite similar in structure to the physical problem of heat diffusion.

In general each individual of a population of individuals is to be allocated into one of a given number of possible states $X \in \Re^I$, given specific activity matrix $A \in \Re^{E \times (I-1)}$ and cost matrix $c \in \Re^{F \times (I-1)}$. An activity matrix A consists in general of E attributes each having *I-1* values, one value for each alternative i=1,...,I-1. A cost matrix *c* consists in general of *F* different cost attributes each having *I-1* values. In our specific problem, analysed in this treatise, the activity matrix is in fact a vector equal to the unity vector. The cost matrix is in fact a vector of the prices of the members of B.

A solution[18] to such a general allocation problem, that is for a given pair of A and *c*, denoted (A,c), is given by a specific population state or microstate ω^0, which implies a population frequency or macrostate $z(\omega^0)$. The solution ω^0 is not known. But, if it is assumed that the same population has a cost efficient behaviour, then the probability of observing different population states $\omega \in \Omega$ can be assessed for a given system (A,c). That is, an assessment of the probabilities that the different elements of Ω are the solutions to the specific problem at hand, i.e. the probabilities $p(\omega)$, for all $\omega \in \Omega$ are in fact the probabilities that the different elements of Ω are solutions to the problem (A,c). Therefore, the problem is actually solved by finding the probabilities $p(\omega)$[19]. The population state ω^0, for which the corresponding probability $p(\omega^0)$ is the greatest of all probabilities is then the most probable true population state; i.e. the most probable solution to the problem (A,c).

The correct probability $p(\omega)$ that a given population state ω is a solution to the problem (A,c) can not be derived directly from the measurable space (X,Γ) defined on X. Since X is a space of individual states, the elements of the measurable space (X,Γ)

---

[18] A solution in this context can be compared to the equilibrium distribution of the particles in the physical problem of heat diffusion, see the discussion above.

[19] Also the physical problem of heat diffusion was solved by finding a particualr probability distribution that told the probability that a particle would be in a particular state.

are the measurable individual states. If the population state ω can be derived from these individual states, and it can be, then a measure on the measurable space (X,Γ) indirectly measure the population state.

Since $p \in P \subset (X,\Gamma,p)$ is only one of H different possible probability distributions there exists H different probability measures of the probabilities $\bar{p}(\omega)$ for each $\omega \in \Omega$. Since only one probability distribution of all those which are elements of P can be said to be the true probability distribution p* over the measurable space given the specific problem (A,c), the question is how to arrive at this distribution among all the P distributions offered for this specific problem.

The answer is: It should be recognized that the probability distributions P represent expressions of the behaviour of the population since the probability that an individual member of the population will be in a particular state says something about the assumed distribution of states among the individuals of the population and thus something about the population. We have assumed that the behaviour of the population is cost efficient. This implies that the true probability distribution p* expressing the true behaviour of the population has to satisfy definition S29, i.e. equation [E25]. The set of probability distributions that do not satisfy this definition can be excluded as possible distributions. The next question is if there exists a unique probability distribution satisfying this principle. It will be shown below that such a unique $p^* \in P$ satisfying the principle does exist.

We will now show the existence of a unique probability distribution p* given a specific problem (A,c).

S30 Definition.

For any matrices $A \in \Re^{E\times(I-1)}$ and $c \in \Re^{F\times(I-1)}$ the pair (A,c) is said to be c-identifiable if and only if for each column of the transpose matrix $c^T=[c_1,...,c_f,...,c_F]$, $c_f \in \Re^{I-1}$, $c_f \notin sp([\sigma, A^T, c_1,...,c_{f-1}, c_{f+1},...,c_F])$ [20].

Definition S30 says that the pair (A,c) is defined as c-identifiable if each column of c^T is not an element in the space of linear combinations[21] of the other columns in the matrix, the unity vector $\sigma^T = [1,...,1]$, and the matrix A^T. In other words: Each column of the cost matrix shall be linearly independent[22] of all the other vectors in the system (A,c), which implies that there exists only one solution, i.e. there exists a unique set of cost coefficients, $\delta = [\delta_f] \in \Re^F$, defining the solution of the system. In the context of assessing consumer choice, the activity and cost matrices are vectors in the Euclidian space $\Re^{1\times(I-1)}$. Thus there exists a unique cost coefficient, $\delta \in \Re$, which yields the solution to the system, or the frequency solution and the macrostate $z(\omega^0)$ implied by the true solution or microstate ω^0.

But there may exist several different sets of activity coefficients, β, which in combination with the unique set of cost coefficients define the solution.

---

[20] Sp = Span: The linear span of a set in a vector space is the smallest linear subspace containing the set.

[21] Linear combination: The sum of the respective products of the elements of some set with constant coefficients. If β_1 and β_2 are constant coefficients to the elements a_1 and a_2 of the set A, then $\beta_1 a_1 + \beta_2 a_2$ is an linear combination.

[22] Linearly independent: There is no linear combination of given elements that equals zero, given that not all the coefficients are equal to zero. a_1 and a_2 of the set A are linearly independent vectors if for the scalars β_1 and β_2 where at least one of the scalars are different from zero, $\beta_1 a_1 + \beta_2 a_2 \neq 0$.

S31 Definition.

If the matrix $[\sigma, A^T, c^T] \in \Re^{(E+F+1)\times(I-1)}$ is of full column rank[23], then (A,c) is said to be fully identifiable[24].

Definition S31 says that if all column vectors in the matrix $[\sigma, A^T, c^T]$ are linearly independent[22], then there exists a unique solution, i.e. there exists a unique set of activity and cost coefficients that give the macrostate solution $z(\omega^0)$.

The following result is given by Erlander & Smith (1990) concerning the cost efficient probability distribution $p^* \in P$ for the system (A,c) given the macrostate solution $z(\omega^0)$:

S32 Theorem.

1. If (A,c) is c-identifiable for a rational matrix A and *c* and any positive probability distribution $p \in P$, then p^* is c-efficient with respect to A if and only if there exists a unique non-negative vector $\delta \in \Re_+^F$ such that for a vector $\beta \in \Re^E$ it is true that

$$p^* = \exp(\alpha \cdot \sigma + A^T\beta - c^T\delta),$$

where $\alpha = -\log[\sigma^T \cdot \exp(A^T\beta - c^T\delta)]$.

2. If (A,c) is fully identifiable then β is unique.

[23] Full column rank: All the columns of the matrix are linearly independent.

[24] Fully identifiable means that there exists a unique solution to the problem, i.e. to (A,c).

We are not going to prove the theorem but will give some indications of the proof. The reader is referred to Erlander & Smith (1990) for a complete proof.

First we have to decide if $p^* = \exp(\alpha \cdot \sigma + A^T\beta - c^T\delta)$ is cost efficient. Let $z(\omega)$, $z(\omega') \in Z(\Omega)$, denoted z and z' respectively, be two possible solutions which are such that $\sigma^T z = \sigma^T z'$, and $Az = Az'$. Define $\log p = \alpha\sigma + A^T\beta - c^T\delta$. Now $z^T \log p - z'^T \log p = z^T(\alpha\sigma + A^T\beta - c^T\delta) - z'^T(\alpha\sigma + A^T\beta - c^T\delta)$, which is equal to $\alpha(\sigma^T z - \sigma^T z') + \beta^T(Az - Az') + \delta^T(cz' - cz) = \delta^T(cz' - cz)$ since $\alpha(\sigma^T z - \sigma^T z') + \beta^T(Az - Az') = 0$ as a consequence of $\sigma^T z = \sigma^T z'$, and $Az = Az'$. We now observe that if $cz \le cz'$, then $z^T \log p - z'^T \log p \ge 0$ since δ is non-negative and therefore $z^T \log p - z'^T \log p = \delta^T(cz' - cz) \ge 0$. Thus $cz \le cz'$, $Az = Az' \Rightarrow z^T \log p \ge z'^T \log p \Leftrightarrow p(\omega) \ge p(\omega')$, and consequently the form $\log p = \alpha\sigma + A^T\beta - c^T\delta \Leftrightarrow p = \exp(\alpha\sigma + A^T\beta - c^T\delta)$ is cost efficient.

The next question is if every cost efficient probability distribution $p^* \in P$ has to be of the form $p = \exp(\alpha\sigma + A^T\beta - c^T\delta)$. Let $z(\omega^0) = z$ be a solution, i.e. the frequency profile of the distribution of states given the population state $\omega^0 \in \Omega$. It is known that the population probability, i.e. the probability of observing this particular population state ω^0, is denoted $p(\omega^0)$ and derived from the true probability distribution p*. It should be recognized from [E24] that the true probability distribution $p^* \in P$ can be derived from $p^*(\omega^0)$ if *z* is known. Thus we only have to find the cost efficient $p^*(\omega^0)$.

It follows from [E24] that $\log p^*(\omega^0) = z^T \log p^*$. If $p^*(\omega^0)$ is to be cost efficient with respect to A for the system (A,c), then $\sigma^T z = 1$, $Az = a$, and $cz \le \eta$, where *a* is the activity level, and η is the lowest level of costs of all other states with an activity level equal to a.The problem of finding the cost efficient $p^*(\omega^0)$ can therefore be formulated as a linear programming problem:

$$\text{[E26]} \quad \begin{aligned} &\min_{z} \log p^*(\omega^0) = z^T \log p^* \\ &\text{subject to } \sigma^T z = 1,\ Az = a,\ cz \leq \eta. \end{aligned}$$

The solution to problem [E26] is a vector $z = z(\omega^0) = [z_1,\ldots,z_i,\ldots,z_{I-1}]$ which is the true frequency distribution if the information in A, c, a, and η is the information which the members of the population are facing. The right hand side values a and c, are in this problem assumed to be known. log p* is only the coefficient of the objective function.

Erlander & Smith (1991) states that if the problem has a positive solution $z \in \Re^{I-1}_{++}$, then a well known duality property of linear programming [Dantzig (1963)] says that then there exists a scalar $\alpha \in \Re$ and a vector $\beta \in \Re^E$ such that $\log p^* \in \Re^{I-1}_{+}$, for a non-negative vector $\delta \in \Re^F_{+}$, can be expressed:

$$\text{[E27]} \log p^* = \alpha \cdot \sigma + A^T\beta - c^T\delta \Leftrightarrow p^* = \exp(\alpha \cdot \sigma + A^T\beta - c^T\delta).$$

The above mentioned duality property follows from the optimality conditions of Kuhn & Tucker (1951). Let $z = [z_1,\ldots,z_{I-1}]$, $c_i \in \Re^F$ be column i in c, $a_i \in \Re^E$ be column i in A, L be the Lagrangian function of problem [E26] where $\lambda_A \in \Re^E$, $\lambda_C \in \Re^F$ and $\lambda_\sigma \in \Re$. Now the first 2I-2 optimality conditions of problem [E26] are:

$$\text{[E28]} \ \frac{\partial L}{\partial z_i} = \log p_i^* - a_i^T\lambda_A - c_i^T\lambda_C - \lambda_\sigma \leq 0, \text{ for all } i=1,\ldots,I-1,$$

[E29] $\frac{\partial L}{\partial z_i} z_i = 0$, for all $i=1,\ldots,I-1$.

If $z_i > 0$ for all $i=1,\ldots,I-1$, then [E28] will have to be equal to zero if both [E28] and [E29] are to be satisfied. Then [E28] will be identical to [E27] where $\lambda_\sigma = \alpha$, $\lambda_A = \beta$, and $\lambda_C = \delta$. In reference to theorem S32 we have thus indicated that every cost efficient probability distribution p* have to be on the form [E27].

To complete the discussion it should be noted that it has to be decided if the system (A,c) is c-identifiable and fully identifiable if a unique solution is to exist.

We observe that in principle the problem of [E26] is similar to problem [E23]. Chiefly the differences are due to the different conditions under which the problems are solved. In [E26] the conditions are the information represented by the cost and activity matrices. In [E23] the conditions are the properties of the system. A reasonable question now is: Why is the solutions of the thermodynamic problem and the problem anlysing the distribution of individuals based on assumptions of population behaviour similar? The answer to this question will be provided in the next section. This answer is also an element in deriving the solution to our distribution problem based on assumptions of individual economic behaviour.

5.5 The answer: Mathematical representation of information and uncertainty.

The central question confronting a member of the producer population B is how to assess the likelihood that a member of the consumer population W will choose his particular offer. Every member of W know exactly which alternative he or she will choose among the *I-1* alternatives available from the members of B given the prices offered. The problem of the member of B is thus a problem of lack of information.

Specifically, the problem facing the i-th member of B is that x_i is capable of assuming two discrete values (1) or (0), i.e. $x = (x_1,...,x_i,...,x_{I-1})$. The corresponding probabilities $p(x_i=1) = p_i$ and $p(x_i=0) = \sum_{j \neq i} p_j$ are not given. The relevant probability concept is the concept of subjective probability since it is the belief of the producer *i* which is of interest when this producer's actions are analysed. Every member of B have thus to decide upon the values of p_i, i=1,...,I-1, as a reflection of their state of knowledge. The test of a good subjective probability distribution is therefore: Does it correctly represent the producer's state of knowledge as to the value of x_i?

Just as in applied statistics the crux of a problem is often how to devise some method of sampling that avoid bias, the problem confronting a member of B is that of finding a probability assignment which avoids bias, while agreeing with whatever information is given. The question we ask is: Is it possible to find any quantity $H(p \in P)$ which measures in a unique way the amount of uncertainty represented by this probability distribution? The great advance provided by Shannon (1948) lies in the discovery that there is a unique, unambiguous criterion for the "amount of uncertainty" represented by a discrete probability distribution, which agrees with our intuitive notions that a broad distribution represents more uncertainty than does a sharply peaked one, and satisfies all other conditions which makes it reasonable.

This advance is reached by the successful specification of conditions for such a measure which ensure both uniqueness and consistency, to say nothing of usefulness. Accordingly it is a very remarkable fact that the most elementary conditions of consistency, amounting really to only one composition law, already determines the function $H(p \in P)$. The three conditions are:

(1) H is a continuous function of *p*.

(2) If all p_i are equal, the quantity $\xi(N) = H(1/N,...,1/N)$ is a monotone increasing function of N, where $N = \sum_{i=1} n_i$, n_i is the number of members of W who choose alternative *i*.

(3) If our information or uncertainty measure is to be consistent, we must obtain the same ultimate uncertainty no matter how the choices are broken down. This last condition refers, to some extent, to the problems that arise when Laplace's principle is applied.

S33 Theorem.

The only quantity which is positive, which increases with increasing uncertainty, and is additive for independent sources of uncertainty, is

$$H(p \in P) = - K \sum_{i} p_i \ln p_i,$$

where K is a positive constant.

<u>Proof:</u> Shannon (1948). From condition (1), it is sufficient to determine H for all rational values $p_i = n_i / \sum_{j=1} n_j$, with n_i, n_j integers, $i,j=1,...,I-1$. But then condition (3) implies that H is determined already from the symmetrical quantities $\xi(N)$. We can regard a choice of one of the alternatives $(x_1,...,x_{I-1})$

as a first step in the choice of one of N equally likely alternatives, the second step of which is also a choice between n_i equally likely alternatives. This could be written

$$H(p\in P) + \sum_i p_i \xi(n_i) = \xi(\sum_i n_i).$$

In particular, we could choose all n_i equal to *m*, whereupon the above expression reduces to

$$\xi(m) + \xi(N) = \xi(mN).$$

Evidently this equation is solved by setting $\xi(N) = K \ln N$, where by condition (2), $K>0$. Substituting $\xi(N) = K \ln N$ into $H(p\in P) + \sum_i p_i \xi(n_i)$

$= \xi(\sum_i n_i)$, we obtain the desired result, $H(p\in P) = k \ln(\sum_i n_i) - k \sum_i p_i \ln n_i$

$= - k \sum_i p_i \ln p_i$. QED.

It is evident that a maximization of $H(p\in P)$ will yield a probability distribution that is unbiased. To use any other distribution would amount to arbitrary assumption of information which the producer does not have.

This principle of maximizing $H(p\in P)$ may be regarded as an extension of Laplace's principle of insufficient reason, to which it reduces in case no information is given except enumeration of the possibilities x_i, $i=1,...,I-1$, with the following essential difference, according to Jaynes (1957): The maximum $H(p \in P)$ distribution may be asserted for the positive reason that it is uniquely determined as the one which is maximally noncommittal with regard to missing information, i.e. uncertainty, instead of the negative one that there was no reason to think otherwise, which is the idea behind the principle of insufficient reason which says that if we do not know anything then every event is equally likely. Thus the concept forwarded above supplies the missing criterion of choice which Laplace needed to remove the apparent arbitrariness of the principle of insufficient reason, and in addition it shows precisely how this principle is to be modified in case there are reasons for "thinking otherwise".

This resulting principle of maximum $H(p \in P)$ which as it is shown is a unique representation of uncertainty or lack of information, is mathematically equivalent to the result that we get when analysing thermodynamic energy diffusion or interestingly, when we are seeking a solution to our problem of the producers' assessment of the probability distribution of the consumers over the *I-1* alternatives of our market. And this is the answer: The problem we seek a solution to in our case of distributing the consumers of W over the alternatives of B is: What is the reasonable probability distribution of the consumers of W assuming they behave according to the listed axioms of choice? The question we ask in the case of thermodynamic diffusion is: What is the reasonable probability distribution of the particles of the system assuming the validity of the 1st and 2nd laws of thermodynamics? Conceptually these questions are logically the same, where the axioms of choice and the laws of thermodynamics, respectively, are the information available in the different cases, and since we in this section have shown that the only unbiased and consistent measure is

$H(p \in P)$ this measure should also appear when we consistently elaborate on these questions.

It should be noted that the result of Shannon (1948) and the result derived when assuming the principle of efficient population behaviour are not connected otherwise than mathematically. As pointed out above the problem solved by Shannon was the technical problem of finding a mathematical expression for information that was unbiased with regard to uncertainty. As such Shannon's measure is a constructed one. The measure derived assuming the principle of efficient behaviour is a measure that complies with the information available. The thermodynamic case is used only to demonstrate the problem of efficient population behaviour and is not in principle an economic problem. It is basically a problem of assessment of behaviour given a particular law of behaviour. The specific result is dependent upon that particular law, but the derivation and the logical structure of the problems is similar.

This gives the starting point of our problem: The assessment of consumer behaviour by the producers. Shannon's measure H(p) is the only unbiased measure of the uncertainty that a probability distribution constitutes, i.e. of the general degree of information comprised in the distribution. Since the producers are to decide upon which of the P possible distributions they will use in calculating the behaviour of the consumers, the measure H(p) is the only consistent measure they can use in evaluating the different distributions. However, the only acceptable distributions which should be submitted to evaluation by H(p) are those which fits the available information. Hence the probability distributions which are eligible for choice to the producers, are those which contain the greatest uncertainty, i.e. H(p), and at the same time accounts for the information in the possession of the producers. These distributions are the only distributions that only take in to account the available information and nothing else. The question is: How to derive the possible distributions that is consistent with these limitations?

As will be shown below, the relevant distributions are derived by formulating and solving a mathematical programming problem where H(p), constituting the objective function, is maximized subjected to a number of constraints where the constraints are formulated with regard to the available information. The solution to this problem will be a unique probability distribution which is in accordance with the available information, and which is unbiased with regard to the information. Hence this distribution will be used by all producers to assess the actions of the consumers.

Chapter 6. Producer behaviour.

6.1 Introduction.

In this chapter we shall discuss the behaviour of the producers, i.e. population B, of the differentiated products set forth in axiom S3. Our basic assumption concerning producer behaviour is that every member of B seeks to maximize his profit. This is done by determining their controlling variable, the price, considering the information they possess.

The question of producer information is the crucial subject of our analysis. The difference between our analysis and the analyses of those using the logit model, is that when using the logit model it is implicitly assumed that the producers have information about the distribution of the preferences of the population W. Our basic point is that it is not reasonable to assume that the members of B have such information. This is unrealistic, and only an assumption of convenience. Hence the fundamental objective of this treatise is to present a model where no such assumptions are stated. That is, we shall put forward an analysis based only on reasonable and accepted assumptions of consumer and producer behaviour.

Thus the problem facing the members of B is their lack of knowledge of anything concerning the members of W except the knowledge of how they make their choices, i.e. the axioms stated so far. Thus the members of B know that for a given price vector $c \in C^{\wedge}$, the choice set of every member of W will be given by $L(c) \in X$, containing exactly $I\text{-}1$ elements, where every element is associated with a member of population B. Furthermore, the members of B know that every member of W will choose that element that has the highest utility to that particular member, and every member of B know how these preferences are formed. But what is unknown to all members of population B is which particular element in L(c) that is preferred by each and every member of population W. That is not all. No member of B knows how the members of population W are distributed as a population over the elements of L(c).

Thus every member of B is confronted with an uncertainty concerning the demand he is facing for a given $c \in C^\wedge$. Since the demand is uncertain the income and the profit will be uncertain too. The question is then: How do a member of population B maximize uncertain profit? The answer to this is our assumption of risk neutrality, i.e. that the member of B is indifferent between receiving the expected profit and the profit itself.

We have thus introduced the notion of expectation, which is based upon the notion of probability. If therefore every member of B is to maximize the expected profit, every member of population B also has to decide upon a probability, a likelihood, that a member of population W will choose the element of L(c) associated with this particular member of B. That is, every member of B has to make an assessment of his subjective distribution $p[L(c) \mid Q] \in \Re^{I-1}$ of population W over L(c) for a given $c \in C^\wedge$, where Q is the information set. This assessment of $p[L(c) \mid Q] \in \Re^{I-1}$ is based on all the information that the members of population B possess.

Then, we shall show that the producers' assessment of expected demand can be uniquely represented by a certain class of probability distributions over the choice set. This means that there is only a unique class of distributions that satisfies the axioms of choice and at the same time is unbiased towards the information that the members of B have access to. Since the producers will gain nothing from using a biased distribution with regard to the information in their possession, they will use an unbiased probability distribution. Since, as we will show, there exists a unique class of distributions that are in accordance with the information held by the members of B, and at the same time is unbiased, this distribution will be used by all members of B since they share the same information. Ultimatly we show that the producers' assessment follows uniquely from the axioms stated above, accounting for a few additional assumptions.

The probability distribution which we are referring to here is not necessarily the same distribution that is derived using the assumptions laid down by the logit model, or other models used to represent demand in problems of this type. But they can be the same if these models satisfy certain assumptions.

We now turn to our characterization of how producers assess consumer demand.

We take the view of a producer, i.e. a member of population B. That is, we will try to say something about which element he expects a consumer to choose. Therefore, the question raised by the producers is: Which element $x^i \in L(c)$, for all $i=1,...,I-1$, will a consumer choose? Obviously a consumer will choose $x^i=D(c)$ of all $i=1,...,I-1$, i.e the greatest element in the choice set $M(c)$. Therefore, from the producer's point of view, the question is how likely it is that $x^i=D(c)$, $i=1,...,I-1$, for an arbitrary member of W, given the information set of the producer.

6.2 Measurement of differentiated product choice behaviour.

This is the problem of the producer of how to measure which choice a consumer is to make from the choice set $L(c) \in X$. This problem has up to now been modelled through the use of random utility and often specified by the logit or probit model. These models assume that the choice behaviour of the consumer is stochastic, or at least the behaviour is modelled as it is stochastic. Thus, the properties of these models have an uncertain economic interpretation. In addition these models, when used to model a market, implicitly assume that the members of B know the distribution of the preferences for the population W as a whole. Such an assumption is not reasonable in our context.

We shall, on the other hand, proceed in describing the properties of our primitive concepts, and making no further assumptions about consumer behaviour, but only, some we believe, easily accepted assumptions about producer behaviour, and no fundamental assumption such as the members of B having any knowledge of the distribution of the preferences of population W. Thus, we will show, that how differentiated product choice behaviour is measured, follows from the properties already set forth for the primitive concepts of our theory.

The problem can be described as follows: We assume that a member of population W, the consumer, is in a situation where the income of each member of population W, and all prices of all the commodities in the economy are given. Then the vector of disposable income adjusted prices $c \in C^{\wedge}$ is given. For a member of population W the choice set $L(c) \in X$ is then given. The question confronting every member of population B is which element in L(c) the member of W will choose, i.e. which element in L(c) is in D(c) for an arbitrarily chosen member of W? The answer to this question can not be provided by the members of B, but every member can make an assessment, a judgment, of how likely it is that a particular element of L(c) will be chosen by a member of W.

We recall that every consumer or member of W in fact chooses an element or consumer point from the choice space X. An element $x \in X$ is a vector $x=[\psi^T, x_I]$, where $\psi \in \Psi \in \Re^{I-1}$, and $x_I \in X_I \in \Re$. We recall that as a consequence of axiom S22, where it is assumed that every choice set is a budget plane, for a given price vector $c \in C^\wedge$ the value of x_I is given if the member of W has made a particular choice. Thus the value of x_I is implied by every choice made from Ψ. Let $x_I^i \in X_I$ be the amount of the general commodity implied by the choice of $\psi^i \in \Psi$. Thus $\psi^i \in \Psi \Rightarrow x^i=[\psi^i, x_I^i] \in X$, and for a given price vector $c \in C^\wedge$ supplied by the population B, every consumers' choice set $L(c) \subset X$ contain at most *I-1* elements that are eligible for choice.

From theorem S15 producers know that every element in X is assigned a utility value by each member of W. It follows from theorem S21, since L(c) is a budget plane, that the element of L(c) for which the utility value is the highest will be chosen by the consumer. Seen from the perspective of the members of population B, they only know the criteria of choice for the members of W from the choice set L(c), not which element they will choose. Consequently the choice has to be measured by each member of B.

Since the members of B have to measure the choices made by the members of W, a measurable space of consumer choices has to exists. We assumed in definition S25 that the choices of the consumers are measured by probailities. Hence every producer will assess the corresponding probability for every element in L(c). It follows from definition S25 that the probabilities are set based on the information available to the producers. From axiom S24 it follows that the information set is equal to all members of the industry B, and that the information set Q contains the principles of producer behaviour, and hence how the producers set their prices, the principles of consumer behaviour, and the market information of the previous market period.

Nothing else is known.

Then since $c \in C^\wedge \Rightarrow L(c) \in \Re^{I-1}$ for all members of population W, it follows that $p[L(c) \mid Q] \in \Re^{I-1}$ for a given $c \in C^\wedge$, also is an discrete distribution. The members of B are therefore seeking a correspondence:

$$[E30]\ \forall c \in C^\wedge,\ p: L(c) \rightarrow p[L(c) \mid Q] \in \Re^{I-1},$$

where $p[L(c) \mid Q] \in \Re^{I-1}$ is the set of probabilities corresponding to the choice set L(c) for any member of B.

Our objective is to characterize the function [E30] which will be identical for all members of B. It follows from [E30] that $p_i(c) \in p[L(c) \mid Q] \in \Re^{I-1}$ corresponds to $x^i \in L(c)$ for a given $c \in C^\wedge$, for all $i=1,\ldots,I-1$. $p_i(c) \in p[L(c) \mid Q] \in \Re^{I-1}$ can be interpreted as the probability that a randomly chosen member of W will choose $x^i \in L(c)$. Theorem S21 implies that the choice probability $p_i(c)$ for all $i=1,\ldots,I-1$, is the probability that the utility value of $x^i \in L(c)$ is greater than the utility values of all other alternatives:

$$[E31]\ p_i(c) = p[x^i = D(c)] = p[u(x^i) > u(x^j)\ \forall x^j \in L(c),\ i \neq j],\ \forall x^i \in L(c),\ \forall c \in C^\wedge.$$

The values $p[L(c) \mid Q] \in \Re^{I-1}$ for a given $c \in C^\wedge$ is found when the question of how probable is it to observe $u[x^i] > u[x^j]$ for all $x^j \in L(c)$, $i \neq j$, is answered.

Equation [E31] is the equivalent of equation [E10] of chapter 4. When we in chapter 4 developed equation [E10] into [E13], and eventually into the explicit expression of [E18], using [E11], [E12], [E14], and [E15] in addition to the axioms of chapter 2, we are now in a position to show that it is possible to arrive at an explicit expression of the choice probability distribution with the use of the axioms of chapter 2 with the addition of a few easily accepted assumptions.

Let us enter the mind of a member of population B. The task of every member of B is to assess the probability distribution $p[L(c^*) | Q] \in \Re^{I-1}$ for a given $c^* \in C^\wedge$. The process of assessing this distribution consists of checking a possible distribution against the information set Q, i.e. one checks if a distribution is consistent with the principles of consumer behaviour. This being the case, one proceeds to check if it also conforms with the given c*, i.e. producer behaviour, and eventually the market information. The set of distributions which are consistent with the information are the distributions which are eligible for choice.

First, let us consider what implications the principles of consumer behaviour have for which distributions that can be used by the members of B. To consider these implications, c* is for a moment neutralized by setting $c_i = V$ for all i=1,...,I-1, where V is a constant. This implies that $x_I^1 = \ldots = x_I^i = \ldots = x_I^{I-1}$.

Now the producers observe that the information in the principles of consumer behaviour imply the following for the distribution $p[L(c^*) | Q]$ for every member of B:

1) Axiom S7 implies that $p_j \geq 0$ for all i,j=1,...,I-1, where $p_j \in p$.

2) The assumption of transitivity in S7 and convexity in S13 imply by themselves nothing concerning the choices of the members of W.

3) Axiom S12 implies that it is more likely to observe the choice of an element

which contains a greater amount of one commodity then all other elements in L(c), and containing no less quantities of other commodities.

Thus axiom S12 alone has any valid implications for the shape of the distribution $p[L(c^*) \mid Q] \in \Re^{I-1}$ for a given c*. Since $c_i = V$ for all i=1,...,I-1, it follows from proposition S23 that every commodity is represented in the same quantity in every element of L(c*). Since every ψ^i serves the same purpose to the consumer and every ψ^i is represented by one unit, i=1,...,I-1, and since $x_I^1 = \ldots = x_I^i = \ldots = x_I^{I-1}$, the problem appears completely symmetric to the producer. Thus every producer will identify $p[L(c^*) \mid Q] \in \Re^{I-1}$ over L(c*) with the uniform distribution.

This argument can be elaborated further. It is known to every member of B that every $x \in L(c^*)$ is different, since from axiom S3 $x^i = [\psi^i, x_I^i] \in L(c^*)$, $\psi^i \in \Re^{I-1}$, where $\psi^i = [x_1 = 0, \ldots, x_i = 1, \ldots, x_{I-1} = 0]$, and $x_I^i \in \Re$, and then at least in principle $\psi^i \neq \psi^j$, $i \neq j$, for all i,j=1,...,I-1. This difference may in principle be of importance to a member of W. Assume for a moment that c* is the price vector that makes $x_I^i = 0$ for all $x^i \in L(c^*)$, i=1,...,I-1. In that case every element of L(c*) contains only the products of the members of B. Although a member for B recognizes that a member of W may prefer an element x^i, in this case x_i, to an element x^j, that is x_j, the reverse situation is also possible. Based on the available information any member of B will in this case recognize that every element could be preferred to any other, and there is no information whatsoever that renders any element more likely to be chosen than any other. Thus every member for B will assign equal probability of being chosen to every element of L(c*). This implies that every member of B views the fact that an element contains a particular alternative x_i, as no information of relevance to the members of B, since no member can in any way use this fact to assess the choice of an arbitrary member of W. Thus the members of B are neutral towards the fact that a particular element contains x_i for all i=1,...,I-1, hence the uniform distribution is specified:

S34 Axiom.
$\forall c \in C^\wedge$, $x_I^1 = \cdots = x_I^i = \cdots = x_I^{I-1} \Rightarrow p[x^i = D(c)] = (I-1)^{-1}$, $x_I^i \subset x^i$, for all $x^i \in L(c)$, $i=1,\ldots,I-1$.

Axiom S34 is similar to Laplace's "Principle of Insufficient Reason". The crucial difference is that a uniform distribution is only assumed if there exists a strong symmetry in the information concerning the different alternatives.

Since the probabilities are subjective for the members of B, it is reasonable to assume that a producer confronted with a situation where it is not possible to say anything about the preferences of the members of W will assign equal probabilities to every alternative in a case where the information concerning each and every alternative is symmetric. Thus axiom S34 has to be seen as very reasonable.

Some may even argue that axiom S34 follows as a theorem from our assumptions and from definition S25. But this is not entirely convincing. In principle every distribution which fits the information set can be assigned to L(c). Thus the question of which distribution will be chosen when there is no relevant information in the information set, has to be answered through an axiom.

Considering every possible price vector $c^* \in C^\wedge$, from axioms S12 and S34 and proposition S23, and definition S25, we now arrive at the following result:

S35 Theorem.

For all $x^i, x^j \in L(c)$, $i,j=1,...,I-1$, $i \neq j$, for a given $c \in C^\wedge$,

$c_i \leq c_j \Rightarrow p[x^i=D(c)] \geq p[x^j=D(c)]$.

<u>Proof</u>: Assume $c_i \leq c_j$. Axiom S34 says, for $x_I^1 = \cdots = x_I^i = \cdots = x_I^{I-1} \Rightarrow$ $p[x^i=D(c)] \in p(c) = (I-1)^{-1}$, $x_I^i \subset x^i$, for all $x^i \in L(c)$, $i=1,...,I-1$. Proposition S23 says, $c_i \leq c_j \Rightarrow x_I^i \geq x_I^j$. Since axiom S12 is contained in Q and says that $x^i \geq x^j \Rightarrow u(x^i) \geq u(x^j)$, it follows from definition S25 that, $c_i \leq c_j \Rightarrow x_I^i \geq x_I^j$ $\Rightarrow p[x^i \in D(c^*)] \geq p[x^j \in D(c^*)]$, for all $i=1,...,I-1$. QED.

Theorem S35 says that since the choice set is in L(c) the price vector *c* determines x_I^i completely. And since the axiom of monotonicity says that $u_i \geq u_j$ if $x_I^i \geq x_I^j$ all other factors equal, this implies that $p(u_i^0) \geq p(u_j^0)$. Therefore, the probability distribution for a consumer over the *I-1* alternatives has to satisfy theorem S35.

The producers are not interested in one consumer's choice alone. Their focus is on the whole population of consumers, denoted W. Theorem S35 is by nature general and valid for all consumers. Thus it is valid for the population of consumers as a whole. In this perspective the probability p_i can be interpreted as the probability that alternative $x^i \in L(c)$ will be chosen by a randomly chosen consumer of the population W.

We have still not answered the question of how the *i*-th producer will set his subjective probability distribution *p* due to the available information. But we noted in theorem S33 that the only unbiased quantity that measures the uncertainty of an arbitrary probability distribution *p* is the measure H(p). In our case where there are several possible distributions P to choose among, the question is: Which distribution is using only the available information, i.e. minimal information, and is at the same time unbaised? The answer is obviously the probability distribution p* that maximizes H(p) and at the same time satisfies all other conditions on *p*, i.e theorem S35. This is so because H(p) is a measure of uncertainty. The higher the value of H(p) is, the greater the uncertainty, and thus the less information is assumed into the distribution. Therefore, the maximum value of H(p) corresponds to the minimal information probability distribution, which is the distribution that uses only the specified information, i.e. the only information available to the producers.

S36 Theorem.

Let the price vector $s \in S$ be rational. Then the probability distribution $p(c) = p[L(c) \mid Q] \in \Re^{I-1}$ satisfies theorem S35 and at maximizes H(p) if and only if it is of log-linear type, i.e. if for some α, $\delta \in \Re$

[E32] $p[L(c) \mid Q] = \exp\{\alpha - \delta \cdot s\}$, $\delta \geq 0 \Leftrightarrow p_i(c) = \exp\{\alpha - \delta \cdot c_i\}$, $\delta \geq 0$, $\forall i$, $s \subset c$,

where c_i is an element of $s \in S$, and $\alpha = -\ln\{\sum_i \exp(-\delta \cdot s_i)\}$.

Proof: It is sufficient to show that [E32] is the unique solution to

[E33] $\{\max_p - K \sum_i p_i \ln p_i$ subject to $\sum_i p_i = 1, \sum_i c_i \cdot p_i \leq \bar{c}\}$, and $c_i \leq c_j \Rightarrow p_i \geq p_j$, and where $\sum_i c_i \cdot p_i \leq \bar{c}$ follows from axiom S24. First we solve [E33]. We state the Lagrangean function for [E33]:

$L = -K \sum_i p_i \ln p_i - \lambda_1(\sum_i p_i - 1) - \lambda_2(\sum_i c_i \cdot p_i - \bar{c})$. The first 2I-2 optimality conditions of the problem are: $\frac{\partial L}{\partial p_i} = K \cdot \ln p_i + 1 + \lambda_1 + \lambda_2 \cdot c_i \geq 0 \wedge \frac{\partial L}{\partial p_i} p_i = 0$, for all i=1,...,I-1. Assume $p_i > 0$ for all i=1,...,I-1. Then $\frac{\partial L}{\partial p_i} = K \cdot \ln p_i + 1 + \lambda_1 + \lambda_2 \cdot c_i = 0$, for all i=1,...,I-1. If we solve for p_i we get: $p_i = \exp(-\lambda_2 c_i/K) \cdot \exp((-1-\lambda_1)/K)$, for all i=1,...,I-1. Since $\sum_i p_i = 1$,

$\sum_i p_i = \exp((-1-\lambda_1)/K) \sum_i \exp(-\lambda_2 c_i/K) = 1 \Leftrightarrow$

$\exp((-1-\lambda_1)/K) = [\sum_i \exp(-\lambda_2 c_i/K)]^{-1}$. Thus we write:

$p_i = \exp(-\lambda_2 c_i/K) \cdot [\sum_i \exp(-\lambda_2 c_i/K)]^{-1}$, for all i=1,...,I-1, which is equal to

[E32] if $\delta = \lambda_2/K$. Next we check that $c_i \leq c_j \Rightarrow p_i \geq p_j$ for all i=1,...,I-1.

Since $0 < p_i \leq 1$, $\frac{\partial p_i(c)}{\partial c_i} = \delta \cdot \{\sum_j e^{-\delta c_j}\}^{-1} \cdot e^{-\delta c_i} \cdot [\{\sum_j e^{-\delta c_j}\}^{-1} \cdot e^{-\delta c_i} - 1] = \delta \cdot p_i \cdot [p_i - 1]$

$\leq 0 \Rightarrow c_i \leq c_j \Rightarrow p_i \geq p_j$, for all i,j=1,...,I-1, i≠j. QED.

It is evident that all members of B will use [E32] as their chosen probability distribution in assessing the choices of the members of W. This follows from axiom S24 where it is specified that the information set is the same for all members of B, i.e. $Q_i = Q$ for all i=1,...,I-1, and from theorem S36, which says that [E32] is a unique representation of the probability distribution over L(c), given $c \in C$^.

Furthermore we recognize that the probability distribution depends upon all prices set by all producers. Thus the probability that a consumer will choose alternative *i* is a function of all prices in the economy as defined here. Observing this we note that

$$[E34]\ \frac{\partial p_i(c)}{\partial c_i} = \delta\cdot\{\sum_j e^{-\delta c_j}\}^{-1}\cdot e^{-\delta c_i}\cdot[\{\sum_j e^{-\delta c_j}\}^{-1}\cdot e^{-\delta c_i} - 1] = \delta\cdot p_i\cdot[p_i - 1] \le 0,$$

since $0 \le p_i \le 1$, and,

$$[E35]\ \frac{\partial p_i(c)}{\partial c_k} = \delta\cdot\{\sum_j e^{-\delta c_j}\}^{-2}\cdot e^{-\delta c_i}\cdot e^{-\delta c_k} = \delta\cdot p_i\cdot p_k \ge 0.$$

The general probability distribution has the properties that the probability that a consumer will choose alternative $x^i \in L(c)$ decreases with increasing price, and increases with the increase in price of any other alternative.

We note that the objective function H(p) in [E33] is an unbiased measure of the general uncertainty in the industry B as to how the final outcome will be for every member of the industry. That H(p) is a measure of uncertainty is obvious since the measure decreases when the probability distribution is sharply peaked, and increases when the probability distribution approaches the uniform distribution.

We observe that the probability of observing the choice of any brand *i* is dependent upon the parameter δ. Assuming K=1, if the expected price level for the whole industry increases marginally, δ is interpreted as the marginal increase in H(p), i.e. the industry uncertainty. Accordingly the uncertainty, i.e. p, will move towards the uniform distribution as $\bar{c}$ increases. This is quite reasonable since the information contained in $\bar{c}$ is of less value to the producers when it implies few restrictions on p than when it implies a strong restriction. Since $0 \leq \bar{c} \leq 1$, a value of $\bar{c}$ close to 1 implies a wider range of possible prices for every member of B than do a value of $\bar{c}$ close to 0. If $\bar{c}$ is close to 1, this implies that any c_j $\forall j$ can be between 0 and 1. If $\bar{c}$ is, say, 0.1 then any c_j will on the average have to lie between 0 and 0.1. Consequently the range of possible probability distributions is narrowed.

Hence we observe that the probability of observing the choice of a particular brand j is the inverse exponential function over the marginal increase in industry uncertainty as a consequence of a marginal increase in the expected price level δ, multiplied by the price of brand j c_j, as set by firm j. Since δ decreases as $\bar{c}$ increases, this interpretation says that the lower the uncertainty is initially, the more sensitive the demand is towards the price level of a particular brand. Finally, the parameter δ can be viewed to encompass both how the present price level of a particular brand will influence the probability of observing the choice of that particular brand, and how the historical market information will influence the present distribution. Hence the parameter δ absorbs some essential information.

In section 8.1 of chapter 8 an extension of the information set will be presented. This extension comprises marketing effort, marketing surveys and quality. Hence these bits of firm specific information are incorporated into problem [E33], and a probability distribution based on this information is added to the information already included in the information set. It should be pointed out that the solution and the equilibrium which is derived for the basic model, will also apply to this extended model.

6.3 A note on logit.

It is evident that the structures of [E32] and [E18] are quite similar. If v_i is assumed to be equal to $-\delta \cdot c_i$ for every i=1,...,I-1, then these models are identical. The crucial question is: Why should anyone using a logit model observe that v_i should be equated with $-\delta \cdot c_i$? And furthermore, as we will show in section 8.1, why should anyone recognize without any axiomatic basis, that v_i should be equal to a specific function over a range of several variables? That this could be the case can not be deduced from the assumptions underlying the logit model. Neither can the value of the parameter δ, and any other paramteres, be found if one should entertain the idea that v_i could be represented by a linear function over price and several other variables. When δ, in principle, can be derived in our model [E32], a corresponding parameter for the logit model has to be determined using additional assumptions of consumer behaviour of some kind. Furthermore, the logit model does not permit any enlargement of the information set to include variables such as marketing effort, quality, and so on. Hence, even though the models [E32] and [E18] are similar in structure, nevertheless they are totally different both regarding the fundamental notions, and their practical use. Consequently these models can hardly be compared when it comes to economic interpretation.

6.4 Expected demand and income.

We stated above that the producer will use his assessment of the consumers expected demand as an input in deciding how to behave. Thus we have to assume the relevance of the concept of expectation.

S37 Definition.

$E(x^i) \equiv p_i \cdot x^i \cdot N$

The expectation that a certain brand *i* will be chosen by an arbitrary consumer is equal to p_i, and the expected transacted quantity is the probability of choice times the number of consumers in population W, $p_i \cdot N$.

The producer is not in a position to know which alternative a consumer of the population W will choose, i.e. he does not know which $x^i \in D(c)$ for all $x^i \in X$, i=1,...,I-1 will be chosen. Therefore, to the producer, the choices of the consumers of W seem random, and *x* can then be said to be a random variable relative to the producers. Since the expectation of a random variable is defined, the producers are now in a position to assess the expected demand from the consumers of W.

Let the population of consumers W be of size N, $W \in \Re^N$. Let $x_n^0(i,c)$ denote the choice of consumer $n \in W$ for a given vector of disposable income prices $c \in C^\wedge$, i.e. that $x^i = D(c)$ for consumer $n \in W$. Let $x^0(i,c)$ denote the total population demand for alternative *i*, and let $x^0(c)$ denote the *I-1* vector with elements $x^0(i,c)$, i=1,...,I-1.

S38 Definition.

The total demand from the population W for an alternative *i* from the set of alternatives i=1,...,I-1, is given by,

$$[E36] \quad x^0(i,c) = \sum_n x_n^0(i,c), \ \forall i, \ x^0(i,c) \in x^0(c).$$

It now follows from definition S37 of expectations that the expected demand vector for the population as a whole for alternative *i*, is given by,

[E37] $E[x^0(c)] = N \cdot p(c) \Leftrightarrow E[x^0(i,c)] = N \cdot p_i(c), \forall\ x(i,c) \in X,$
$E[x^0(i,c)] \in E[x^0(c)],$

where p(c) denotes the probability distribution of the choices for the consumers of W for given prices $c \in C^\wedge$.

[E37] says that the expected demand is equal to the total size of the population times the probability that a consumer of W will choose that particular alternative.

We stated in the representation theorem S? that the distribution for the choice probabilities *p* is represented by [E32], which implies that [E37] can be written,

[E38] $E[x^0(c)] = N \cdot \exp\{\alpha - \delta \cdot s\},$

or

[E38a] $E[x^0(i,c)] = N \cdot \exp\{\alpha - \delta \cdot c_i\}, \forall i=1,\ldots,I-1, E[x^0(i,c)] \in E[x^0(c)].$

We may now conclude that the producers in the differentiated market assess the expected demand as stated in [E38].

The vector of expected producer incomes are calculated as follows,

[E39] $EI[x^0(c)] = E[x^0(c)]^T \cdot s$,

or

[E39a] $EI[x^0(i,c)] = E[x^0(i,c)] \cdot c_i$, $\forall i=1,...,I-1$,

where $EI[x^0(c)]$ denotes the expected income vector for all the *I-1* producers in the market, and $EI[x^0(i,c)]$ denotes the expected income for producer of alternative *i*.

6.5 Producer costs.

We will now introduce the concept of costs. We simply state that every producer i=1,...,I-1 in the market supplying the population W, has a cost function. The *I-1* vectors of costs are given by,

S39 Definition.

[E40] $K = \{\kappa + \nu \cdot x^s\}$,

where $\kappa \in \Re^{I-1}$ is the vector of fixed costs, $\nu \in \Re^{I-1 \times I-1}$ is the diagonal matrix of variable costs, and $x^s \in \Re^{I-1}$ is the production vector.

It should be noted that the producer costs K are indexed in the same way as prices, i.e. K is the income deflated producer costs.

For each producer we may state,

[E40a] $K_i = \kappa_i + \nu_i \cdot x_i^s$, $\forall i=1,\ldots,I-1$,

where x_i^s denotes the quantity produced by producer *i*.

S40 Axiom.

For a given $c \in C^\wedge$, $x^s = E[x^0(c)]$.

Axiom S40 implies that $x_i^s = E[x^0(i,c)]$ for all $i=1,\ldots,I-1$. This assumption is stated for the sake of convenience. Assuming instant production and thereby we avoid introducing the problem of inventory. We write [E40] as follows,

[E41] $K(c) = \{\kappa + \nu \cdot E[x^0(c)]\}$,

or

[E41a] $K_i(c) = \kappa_i + \nu_i \cdot E[x^0(i,c)]$, $\forall i=1,\ldots,I-1$,

which gives the expected costs.

The introduced cost structure is quite simple. This is done to avoid complicating the main issue. Introducing a more complex cost structure, accounting for such problems as capacity etc., would make the model more realistic. However, such an extension is left for further research.

6.6 Profit maximization.

We will now turn to the behaviour of population B, initially defining and state their strategy space. Through an interpretation of this strategy space and the implications of these interpretations, and from our assumptions about the behaviour of the members of population B, a complete description of producer behaviour is provided.

S41 Axiom.

The members of population B can only make a choice from the strategy space

$$S = C^{1\wedge}\times\cdots\times C^{i\wedge}\times\cdots\times C^{I-1\wedge}.$$

The strategy space of the members of the producer population B is denoted $S\in\Re^{I-1}$, which coincides with the price space corresponding to $\Psi\subset X$. Initially the strategy space is made up by the Cartesian product of the sets $C^{i\wedge}$ for all $i=1,\ldots,I-1$, the set of possible prices for the members of B, i.e. $S = C^{1\wedge}\times\cdots\times C^{I-1\wedge}$. As it is possible that the strategy sets of each member of B may differ from the same members' sets of possible prices, we write

$$[E42]\ S = S_1\times\cdots\times S_i\times\cdots\times S_{I-1},$$

where $S_i\subseteq C^{i\wedge}$ denotes the strategy set of the i-th producer.

The strategy space is a vector space where $s^g = [s_1^g,\ldots,s_i^g,\ldots,s_{I-1}^g]$ is a strategy vector and s_i^g is the strategy selected by the i-th producer, g=1,...,G, each producer choosing from G possible strategies.

Note that $C^\wedge = S \times C^{I\wedge}$. Thus the elements of $C^\wedge$ could have been the arguments of the profit function, but since the elements of C^I are influenced neither by the members of population W nor the members of population B, these elements are dropped as any part of the arguments of both the income and the cost functions.

It does follow that every strategy vector implies a set of elements denoted $\pi(s^g) \in \Re^{I-1}$, where the elements are *I-1* vectors of values called producer profit. Let $I(s) \in \Re^{I-1}$ denote the vector of deterministic incomes for a given strategy or price vector s.

S42 Definition.

$\pi(s^g) = I(s^g) - K(s^g)$.

The profit space is now denoted $\pi(S)$ and has as elements every element which is a member of a profit set $\pi(s^g)$ implied by a strategy vector s^g for all g=1,...,G.

It follows from definition S42 that the profit vectors of the set $\pi(s^g)$ are derived from the respective income functions and cost functions which both contain an element of assessment of consumer behaviour. This also forms the connection between population B and population W. If this assessment of consumer behaviour is the same for all members of the population B, then there is only one element of every $\pi(s^g) \subset \pi(S)$, i.e. every $\pi(s^g)$ becomes a profit vector.

A particular strategy selected by a member of the population B may have several implications, but only the profit implication is interesting to a member of B. This follows from the fact that both population B and population W are populations of individuals or associations of individuals. In chapter 2 we described the behaviour of the members of population W. This is a general description of human economic behaviour in a given context. Thus it also applies to the behaviour of the members of population B if the context allows it. We wish to describe how the members of B make their choices among the possible elements of their strategy set. Consequently we recognize that the only economic implication of a chosen strategy is the resulting profit, since this profit is in fact an income to this member. We take it that a member of B is a member of at least one consumer population similar to W:

S43 Axiom.

Every member of B is also a member of at least one population in the same class of populations represented by W.

Axiom S43 implies that an increase in income will result in an increase in the consumption of the general commodity, and by the axiom of monotonicity, and yield increased utility. Thus higher income is preferred over lower, which implies that higher profit is preferred over lower profit. Since the only economic implication of a strategy for a member of B is the amount of profit he receives, he will choose that strategy which maximizes his profit. Let $s_{-i}*$ denote the vector of strategies for the *I-2* other members of B than the i-th.

S44 Theorem.

Every producer i=1,...,I-1 of B, sets the price such that $\max_{s_i \in S_i} \pi_i(s_i, s_{-i}*)$.

Proof: The only economic implication of the strategy space S is the profit space $\pi(S)$. From axiom S43 it follows that every member of B is a consumer in the other capacity . The profit then becomes income. It is evident that a larger income provides the basis for a larger consumption. Axiom S12 of monotonicity states that larger consumption is preferred to lower consumption. Consequently higher profit is preferred to lower profit, and every member of B will maximize his profit so he can consume more in his other capacity as a member of a consumer population. QED.

From definition S37 of expectations, it follows that expected profit can be written,

$$[E43]\ E[\pi(s)] = EI[x^0(s)] - K(s),$$

or

$$[E43a]\ E[\pi_i(s)] = (s_i - \nu_i)\cdot E[x^0(i,s)] - \kappa_i,\ \forall\ i = 1,\ldots,I\text{-}1.$$

S45 Axiom.

Every member of B, i=1,...,I-1, $E[\pi_i(s)]R(E)\pi_i(s)$.

Axiom S45 states that every producer in the market is risk neutral, i.e. that the producers are indifferent between receiving the profit itself or the expectation of the profit.

S46 Theorem.

Every member of B, i=1,...,I-1, sets his price such that, $\max_{s_i \in S_i} E[\pi_i(s_i, s_{-i}*)]$.

<u>Proof</u>: $E[\pi_i(s)]R(E)\pi_i(s)$ for all producers, i=1,...,I-1,

$\Rightarrow \{\max_{s_i \in S_i} E[\pi_i(s)]\}R(E)\{\max_{s_i \in S_i} \pi_i(s)\}$for all i=1,...,I-1. QED.

The producer behaviour is now completely described. We next turn to the discussion of market equilibrium.

Chapter 7. Price equilibrium.

We will now analyse market equilibrium in a differentiated market as defined above. We assume that the different producers, $i=1,...,I-1$, compete between themselves in serving the consumers. Consequently the situation should be described as a game where the producers are the players.

7.1 The game.

There are three ways to characterize a game: The normal form, the characteristic function form and the extensive form. The normal form corresponds to describing the game through strategies, and we consequently choose this form.

There are four parts to the definition of games in normal form: Players, strategies, payoff or profit functions, and additional rules. The first three parts may be symbolized by (B,S,π), where $B=\{1,...,I-1\}$ is the set of players, S is the strategy space of the game and π is the vector of payoff. We denote the strategy set of the i-th player by S_i, and its elements are any possible c_i. A price value that is an element of the strategy set of the i-th player is denoted s_i. A strategy vector is written $s = [s_i,s_{-i}]$ where s_{-i} is the vector of strategies of the $I-2$ players other than the i-th. The payoff or profit function of the i-th player is denoted, as above, $\pi_i(s)$, which is equal to $\pi_i(c)$ as long as $c \in S$. It follows from our elaboration above that $\pi_i(s)$ can be substituted by the expected profit function $E\pi_i(s)$. The profit function is, as noted above, scalar valued. Corresponding to the notation for the strategy sets, a profit vector is written $\pi(s) = [\pi_1(s),...,\pi_{I-1}(s)] = [\pi_i(s),\pi_{-i}(s)]$.

In our model the only additional rule we take in to account says that no binding agreements can be made between the players. Thus our game is non-cooperative.

7.2 An equilibrium concept.

To establish an equilibrium we need a conceptual definition. The non-cooperative equilibrium is characterized by two conditions, see among others Friedman (1977). The first states that the equilibrium strategies, denoted by $s^* = [s_1^*,\dots,s_{I-1}^*]^T$, are actually in the strategy sets of the players. The second condition is that no players could increase his own profit by deviating from his equilibrium strategy, given that the other players use their equilibrium strategies. Formally,

S47 Definition.

A price equilibrium is a vector $s^* = [s_1^*,\dots,s_{I-1}^*]^T \in S$ such that

$E\pi_i(s_i^*,s_{-i}^*) \geq E\pi_i(s_i,s_{-i}^*)$, $\forall s_i \in S_i$ and all $i=1,\dots,I-1$.

A price equilibrium is a vector $s^* = [s_1^*,\dots,s_{I-1}^*]^T$ such that each producer $i=1,\dots,I-1$, maximizes his profit with respect to s_i at s_i^* conditional upon prices s_{-i}^* chosen by the other producers. This price equilibrium is called the Nash equilibrium in pure strategies.

7.3 The existence of a unique price equilibrium.

We now turn to the question if there exists a unique price equilibrium. The original theorem of the existence of non-cooperative equilibrium was proved by Nash (1951). Here we will build on the contraction mapping theorem first stated by Banach (1922), and we shall show that there exists a unique price equilibrium in pure strategies for the designed game.

Let us first define the best reply function of the i-th player.

S48 Definition.

The best reply function of producer *i* is

$R_i(s_{-i}) = \text{argmax } E\pi_i(s_i, s_{-i}),\ i=1,\ldots,I\text{-}1.$

The best reply function is the strategy the *i*-th player will choose given all the other players' choices. The strategic variable in our model is price, that is the price the *i*-th player will choose faced with the choices of all the other *I-2* players. Thus $R_i(s_{-i})$ is a function valued over s_{-i}.

Due to definition S48 the best reply function for the *i*-th player can be specified as follows:

[E44] $\frac{\partial E\pi_i}{\partial c_i} = 0 \Rightarrow c_i = R_i(s_{-i})$, for all i = 1, ..., I-1.

We calculate:

[E45] $\frac{\partial E\pi_i}{\partial c_i} = p_i - \delta(\pi_i p_i - \pi_i p_i^2) = 0$, where $\pi_i = c_i - v_i$, $p_i = p_i(c_{-i}) = p_i(s_{-i})$, $\forall i$.

The expected profit to the *i*-th player if an arbitrary consumer chooses his brand is $E(\pi_i) = \pi_i p_i$, since the *i*-th producer makes a profit only if the brand produced by the *i*-th player is chosen by the consumer.

The same producer will experience the following variance in his individual profit:

[E46] $Var(\pi_i) = E[(\pi_i - E\pi_i)^2] = p_i(\pi_i - \pi_i p_i)^2 + (1-p_i)\cdot(0 - \pi_i p_i)^2$
$= \pi_i^2 p_i - \pi_i^2 p_i^2$.

If [E45] is multiplied by π_i on both sides we get:

[E47] $\pi_i p_i - \delta(\pi_i^2 p_i - \pi_i^2 p_i^2) = 0 \Leftrightarrow E\pi_i - \delta Var(\pi_i) = 0$, for all i = 1,...,I-1.

Hence in equilibrium

[E48] $\delta = \frac{E\pi_1}{Var(\pi_1)} = \cdots = \frac{E\pi_i}{Var(\pi_i)} = \cdots = \frac{E\pi_{I-1}}{Var(\pi_{I-1})}$.

[E48] says that the marginal increase in the function H(p), δ, as a consequence of a marginal increase in the expected price level for the whole industry, $\bar{c}$, shall at the equilibrium be equal to expected profit relative to the variance of the profit. We observe that this implies that if the price level is expected to rise from one period to the next δ will be reduced - implying among other things that a rise in prices will affect demand less then earlier. However, from [E48] such a decrease in δ also implies that the relative figure $\frac{E\pi_i}{Var(\pi_i)}$ will decrease for all *i*. The economic interpretation of this is that uncertainty increases, i.e. the variance shows a relative increase. This is so because the effect of known quantities on choice probabilities such as prices decreases in importance relative to unknown quantities. Hence the probability distribution over L(c) will be wider, i.e. the probability of observing a high priced brand being chosen has increased relative to the probability of observing the choice of a low priced brand. Since price is assumed to be the essential information determining the probability distribution, such a development implies greater uncertainty with regard to the outcome based on the available information.

From [E45] the best reply function can be deduced:

$$[E49]\ R_i(c_{-i}) = v_i + \delta^{-1}(1-p_i)^{-1} \Leftrightarrow R_i(c_{-i}) = v_i + \delta^{-1} + \delta^{-1}\cdot e^{-\delta R_i}\cdot[\sum_{j\neq i} e^{-\delta c_j}]^{-1},$$

for all $i = 1,\ldots,I-1$.

Equation [E49] is a functional equation which has no explicit solution for $R_i(c_{-i})$. Hence we can only obtain an approximation of the best reply function. It is reasonable to approximate [E49] at the point $c_j = \bar{c}\ \forall j=1,\ldots,I-1$, since it is assumed that $\sum_j c_j\cdot p_j \leq \bar{c}$. Thus it is reasonable to expect that c_j, $j=1,\ldots,I-1$, will fluctuate around $\bar{c}$. Consequently $\bar{c}$ is a more reasonable evaluation point than for example $c_j = 0$ $\forall j=1,\ldots,I-1$.

From [E49] we deduce that if $c_j = R_i = \bar{c} \Rightarrow$

$$[E50]\ \bar{c} = v + \delta^{-1} + \delta^{-1}(I-2)^{-1},$$

where $v = v_i = v_j\ \forall i,j$, since $\bar{c} = v_i + \delta^{-1} + \delta^{-1}e^{-\delta R_i}(\sum_{j\neq i} e^{-\delta c_j})^{-1} \Rightarrow$

$\bar{c} = v_i + \delta^{-1} + \delta^{-1}e^{-\delta\bar{c}}(\sum_j e^{-\delta\bar{c}})^{-1}$. But since $\bar{c}$, δ, and $(I-2)$ is equal to all players, $v = v_i = v_j\ \forall i,j$ if $c = \bar{c}$, hence $\bar{c} = v + \delta^{-1} + \delta^{-1}e^{-\delta\bar{c}}(e^{\delta\bar{c}}(I-2)^{-1})$. Thus for $c = \bar{c}$ we set

$v = 0.$

Since $\bar{c}$ is a known quantity, we solve [E50] for the unknown parameter δ in the point $c = \bar{c}$:

[E51] $\delta = \dfrac{1 + (I-2)^{-1}}{\bar{c}}.$

We note that δ can be estimated as a function over the range of competitors represented by the number I-2, and the historical information represented by the expected price level for the industry as a whole, $\bar{c}$. Furthermore, we observe that δ will approach $\frac{1}{\bar{c}}$ when (I-2) approaches any large number. Consequently δ will decrease with increasing number of competitors and increasing industry price level. Intuitively this seems reasonable. We recall that if $\bar{c}$ is increased marginally, δ is interpreted as the marginal increase in the objective function [E33], i.e. in the industry uncertainty. Hence it is reasonable to assume that δ will decrease as the uncertainty increases. Obviously the uncertainty will increase with increasing numbers of competitors, since the market power of each player is expected to decrease with increasing numbers. As stated above an increase in the maximum acceptable industry price level $\bar{c}$ will slacken the restrictions on the range of individual prices, and hence the probability distribution *p*. Accordingly the uncertainty increases. It is, therefore, reasonable to observe that the marginal increase in the uncertainty is positive, and decreases with increasing $\bar{c}$. Consequently it is plausible that $\delta \rightarrow \frac{1}{\bar{c}}$ for increasing number of competitors.

In $c = \bar{c}$ the first and second partial derivatives are:

[E52] $\frac{\partial R_i}{\partial c_j} = e^{-(1 + (I-2)^{-1})}(I-2)^{-1}(1+(I-2)^{-1})^{-1} = f_1$, for all i,j=1,...,I-1, i≠j.

[E53] $\frac{\partial^2 R_i}{\partial c_j^2} = \frac{2(I-2)^{-1} - 1}{(I-2)^2 \cdot \bar{c}} = f_2$, and

[E54] $\frac{\partial^2 R_i}{\partial c_j \partial c_k} = \frac{2}{(I-2)^3 \cdot \bar{c}} = f_3$, for all i,j,k=1,...,I-1, j≠i,k, k≠i.

Assuming that the variable cost is positive, the approximation of equation [E49] in $c = \bar{c}$ is:

$$[E55]\ R_i(c_{-i}) = v_i + \bar{c}\{1 - (1 + (I-2)^{-1})e^{-(1+(I-2)^{-1})} + (I-2)^{-1}[(I-2)^{-1} - \tfrac{1}{2}]\}$$
$$+ f_1\sum_{j\neq i} c_j + \tfrac{1}{2}f_2\sum_{j\neq i} (c_j^2 - \bar{c}c_j) + \tfrac{1}{2}f_3\sum_{j\neq i,k}\ \sum_{k\neq i,j} (c_jc_k - \bar{c}(c_j+c_k) + \bar{c}^2),$$

assuming that terms of higher order are equal to zero

We observe that [E55] is only dependent upon known quantities, that it will increase with $\bar{c}$, v_i, and c_j for all j≠i, j=1,...,I-1, and decrease with increasing number of competitors, I-2.

The approximation [E55] of the reply function is presented only in order to research some of the properties of this important relationship further. However, when it comes to our objective of establishing a unique price equilibrium for the game in question, we will have to use equation [E49] directly. We now turn to this problem of establishing a unique price equilibrium.

Let us first define the term contraction mapping:

S49 Definition.

The mapping R(c) is a contraction if and only if there exists a number $0 < k < 1$ such that

$$d(R(c^1), R(c^2)) \leq k \cdot d(c^1, c^2) \text{ for all } c^1, c^2.$$ [4]

We now recall a useful property of metric spaces:

S50 Lemma.

R: $S \to S$ is a continuous contraction mapping that associates a point R(s) of S with a point *s* of S, where S is a complete metric space. Then there exists a unique fixed point s^*: $s^* = R(s^*)$.[25]

Proof: Existence: Choose an arbitrary point *c* in C. We have:
$d(R^n(c), R^{n+1}(c)) \leq k \cdot d(R^{n-1}(c), R^n(c)) \Leftrightarrow d(R^n(c), R^{n+1}(c)) \leq k^n \cdot d(c, R(c))$
If $n \to \infty$, then $d(R^n(c), R^{n+1}(c)) \to 0$, since $k^n \to 0$. Thus there exists a fixed point for $R^n(c) = R^{n+1}(c)$, $n \to \infty$. Uniqueness: Assume the existence of two fixed points, *z* and *w*, for the contraction mapping R. Since R is a contraction $d(R(z), R(w)) \leq k \cdot d(z,w)$. But, since *z* and *w* are fixed points, the following is true: $d(z,w) = d(R(z), R(w))$, since $R(z) = z$, $R(w) = w$. It follows that the contraction $d(R(z), R(w)) \leq k \cdot d(z,w) \Rightarrow d(z,w) = 0 \Rightarrow z = w$. Thus any fixed

25 Definition: A fixed point is a point that is mapped onto itself by a given transformation.

point of a contraction is unique. QED.

This lemma was first stated by Banach (1922) and is called the contraction mapping fixed point theorem. We now state:

S51 Theorem.

The game (B,S,π), where the best reply function for the *i*-th player is represented by [E49], has a unique fixed point.

Proof: We will show that the best reply function [E49] is a contraction mapping for all $i = 1,\dots,I-1$. Lemma S50 says that if the following two conditions are satisfied then the game has a unique equilibrium point:

1. R_i is a contraction mapping for all $i = 1,\dots,I-1$.

2. The strategy space is a complete metric space.

We start by proving that R_i is a contraction mapping. Since the best reply function is an unknown function specified by a functional equation, we have to research the properties of R_i. It is known that if $\frac{\partial R_i}{\partial c_k} < 1$ for all $k=1,\dots,I-1$, $k \neq i$, then R_i is a contraction mapping over the strategy space S.

$$\frac{\partial R_i}{\partial c_k} = -[1-p_i]^{-2} p_i \{(1-p_i)\frac{\partial R_i}{\partial c_k} - p_k\} \Leftrightarrow \frac{\partial R_i}{\partial c_k} = \frac{p_i p_k}{[1-p]^2 + [1-p_i]p_i}, \text{ for all } i,\ k \neq i.$$

R_i is a contraction if $\frac{\partial R_i}{\partial c_k} < 1$, for all i, $k \neq i$. Thus we check:

$\frac{\partial R_i}{\partial c_k} = \frac{p_i p_k}{[1-p]^2+[1-p_i]p_i} < 1 \Leftrightarrow p_i(1+p_k) < 1$, for all i, k≠i.

We know that p_k for all k=1,...,I-1, k≠i, cannot be greater than $1-p_i$. This implies that if $p_i(1+p_k) < 1$ is true for $p_k = 1-p_i$, then $p_i(1+p_k) < 1$ holds for every $0 \leq p_k \leq 1-p_i$, for all k=1,...,I-1, k≠i. If we set $p_k = 1-p_i$ we get:

$1-p_i < p_i^{-1} - 1 \Leftrightarrow p_i < 1.$

Thus, if $0 < p_i < 1$ for all i=1,...,I-1, then $\frac{\partial R_i}{\partial c_k} = \frac{p_i p_k}{[1-p]^2+[1-p_i]p_i} < 1$ for all i,k, k≠i. Since $p_i = \exp(-\delta c_i)\cdot(\sum_j \exp(-\delta c_j))^{-1}$ and $0 \leq c_i \leq 1$ for all i=1,...,I-1, then $0 < p_i < 1$ will always be true. Thus R_i is a contraction mapping for all i=1,...,I-1.

The second condition is that the strategy space S constitutes a complete metric space. No player will set his price lower than the normalized cost, and no higher than the level of the normalized income of the producers. Therefore, $v_k \leq c_k \leq 1$, for all k, and thus the strategy set $S_k = [v_k,1]$ for all k = 1,...,I-1, which is a compact and convex set of $\Re_+$. Thus the strategy space S is a complete metric space since every compact space is a complete space. QED.

In proving theorem S51 we have established that the subjective probability distribution of every member of B is represented by [E32]. We may therefore conclude that all members of B will assess the demand from the population of consumers W using the indicated probability distribution as stated by [E37]. Thus the demand equation that the member of B faces is stochastic, which it always will be since it represents a forecast on the hand of the producer.

The equilibrium established in theorem S51 is the expected price equilibrium and is as such used by the producers in setting their offer to the market. The subsequent market equilibrium will be a consequence of the expected market equilibrium and the unknown demand from the members of W.

The offer presented by member *i* of B can be summarized by $c_i^* = R_i(c_{-i}^*)$ and $x_i^{s*} = N \cdot p_i(c^*)$. The number of individual members of population W who would want to choose alternative *i*, for a given price vector c^*, i.e. number of consumers for who $D(c^*) = x^i$, is denoted by $n_i(c^*)$.

S52 Axiom.

If $x_i^{s*} < n_i(c^*)$ for any $i=1,\ldots,I-1$, then those members of W who do not receive their choice will not choose any other alternative from Ψ.

Axiom S52 states that any surplus demand for a particular alternative is not allocated to any other alternative. This assumption is stated to avoid complicating the analysis with reallocation of surplus demand.

It follows that since the price level is assumed to be fixed by the producers at each incremental moment, the quantity transacted at each moment will be $\min\,[x_i^{s*}, n_i(c^*)]$. Denote the transacted quantity in equilibrium for alternative x^i by x_i^*. We state:

S53 Theorem.

c^* and $x_i^* = \min\,[x_i^{s*}, n_i(c^*)]$ represent the unique market equilibrium for the market defined by W and B.

<u>Proof</u>: Due to theorem S51 $c^* \Rightarrow c_i^*$ for all $i=1,\ldots,I-1$. Due to axiom S40 $x_i^s = N\cdot p_i$ for all $i=1,\ldots,I-1$. $c^* \Rightarrow x_i^{s*} = N\cdot p_i(c^*) \wedge n_i(c^*)$. In virtue of axiom S52 $x_i^{s*} = N\cdot p_i(c^*) \wedge n_i(c^*) \Rightarrow x_i^* = \min\,[x_i^{s*}, n_i(c^*)]$. QED.

Theorem S53 concludes our analysis by stating the existence of a unique market equilibrium. This equilibrium is achieved by recognizing that the producers initially have to analyze their own situation, and then market a fixed offer. These two assumptions are the basic practical assumptions of the presented theory.

Chapter 8. Extensions and reflections.

8.1 An extened model.

The model presented in this treatise on producer assessment of demand, is based on the information included in Q, i.e. some historical data, and the general principles of economic behaviour. We will in the subsequent analysis discuss how the information set should be enlarged and how such an extension affects the results of the basic model presented above.

Market research is quite often used by firms when an assessment of future demand is done. One naturally ask how such information should be used in the presented model. This maket information could be comprised in the model by imposing restrictions on the probability distribution, p, i.e. I-1 additional restrictions are included in the problem denoted [E33]. Assume that θ_j, for all $j=1,...,I-1$, is the fraction of the respondents in a market survey which said they would choose brand j. One could then impose the restriction $p_j \geq \theta_j - \frac{\theta_j}{I-2} \Leftrightarrow \frac{I-2}{I-3}p_j \geq \theta_j$, for all $j=1,...,I-1$, $I-2 > 0$. This restriction says that the probability of observing that brand j is chosen should be greater than the observed fraction adjusted for some expression which indicates that the probability p_j should approach the fraction θ_j when the number of alternatives open to the consumers increases. A tentative suggestion of such an expression is $\frac{\theta_j}{I-2}$. Such a model would comprise both the uncertainty in the situation and the available information, and at the same time be unbiased towards the information used.

Eventually the model could be given an economic interpretation. Denote the marginal change in the objective function in [E33] of an increase in θ_j by μ_j,

$j=1,\ldots,I-1$. Then μ_j can be interpreted as an expression of the marginal reduction in the total industry uncertainty, and hence also the uncertainty to the producer of brand *j*, if θ_j is increased.

The information gained from a market survey could also encompassed in the restriction $\sum_j c_j \cdot p_j = \bar{c}$ in [E33], where $\bar{c}$ could be constructed such $\sum_j c_{j,t-1} \cdot \theta_j$.

An other category of information which should be considered relevant is represented by macroeconomic quantities such as expected economic growth. It is in most cases likely that any expected general economic growth will reduce the effect of high or low prices on the choice probability, since increased income will make the consumers less price conscious. Denote the expected growth rate in the economy for the present period by g. If g is incorporated in the expected price level $\bar{c}$, so that the expected price level is somewhat increased as a function of g, then δ will decrease and the effect of the level of any brand's price on demand will be less. Since it is plausible that prices in general will rise as a consequence of positive economic growth, such a incorporation of g into $\bar{c}$ is quite tenable. Hence $\bar{c}$ in the problem [E33] can be modified so that: $\bar{c} = F(g) \cdot \sum_j d_{j,t-1} \cdot c_{j,t-1}$, where $F(g)$ is a function over g.

It is generally accepted that increased quality, everything else being equal, will increase a brand's chances of being chosen. The problem is of course to measure quality in any objective way. Firstly, quality can be distinguished into two different categories: Horizontal and vertical quality. Horizontal quality deals with such features as colour, style, design, texture, flavour, and so on. Vertical quality is higher standard, superior, inferior, octane content in petrol, numbers of doors in a car, hours of light of an electric bulb, and so on. Any objective measure of horizontal quality can hardly be defined. However, one should be able to measure vertical quality. Denote q_j $\forall j$ as some public and official normalized measure of the vertical quality of brand *j* in industry B. We would expect the probability that *j* is chosen to increase with q_j, everything else constant. Hence if vertical quality is included in the information set, then $\frac{\partial p_j}{\partial q_j} \geq 0$ $\forall j$. The question is how quality is to be introduced into our assessment problem. Obviously the quality index will have to be included in some constraint in the problem [E33] of theorem S36. The question is how this is to be done to be in accordance with reasonable economic behaviour with regard both to consumers and producers. A suggestion is that the following constraint is added to the problem of [E33]: $\sum_j p_j \cdot q_j \geq \bar{q}$, where $\bar{q}$ is the industry minimum average quality level, which is either known to all producers, or assumed to be a sort of consensus level which the industry as a whole does not want to fall below. In the car industry for the segment of small cars, the quality index can be measuring the number of doors and the level of extra equipment such as radio, heating, ventilation, security systems and so on. The minimum average level can be some level that most producers will be above, and only the rather low quality cars will be below. Consequently $\bar{q}$ can be set to be equal to the critical value which classifies a brand as low quality if below, and as a brand of medium and higher quality if above.

Most economists believe that active <u>marketing</u> influence demand and preferences. Some argue that increased marketing efforts increase demand. In our context that implies that increased marketing effort increases the probability of choice.

How tenable this theory is can be questioned. However, it has some theoretical and empirical merit. Let b_j denote the marketing effort of firm j, j=1,...,I-1, e.g. a quantity measuring the resources used on advertising during the previous period. How b_j, $\forall j$, is to be comprised in our model has to be researched further. However, a suggestion is forwarded: Recognizing that b_j in general should effect p_j, so that an increase in the effort b_j should increase p_j at a decreasing rate, i.e. $\frac{\partial p_j}{\partial b_j} \geq 0$, $\frac{\partial^2 p_j}{\partial b_j} \leq 0$, an additional constraint in problem [E33] can be formed: $\sum_j p_j \cdot b_j \geq \bar{b}$, where $\bar{b}$ is some agreed averaged minimum level of marketing effort for the industry as a whole. Hence $\bar{b}$ should be interpreted as a kind of critical value for a company's marketing effort. If the effort is below this value, then the company could be regarded to be without brand profile. The company does not seek to distinguish itself from the other producers in the industry. This value $\bar{b}$ could therefore be set through a survey or analysis of which level of marketing effort is regarded as the critical value distinguishing non-profile and profiled companies. Such a constraint says that the probability distribution p should be such that at least $\bar{b}$ on the average is observed for the industry as a whole. Consequently, everything else equal, we should be more likely to observe the choice of brand j than brand i if $b_j > b_i$. This is in accordance with the theory saying that the greater the marketing effort the higher the demand or the probability of choice.

To account for the suggestions above with regard to an extended information set, the information problem [E33] can be remodelled:

$$\max -\sum_j p_j \cdot \ln p_j$$

[E56]

subject to

$$\sum_j p_j = 1,\ \sum_j p_j \cdot c_j \le \bar{c},\ \sum_j p_j \cdot q_j \ge \bar{q},\ \sum_j p_j \cdot b_j \ge \bar{b},\ \tfrac{I-2}{I-3} p_j \ge \theta_j\ \forall j.$$

The Lagrangean function is:

$$L = -\sum_j p_j \cdot \ln p_j - \lambda_1(\sum_j p_j - 1) - \lambda_2(\sum_j p_j \cdot c_j - \bar{c}) + \lambda_3(\sum_j p_j \cdot q_j - \bar{q})$$

$$+ \lambda_4(\sum_j p_j \cdot b_j - \bar{b}) + \tfrac{I-2}{I-3}\sum_j \mu_j(p_j - \theta_j),$$

where λ_1, λ_2, λ_3, λ_4, and $\mu_j\ \forall j$, are the Lagrangean multipliers.

The 2I-2 optimality conditions are:

$\frac{\partial L}{\partial p_i} = \ln p_i + 1 + \lambda_1 + \lambda_2 \cdot c_i - \lambda_3 \cdot q_i - \lambda_4 \cdot b_i - \frac{I-2}{I-3}\mu_i \ge 0 \wedge \frac{\partial L}{\partial p_i} p_i = 0$, for all $i=1,\ldots,I-1$. Assume that $p_i > 0\ \forall i$, then

$\ln p_i + 1 + \lambda_1 + \lambda_2 \cdot c_i - \lambda_3 \cdot q_i - \lambda_4 \cdot b_i - \frac{I-2}{I-3}\mu_i = 0 \Rightarrow$

$p_i = e^{-\lambda_1 - 1 - \lambda_2 \cdot c_i + \lambda_3 \cdot q_i + \lambda_4 \cdot b_i + \frac{I-2}{I-3}\mu_i}$, where $e^{-\lambda_1 - 1} = [\sum_j e^{-\lambda_2 \cdot c_j + \lambda_3 \cdot q_j + \lambda_4 \cdot b_j + \frac{I-2}{I-3}\mu_j}]^{-1}$. Hence:

[E57] $p_i = e^{-\delta c_i + \rho q_i + \gamma b_i + \frac{I-2}{I-3}\mu_i}[\sum_j e^{-\delta c_j + \rho q_j + \gamma b_j + \frac{I-2}{I-3}\mu_j}]^{-1}$, where $\delta = \lambda_2$, $\rho=\lambda_3$, $\gamma = \lambda_4$.

In equation [E57] δ has the same interpretation as earlier. The parameter ρ is interpreted as the marginal decrease in industry uncertainty, i.e. the narrowing of the probability distribution *p* as a consequence of a marginal increase in the minimum average quality level. That the industry uncertainty should decrease when the minimum acceptable quality level is increased is reasonable, since we would expect such an increase to induce consumers to turn from the low quality brands towards brands of higher quality. Hence the probability that the consumer will choose any high quality brand increases, and subsequently the industry uncertainty is reduced since stronger restrictions are introduced. Furthermore, ρ will increase at a decreasing rate with increasing $\bar{q}$, and hence the higher the minimum level is set, the more probable it will be to observe the consumers choosing a high-quality brand. If one assume that $\bar{q}$ is set by a survey in the market as to quality attributes, one will typically observe that $\bar{q}$ will be low in areas where the standard of living is low, and high in developed areas.

γ is interpreted as the marginal decrease in industry uncertainty if the average minimum marketing level is increased marginally. Obviously $\frac{\partial p_i}{\partial b_i} = -\gamma \cdot p_i[p_i-1] \geq 0$, and $\frac{\partial^2 p_i}{\partial b_i^2} = -2\gamma^2 p_i^2 \leq 0$. This is tenable if $\bar{b}$ is interpreted as the critical level of marketing which a company has to be above to gain attention from consumers, since an increase in the minimum level will weaken the market power of the low spending fimrs, to the benefit of the high spendig firms. Hence, if the marketing effort constraint is to be respected the demand for the brands of the stronger marketing companies have to increase at the cost of the firms which are weaker as to marketing.

μ_i $\forall i$ is interpreted as the marginal decrease in industry uncertainty as a consequence of an increase in the market share, measured for the i-th firm in a recent market survey conducted for the whole industry. We note that $\frac{\partial p_i}{\partial \mu_i} = \frac{I-2}{I-3} p_i(1-p_i) > 0$, which says that the marginal influence of brand *i* on the industry uncertainty will be at its greatest when the probability of observing brand *i* is equal to 0.5. This is reasonable since in that case it will be as likely to observe the choice of *i* as that of all other brands taken together, and a marginal change in for example θ_i, will therefore have a strong effect.

Based on the expression [E57] of the choice probability we may now deduce the best reply function for firm *i* for an extended information set:

$$[E58]\ R_i(c_{-i}) = \nu_i + \delta^{-1} + \delta^{-1}\, e^{-\delta R_i + \rho q_i + \gamma b_i + \frac{I-2}{I-3}\mu_i} [\sum_{j \neq i} e^{-\delta c_j + \rho q_j + \gamma b_j + \frac{I-2}{I-3}\mu_j}]^{-1}\ \forall i.$$

Equation [E58] is the general reply function, assuming that the producers use a particular range of information, and that the variables q_j and b_j are fixed and cannot be altered by the players, i.e. price is the only strategic variable of the producers.

Some further research has to be directed towards an extension of the model to several strategic variables, e.g. q_j and b_j in addition to c_j. In this case the best reply function will be the solution to the following differential equation:

$$dE\pi_i = \frac{\partial E\pi_i}{\partial c_i} \cdot dc_i + \frac{\partial E\pi_i}{\partial q_i} \cdot dq_i + \frac{\partial E\pi_i}{\partial b_i} \cdot db_i \Rightarrow$$

$$[E59]\ \frac{dE\pi_i}{dc_i} = \{p_i - \delta \cdot \pi_i \cdot p_i[1-p_i]\} + \rho \cdot \pi_i \cdot p_i[1-p_i] \cdot \frac{dq_i}{dc_i} + \gamma \cdot \pi_i \cdot p_i[1-p_i] \cdot \frac{db_i}{dc_i}\ \forall i.$$

We observe that the best reply with regard to price, c_i*, quality, q_i*, and marketing effort, b_i*, will be found when solving equation [E59], and subsequently all these variables depend on each other. Since it should be of considerable interest to derive an expression and a model which allows more than one strategic variable to influence firm behaviour, the implications of equation [E59] should be a subject of further research.

Returning to equation [E58] we recognize that theorem S51 applies to this extended behavioural equation as well as to the basic relationship of equation [E49]. Hence all the fundamental points made in relation to the basic model represented by equations [E32] and [E49], are valid for the extended model represented by equations [E57] and [E58], if price is kept as the only strategic variable. Subsequently theorem S53 also applies to the extended problem, and an equilibrium is assured.

Consequently we can conclude that a theory comprising how producers in differentiated markets should deal with uncertainty and a wide range of information, but not complete information, is forwarded in this treatise.

Included in this theory is a detailed analysis which explicitly points out how demand functions are to be constructed, and which interpretations the parameters of such models should be subjected to in differentiated markets. That is, any demand function should be on the log-linear form. One should thus not be surprised to observe that in quite a few econometric analyses the log-linear specification is the most appropriate. Observing this econometric implication, we turn in section 8.3 to a discussion of some of the econometric implications of the theory.

8.2 A note on competition.

Competition as such has not been the main subject of this treatise. However, it is obvious that the model forwarded should eventually be used to analyse the competition between firms in the relevant context. Specifically, the presented model should enjoy some significance in analysing the extent of differentiation in an industry with respect to such variables as quality, marketing effort, etc.

The Lerner index L_i is defined as the margin between equilibrium price and marginal cost of firm *i*.

$$[E60]\ L_i = \frac{R_i - v_i}{R_i} = [1 + v_i \cdot \delta \cdot (1-p_i)]^{-1}\ \forall i.$$

The Lerner index measures the degree of market power of firm *i* compared to all other firms in the industry. The firm having the highest index value has the highest profit per unit produced and has the most market power.

Relation [E60] shows that in our context and in a Nash price equilibrium, market power increases with market share, i.e. p_i, and it decreases with higher marginal cost, v_i. From relation [E57] it is known that the market share p_i increases with an increase in quality, and an increase in marketing effort, and that it decreases with increasing price. The effect of these quantities upon p_i depends on the parameters, ρ, γ, and δ, which depends upon industry measures of critical levels. Hence the measure of industry uncertainty H(p) should be recognized as a measure of differentiation within the industry, and subsequently as a measure of competition in the industry. At this point we observe that Shannons (1948) measure H(p) has been proposed as a measure of concentration, i.e. competition, by several authors, e.g. Theil (1967), Horowitz & Horowitz (1968). See also Jacquemin (1987). However, these proposals have been rather ad hoc. The proposals have been forwarded not based on any axiomatic analysis but out of a need to measure concentration within a industry. Thus the precise measure of industry concentration should be the aggregation of the individual Lerner indexes, weighted by market share p_i:

$$[E61]\ L = \sum_i p_i \cdot L_i = \sum_i \frac{p_i}{1 + v_i\delta(1-p_i)}.$$

The difference between H(p) and L is not crucial, but in principle H(p) measures the industry uncertainty and the level of differentiation when L measures the level of concentration within the industry. Hence in practice there are no difference.

Returning to [E60] we observe that the market power of firm i is increased when the quality of the firm's brand is increased, and when the marketing effort of the firm increases. This is in accordance with what one would expect if these variables are given significance as information. Hence the producer assessment model presented could be used, in principle, to analyse complex situations where competition relies not only on price and capacity but also on a diverse range of other variables. How the presented model should be used in an analysis involving a more complex production structure, is left to further research.

8.3 A note on econometrics.

Most models of discrete choice problems are specified as econometric models. This is the case with Börsch-Supan (1987), and McFadden (1978), to mention a few.

Most analysts state, as do Börsch-Supan (1987), that such artificial assumptions as random utility are necessary to derive an econometric model. They recognize that the assumption is questionable, but, nevertheless make the assumption out of convenience. The model set forth in this treatise is not an econometric model as such. It belongs to the domain of economic theory. However, every economic theory has econometric implications, as does this particular theory. We shall not meticulously discuss all relevant problems which arise when an econometric specification based on our model is made. Instead we shall bring up some ideas for further research on these topics.

Quite often an econometric analysis is concerned with the specification and estimation of a demand function. In short this is the problem of identifying the empirical demand relation based on empirical observations that represents equilibrium situations between the demand and the supply relations. One problem which has to be considered in such a specification, is the problem of autocorrelation.

The autocorrelation problem for an econometric specification derived from the theory set forth in this treatise, illustrates at the same time that an econometric specification cannot be made in isolation from the underlying economic model. Assume that the members of B have offered to the market their products four different times. Then the economic agents have gained some experience which is included as historical data in the producer information set Q. Each time the producers have assessed demand and offered their quantities and prices represented by $x_i^{s*} = N \cdot p_i(c^*)$ and c_i^* for every i=1,...,I-1. The consumers have responded by demanding $n_i(c^*)$. Since the price level c^* is decided upon and fixed before the actual demand is known to the producers, the observed quantity is $x_{i,t}^* = \min\,[x_{i,t}^{s*}, n_{i,t}(c^*)]$ for every i=1,...,I-1 of B. Assume that for producer *i*, $x_{i,t}^* = n_{i,t}$, $x_{i,t+1}^* = x_{i,t+1}^{s*}$, $x_{i,t+2}^* = x_{i,t+2}^{s*}$, and $x_{i,t+3}^* = n_{i,t+3}$. We note that the actual demand from population W is observed for *t* and t+3.

However, for t+1 and t+2 the expected demand constitute the observations. The figure below illustrates the situation. The thin line indicates the expected demand as it is assessed by the producers. The thick line indicates actual demand. The points that will be observed are indicated by the notation *t* to t+3.

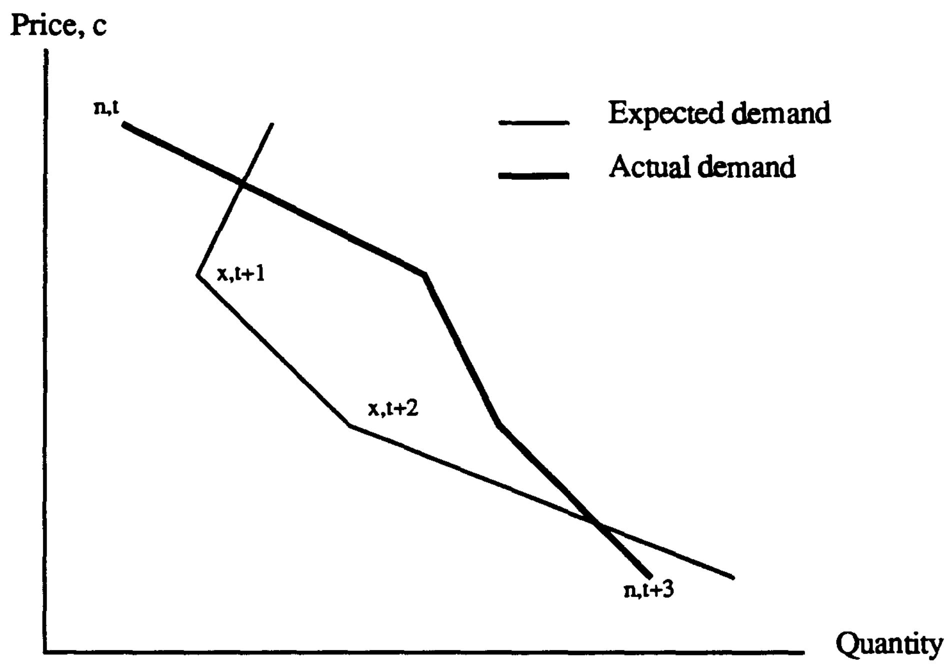

Figure: Expected and actual demand.

It is obvious that the transactions observed for a given price of brand *i* do not necessarily coincide with the actual demand. This is due to axiom S52, and an underlying assumption implying that all producers present their offers in the market simultaneously. The latter assumption, which is not unreasonable in a pure theoretical context, may be deemed invalid in an econometric context since producers usually interact continuously with the market. However, the real econometric problem of a specification based on our theory is due to the fact that the expected relationship used by the producer to set his offer in the market is based on historical data. As such the offer denoted by x_{t+1} in the figure above is derived using the information gained in period *t*, i.e. the information set Q_{t+1} contains the information c^*_t and n_t. Hence the observation at t+1 is not independent of the previous observation at *t*. Neither of the values of the variables x^*, x^{s^*}, n, and c^* is independent of the previous values. Recall that the actual transacted quantity for brand *i* at time *t* in principle can be expressed

[E62] $x^*_{i,t+1} = f^*(x^{s^*}_{i,t+1}, n_{i,t+1})$.

The price of the i-th brand is

[E49] $c^*_{i,t+1} = f_c(c^*_{i,t}, c^*_{-i,t}, c^*_{-i,t+1}, x^{s^*}_{i,t}, x^*_{i,t}, Q^0)$,

which is equal to the best reply function to the i-th player, where Q^0 is the information set excluded for $c^*_{i,t}$, $c^*_{-i,t}$, $c^*_{-i,t+1}$, $x^{s^*}_{i,t}$, $x^*_{i,t}$.

The actual demand is expressed

[E63] $n_{i,t+1} = f(c^*_{i,t+1})$.

The quantity offered in the market, i.e. the supply, is expressed

[E64] $x^{s^*}_{i,t+1} = f^{s^*}(c^*_{i,t+1})$.

It is obvious from equations [E49] and [E62] to [E64], that $x^*_{i,t+1}$ is dependent upon $x^*_{i,t}$, $c^*_{i,t+1}$ is dependent upon $c^*_{i,t}$, $n_{i,t+1}$ is dependent upon $n_{i,t}$, and $x^{s*}_{i,t+1}$ is dependent upon $x^{s*}_{i,t}$. Hence in principle autocorrelation will be present in any attempt to estimate these relationships. When autocorrelation is observed in an econometric model this is not necessarily due to any mis-specification. On the contrary autocorrelation is a phenomenon which quite often should be observed since economic behaviour is autocorrelated. Consequently econometric models which includes autocorrelation should not necessarily be rejected. Instead one should seek the answer to this econometric problem in the underlying economic model, and solve the problem based on an explicit recognition of autocorrelation as an economic phenomenon and not only as an econometric problem.

Therefore, when econometric models are specified based on the assumption that the observations that are observed are formed in the equilibrium between actual demand and supply, without accounting for the assessment process which is described in this treatise, it seems natural to assume that the model's residuals are non-correlated. It should not come as a surprise that autocorrelation is observed when the model is fitthe ted to data. When this is the case most analysts reject their specification, but keep their underlying economic model. As illustrated above, it is not necessarily the specification which should be rejected. It is surely the underlying economic model, i.e. the assumption that the actual observations are not themselves determinants of economic behaviour and later observations, which should be rejected. The problem of autocorrelation is derived directly from an economic phenomenon. A phenomenon which should not be assumed away. It should be given an important position in any economic models leading to econometric specifications of relationships in differentiated markets.

Consequently most econometric specifications dealing with economic behaviour in differentiated markets at time *t* should include as exogenous variables the observations of previous periods, i.e. t-1, t-2, and so forth. However, one should not expect autocorrelation to go away due to such a specification, since it should be evident that the structure of the underlying economic situation will be far more complex than any econometric model can comprise. Hence purely econometric techniques such as Cochrane & Orcutt's (1949) method will have to be applied even if the model is correctly specified.

8.4 Summary.

In several models involving the actions of a producer population in relation to the behaviour of a population of consumers, the behaviour of both the producers and the consumers is assumed to be based on full information. It should be common knowledge that this is never the case in principle. Since most producers have some influence on the behaviour of the consumers, at least the actions of these producers will almost always be based on assumptions and assessments of what the consumers will do given a specific action at the hand of a particular producer. In particular the demand relation that every producer has to use in determining his price, will represent a forecast, not a known truth.

Let us consider the problem of measuring demand from the point of view of a member of a producer population. If demand is defined as the expressed desire of the consumers, then demand will never be observable to any producer until the producer has presented his offer to the market. Therefore, demand by nature is something that occurs as a consequence of, among other things, the actions of the producer. The producer will accordingly have to present his offer to the market without knowing the demand he will face. It follows that the demand is some future event and hence uncertain. Since the producer will have to set his price based on some notion as to how the consumers will react, he will have to measure or assess the demand based on limited information. In assuming a deterministic demand relation this fact is not always apprehended by analysts.

This treatise researches some of the implications of assuming that the producers face such uncertainties in the context of differentiated markets. Accordingly the fundamental objective of this treatise is to point out that every demand relation which any producer is faced with will always be a forecast, and therefore uncertain. The main objective of the treatise is to establish how this uncertainty is to be accounted for in an economic model of market behaviour in discrete and differentiated markets.

We have established that the producers view the demand from the consumers as uncertain, and that this uncertainty, by all producers involved, is measured through the use of choice probabilities. We have pointed out that the only relevant information which should be assumed to be in the possession of the producers relates to the principles of economic behaviour, and records of historical market data. Furthermore, we have established that among the numerous possible probability distributions that can be assigned to measure this uncertainty, only one particular class will be used by the producers. The probability distribution employed by all members of the producer population will be the unique distribution that satisfies the fundamental axioms of choice, and the information available to the producer.

This fact being established we proceed to show how in such a market of differentiated products a unique equilibrium point can be assured, assuming that price is the only strategic variable of the producers. Thus a basic theory of producer assessment of demand and market equilibrium in a discrete and differentiated market is presented.

We have shown above that our basic model, based only on price as a discriminating variable, can easily be extended to include any other information such as marketing surveys and marketing efforts, and any other relevant data. What data is relevant has to be decided by the analyst. However, the basic concept is the same. An unbiased measure for the industry uncertainty is constructed. As Shannon (1948) have shown in theorem S33 the only unbiased measure is H(p). Hence the relevant information has to be formulated as constraints to the problem of maximizing H(p). H(p) behaves so that it reaches its maximum when uncertainty is total. Subsequently H(p) is maximized subject to the relevant constraints since the minimal information *p* is what we are seeking.

We have shown that in a discrete and differentiated market the producers do face uncertainty as to which choices the consumers will make. Accordingly the producers have to make some sort of assessment of demand. Furthermore, we have shown that this assessment will have to be carried out in the way we have devised in this treatise. We have also shown that the total assessment problem includes an assessment of the market price equilibrium as well, and based on this price equilibrium the producers set their offers in the market. The inclusion of additional information such as information gained in a market survey, or the inclusion of marketing efforts, do not change this process, and the principal result. However, the inclusion of such additional information, as well as positive and different variable costs, do change the actual result. If the variable costs are assumed to be equal all through the industry, the equilibrium will be achieved when all producers have the same price and produce the same quantity. This follows from equation [E49]. If the variable costs, v_j $\forall j$, as well as the marketing effort, b_j $\forall j$, the quality, q_j $\forall j$, and the information gained by a market survey, $\theta_j \; \forall j$, is different for most firms, then there will still exist a price equilibrium, but now both prices and quantities will vary. Hence the inclusion of additional information makes the model more realistic and practical, but does not change the fundamental results.

Consequently we may conclude that we have arrived at a general theory of how relevant information should be comprised into a tenable model of economic behaviour in discrete and differentiated markets. This has to be considered a step forward in the microeconomic analysis of differentiated markets, since most models used to model such markets are not able to account for a broad range of relevant information in a reasonable way and at the same time achieve an equilibrium solution. In fact, reasonable models accounting for firm specific information, which simultaneously models economic human behaviour in a consistent way, have not so far been available when it comes to differentiated markets. To the extent that consistent models of producer behaviour in differentiated markets, accounting for a wide range of firm specific and industry related information, but not complete information, is of interest, the model presented in this treatise should satisfy a need.

Bibliography

Aczel, J. Lectures on Functional Equations and Their Applications. Academic Press, New York, 1966.

Adler, T. & Ben-Akiva, M. A Joint Frequency, Destination, and Mode Choice Model of Shopping Trips. Transportation Research Record 569, 1975.

Akin, J.S., D.K. Guilkey & R. Sickles. A Random Coefficient Probit Model with an Application to a Study of Migration. Journal of Econometrics, pp. 233-246, 1979.

Amemiya, T. On a Two-Step Estimation of Multivariate Logit Models. Journal of Econometrics, Vol. 8, No. 1, pp. 13-21, 1978.

Amemiya, T. & F.C. Nold. A Modified Logit Model. Review of Econ. Statist., May, pp. 255-257, 1975.

Amemiya, T. Qualitative Response Models: A Survey. Journal of Economic Literature, Vol. 19, pp. 1483-1536, 1981.

Anas, A. Discrete Choice Theory, Information Theory and Multinominal Logit Gravity Models. Transportation Research 17B, pp. 13-33, 1983.

Anas, A. Residential Location Markets and Urban Transportation. Academic Press, New York, 1982.

Anderson, S.P., A. de Palma, J.-F. Thisse. A representative consumer theory of the logit model. Int. Econ. Rev., 29, No. 3, pp. 461-466, 1988.

Anderson, S.P., A. de Palma, J.-F. Thisse. Discrete Choice Theory and Product Differentiation. MIT Press, 1992.

Arizono, I. & H. Ohta. Naval Research Logistics Quarterly, Vol. 33, pp. 251-260.

Banach, S. Sur les opérations dans les ensembles abstrais et leur application aux équations intégrals. Fundamenta Math., 3, pp. 133-181, 1922.

Bartel, A. The Migration Decision: What Role Does Job Mobility Play? Amer. Econ. Rev., Dec., pp. 775-786, 1979.

Basman, R.L., D.J. Slottje, K. Hayes, J.D. Johnson & D.J. Molina. The generalized Fechner-Thurstone direct utility function and some of its uses. Lecture Notes in Economics and Mathematical Systems, 316, Springer-Verlag, New York, 1988.

Batten, D. F. Entropy, information theory, and spatial input-output analysis. Umeå economic studies, Umeå universitet, Umeå 1981.

Beggs, S & N.S. Cardell. Choice of Smallest Car by Multi-Vechile Households and the Demand for Electric Vehicles. Transportation Research, Vol. 14, No. 5-6, pp. 389-404, 1980.

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

Ben-Akiva, M. & T. Watanatada. Application of a Continuous Spatial Choice Logit Model, in C. Manski & D. McFadden (eds.), Structural Analysis of Discrete Data with Econometric Applications, MIT Press, 1981.

Ben-Tal, A. The entropic penalty approach to stochastic programming. Mathematics of Operations Research, Vol. 10, No. 2, pp. 263-279, 1985.

Ben-Tal, A, A. Charnes & M. Teboulle. Entropic Means. Report CCS 568, University of Texas, 1987.

Ben-Tal, A. & M. Teboulle. Penalty functions and duality in stochastic programming via -divergence functionals. Mathematics of Operations Research, Vol. 12, No. 2, pp. 224-240, 1987.

Berkovec, J. & J. Rust. A Nested Logit Model of Automobile Holdings for One Vehicle Households. Transportation Research, 1985.

Blomquist, G. Value of Life Saving: Implications of Consumption Activity. Journal of Political Economy, June, pp. 540-458, 1979.

Boltzman, L. Wien. Ber. 63, Wien 1871.

Borch, K. H. The Economics of Uncertainty. Princeton University Press, 1968.

Bos, G.G.J. A Logistic Approach to the Demand for Priced Cars. Netherlands, Tilburg University Press, 1970.

Boskin, M.J. A Conditional Logit Model of Occupational Choice. Journal of Political Economy, Mar.-Apr., pp. 389-398, 1974.

Brottemsmo, J.A. & E. Moxnes. Automobile Choice in Norway: An Aggregate Level Study. Unpublished Paper, CMI, Bergen, 1991.

Buck, B. & V. A. MacAulay (Eds.). Maximum Entropy in Action. Oxford Science Publications, Oxford 1991.

Börsch-Supan, A. Econometric Analysis of Discrete Choice. Lecture Notes in Economic and Mathematical Systems, No. 296, Springer-Verlag, Heidelberg 1987.

Börsch-Supan, A. On the Compatibility of Nested Multinominal Logit Models with Utility Maximization. Journal of Econometrics, 1987.

Börsch-Supan, A. Tenure Choice and Housing Demand, in Stahl, K. & R. Struyk (eds.), U.S. and German Housing Markets: Comparative Economic Analysis, The Urban Institute Press, Washington D.C., 1985.

Börsch-Supan, A. Household Formation, Housing Prices, and Public Policy Impacts. Journal of Public Economics, Vol. 25, 1986.

Börsch-Supan, A. & J. Pitkin. On Discrete Choice Models of Housing Demand. Journal of Urban Economics, Vol. 21, 1987.

Caplin, A. S. & B. J. Nalebuff. Multi-Dimensional Product Differentiation and Price Competition. Oxford Economic Papers, 38 Suppl., pp. 129-145, 1986.

Caplin. A. S. & B. J. Nalebuff. Aggregation and Imperfect Competition on the

Existence of Equilibrium. Econometrica, 59, pp. 25-59, 1990.

Cameron, T.A. A Nested Logit Model of Energy Conservation Activity by Owners of Existing Single Family Dwellings. Review of Economics and Statistics, 1984.

Carnot, S., 1824 in Magie, W.F. The Second Law of Thermodynamics, New York, 1899.

Chamberlin, E. H. The Theory of Monopolistic Competition. Havard University Press, 1933.

Chow, G. Demand for Automobiles in the United States. Amsterdam, North-Holland, 1957.

Clark, C. The Greatest of a Finite Set of Random Variables. Operations Research, Mar-Apr., pp. 145-162, 1961.

Cochrane, D. & G.H. Orcutt. Application of least squares to relationships containing auto-correlated error terms. J. Am. Stat. Assoc., Vol. 44, 1949.

Cragg, J.G. & R.S. Uhler. The Demand for Automobiles. Can. J. Econ, Aug., pp. 386-406, 1970.

Cramer, J.S. Econometric applications of Maximum Likelihood methods, Cambridge University Press, Cambridge, 1986.

Da Vanzo, J. Does Unemployment Affect Migration?-Evidence from Micro Data. Rev. Econ. Statist., Nov., pp. 504-514, 1978.

Daganzo, C., F. Bouthelier & Y. Sheffi. Multinominal Probit and Qualitative Choice: A Computationally Efficient Algorithm. Transportation Science, 11, pp. 338-358, 1977.

Daganzo, C. Multinominal Probit: The Theory and Its Application to Demand Forecasting. New York, Academic Press, 1979.

Dantzig, G.B. Linear programming and extensions. Princeton University Press, Princeton, 1963.

David, J.M. & W.E. Legg. An Application of Multivariate Probit Analysis to the Demand for Housing: A Contribution to the Improvement of the Predictive Performance of Demand Theory, Preliminary Results. Amer. Statist. Assoc. Proceedings of the Bus. and Econ. Statist. Section, Aug., pp. 295-300, 1975.

Davis, H.T. The Theory of Econometrics. Principa Press, Bloomington 1941.

Debreu, G. Review of "Individual Choice Behavior" by R. Luce. Amer. Econ. Rev., Dec., pp. 186-188, 1960.

Diewert, W.E. Symmetry conditions for market demand functions. Rev. Econ. Stud., 47, pp. 595-601, 1980.

Diewert, W.E. Hicks Aggregation Theorem and the Existence of a Real Value Added Function. Mimemo, 1976.

Diewert, W.E. Applications of Duality Theory, in M. D. Intriligator & D.A. Kendrick (eds.) "Frontiers of Quantitative Economics II", Amsterdam 1974.

Dixit, A. & J. E. Stiglitz. Monopolistic Competition and Optimum Product Diversity. American Economic Review, 67, pp. 297-308, 1977.

Domencic, T.A. & D. McFadden. Urban Travel Demand. North-Holland, Amsterdam, 1975.

Dubin, J. & D. McFadden. An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. Econometrica, 52, 1984.

Duncan, G.M. & F. Stafford. Do Union Members Receive Compensating Wage Differentials? Amer. Econ. Rev., June, pp. 355-371, 1980.

Duncan, G.M. Formulation and Statistical Analysis of the Mixed, Continuous/Discrete Dependent variable Model in Classical Production Theory. Econometrica, May, pp. 839-852, 1980.

Eisen, M. Introduction to Mathematical Probability Theory. Prentice-Hall, New Jersey 1969.

Epple, D. Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products. Journal of Political Economy, Vol. 95, No. 1, 1987.

Eriksson, E.A. & P.O. Lindberg. Equilibria in Additive Random Utility Models. Royal Institute of Technology, Sweden, 1990.

Erlander, S. On the principle of monotone likelihood and log-linear models. Mathematical Programming Study, Vol. 25, pp.108-123, 1985.

Erlander, S. Accessibility, Entropy and the Distribution and Assignment of Traffic, Working Paper, Linkøping University, 1976.

Erlander, S. & T.E. Smith. General Representation Theorems for Efficient Population Behavior, Applied Mathematics and Computation, Vol. 36, pp. 173-217, 1990.

Erlander, S. & N. F. Stewart. The Gravity Model in Transportation Analysis - Theory and Extensions. VSP, Utrecht, 1990.

Evans, A.W. The Calibration of Trip Distribution Models with Exponential or Similar Cost Functions. Transportation Research, Vol. 5, pp. 15-38, 1971.

Fields, G.S. Place to Place Migration: Some New Evidence. Rev. Econ. Statist., Feb., pp. 21-32, 1979.

Flinn, C.J. & J.J. Heckman. New Methods for Analyzing Structural Models of Labor Force Dynamics. Journal of Econometrica, 1982.

Foley, D. K. A Statistical Equilibrium Theory of Markets. J. Economic Theory, 62, pp. 321-345, 1994.

Fridstrøm, L. "Stated preference" - eller økonomi som eksprimentalvitenskap. Sosialøkonomen nr.2, pp. 18-23, 1992.

Friedman, J. Oligopoly and the Theory of Games. North-Holland, 1977.

Friedman, J. Oligopoly Theory. Cambridge University Press, 1983.

Friedman, J. Game Theory with Applications in Economics. Oxford University Press, 1986.

Georgescu-Roegen, N. The Entropy Law and the Economic Process, Harvard University Press, 1971.

Gibbs, J.W. Elementary Principles in Statistical Mechanics. Longmans Green and Company, New York, 1902, 1928.

Giles, R. Mathematical Foundations of Thermodynamics. Pergamon Press, Oxford, 1964.

Grandy, W.T Jr. & L.H. Schick. Maximum Entropy and Bayesian Methods. Fundamental Theories of Physics, Kluwer, Dordrecht 1991.

Gronau, R. The Allocation of Time of Israeli Women. J. Polit. Econ., Aug., pp. 201-220, 1976.

Guilkey, D.K. & P. Schmidt. Some Small Sample Properties of Estimators and Test Statistics in the Multivariate Logit Model. J. Econometrics, April, pp. 33-42, 1979.

Gumbel, E.J. Bivariate Logistic Distributions. J. Amer. Statist. Assoc., June, pp. 335-349, 1961.

Gunderson, M. Retention of Trainees: A Study with Dichtomous Dependent Variables. J. Econometrics, May, pp. 79-93, 1974.

Hall, R. Turnover in the Labor Force. Brookings Papers on Economic Activity, Vol. 3, 1970.

Hanushek, E. A. & J.M. Quigley. An Explicit Model of Residential Mobility. Land Economics, Vol. 54, 1978.

Hartley, R.V.L. Transmission of Information. Bell System Technical Journal, VII, pp. 535-544, 1928.

Hausman, J. A. & D.A. Wise. A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences. Econometrica, Vol. 48, No. 2, pp. 403-426.

Hausman, J.A. Exact Consumer's Surplus and Deadweight Loss. American Economic Review, Vol. 71, No. 4, pp. 662-676, 1981.

Hausman, J.A. Individual Discount Rates and the Purchase and Utilization of Energy-Using Durables. Bell J. Econ. Manage. Sci., Spring, pp. 33-54, 1979.

Hausman, J. A. & D. McFadden. Specification Tests for the Multinominal Logit Model. Econometrica, Vol. 53, 1984.

Hausman, J.A. Specification Tests in Econometrics. Econometrica, Nov., pp. 1251-1272, 1978.

Heckman, J. The Incidential Parameters Problem and the Problems of Initial Conditions, in C. Manski & D. McFadden (eds.), Structural Analysis of Discrete Data with Econometric Applications, MIT Press, 1981.

Heckman, J. Statistical Models for Discrete Panel Data, in C. Manski & D. McFadden (eds.), Structural Analysis of Discrete Data with Econometric Applications, MIT Press, 1981.

Heckman, J. & R. Willis. A Beta Logistic Model for the Analysis of Sequential Labor Force Participation of Married Women. Journal of Political Economy, Vol. 85, 1977.

Heckman, J.J. & R.J. Willis. Estimation of a Stochastic Model of Reproduction: An Econometric Approach, in Household production and consumption, N.E. Terleckyj, New York, NBER, pp. 99-138, 1975.

Heckman, J.J. Simultaneous Equations Models with Continuous and Discrete Endogenous Variables and Structural Shifts, in Studies in nonlinear estimation, S.M. Goldfeld & R.E. Quandt (eds.), Cambridge, Ballinger, pp. 235-272, 1976.

Heckman, J. Dummy Endogenous Variables in a Simultaneous Equation System. Econometrica, Vol. 46, 1978.

Heckman, J. Sample Selection Bias as a Specification Error. Econometrica, Vol. 47, 1979.

Heckman, J. Statistical Models for Discrete panel Data, in C. Manski & D. McFadden (eds.), Structural Analysis of Discrete Data with Econometric Applications, MIT Press, 1981.

Henderson, J. M. & R.E. Quandt. Microeconomic Theory; A Mathematical Approach. 2.ed., McGraw-Hill, New York 1971.

Herniter, J.D. An Entropy Model of Brand Purchase Behavior. Journal of Marketing Research, Vol. X, pp. 361-75, 1973.

Hicks, J.R. A Revision of Demand Theory. London 1956.

Hicks, J.R. Value and Capital. Oxford 1939.

Hobson, A. A new theorem of information theory. Journal of Statistical Physics, 1, pp. 383-391, 1969.

Horowitz, A.R. & I. Horowitz. Entropy, markov processes and competition in the brewing industry. J. Industrial Econ., Vol. 16, pp. 196-211, 1968.

Horowitz, A.R. & I. Horowitz. The real and illusory virtues of entropy-based measures for business and economic analysis. pp. 121-136, 1976.

Hotelling, H. Stability in Competition. Economic Journal, 39, pp. 41-57, 1929.

Hughes, G.A. On the Estimation of Migration Equations. Mimeograph, Faculty of Economics, Cambridge University, 1980.

Jacobsen, S.K. On the solution of an entropy maximizing location model by submodularity. Research Report 6/88, Technical University of Denmark, 1988.

Jacquemin, A. The New Industrial Organization. Clarendon Press, Oxford, 1987.

Jaynes, E.T. Where Do We Stand on Maximum Entropy in The Maximum Entropy Formalism, Levine, R.D and Tribus, M (eds.), MIT 1978.

Jaynes, E.T. Information Theory and Statistical Mechanics. Physics Review, Vol. 106, No. 4, pp. 620-630, 1957.

Johnson, N. & S. Kotz. Distributions in Statistics; Continuous Multivariate Distributions. New York, 1973.

Jörnsten, K.O. & J.T. Lundgren. An Entropy-Based Modal Split Model. Transportation Research, Vol. 23B, No. 5, pp. 345-359, 1989.

Kapur, J.N., C.R. Bector & U. Kumar. A Generalization of the Entropy Model for Brand Purchase Behavior. Naval Research Logistics Quarterly, Vol. 31, pp. 183-198, 1984.

Kapur, J.N. Twenty-five years of maximum-entropy principle. J. Math. Phy. Sci., Vol., 17, No. 2, pp. 103-156, 1983.

Kohn, M.G., C.F. Manski & D.S. Mundel. An Empirical Investigation of Factors which Influence College-Going Behavior. Ann. Econ. Soc. Measure, Fall, pp. 391-420, 1976.

Kuhn, H. W. & A. W. Tucker. Nonlinear Programming. Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, ed. Neyman, J. University of California Press, 1951.

Kumar, U, J.N. Kapur & C.R. Bector. Renyi's Entropy Model for Brand Purchase Behavior. Journal of Information & Optimization Sciences, Vol. 6, pp. 223-242, 1985.

Kumar, U & C.R. Bector. Maximum Entropy Approach for Some Brand Switching Models with Different Mathematical Structures. Journal of Information & Optimization Sciences, Vol. 8, 1987, pp. 65-75.

Kumar, U. A One Parameter Entropy Model for Brand Purchase Behavior. Journal of Information & Optimization Sciences, Vol. 10, 1989, pp.519-534.

Lamond, B. & N.F. Stewart. Bregman's balancing method. Transportation Research, Vol. 15B, pp. 339-248, 1981.

Lave, C.A. The Demand for Urban Mass Transportation. Rev. Econ. Statist., Aug., pp. 320-323, 1970.

Lave, C & K. Train. A Disaggregate Model of Auto-Type Choice. Transportation Research, Vol. 13A, No. 1, pp. 1-9, 1979.

Lee, L.F. Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables. Int. Econ. Rev., June, pp. 415-434, 1978.

Leonardi, G in Thisse, J.F. & H.G. Zoller (eds.). Locational Analysis of Public Facilities. North-Holland, 1983.

Lerman, S. & C. Manski. On the Use of Simulated Frequencies to Approximate Choice Probabilities, in C. Manski & D. McFadden (eds.), Structural Analysis of Discrete Data with Econometric Applications, MIT Press, 1981.

Li, M.M. A Logit Model of Homeownership. Econometrica, Vol. 45, 1977.
Lindberg, P.O., E.A. Eriksson & L.G. Mattsson. Generalized Extreme Value Choice Models and Invariance Property. Working Paper, 1990.
Lundgren, J.T. Optimization Approaches to Travel Demand Modelling. Linkøping Studies in Science and Technology, Dissertations No. 207, Linkøping University, Linkøping 1989.
Maasoumi, E. Information Theory, in The New Palgrave Econometrics, J. Eatwell, M. Millgate and P. Newman (eds.), Norton, New York 1990.
Manski, C.F. Maximum Score Estimation of the Stochastic Utility Model of Choice. J. Econometrics, Aug., pp. 205-228, 1975.
Manski, C. & L. Sherman. An Empirical Analysis of Household Choice Among Motor Vehicles. Transportation Research, Vol. 14A, No. 5-6, pp. 349-366, 1980.
Mattsson, L.G. & P.O. Lindberg. A Note on the Convexity of Certain Homogeneous Functions Including Generators for Generalized Extreme Value Choice Models. Royal Institute of Technology, Sweden, 1984.
McEliece, R. J. The Theory of Information and Coding. A Mathematical Framework for Communication. Addison-Wesley, 1977.
McFadden, D. In C.F. Manski & D. McFadden (eds.), Structural Analysis of Discrete Data with Econometric Applications. MIT press, Cambridge, Mass., 1981.
McFadden, D. Quantal Choice Analysis: A Survey. Ann. Econ. Soc. Measure, Fall, 363-390, 1976.
McFadden, D. Modelling the Choice of Residential Location, in A. Karquist et al (eds.), Spatial Interaction Theory and Planning Models, Amsterdam, North-Holland, 1978.
McFadden, D. Conditional Logit Analysis of Qualitative Choice Behavior, in P. Zarembka (ed.), Frontiers in Econometrics, New York, Academic Press, 1973.
McFadden, D. On Independence, Structure and Simultaneity in Transportation Demand Analysis. Working Paper No. 7511, University of California, Berkeley, 1975.
McFadden, D., C. Winston & A. Børsch-Supan. Joint Estimation of Freight Transportation Decisions under Non-Random Sampling, in Daughty, A, Analytical Studies in Transport Economics, Cambridge University Press, 1985.
McFadden, D. Quantitative Methods for Analysing Travel Behavior of Individuals: Some Recent Developments, in Hensher, D & P. Stopher (eds.), Behavioral Travel Modeling, Croom Helm, London 1979.
McFadden, D. The Measurement of Urban Travel Demand. Journal of Public Economics, Vol. 3, pp. 303-328, 1974.
McGillivray, R.G. Binary Choice of Urban Transport Mode in the San Francisco Bay Region. Econometrica, Sept., pp. 827-848, 1972.
Medoff, J.L. Layoffs and Alternatives Under Trade Unions in U.S. Manufacturing.

Amer. Econ. Rev., June, pp. 380-395, 1979.
Moore, W.J., R.J. Newman & R.W. Thomas. Determinants of the Passage of Right-to-Work Laws: An Alternative Interpretation. J. Law Econ., Apr., pp. 197-311, 1974.
Morrison, S.A. & C. Winston. An Econometric Analysis of the Demand for Intercity Passenger Transportation, from T.E. Keeler, Research in Transportation Economics, Vol. 2, 1985.
Murtaugh, M. & H. Gladwin. A Hierarchal Decision-Process Model for Forecasting Automobile Type-Choice. Transportation Research, Vol. 14A, No. 5-6, 1980.
Nash, J. F. Noncooperative Games. Annals of Mathematics, 45, pp. 286-295, 1951.
Nickell, S. Education and Lifetime Patterns of Unemployment. J. Polit. Econ., Oct., pp. s117-131, 1979.
Oniescu, O. Éergie informationelle. Comptes Rendus de l'Académie des Sciences, Series A, CCLXIII, 1966.
Osten, S. Industrial Search for New Locations: An Empirical Analysis. Rev. Econ. Statist, May, pp. 288-292, 1979.
Parks, R.W. Determinants of Scrapping Rate for Postwar Vintage Automobiles. Econometrica, July, pp. 1099-1116, 1977.
Parks, R.W. On the Estimation of Multinominal Logit Models from Relative Frequency Data. J. Econometrics, Aug., pp. 293-304, 1980.
Parsons, D.O. Racial Trends in Male Labor Force Participation. Amer. Econ. Rev., Dec., pp. 911-920, 1980b.
Parsons, D.O. The Decline in Male Labor Force Participation. J. Polit. Econ., Feb., pp. 117-134, 1980a.
Pencavel, J.H. Market Work Decisions and Unemployment of Husbands and Wives in the Seattle and Denver Income Maintenance Experiments. Mimeographed, April, 1979.
Philippatos, G.C. & N. Gressis. Conditions of equivalence among E-V, SSD, and E-H portfolio selection criteria: The case for uniform, normal and lognormal distributions. Man. Sci., Vol. 21, No. 6, pp. 617-625, 1975.
Philippatos, G.C. & C.J.Wilson. Entropy, market risk, and the selection of efficient portfolios. Applied Economics, Vol. 4, pp. 209-220, 1972.
Powers, J.A. , L.C. Marsh, R.R. Huckfeldt et al. A Comparison of Logit, Probit and Discriminant Analysis in Predicting Family Size. Amer. Statist. Assoc. Proceedings of the Soc. Statist. Section, Aug., pp. 693-697, 1978.
Quandt, R.E. Estimation of Modal Splits. Transportation Research, pp. 41-50, 1968.
Radner, R. & L.S. Miller. Demand and Supply in U.S. Higher Education: A Progress Report. Amer. Econ. Rev., Papers and Proceedings, May, pp. 326-334, 1970.
Raj, S.J. Journal of Marketing, Vol. 49, 1985, pp. 53-59.

Rosen, H.s. & K.T. Rosen. Federal Taxes and Homeownership: Evidence form Time Series. J. Polit. Econ., Feb., pp. 59-75, 1980.

Roy, R. La Distribution durevenue entre les divers biens. Econometrica, July, pp. 205-225, 1947.

Sakai, Y. Revealed favorability, indirect utility, and direct utility. J. Econ. Theory, 14, pp. 113-129, 1977.

Savage, L.J. The foundations of Statistics. John Wiley and Sons, 1972.

Schmidt, P. & R.P. Strauss. The Prediction of Occupation Using Multiple Logit Models. Int. Econ. Rev., June, pp. 471-486, 1975.

Schmidt, P. & R.P. Strauss. The Effects of Unions on Earnings and earnings on Unions: A Mixed Logit Analysis. Int. Econ. Rev., Feb.,pp. 204-212, 1976.

Schrödinger, E. Statistical Thermodynamics, Cambridge University Press, 1948.

Shannon, C.E. A Mathematical Theory of Communication. The Bell System Technical Journal, Vol. XXVII, No. 3, pp. 379-456, 1948.

Shiller, B.R. & R.D. Weiss. The Impact of Private Pensions on Firm Attachment. Rev. Econ. Statist., Aug., pp. 369-380, 1979.

Shimony, A. Synthese, No. 63, pp. 35-53, 1985.

Silberman, J.I. & Durden, G.C. Determining Legislative Preferences an the Minimum Wage: An Economic Approach. J. Polit. Econ., Apr., pp. 317-329, 1976.

Small, K. A. The Scheduling of Consumer Activities: Work Trips. American Economic Review, Vol. 72, 1982.

Smith, T.E. A Cost-Efficiency Principle of Spatial Interaction Behaviour. Regional Science and Urban Economics, Vol. 8, pp. 313-333, 1978.

Smith, S.A. IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-4, No. 2, 1974, pp. 157-163.

Smith, J.P. The Distribution of Family Earnings. J. Polit. Econ., Oct., pp. s163-192, 1979.

Snickars, F. & J.W. Weibull. A Minimum Information Principle, Theory and Practice. Reg. Sci. and Urban Econ., 7, pp. 137-168, 1977.

Theil, H. Economics and Information Theory. North-Holland, 1967.

Thurstone, L. A Law of Comparative Judgment. Psych. Rev., 34, pp. 273-286, 1927.

Tobin, J. The Estimation of Relationships for Limited Dependent Variables. Econometrica, Jan., pp. 24-36, 1958.

Tollefson, J.O. & J. A. Pichler. A Comment on "Right-to-Work" Laws: A Suggested Economic Rationale. J. Law Econ., Apr., pp. 193-196, 1974.

Train, K. Qualitative Choice Analysis. MIT Press, Cambridge 1986.

Train, K. A Structural Logit Model of Auto Ownership and Mode Choice. Review of Economic Studies, Vol. XLVII, pp. 357-370, 1980.

Train, K. & D. McFadden. The Goods/Leisure Tradeoff and Disaggregate Work Trip

Mode Choice Models. Transportation Research, Vol. 12, pp. 349-353, 1978.
Uhler, R.S. The demand for housing: An Inverse Probability Approach. Rev. Econ. Statist., Feb., pp. 129-234, 1968.
Warner, S.L. Stochastic choice mode in urban travel: A study in binary choice. Evanston, Northwestern University Press, 1962.
Warren, R.S. & R.P. Strauss. A Mixed Logit Model of The Relationship Between Unionization and Right-to-Work Legislation. J. Polit. Econ., June, pp. 648-655, 1979.
Watson, P.L. & R.B. Westin. Transferability of Disaggregate Mode Choice Models. Reg. Sci. and Urban Econ., May, pp. 227-249, 1975.
Weddepohl, H. N. Axiomatic Choice Models. Rotterdam University Press, 1970.
Westin, R.B. Predictions from Binary Choice Models. J. Econometrics, May, pp. 227-249, 1975.
Wiener, N. (1948), Cybernetics. 2nd. ed., New York 1961.
Wilensky, G.R. and L.F. Rossiter. OLS and Logit Estimation in a Physician Location Study. Amer. Statist. Assoc. Proceedings of the Soc. Statist. Section, Aug., pp. 260-265, 1978.
Williams, H. On the Formation of Travel Demand Models and Economic Evaluation Measures of User Benefit. Environment Planning A.9, 1977.
Willis, R.J. & S. Rosen. Education and Self-Selection. J. Polit. Econ., Oct., pp. S7-36, 1979.
Winston, C. Conceptual Developments in the Economics of Transportation: An Interpretive Survey. Journal of Economic Literature, Vol. 23, 1985.
Wu, D.M. An Empirical Analysis of Household Durable Goods Expenditure. Econometrica, Oct., pp. 761-780, 1965.
Zellner, A. Bayesian Methods and Entropy in Economics and Econometrics, in Grady, W.T. & L.H. Schick, Maximum Entropy and Bayesian Methods, Kluwer, Dordrecht 1991.
Zipf, G.K. Human behavior and the principle of least effort, Addison-Wesley, Cambridge MA, 1949.