Food Retailing and Consumer Behaviour in Waterloo, Ontario

Wayne Dennis Hilmo
Wilfrid Laurier University

Follow this and additional works at: http://scholars.wlu.ca/etd
Part of the Human Geography Commons

Recommended Citation
Hilmo, Wayne Dennis, "Food Retailing and Consumer Behaviour in Waterloo, Ontario" (1971). Theses and Dissertations (Comprehensive). 1517.
http://scholars.wlu.ca/etd/1517

This Thesis is brought to you for free and open access by Scholars Commons @ Laurier. It has been accepted for inclusion in Theses and Dissertations (Comprehensive) by an authorized administrator of Scholars Commons @ Laurier. For more information, please contact scholarscommons@wlu.ca.
FOOD RETAILING AND CONSUMER BEHAVIOUR

IN WATERLOO, ONTARIO.

by

Wayne Dennis Hilmo

Submitted in partial fulfillment of the requirements for the M.A. Degree in Geography

Department of Geography
Waterloo Lutheran University
Waterloo, Ontario

1971
TABLE OF CONTENTS

LIST OF TABLES .......................................................... i
LIST OF MAPS ............................................................. ii

Chapter

I. INTRODUCTION ..................................................... 1
   Reasons for the Study ............................................. 2
   Interurban and Intraurban Studies ............................. 2
   The Behaviour of the Entrepreneur and the Consumer .... 3
   Marketing Geography, Central Places and Individual Behaviour ........................................... 4
   The Prediction of Consumer Spatial Choice ............... 5
   Groceries as a Study Commodity .............................. 7
   The Structure of the Study ...................................... 9

II. REVIEW OF THE LITERATURE ................................. 11
   Central Place and Location Theories ....................... 12
   Central Place Theory ............................................. 12
   Location Theory .................................................. 14
   Recent Research on Consumer Spatial Choice ............ 15
      Baumol and Ide ................................................ 15
      Huff ............................................................. 18
      Marble ........................................................ 20
      Nystuen ....................................................... 21
      Rushton ....................................................... 23
      Clark .......................................................... 25
      Golledge ...................................................... 26
Chapter

III. THE CONCEPTUAL MODEL AND ITS PROBLEMS ........ 30

Problems of the Distance Variable ............ 30
Problems of the Utility Variable ............ 33
Problems in Relating the Two Variables .... 36
Summary ........................................ 37

IV. THE BASIC CONSUMER ALLOCATION MODEL ....... 38

The Basic Structure of the Consumer Allocation Model ........ 38
Results of the Basic Consumer Allocation Model .................. 42
Summary ........................................ 42

V. ALTERATIONS TO THE BASIC ALLOCATION MODEL ... 44

Transformations of the Original Utility Values .................. 44
An Alternate Means of Deriving Store Utilities .............. 45
An Alternate Measure of Distance ............... 46
An Alternate Relationship Between the Two Variables .... 47
Summary ........................................ 48

VI. AN ANALYSIS OF THE RESULTS OF THE MODEL .... 50

Maps of Actual and Predicted Consumer Choices .................. 50
The Correct and Incorrect Allocations Of the Model .................. 56
The Consumers Allocated to the Various Stores .... 58
Summary ........................................ 63
LIST OF TABLES

Table

1. Consumers Correctly Allocated to Each Store ..... 58
2. Results of the Tukey's Q Test ....................... 61
LIST OF FIGURES

Map

1. The Location of the Stores .......................... 40
2. Stores Chosen by the Consumers ....................... 52
3. Stores Predicted by the Model ........................ 55
CHAPTER I

INTRODUCTION

The main objective of this study is to develop a computer model to predict the food stores which individual consumers choose for the major part of their grocery purchases. It is worthwhile to develop such a predictive model for two reasons. In the first place, it is useful to be able to predict consumer store choices for purposes of retail planning. Secondly, a working predictive model illuminates the main bases of the consumer spatial choice process and thus adds to our understanding of urban spatial organization.

A basic idea of model building is that a model is a simplification of reality in which a small number of variables may serve with great economy of data to capture and make clear the essence of the processes involved. As Haggett suggests, "Successful models are those which manage a considerable amount of simplification without introducing extraneous noise." 1 If it is possible to design a model which predicts consumer store choice on the basis of a few simple variables, the resulting model is of greater value than a model in which so many variables are introduced that the result approaches a duplication of reality rather than an enlightening simplification and abstraction.

Reasons for the Study.

This study was undertaken because of dissatisfaction with work already published on retail structure and the consumer spatial choice process. It is an attempt to fill several gaps that the author feels exist in existing research.

Interurban and Intraurban Studies.

Although much work has been done by central place theorists to develop and refine models dealing with retail activities in the interurban case, little work has been done to apply these models to the intraurban case or to develop alternate models which effectively deal with intraurban retail structure. As will be demonstrated in the next chapter, the published work on intraurban consumer behavior is either of a very preliminary nature or is almost entirely theoretical. Various researchers have considered the nature of the consumer spatial choice process in intraurban areas, but none of them have produced an operational model to predict the actual store choices made by individual consumers. They have considered the nature of the consumer spatial allocation process at length, but they have not brought their conceptual models to the stage where these models could be utilized to predict the store at which a particular consumer shops. Only Rushton has successfully operationalized a predictive model of consumer behavior and his work was not done in the intraurban case.

---

In this study, the development of a conceptual model of the consumer spatial choice process is viewed as only a first step in the development of a working predictive model. The conceptual model must be translated into an operational framework and then tested against reality.

The Behaviour of the Entrepreneur and the Consumer.

There are two groups of individuals whose behaviour is of interest both in marketing geography and in central place studies. These groups are the entrepreneurs and the consumers. Of the two, the behaviour of the entrepreneur has received the most attention from researchers.

In marketing geography, much research has been carried out in an effort to aid the entrepreneur in choosing profitable locations for stores. The behaviour of the consumer is considered in this research, but only in the aggregate. There is no concern with the store that is chosen by an individual consumer in a particular spatial situation, but only with the aggregate choices that collectively determine the success or failure of a store.

In central place theory, the emphasis is also on the behaviour of the entrepreneur: what effect does the behaviour of the entrepreneur have on the development of the central place hierarchy? When the purchasing power of consumers in the area tributary to a central place exceeds the threshold value necessary to support a store selling a good, an entrepreneur will open a store to sell that good. Conversely, when the purchasing power in the tributary area falls below the threshold for the good, the entrepreneur will

---

no longer offer that good for sale. The assumption is made in central place theory that consumers travel to the nearest store or central place offering the good they require.

In both marketing geography and central place theory, the behaviour of the individual consumer has been a neglected area of study. Because of this neglect, the study of the consumer spatial choice process could well yield some valuable insights into the spatial nature of retailing in intra-urban areas. Also, such research is of interest because the survival of a store at a particular location is a joint function of the entrepreneur's original locational decision and of the store choices of individual consumers.

Marketing Geography, Central Places and Individual Behaviour.

Central place theory and marketing geography are two areas of geography that are concerned with the spatial aspects of retailing. In a sense, they occupy two ends of a spectrum of possible approaches to the geography of retailing. At one end of this spectrum lies the empirical field of marketing geography, which is concerned with the solution of practical problems in retailing. Its researchers are concerned with consumer behaviour only in the aggregate. They do not need to know which consumers shop at a given store as long as they are able to make certain that the stores with which they deal are located in such a way that they draw the maximum numbers of consumers. At the other end of this spectrum is the theoretical study of central place systems. The theory assumes that consumers patronize the nearest central place which offers the good that they require. Therefore, the behaviour of the individual consumer is of no real concern, since all consumers behave the same way.
There is a need for a third approach to the geography of retailing. This approach would be concerned with examining the behaviour of the individual consumer. It departs from marketing geography in that it considers the consumer as an individual instead of consumers in the aggregate. It also departs from central place research in that it does not assume that all consumers behave in the same way. Many consumers are dominated by the friction of distance and choose the nearest store for their grocery purchases, but others are more strongly influenced by the character of different stores and will travel greater distances to reach the store of their choice. There are even some consumers whose store choice appears to be almost irrational.

The focus of this study will be on the behaviour of the individual consumer. Such a focus is appropriate to the current view of geography as a subject concerned with the development of theory explaining the location of various phenomena in space and the interrelationship of these phenomena in spatial systems. Only by examining individual components of the retail subsystem will it be possible to fully understand its nature and operation. This study looks at the individual consumer as one component in the retail subsystem. A full understanding of each of the subsystems -- retail, residential and employment -- which comprise the urban system is a necessary step on the path towards a full comprehension of the entire system.

The Prediction of Consumer Spatial Choice.

It was originally felt by this researcher that an adequate model to predict the store choices of individual consumers would have to consider a large number of variables if it was to have any success in predicting such choices. Even casual observation indicates that consumers do not
invariably go to the nearest store to shop for their groceries. When the relevant literature was examined in detail, it was discovered that discussions of the problem of predicting the stores chosen by consumers usually considered two basic variables. One of these variables was some measure of the cost of overcoming the friction of distance and the other was a variable giving an indication of the utility that different stores have for consumers.

The utility value is a surrogate measure indicating the relative worth of different stores to consumers. It is a summary of a large number of variables which collectively determine how much consumers like a particular store. Some of these variables are congestion, ease of parking, quality of service, quality of meat or produce and the number of different products for sale. The utility value serves as a substitute for this group of reciprocally associated variables. A major problem of this study is to develop a valid measure of a store's utility. Chapter III will consider some problems with the use of a utility variable and will suggest which of the ways that have been utilized to measure it will be most fruitful.

The observation that the published literature on consumer store choice suggests a simple, two-variable predictive model led to a reconsideration of the need to incorporate a larger number of perceptual and behavioural variables. Since no researcher has published a predictive model of consumer store choice in an intraurban area, no conclusions may be drawn as to the value of this simple conceptual model. In this study, the simple conceptual model will be made operational and some conclusions will be drawn regarding its value as a predictive model and regarding the
need and desirability of utilizing further variables in later research.

**Groceries as a Study Commodity.**

Consumer goods can conveniently be divided into three groups on the basis of frequency of purchase and the willingness of consumers to travel to purchase the good. These three groups are convenience goods, shopping goods and specialty goods. Convenience goods are purchased regularly and often; they are goods that are in daily usage. As a result, the consumer is not usually willing to travel far to acquire them. Therefore, the friction of distance is likely to be a very significant variable in the examination of the consumer spatial decision process with regard to such goods. Convenience goods stores such as grocery stores tend to locate in such a way that they are separated from similar stores and can achieve dominance in market areas by taking advantage of the reluctance of consumers to travel very far to reach such stores.

Shopping goods are purchased on a less frequent basis than convenience goods. Since they are usually not standardized in either price or quality, such shopping goods as clothing, shoes and furniture are generally purchased after a period of comparison shopping or 'shopping around'. Because they are required less frequently than convenience goods, the consumer is willing to travel a considerable distance to buy them and to engage in comparison shopping. Stores selling shopping goods often tend to cluster together because of entrepreneurial desire to take advantage of consumer comparison shopping for such goods. Thus, women's clothing stores are often located in fairly close juxtaposition to one another.

Specialty goods are unique or rare items for whose purchase the consumer may be willing to almost totally ignore distance. For example, one painting cannot effectively be substituted for another; one cannot
shop around to get the best price on a unique item. For this reason, consumers will travel great distances if necessary to obtain specialty goods and the stores which sell such items may locate almost anywhere in the urban area without reducing their volume of business.

Since convenience goods are purchased more frequently than either shopping or specialty goods, the consumer is more likely to have planned his shopping trips for such goods in such a way as to minimize the travel costs involved. Also, because the consumer shops more often at convenience goods stores, he is more likely to be able to recall details of his shopping trips to such stores. This fact is significant since much of the data for this study will be gathered from consumers by questionnaires. These are two important reasons for selecting groceries as the study commodity in the examination of the consumer's spatial choice process.

Rushton suggests that:

"The respondent, moreover, is likely to know well the places where this commodity is sold, and since such a large proportion of his expenditures is allocated for this commodity, he is likely to consciously choose and plan his consumption pattern for it." 4

This quotation indicates another powerful reason for the use of groceries as the study commodity. Groceries represent a large proportion of all retail transactions. In the Waterloo study area, grocery sales represent 30 per cent of all retail sales. The understanding of the consumer's choice process with regard to the choice of store for purchases of such

---

4 Rushton, p. 12.
magnitude contributes more to our comprehension of the entire retail subsystem than a commodity representing less of total consumer expenditures.

Finally, even though one purpose of the study is to examine consumer behaviour in a geographical rather than a marketing context, a study of the consumer's spatial choice process in regard to choosing a grocery store could also have practical significance because "the supermarket operator stands to lose heavily if he cannot tempt enough people into his shop and induce them to buy." 5

The Structure of the Study.

The remainder of this study is composed of six chapters. The following is a brief consideration of the basic form and purpose of each of these chapters.

Chapter II. The second chapter is a review of the literature of greatest relevance to the problem of predicting the grocery stores which consumers choose for their major food purchases. The object of the chapter is not merely to summarize the articles, which are readily available, but rather, to identify the contribution that each researcher has made towards the isolation of the important variables in the consumer spatial choice process and towards the explication of the relationships between these variables.

Chapter III. Examination of the literature reveals that most of the researchers in this area share a common conceptual model. In Chapter III, this conceptual model is discussed and elaborated and the particular problems involved in operationalizing each of the main variables and

---

in expressing the relationship between these variables will be considered.

Chapter IV. In Chapter IV, a computer model is developed which allocates consumers to stores on the basis of the significant variables identified in Chapter II. To do this, utility values for each of the stores in the Kitchener-Waterloo area are calculated from 240 consumer interviews done at random in Waterloo, Ontario in June of 1971. These utility values are used in combination with the relative location of consumers to predict the stores that consumers choose for their major grocery purchases.

Chapter V. To improve the level of prediction, a number of alterations to the basic consumer allocation model are considered: alterations to the utility values, alterations to the distance variable and alterations to the formula relating the two main variables. The results of these alterations are discussed.

Chapter VI. The results of the best version of the consumer allocation model are analysed in this chapter. Two "desire-line" maps showing where consumers actually went to shop for their groceries and where the model predicted that they would go are compared and contrasted. In addition, an attempt is made to isolate the important differences between the consumers that the model was able to correctly predict and those that it was not able to correctly predict and between the consumers allocated to each of the stores. These differences suggest some possible ways of improving the level of prediction of this consumer allocation model in future research.

Chapter VII. In the concluding chapter, the study is briefly reviewed and some conclusions are drawn as to the validity of this approach to the understanding of the consumer's spatial choice process.
CHAPTER II

Review of the Literature

It was suggested in Chapter I that one way of approaching the development of a model to predict the stores which consumers choose for their major grocery purchases would be to examine the relevant geographical literature and isolate the variables that other researchers have considered of value for predictive purposes. This chapter will consider the published work of several researchers and how their ideas relate to the problem of this study.

In spite of the seeming variety of approaches to the prediction of consumer spatial choice, only a few basic variables appear in the studies that have been made and they are recognized by all of the researchers as being significant. Let us, then, examine the major geographical works dealing with the consumer's choice of store and extract from them the conceptual model on which they are based.

Accordingly, this chapter will consider the ideas of the several researchers who have worked on the problem of predicting the stores that individual consumers choose for their grocery shopping. The articles will not be summarized but will be critically examined for their specific contributions to the isolation of variables critical to a predictive model and to the explication of the nature of the relationship between these variables.
Central Place and Location Theories.

Two bodies of geographical theory contribute to an understanding of the consumer's spatial choice and its prediction. They are: central place theory and location theory.

Central Place Theory.

In most central place studies the emphasis has been on isolating the hierarchy of central places and then examining the functioning of the central place system; the behaviour of the consumer has received very little attention. Such limited consideration of the role of the consumer in the development and maintenance of the central place landscape has been recognized by some central place researchers as a major impediment to further progress in their work. 6

Most central place studies are undertaken from the viewpoint of the entrepreneur, with the spatial behaviour of the consumer appearing in the form of a postulate or assumption which is rarely tested in any explicit fashion. Basically the assumption is made that the behaviour of the consumer is solely a function of distance. As Rushton, Golledge and Clark suggest "the principal postulate of central place theory is that consumer expenditure patterns conform to a particular lawfulness, namely that consumers visit the nearest place in which a good is offered." 7

---


Christaller identified four basic factors which he felt influenced the travel behaviour of consumers. These factors were "the size and importance of the central place, the price willingness of the purchaser, the subjective economic distance, and the type, quality, and price of a good." However, because his emphasis was not on the consumer and because of the need to make some simplifying assumptions to facilitate the development of his central place theory, he did not incorporate all these factors into his analysis.

As Clark and Rushton state:

"Christaller, particularly, created much confusion by indulging in an extended discussion of how consumers were likely to behave in space, and then following this with a theoretical derivation based on the premise that consumers always patronize the nearest place which offers the required good." 9

Distance in central place theory is therefore the variable that determines the store that the consumer chooses.

---


Location Theory.

The most developed aspects of location theory have been concerned with the productive sectors of the economy. Location theorists have tended to ignore the influence of the consumer in their analyses; they have been content to sweep away the behaviour of the consumer with various simplifying assumptions in much the same way as central place theorists.

Losch, like Christaller, assumes that the consumer is dominated in his spatial choice behaviour by the friction of distance. In the development of his market area concept, he assumes that all consumers possess identical characteristics and therefore have identical demand curves, decreasing regularly with distance. 10

Isard observes that consumers generally behave in a spatially rational fashion, but he also notes that distance influences the behaviour of different individuals in different ways. 11 He introduces the concept of 'space preferences', which also appears in later formulations of the problem of predicting consumer spatial choice behaviour. 12 According to

---


12 Isard, W., pp. 81-88.
this concept, space is perceived by different people in different ways, and so the effect of the friction of distance may vary. Some people react to distance in a highly negative way and tend to keep their movements at a minimum. They fit very well into the view of the consumer as a distance minimizer or as spatially rational man. On the other hand, there are other people for whom the friction of distance does not seem to have much meaning and who regularly travel greater distances than is necessary. Their choice of store is conditioned by other factors.

In location theory the consumer is viewed as a perfectly rational being who has complete knowledge of the choices open to him and the costs that are associated with choosing each alternative. Since he is fully aware of his purchase opportunities, his choice of store may be conceptualized as a completely deterministic one. He invariably chooses that store at which he is able to minimize the costs associated with shopping. As with the central place construct, the consumer's decision is seen to be dominated by the friction of distance, which is the major cost in choosing an alternative.

Recent Research on Consumer Spatial Choice.

Several other researchers have been specifically concerned with the consumer's spatial choice process. We will first consider a general treatment of the problem of predicting consumer store choices by two economists and then several more specific works by geographers.

**Baumol and Ide.** These two economists reject one of the limiting features of the location theoretic treatment of consumer spatial behaviour.

---

They reject the assumption that the consumer possesses complete knowledge of all possible shopping opportunities, the relative location of each opportunity and the costs associated with choosing each alternative. They feel that consumers are not even aware of all their shopping opportunities and probably never have a complete and perfect knowledge of the costs associated with shopping at each store. For this reason they feel that one cannot adequately treat the consumer spatial decision process deterministically and they use a probabilistic approach.

Baumol and Ide suggest that consumers gain their knowledge of their shopping environment by making observations concerning the size of stores and the nature of different classes of stores. These observations give the consumer a knowledge of the variety of goods which he may expect to find in a store of a certain size, but this knowledge is incomplete. Nevertheless, the knowledge derived from such observations enables the consumer to derive an estimate of his probability of success in finding a particular good at a given store. Baumol and Ide call these estimates "subjective probabilities" and represent them by the term \( p(N) \). They assume that this probability is a function of \( N \), the number of different items stocked by a store.

They suggest that a consumer's choice of a store is influenced by three other factors in addition to the subjective probability of satisfaction. These three factors are all costs that are associated with shopping at a particular store. The first of these factors is the cost of travelling to the store, which they consider to be directly proportional to the distance which has to be travelled. Their second factor is the cost incurred while shopping due to crowding and congestion. They feel that this cost increases as the square root of the number of different items stocked by a store. Their final
factor is an opportunity cost, which they describe as the cost to the consumer of making a particular shopping trip instead of undertaking some other activity. They represent these three cost factors as $C_dD$, $C_n\sqrt{N}$ and $C_i$ respectively and view the total cost of making a shopping trip as:

$$TC = C_dD + C_n\sqrt{N} + C_i$$

This total cost is related to the subjective probability of satisfaction in the following formula:

$$F(N,D) = w p(N) - v \left( C_dD + C_n\sqrt{N} + C_i \right)$$

where $F(N,D)$ is the consumer's expected benefit from shopping at a particular store and $w$ and $v$ are subjective weights assigned by the consumer. They do not go into the derivation of these weightings.

This formula would be evaluated for each consumer in relation to each store in an area. The total cost of shopping at a given store would be subtracted from the subjective probability of satisfaction for that store. When this calculation had been made for all stores, the consumer would be allocated to that store at which his expected benefit was highest.

In the work of Baumol and Ide, we can see the first version of the conceptual model which is the basis of most of the work done on the consumer spatial decision problem to the present time. The consumer's choice may be considered in a two attribute model which relates a characteristic of the stores to an attribute indicating the cost to the consumer of shopping at different stores. The subjective probability of satisfaction is an indication of the utility that different stores have for consumers. It indicates that consumers prefer to shop at some stores instead of others because of the different variety of goods carried at each store. The total cost of shopping is largely a measure of the cost of overcoming the friction of distance, which they consider to be the most important component of total cost.
They do not attempt to operationalize their model in order to predict the stores which consumers choose for any particular class of goods, but their discussion is of value because it is an early attempt to explicitly examine the consumer spatial choice problem. This had not previously been done.

Huff. Huff is a market analyst who has published a series of articles which attempt to make the conceptual model of Baumol and Ide operational. He was specifically concerned with building a model of consumer spatial behaviour which would be probabilistic in nature; like Baumol and Ide, he felt that the consumer makes his decision under conditions of uncertainty and his choice could not easily, as a result, be set in a deterministic framework.

As with the model of Baumol and Ide, Huff’s model basically involves only two variables. He uses travel time from home to store or shopping centre as a measure of the friction of distance, while Baumol and Ide utilize a cost function of which distance is a dominant factor. His other variable is the size of the store or shopping centre which he uses as a surrogate measure for the utility of the store or centre. These two

---


variables are related in the context of a modified gravity model in which the probability, \( P_{ij} \), of a consumer at location \( i \) choosing the store at location \( j \) is directly proportional to the size of the store, \( S_j \), and inversely proportional to the time distance between the store and the consumer, \( T_{ij} \). The size of the store is thus the 'mass' in the gravity model. His formulation is as follows:

\[
P_{ij} = \frac{S_j}{T_{ij}^e} \left/ \sum \left( \frac{S_j}{T_{ij}^e} \right) \right.
\]

where \( e \) is an empirically determined exponent to which distance is raised to account for the magnitude of its influence on the shopping trip.

Unlike Baumol and Ide, Huff sets his key variables in a workable framework detailing the relationships between them. For this reason, Huff's model is operationally the most developed of any to be considered so far and could be utilized in an attempt to predict consumer choice of store. The fact that Huff has not used his model to predict the store choices of individual consumers may be due to his greater concern with the delimitation and analysis of market areas for retail planning or may indicate that his later research has been done privately for a large retailer and has not been published.

The greatest value of Huff's work for the present study lies in the manner in which he clearly sets forth the elements of a consumer allocation model and suggests a means of relating these variables in an operational framework.
Marble suggested that it might be feasible to operationalize their model using game theory. He later went on to follow his own suggestion and presented a game theoretic approach to the analysis and prediction of consumer spatial choice. He considered that the potential returns in a spatial choice situation were a function of the relative location of the decision maker and the ease with which the decision maker is able to undertake movements within the system. But, like Isard, he makes the point that the space preferences of individuals differ greatly and that the influence of this distance factor will therefore not be constant.


16 Ibid., p. 37.


18 Ibid., p. 37.
Marble then sets up a game theoretic matrix in which the consumer is assumed to be operating under conditions of risk, not knowing what state of nature (that is, the goods offered and the price levels at each store) may prevail. The payoffs in the game matrix represent the net benefit to the consumer when a certain store is chosen and a given state of nature is found to prevail. The consumer chooses his store in such a way as to maximize the net benefit which he will receive no matter what state of nature actually exists.

This treatment shows once again that the consumer's spatial choice problem essentially involves a distance function and a utility function. His distance function is a measure of the costs incurred in overcoming the friction of distance and his utility function is surrogated by the number of different goods which a store sells and the price level of the store. Setting the problem in a game theoretic framework may help to clarify the basic nature of the decision which confronts the consumer, but it does not really help in the development of a model to predict consumer choice of store because of the difficulty of arriving at reasonable estimates of the payoffs for every combination in a multiple choice and multiple state of nature game. This conceptualization does, however, reinforce the emerging picture of the consumer's spatial choice as resulting from a trade off in which the disutilities involved in overcoming the friction of distance are balanced against the utility of shopping at a particular store.

Nystuen. The central concern of Nystuen's article is the nature

---

of and the reasons for the multiple-purpose shopping trip. In the process of introducing his simulation model of such trips, he comments on the consumer spatial decision problem from the consumer's point of view. "When the economic agent is not the firm, but the customer who decides how he is to obtain products, a reasonable assumption is that the customer attempts to maximize his net travel return." It is Nystuen's contention that a range of utilities are available at various retail stores and that net travel return is a function of the satisfaction these utilities offer the consumer, minus the costs associated with travel to a store. The costs of travelling to a store arise from the friction of distance, the mode of travel and the level of congestion. In addition, he incorporates an opportunity cost into his cost function and suggests that "the traveler must also deduct for loss of satisfaction which would have been available had he remained at home."  

Since distance is the most important component of his cost function, Nystuen's conceptualization of net travel return as a function of store utility less the costs of travel and lost opportunity reduces quite readily to the two-attribute model relating store utility and distance. However, he does not specify the relationship between his two key variables in such a way that the model can be readily made operational. He is content

20 Nystuen, p. 57.

21 Nystuen, p. 56.
to say that "To induce a trip and to have a trip continue, the anticipated net return must be positive."²² He does not seriously consider the problems of how to derive utility values for stores, of how to incorporate the friction of distance or of how these variables are to be related in a predictive format. One reason may be that his work is primarily directed at the simulation of the number and type of stores chosen on the multiple-purpose shopping trip and the temporal extent of such trips. He sets up a detailed conceptual model of the consumer spatial choice process, but unfortunately he does not carry on to develop this conceptual model into a predictive model of consumer store choice.

Rushton.²³ Although Rushton is concerned with the allocation of rural shoppers to shopping towns rather than the intraurban allocation problem, his work was the first to be examined which has the prediction of the consumer's choice of food store as its specific concern. He indicates in his study the intuitive belief that people who are in similar spatial situations tend to allocate their expenditures in similar ways.

²² Nystuen, p. 56.

He also considers two primary variables in his analysis. In the first place, he considers the distance between a consumer and a town to be a "prime factor in determining whether or not that town will be patronized by a particular customer." He uses town size as his second variable and views it as the measurable characteristic of a town which best represents its utility to consumers.

He relates his variables by computing an "attractiveness index" for each of the possible size and distance combinations. This index gives an indication of that size-distance combination's attraction for the consumers.

These attractiveness indices are then used to develop a series of indifference curves which connect all size and distance combinations which are of the same value to the consumer or to which the consumer is indifferent. Rushton suggests that "an indifference curve presents a graphic picture of consumer preferences between the advantages of the size of the chosen town against the disadvantage of the distance traveled to reach the town." 25

All towns within 25 miles of a consumer are plotted on an indifference curve with town size on the vertical axis and distance on the horizontal axis. The model then allocates the consumer's major grocery purchases to that town which lies on the highest indifference curve.

24 Rushton, p. 21.
25 Rushton, p. 37.
Using this technique, Rushton was able to correctly predict 60% of the towns chosen by consumers for their major grocery purchases. Since this is the only work encountered which actually utilized a model for predictive purposes, a level of prediction of 60% will serve as a goal for the predictive model to be developed later in this study.

Clark. 26 Clark casts his study as a test of some of the basic postulated of central place theory. Working in Christchurch, New Zealand, he interviewed a random sample of 521 consumers and asked each to indicate his source of supply of three convenience goods (groceries, meats and vegetables), three shopping goods (clothing, furniture and appliances) and three services (dry cleaning, beauty shops and banking).

He first sought to test the hypothesis that consumers in urban areas travel to the nearest centre for their purchases of groceries, meat and vegetables. He tested this postulate in two ways. One way was to calculate the percentage of consumers who went to the nearest centre and then compare that percentage with the percentage who went to alternate centres. The other way was to compare the distances that would be involved in travelling to the nearest centre with the distance to the centres that were actually chosen.

In this test he utilized the notion of a zone of indifference. If two stores differ by less than 220 yards in their distance from a consumer,

---

he assumes that the stores are insufficiently different in distance for
the consumer to choose one or the other on the basis of distance alone.
With this allowance, he found that 63% patronized the nearest store for
their groceries. His conclusion was that the nearest store hypothesis
should be rejected and that an alternate hypothesis should be put forward.
However, this researcher would conclude that if as many as 63% of consumers
choose the nearest store, then Clark has shown that the friction of distance
is a major variable to be incorporated into any model which attempts to
predict consumer choice of food store.

Clark suggests replacing the nearest centre hypothesis with an alter­
native approach which would involve a measure of the relative attractiveness
of a store. 27 He notes that such an approach has already been success­
fully taken in the work of Rushton. Since it has already been noted that
Rushton's model is based on a trade off between distance and town size,
Clark's ultimate conclusion is that consumer spatial choice is a function
of the familiar distance and store utility variables, even though he
appears to discount the influence of distance.

Golledge. 28 To a greater extent than any other researcher, Golledge

27 Clark, p. 396.


---


has attempted to establish an extensive conceptual framework for the study of consumer spatial behaviour and consumer spatial choice. His work may perhaps be validly criticized for using, without serious questioning, techniques and models derived from such fields as psychology. Some of these techniques and models still present major problems in the fields in which they were first developed. 29 Nevertheless, his work is of interest both because it suggests some fruitful new directions that the geographic study of consumer choice processes might take and because it undertakes the classification of the various types of models that have so far been utilized in such studies. He also suggests how different types of consumers may restrict the predictive value of such models.

In a recent article, he considers the main types of models that have been used in the study of consumer behaviour. 30 Some of the models that he discusses are relevant to the problem of predicting the consumer's choice of store for major grocery purchases. He first considers what he


terms a Place Loyalty Model. This is a model in which consumers "develop a strong preference or habit for patronizing one place to the exclusion of others." This model is basically a deterministic one in which economically and spatially rational consumers find the store which involves the least cost of movement and invariably chooses this nearest store to the exclusion of all others. He considers the central place model to be of this type.

He then discusses what he terms Single Vector Market Share Models. These are probabilistic models in which probabilities are developed which indicate the proportion of times each store is patronized. Such a model can be constructed for aggregates of consumers by averaging the proportion of times people choose each store. What Golledge seems to be considering here is a method for arriving at a utility value for each store relative to all other stores. These values are transformed so that they are probabilities which sum to one.

These first two models which Golledge considers are based on the distance variable and the utility that a store has for the consumer. In neither case does he operationalize the model to actually predict the stores which consumers choose.

Finally, he presents a Stochastic Perceptual Model. Through an interview process, information is gathered on the variables which are important to consumers in their choice of store. Each consumer is asked

---

31 Ibid., p. 419.
to rank the variables that influence his choice. The number of times that each variable is ranked first is summed and the results are transformed to create a probability vector. This probability vector represents the probability that a variable will be important in the choice processes of consumers. At the same time the consumers are asked to indicate which stores they feel rank highest in each variable. A matrix is formed with this information which gives an indication of the relative attractiveness of each store with regard to each of the variables. These values are transformed and expressed as probabilities. If the transformed matrix and the probability vector are multiplied, the result is a vector which gives the probability of consumers going to each of the stores. These probabilities may be viewed as utility values and used in combination with distance to predict consumer choices.

The main deficiency in Golledge's work is one that is common to most of the research considered in this review. He does not attempt to operationalize his models and utilize them to predict the stores which consumers choose for their major shopping trips.

The greatest weakness of most of the studies considered in the review of the literature is that they did not attempt to convert their conceptual models to operational ones. In the next chapter, we will see how this deficiency can be overcome and an operational predictive model developed.
CHAPTER III.

The Conceptual Model and Its Problems

The most difficult part of devising an adequate model for predicting the consumer's choice of grocery store is, as we have been in earlier studies, the step of making a theoretical model operational. This chapter is concerned with the translation of an abstract conceptual model into a working predictive model which can effectively allocate consumers to grocery stores.

The conceptual model which has emerged from the examination of the existing literature involves two major variables. One of these is the negative effect that increasing distance has on the likelihood of choosing a particular store, while the other is the utility that a particular store has for consumers.

In this chapter, we shall consider the nature of these two variables, ways in which they may be related and the problems associated with utilizing this conceptual model in an operational format to predict store choices. Such a discussion now will make it possible to develop an adequate predictive model later.

Problems of the Distance Variable.

In most geographical discussions of individuals in a spatial choice situation, distance is considered to be a variable of major importance. Any conceptual treatment of the consumer's decision to allocate his major grocery purchases to a particular supermarket must necessarily consider the distance between that consumer and the various supermarkets available for
him to choose. This distance is a primary determinant in the consumer's choice process. Since this variable is widely recognized and accepted by geographers as important in the explication of the type of problems with which they deal, it might be imagined that all the major difficulties associated with the use of distance as a variable have by now been overcome. This is not the case. Like the concept of intelligence among psychologists, there is little agreement among geographers on the exact nature of distance, even though most of them feel it to be very important.

In the first place, there is little agreement as to the system of measurement most appropriate for the distance variable. There are many ways of viewing and measuring distance. For example, the distance between a consumer and a store may be measured as the shortest distance between these two locations on the surface of the earth. This straight line distance is the easiest form of distance to understand and work with.

But "city-block metric" distance is another way of measuring the distance between a consumer and a store and it takes into account the fact that one often cannot travel in a straight line between two locations in an urban area. Thus, the shortest possible route between two locations may require an individual to travel three blocks to the west and two blocks to the north, because the shorter, straight line route is blocked by buildings.

A third way of measuring the physical distance between two locations within a city involves measurement along the best roads connecting the two places. The assumption is made that people will choose to travel along the highest quality of road available and that distance measured along such roads is a more realistic indication of the friction of distance than other measures of physical distance.
While various researchers have shown a preference for one or another of these physical measures of distance, other researchers have suggested that no system of measurement in direct physical units can accurately reflect the impact of distance on an individual's behavior in a spatial choice situation. For this reason they have substituted the time distance between locations as a more valid measure. This time distance is often found by taking the average driving time along important roads at various times of the day.

Then too, it has been suggested, in the literature of environmental perception, that perceived distance is a more meaningful measure than either physical distance or time distance. According to this view, the friction of distance is best indicated by the distance that the consumer perceives between his location and that of various stores, rather than the distance which actually exists. The consumer makes his decision on the basis of the distance that he feels exists between the two locations and the perceived distance often is not the same as the actual distance.

In spite of the fact that researchers have suggested a variety of approaches to the measurement of the friction of distance, none of them have attempted to develop an operational model to predict the consumer's choice of food store in an intraurban area. For this reason, our first version of the model will utilize the straight line distance between a consumer and a store since it is simple and easy to work with.

A second problem with the distance variable arises from the fact that not everyone reacts to distance in the same way. As mentioned earlier, some individuals are very strongly influenced in their store choices by
the costs and effort associated with overcoming the friction of distance, while others seem to be both willing and able to travel greater distances for the same goods. In the first version of the model it will be assumed that all individuals react in a similar fashion to distance. If subsequent analysis of the results of the predictive model show that clearly identifiable groups of consumers have different space preferences, this fact may then be introduced into a revised model to improve the predictive ability of the model.

Problems of the Utility Variable.

The greatest difficulty with using the utility that various stores have for consumers arises from the need to decide which of several ways of measuring utility is the most valid.

Two different approaches have been taken to the derivation of utility values. One of these uses physical measures as surrogates for utility. For example, Huff uses the size of a store in terms of retail floor area as his measure of a store's utility and Baumol and Ide use the number of different items sold by a store. To use physical surrogates for the utility of a store requires the naive assumption that physical size is the most important measure of a store's relative worth to consumers. This assumption ignores the fact that other factors such as the quality of service or the price level of a store may also be important in determining the utility which a store has for consumers, even though these factors are not necessarily correlated with the size of the store.

A second approach to the determination of store utility values has been taken by some researchers. This second approach involves interviewing consumers to determine directly the utilities that different stores have
for them. One way to derive utilities through consumer interviews is to ask each consumer to rank the three stores which he considers 'best' in the area. If the term 'best' is not interpreted for the consumer during the interview, it may be assumed that the consumer decides to rank stores on the basis of those factors that are most important to him in his choice of a store. The reason for ranking only three stores is that consumers probably do not have accurate knowledge of all stores in an area. For this reason, a forced ranking of all stores would generate considerable noise in the form of meaningless answers and guesses.

These rankings could then be converted to utility values by utilizing one of the psychological scaling programmes such as developed by Torgerson and Shepherd. In the present research, an attempt was made to gather the rankings necessary for applying one of these psychological scaling models. It was found, however, that a large number of the consumers who were interviewed could not rank the three best stores in the study area. Many consumers did all of their shopping at one store and had no opinions about other stores. For this reason, the approach which utilizes psychological scaling models in the derivation of store utilities could not be

---


A second approach to the derivation of store utilities by consumer interviews utilizes the Stochastic Perceptual Model which Golledge has presented and which has already been considered above. This method involves gathering data from the consumers on the importance of different variables in their choice of stores and on the ranking of each store with regard to each of the variables. The end result of this process is a vector indicating the relative utility of each store for consumers if each of the variables is weighted according to its contribution to the total utility of a store.

This researcher found that it was difficult to gather the data required for this method of deriving utilities. Consumers were not usually able to rank stores with regard to the different variables and often did not understand what was required of them in this part of the questionnaire. The Stochastic Perceptual Model, thus, has proven too difficult to utilize, even though it looks good in theory.

Because of the difficulty of gathering data from consumers for the Stochastic Perceptual Model and for psychological scaling models, this researcher feels that neither approach will generate utility values which adequately reflect the actual worth of different stores to consumers. Also, it is felt that both methods are unduly complex and probably would tend to obscure the consumer's spatial choice process, instead of clarifying it.

For these reasons, the first version of the predictive model will utilize a simple count of the number of times each store was ranked as best by consumers. These simple utility values will be incorporated with
the distance variable in a framework to be discussed in the next part of this chapter.

Problems in Relating the Two Variables.

Having conceptually isolated distance and store utility as the significant variables to be used in a model to predict consumer store choices for major grocery purchases, the nature of the relationship between these variables must now be considered. Review of the literature indicated that there have been two basic ways of expressing this relationship. It is appropriate now to consider and assess each of these.

The models of Baumol and Ide and of Nystuen both suggest subtracting the distance variable from the utility variable to find the attractiveness of a particular store for a given consumer. When this calculation has been made for the consumer with regard to each of the stores, the consumer is then allocated to that store which has the highest attractiveness index. The problem with this model is that the utility and distance variables are not measured in the same units and, as a result, the meaning of the attractiveness index is rather unclear.

A more valid way of expressing the relationship between the two variables found in the work of Huff, Rushton and Clark, who all use modified gravity models. In such models, the attractiveness of a store is directly proportional to the utility of that store and inversely related to the distance between that store and the consumer. The distance variable may be raised to an exponent in order to modify the influence that the two variables each have on the attractiveness. The formula is: \( U/D^e \), where \( U \) is the utility of a store, \( D \) is the distance between that store and the consumer.
and e is an exponent to which distance is raised. The appropriate value of the exponent is found by running the computer predictive model with a variety of different exponents until that exponent which gives the highest level of prediction is reached.

Summary.

In this chapter, a conceptual model has been discussed which was derived from the literature reviewed in Chapter II. A number of barriers to the operationalization of the conceptual model have been considered. The conceptual model finally accepted is a gravity model using the simplest measures of distance and store utility. These measures are the straight line distance between the store and the consumer and the number of times each store was ranked as the best store in the area.

The next chapter will consider the computer allocation model which has been designed to predict the consumer's choice of store. It utilizes the simple conceptual model as its basis.
CHAPTER IV

The Basic Consumer Allocation Model

In this chapter, a computer model will be described which operationalizes the conceptual model of the consumer's choice process that was presented in Chapter III. The first part of the chapter discusses the computer programme at a general level in order to familiarize the reader with its basic structure. The reader may also wish to read the more detailed consideration of the function of each of the statements in the programme which appears in Appendix I.

The second part of this chapter presents the results of the basic consumer allocation model. This basic model utilizes store utilities that were derived by totaling the number of times each store was selected by a consumer as the 'best' store in the study area. These simple store utilities are used with the straight line distance between stores and consumers in the framework of a simple gravity model to compute the attractiveness indices necessary for the allocation of consumers to particular stores for their food shopping.

The Basic Structure of the Consumer Allocation Model.

The programme first reads a data card for each of the stores involved in the allocation process. In this study, a store is considered only if at least one of the consumers interviewed rated it as the best store in the Kitchener-Waterloo area. Since utility values are based on this rating of the best stores in the areas, stores not rated best at least once have no utility and therefore have no chance of being selected by the computer allocation model. Such stores are not, therefore considered in the computer
allocation model. On this basis, fourteen stores are considered in the allocation model. They include all the supermarkets which are located in Waterloo and five supermarkets which are located in Kitchener. The data card for each store contains a code number for that store, a location for the store expressed as X and Y co-ordinates on a grid covering the study area, and a utility value for that store derived by interviewing consumers to find out which stores they consider to be the best stores in the area.

Next the programme reads a data card for each of the consumers who were interviewed. The consumers were interviewed at their homes. Using the map on which the store location grid had been drawn, 240 sample points were selected at random in the Waterloo study area. Interviews were carried out at the households nearest to these sample points. The 240 interviews represent approximately 39% of the households in the Waterloo study area. The data cards for the 240 consumers each contain a code number representing the consumer, a location for the consumer expressed as X and Y co-ordinates on the location grid, and the code number of the store from which the consumer actually purchased his groceries.

Having read and stored data on both the stores and the consumers, the main computational section of the programme then uses this data to calculate an attractiveness index for each consumer in regard to each of the stores. Each consumer is then allocated to that store for which the attractiveness index is highest. In particular, this allocation is carried out by two DO-loops, one nested within the other. The outer DO-loop successively calls the data stores for each consumer from the computer's memory. Then the inner DO-loop utilizes the data provided on a consumer by the outer DO-loop to calculate the distance between that consumer and
THE LOCATION OF THE STORES

STORE CODE NUMBER

1. ZEHR'S, Parkdale Plaza.
2. ZEHR'S, County Fair Plaza.
3. WAREHOUSE MARKET, King St.
4. ZEHR'S, Tower's Plaza.
5. DOMINION, Westmount Place.
6. ZEHR'S, Waterloo Square.
8. BUSY B, Waterloo Square.
9. WAREHOUSE MARKET, Union St.
10. ZEHR'S, Belmont St.
11. CITY MEAT MARKET, King St.
12. DUTCH BOY, Margaret St.
16. DUTCH BOY, King St.
25. HIWAY MARKET, King St.
each of the stores and to calculate the attractiveness index which each store has for that consumer. This attractiveness index related the straight line distance between a store and a consumer and the utility of a store in the context of a gravity model. The utility of a store is substituted for mass in the gravity formula and the result is an index of the relative attractiveness of different stores for a particular consumer. The programme then allocates the consumer to that store for which this computed attractiveness index is highest.

Returning to the outer DO-loop, the programme compares the store to which the consumer is allocated by the model with the store at which the consumer actually shopped. If the two store numbers are the same, the programme increments a counter by one. When all the consumers have been considered, this counter contains the number of consumers that the model was able to correctly allocate. After printing the code number of the consumer, the code number of the store at which he actually shopped and the code number of the store to which the computer model allocated him, the programme returns to the outer DO-loop and another consumer.

When all consumers have been allocated, the programme leaves the outer DO-loop and goes on to divide the number of consumers correctly allocated by the total number of consumers. When this fraction is multiplied by 100, the result is the percentage of consumers which the model was able to correctly allocate. This percentage is printed and the programme comes to an end.

Appendix I contains a more detailed discussion of the individual statements which comprise this computer allocation model.
Results of the Basic Consumer Allocation Model.

The basic model utilized simple totals of the number of times each store was ranked as best by the consumers as the store utility values. In the next chapter, some other approaches to store utilities are taken in an effort to improve the predictive ability of the model. The original utility values are subjected to logarithmic and square root transformations and an alternative means of deriving utilities is also used. For this reason, it was necessary to convert different utility measures to common units so that the results of different approaches to utility would be of comparable magnitude. This was done by setting the highest utility value resulting from each approach equal to 100 and then considering all other utility values to be proportions of that value. For example, if the highest utility value was 60 and the next highest value was 40, their converted values become 100 and 40/60 times 100 = 66.7.

When the basic model was run on the computer, it correctly allocated 57.50% of the consumers when the distance exponent was set at two. Since only 47.50% of the consumers shopped at the stores nearest to their homes, this model represents a significant improvement over one based upon consumer distance minimization alone.

Summary

This basic model utilizing simple conceptualizations or both distance and store utilities in a gravity model with an exponent of two is accepted as a valid predictive model. It predicts 57.50% of consumer store choices and conforms well to the classic gravity model. However, it does not reach the goal of a 60% level of prediction set earlier in this study. One reason for this failure is that the model
tends to overallocate consumers to stores with high utility values. Could the level of prediction of the model be increased by making modifications in the operational model? We will consider this possibility in the following chapter.
CHAPTER V

Alterations to the Basic Allocation Model

Although the original model, using the gravity formula and simple measures of distance and store utilities, was found to correctly predict 57.50% when an exponent of two was used, it was felt that the level of prediction could be improved. The present chapter considers three types of changes that could be made in the basic model without introducing other variables. The distance variable could be modified, the store utility variable could be altered or a different formula could be used to relate the two variables.

Transformations of the Original Utility Values.

Results of the model considered in the last chapter showed that the model tended to overallocate consumers to stores with high utility values. For example, the Zehr's supermarket at the Towers Shopping Centre (4) has a very high utility value and the model allocated many consumers to it that should have been allocated to other stores. The tendency of the model to overallocate consumers to such stores may indicate that the utility values are lognormally distributed and not normally distributed. Therefore, a logarithmic transformation of the original utility values may improve the level of prediction of the model. Such a transformation has the effect of changing the relative size of the utility values, increasing the magnitude of the lower utility values in relation to the higher utility values.
In an effort to improve the predictive level of the model, the programme was run using utility values that had been subjected to logarithmic and square root transformations in the effort to decrease the dominance of the stores with high utility values. The logarithmically transformed utilities enabled the model to predict 60.63% of consumer choices correctly. This is 13% more than the 47.50% that could be predicted by using distance alone and turns out to be the best level of prediction resulting from any of the variations of the model.

The square root transformation of the utility values also improved the level of prediction of the consumer allocation model, but the improvement was not as great as the improvement made possible by using the logarithmic transformation. When the exponent of distance was two and the square root transformation was used, 60% of consumer choices were correctly predicted.

An Alternate Means of Deriving Store Utilities.

Since it was observed that stores with high utility values were distorting the model, an alternate approach was taken to the derivation of utility values for the stores. Each consumer had been asked during the interviews to rank the three best stores in the Kitchener-Waterloo area. The first approach to the derivation of store utilities used the number of times that a store was ranked as the best store in Kitchener-Waterloo as an indication of the utility of that store for consumers. In the alternate approach, each store gains a utility of three if it is ranked as the best store in the area. It gains a utility of two if it is ranked as third best. As with the original utility values, however, these
revised utilities also caused the model to overallocate consumers to stores with high utility values. For this reason, the revised utilities were also subjected to logarithmic and square root transformations.

The revised utility values enabled the model to predict 58.75% of consumer choices when the distance exponent was set at two. This is a good level of prediction, but it is not as high as the level that is made possible by the logarithmic transformation of the original utility values. When these revised utility values are logarithmically transformed, the level of prediction remains at 58.75%.

The level of prediction rises to 60.83% when the revised utility values are transformed by taking their square roots. This level of prediction is equal to that achieved when using a logarithmic transformation of the original utility values. The alternate approach to the derivation of store utilities does not appear to offer any improvement over the original utilities. They are also conceptually more difficult to justify because the weightings of rankings of best, second best and third best stores in the area are arbitrary and could have been treated in a variety of ways.

An Alternate Measure of Distance.

By using a logarithmic transformation of the original utility values, the level of prediction of the model was increased from 57.50% to 60.83%. One might also try to improve the level of prediction of the model by using a different measure of the friction of distance. In Chapter III, it was decided that straight line distance would probably be the most meaningful measure of the friction of distance between a consumer and a store. Since this straight line conceptualization of distance resulted in
high levels of prediction, the researcher decided not to test a variety of other measures of distance. One objective of this study was to develop a simple model which would be able to achieve a high level of prediction with data that could be gathered quickly and inexpensively. It was felt that any improvement in the level of prediction that might have been made possible by the use of time distance or perceived distance would be more than offset by the increase in the complexity of the data required for their use.

However, 'city-block metric' distance is not subject to this cost constraint since it may be calculated from the grid locations of the consumers and the stores as easily as straight line distance. Therefore it was decided to run the model utilizing 'city-block metric' distance as the distance variable. This alternative to straight line distance was used in the model with the original utilities and with the logarithmic and square root transformations of the original utilities. In all three cases, the best level of prediction that could be achieved was 58.33%. Clearly, this alternate approach to the friction of distance does not add to the predictive power of the model.

An Alternate Relationship Between the Two Variables.

When considering the conceptual model in Chapter III, it was noted that the two variables had been related to each other in two main ways in previous works. In the gravity formula, the store utilities are divided by distance raised to an exponent, while in the alternate approach, distance is subtracted from the store utilities. In other words, store utility is viewed as the benefit to be gained from shopping at a particular store and distance is viewed as the cost of shopping at that store.
The results of using this alternative to the gravity model were very poor. Using the original utility values, the best level of prediction that could be achieved by varying the distance exponent was 29.16%. With logarithmically transformed utilities, the level of prediction rises to 33.75%, but when the original utilities are transformed by taking their square roots, the level falls back to 29.16.

It is clear that the gravity model is a more appropriate way of relating these two variables. Perhaps if a researcher could get accurate measures of the costs and benefits associated with shopping at different stores, the alternate relationship would be more meaningful, but with the data available, it is not a useful way of relating the two variables.

Summary.

The modifications to the basic consumer allocation model that have been considered in this chapter were able to raise the predictive level of the model from 57.50% to 60.83%. This level of prediction is reached when straight line distance and a logarithmic transformation of the original utility values are related in a gravity formula.

The model may be considered successful for two reasons. In the first place, it was applied in an intraurban setting which is more complex than the interurban setting of Rushton's study. Secondly, it has a level of prediction as high as that reached in Rushton's study, but at the same time is somewhat simpler and easier to utilize. 33

---

33 Rushton, p. 41.
In the next chapter, the results of the most successful version of the model will be analysed in an effort to uncover regularities in socioeconomic and trip characteristics of different groups of consumer. Such regularities might be utilized in future efforts to improve the model.
CHAPTER VI

An Analysis of the Results of the Model

In this chapter, the results of that version of the consumer allocation model which gives the best level of prediction will be analysed. This version of the model relates two variables in a gravity formula. The variables are straight line distance and store utilities. The store utilities are derived by totalling the times that each store was ranked as best by the consumers and then logarithmically transforming these values. This revised version of the model was able to correctly allocate 60.83% of the consumers.

This chapter is divided into four sections. In the first section, two desire line maps are presented which show where the consumers actually shopped and where the model predicted that they would shop. The second section discusses the results of a comparison of a variety of shopping trip and socioeconomic characteristics of those consumers correctly allocated by the model and those not correctly allocated. In section three, the same characteristics are considered for the consumers that the model allocated to each store. Finally, some conclusions are drawn with regard to these analyses and some suggestions are made regarding the use of such data on socioeconomic and trip characteristics in future research designed to improve the level of prediction of the model.

Maps of Actual and Predicted Consumer Choices.

Figure 2 is a desire line map which shows where consumers actually went to purchase the major part of their groceries. On such a map, a line
STORES CHOSEN BY CONSUMERS

STORE CODE NUMBER

1. ZEHR'S, Parkdale Plaza
2. ZEHR'S, County Fair Plaza
4. ZEHR'S, Tower's Plaza
5. DOMINION, Westmount Place
6. ZEHR'S, Waterloo Square
7. A & P, Waterloo Square
8. BUSY B, Waterloo Square
9. WAREHOUSE MARKET, Union St.
10. ZEHR'S, Belmont St.
11. CITY MEAT MARKET, King St.
12. DUTCH BOY, Margaret St.
16. DUTCH BOY, King St.
25. HIWAY MARKET, King St.
is drawn between each consumer and the store with which he is associated. A careful examination of Figure 2 will reveal a number of significant facts.

First, it will be noticed that this map of the actual choices of consumers shows distance effects at two distinct levels. Although there are a number of unnecessarily long trips, most of the trips are to stores that are relatively close to the consumers. As was mentioned earlier, 47.50% of consumers actually choose the nearest store for their major grocery purchases. At another scale, it will be noticed that most of the consumers shop either at supermarkets in Waterloo or at one of the three nearest supermarkets in Kitchener. Only two consumers out of 240 travel greater distances to shop.

Consumers actually travelled an average of 0.74 miles to do their shopping. This is almost one fifth of a mile greater than the average distance predicted by the model. This fact indicates that consumers actually travel greater distances than necessary for their grocery purchases. Examination of Figure 2 shows that there are a large number of trips on which the consumer passes one store on the way to another.

The consumer questionnaires suggest a number of reasons for these unnecessarily long trips. A considerable number of those consumers incorrectly predicted by the model are, nevertheless, strongly influenced by the friction of distance because they shop on their way to or from work. Eighteen of the incorrectly allocated consumers stated in the interviews that they shopped along their route from work and went only a short distance out of their way to do so. This fact helps to explain why the model predicted only 60.83% of consumer choices. The group of consumers that shop on the
way from work could not possibly be allocated by a model which utilizes the distance between the home of the consumer and the store as the distance variable, since the most important distance to them is the distance that a store is from their normal route to work.

Other consumers travel greater distances for their grocery shopping because they usually visit other stores while on their grocery shopping trip. Twenty of the incorrectly allocated consumers stated that they normally shopped at other stores while on their grocery trip. The most frequently mentioned other store was a department store. Since not all supermarkets are near to department stores, such consumers are forced to travel farther because of their desire to fulfill several shopping objectives on one shopping trip.

Finally, there are two groups of individuals who have considerable spare time to spend and do not seem to feel the need to minimize the distance that they travel in their grocery shopping. Ten of the incorrectly allocated consumers were retired. Many such consumers suggested in the interviews that they considered shopping to be an opportunity to get around and see more of the city. The second group is composed of consumers who follow sales at various supermarkets and regularly shop at four or more different stores. Twelve of the incorrectly predicted consumers divided their patronage in such a fashion. It can only be accident if the model correctly allocates such a consumer to that store in which he does the greatest percentage of his grocery shopping.

The discussion of the map of actual consumer store choices suggests a number of ways to partially account for the fact that some consumers
travel considerably farther in reality than they do in the predictive model. Figure 3 is a desire line map which shows the stores to which the model allocated each consumer. A solid line connects each consumer with the store to which he was allocated by the model.

It will be noticed that some of the stores where consumers actually shopped received no consumers as a result of the operation of the consumer allocation model. In the case of the Zehr's (6), the store utility value was so small that even consumers who lived very close to the store were not allocated to it. This store is too small to be considered a supermarket and should be grouped with smaller stores. No consumer actually shopped in any of the other small grocery stores for the major part of his groceries. The Warehouse Market (3) had been doing very poorly and went out of business during the time the interviews were being carried out. None of the consumers interviewed before that time did the major part of their shopping at that store. In addition, the Dutch Boy (16) and the Highway Market (25) were allocated no consumers by the model, even though they received some in reality. They are both too distant from Waterloo.

The consumers are allocated to the ten remaining supermarkets. From Figure 3, it will be obvious that the allocation model results in shorter consumer trips than took place in reality. The friction of distance is a very strong factor in the model. The map of predicted consumer choices is much less confused than the map of actual consumer choices. Definite market areas are created around each of the stores to which the model allocates consumers. There is a 'watershed' between stores at which all consumers on one side go to another store. The goal of this study has not been the development of such market areas, but their emergence is an
STORES PREDICTED BY THE MODEL

STORE CODE NUMBER
1. ZEHR'S, Parkdale Plaza.
2. ZEHR'S, County Fair Plaza.
3. ZEHR'S, Tower's Plaza
4. DOMINION, Westmount Place.
6. BUSY B, Waterloo Square.
7. WAREHOUSE MARKET, Union St.
8. ZEHR'S, Belmont St.
9. CITY MEAT MARKET, King St.
10. DUTCH BOY, Margaret St.
One way to bring the map of predicted store choices closer to that of actual store choices would be to utilize socioeconomic characteristics of the consumer or trip characteristics as additional variables in the model. The second part of this chapter will compare such characteristics for those consumers that the model was able to correctly predict and those that it was not able to correctly predict.

The Correct and Incorrect Allocations of the Model.

Difference-of-means tests are used in this section to examine differences in various characteristics of those consumers correctly allocated by the model and those consumers incorrectly allocated. These tests are designed to determine if the difference between the means of two samples can be attributed to chance alone or if the difference is statistically significant. In all the tests used in this chapter, the level of significance is .05.

When these tests were applied to data on the consumer's shopping habits, a number of statistically significant differences were found. The number of shopping trips each week was found to be significantly higher for those consumers who were correctly allocated by the model. They made an average of 1.84 trips a week, while the incorrectly allocated consumers only made 1.05 trips weekly. Their trips may have been more frequent because they perceived their stores as being significantly closer, in travel time, than

---

the incorrectly allocated consumers did. They felt that it took an average time of 4.86 minutes to reach their chosen store; the incorrectly allocated consumers felt that the same trip would take 6.43 minutes. In addition, the consumers who were correctly allocated travelled shorter distances than the incorrectly allocated consumers to reach their shopping stores. They travelled an average of .57 miles, whereas the consumers who were not correctly allocated travelled an average of 1.02 miles. Differences in the size of the weekly grocery bill were not found to be statistically significant.

Two socioeconomic characteristics exhibited statistically significant variation when the difference-of-means test was applied to them. The average annual income of those consumers who were correctly allocated was $11,512.00 and the average annual income of those incorrectly allocated was $8,741.00. Thus, the correctly allocated consumers have significantly greater incomes. On the other hand, they have only lived in their present homes for an average of 8.5 years as compared to the incorrectly allocated consumers who have occupied their homes for an average of 12.7 years. This difference is also significant at the .05 level. Neither the number of members in the family nor the age of the head of the household were found to vary significantly.

The consumer whose store choice is correctly predicted by the model is therefore different from the consumer whose choice was incorrectly predicted in several ways. He has lived a shorter time in his present house than the incorrectly allocated consumer, but he has a higher average annual income. He does not travel as far as the incorrectly allocated consumer to do his shopping, but he makes a greater number of shopping trips each week and he perceives his trips as being of short duration.
The Consumers Allocated to the Various Stores.

In a further attempt to isolate socioeconomic or shopping trip characteristics that might be used to improve future versions of the consumer allocation model, the consumers who were allocated to each of the stores are now considered. An examination of Table 1 reveals that different stores have wide variations in the percentage of the consumers allocated to them that were correctly allocated.

TABLE 1
CONSUMERS CORRECTLY ALLOCATED TO EACH STORE

<table>
<thead>
<tr>
<th>Store Code</th>
<th>Store Name</th>
<th>Number of Consumers Allocated</th>
<th>Percentage Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Zehr's</td>
<td>31</td>
<td>83.9</td>
</tr>
<tr>
<td>2</td>
<td>Zehr's</td>
<td>17</td>
<td>88.2</td>
</tr>
<tr>
<td>4</td>
<td>Zehr's</td>
<td>77</td>
<td>63.4</td>
</tr>
<tr>
<td>5</td>
<td>Dominion</td>
<td>51</td>
<td>62.8</td>
</tr>
<tr>
<td>7</td>
<td>A &amp; P</td>
<td>3</td>
<td>33.3</td>
</tr>
<tr>
<td>8</td>
<td>Busy B</td>
<td>16</td>
<td>50.0</td>
</tr>
<tr>
<td>9</td>
<td>Warehouse Market</td>
<td>15</td>
<td>6.7</td>
</tr>
<tr>
<td>10</td>
<td>Zehr's</td>
<td>12</td>
<td>33.3</td>
</tr>
<tr>
<td>11</td>
<td>Central Meat Market</td>
<td>12</td>
<td>41.7</td>
</tr>
<tr>
<td>12</td>
<td>Dutch Boy</td>
<td>7</td>
<td>85.5</td>
</tr>
</tbody>
</table>
If one relates this table to the locations of the various stores as shown in Figure 1, a fairly clear pattern emerges. Those stores that have a percentage of correct allocations higher than the 60.83% of the model as a whole are almost all located peripherally in such a manner that there is no other supermarket between them and the outskirts of the city. On the other hand, almost all of those stores which have a lower percentage of correct allocations than the model as a whole are in core locations and have other stores located between them and the outskirts of the city. This group includes Warehouse Market (3), Zehr's (6) and Dutch Boy (16) to which no consumers were allocated, but does not include Zehr's (10), whose location is not clearly one of the core or the periphery.

This distinction between the percentage of correct allocations to core and peripheral stores suggests that there may be a fundamental difference between consumers in these two distinct spatial situations. If there are such differences, it may be possible to incorporate them into future versions of the consumer allocation model. This would give support to Isard's contention that different people react differently to the friction of distance. 35

In order to examine possible differences in the characteristics of the consumers allocated to different stores, a number of Analysis of Variance tests were carried out on the socioeconomic and shopping trip data. 36 This test looks for differences between samples by a process of

35 Isard, p. 81-88.

36 Blalock, p. 242.
comparing the variance within different samples with the variance between them.

The same group of trip and socioeconomic characteristics as were considered in the examination of the differences between those consumers correctly allocated and those who were not correctly allocated are tested. Among the trip characteristics, it was found that there were not significant differences between stores in terms of the size of the weekly grocery bill, the number of shopping trips each week or the time estimated to reach the store. Among the socioeconomic characteristics, family size and the age of the head of the family were found to exhibit no statistically significant variations.

However, significant differences were found between stores in terms of the distances that the model predicted, the average annual income and the number of years that the consumer has occupied his present home.

It is not enough, however, just to know that there are significant differences between the consumers who were allocated to different stores by the model. It is also necessary to know the stores between which these differences lie. For this reason, a statistical test called Tukey's Q test was applied to the data of the three characteristics that showed significant variation in the Analysis of Variance tests. This test enables the researcher to determine the samples between which the significant differences lie. Table 2 illustrates the results of this test.

---

### TABLE 2
**RESULTS OF THE TUKEY'S Q TEST**

<table>
<thead>
<tr>
<th>Annual income.</th>
<th>5</th>
<th>2</th>
<th>10</th>
<th>1</th>
<th>4</th>
<th>12</th>
<th>9</th>
<th>11</th>
<th>8</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Years at location.</th>
<th>1</th>
<th>2</th>
<th>12</th>
<th>5</th>
<th>4</th>
<th>9</th>
<th>10</th>
<th>8</th>
<th>11</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>few</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>many</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Predicted distance.</th>
<th>1</th>
<th>2</th>
<th>(4)</th>
<th>5</th>
<th>(12)</th>
<th>11</th>
<th>10</th>
<th>8</th>
<th>(9)</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>long</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>short</td>
</tr>
<tr>
<td>short</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The straight lines between the store numbers indicate that Tukey's Q test has shown a significant difference between those two stores on that particular characteristic. Store numbers not separated by lines form groups of stores within which there is no significant variation. The stores with circles around them are peripheral stores, while those within a square are core stores. The Dutch Boy (12) is basically a peripheral store but its locational situation is somewhat less clear than the four other peripheral stores.

In each case, the core stores all lie at one end of the range of values and the peripheral stores lie at the other end. There is invariably a significant difference between the core stores and the peripheral stores.

The consumers who were allocated to the core stores have lived for a longer time in their present homes and have lower incomes than those
allocated to the peripheral stores. When they are allocated by the model, they are sent shorter distances than those consumers allocated to peripheral stores.

It may be that the level of prediction in peripheral stores is higher because these stores are in planned shopping centres which were designed and located so as to dominate certain market areas. Such stores are usually found in locations which are clearly the closest to consumers in a large area. In addition, the presence of other stores and ample parking attracts the multiple purpose shopper. Because peripheral stores dominate their market areas, they do not require market areas that are large in population.

The core stores do require market areas with fairly large populations because they are not able to dominate market areas on the basis of the friction of distance. It is usually less clear to a consumer in a core location which is the nearest store. One of the reasons that some core supermarkets are failing is that they cannot attract enough consumers from the immediate area to survive and they are often unable to draw consumers from greater distances.

Habit patterns seem to play a significant role in determining the consumers that actually choose core stores and those that actually choose peripheral stores. Consumers located near to core stores have often lived at the same location for many years and have gotten the habit of shopping at certain stores, even though conditions have changed over the years and these stores are no longer in superior locations. For this reason, their store choices may not appear to be particularly rational and may not be easily duplicated in a simple predictive model.

Peripheral consumers have formed their shopping habits at a later time,
both because they have lived at their present locations for fewer years and because the stores in peripheral locations have not been in existence for very long. As a result, their store choices are the manifestation of a relatively recent spatial choice process and are more likely to be spatially rational in terms of the present distribution of supermarkets.

Summary.

In this chapter, the maps showing the actual choices of consumers and the choices made by the consumer allocation model were first discussed. Then attempts were made to isolate statistically significant differences between those consumers who were successfully predicted and those who were not successfully predicted and between the consumers who were allocated to the various stores.

There are clear differences between the consumers who were correctly allocated by the model and those who were not correctly allocated. It is possible that these differences could be utilized in a future attempt to improve the predictive level of the model.

Further significant differences emerged from the consideration of the variations between the consumers allocated to different stores. There are clear spatial, temporal and economic differences between the consumers who were allocated to the core stores and those who were allocated to peripheral stores. The core consumers have lived at their present houses for a longer period, they have lower incomes than peripheral consumers and the model predicts that they will travel shorter distances to shop.

It is not unreasonable to conclude that these differences between core and peripheral consumers could be utilized in future research to improve the level of prediction of the consumer allocation model.
The basic objective of this study has been to develop and test a model to predict the stores which consumers choose for their major grocery purchases in an intraurban area. This objective has been approached by a process of examining the existing literature on consumer store choice and attempting to isolate the key variables in the formulations of other researchers. No researcher has published a predictive model of consumer store choice for an intraurban area, but many have considered the nature of the consumer spatial choice process. Several of their works were reviewed in this study and it was found that they all shared a basic conceptual model. They all considered distance to be a key variable and they all utilized some measure of the utility that stores have for the consumers as another variable.

Various problems related to making the conceptual model operate to predict the stores that consumers choose were then considered. It was concluded that the straight line distance between a consumer and a store would be the best measure of the friction of distance and that the utility that each store has for consumers could be derived by asking each consumer which store he thought was best in the Kitchener-Waterloo area. The utility of a store would then be the number of times it was considered the best store in the area.

A consumer allocation model was developed which related these two main
variables in a gravity formula. The computer programme used this formula to calculate an attractiveness index for each consumer in relation to each store. The consumer was then allocated to that store for which this attractiveness index was highest. The results of this basic model were encouraging, but it was felt that the level of prediction of 57.50% could be improved upon by making some modifications to this basic model.

Modifications were made to the distance variable, the store utility variable and to the formula relating these two variables. The utilization of a logarithmic transformation of the original store utility values resulted in the greatest improvement of any of the modifications. With this modification, the model was able to correctly allocate 60.83% of the consumers.

Finally, the results of this modified model were analysed in an effort to discover regularities in various socioeconomic and shopping trip characteristics that might be used to improve the model in future research. Statistically significant differences were found between those consumers correctly allocated by the model and those who were not correctly allocated by the model. In addition, an examination of the socioeconomic and shopping trip characteristics of the consumers who were allocated to each of the different stores revealed a striking difference between those consumers who were allocated to stores located in the core area and those who were allocated to stores on the periphery.

In conclusion, it may be said that the study has been a success. A model has been developed which is able to correctly predict the stores chosen by 60.83% of the consumers with the use of data that is simple and
easy to gather. This research is, however, only a first step towards a full understanding of the behaviour of the consumer in spatial choice situations. Future research should consider the bases of the behaviour of those consumers that the model could not correctly allocate. Their spatial behaviour may well be rational at a different level than that considered in this study.
APPENDIX I

The basic computer model was explained briefly in the main body of the study. The following is a detailed discussion of the individual statements that make up this programme. A copy of the basic programme is shown below.

```
$JOB WATFIV DOOOUOW, KP = 29
C WAYNE HILMO
DIMENSION NCODE (14), XSTORE(14), YSTORE(14), UTILITY(14), NMXCOD(60), X
  1 ONS(60), YCONS(60), NACT(60), BIG(60), AIRDST(14), VALUE(14), NPRED(60)
  2 XMILES (14)
50 FORMAT (12,F5.0,F4.0,F6.2)
51 FORMAT (13,F5.0,F4.0,12)
60 FORMAT (315)
61 FORMAT (F34.2)
N=14
NMX =60
SAME =0.0
DO 1 I = 1,N
READ(5,50)NCODE(I), XSTORE(I), YSTORE(I), UTILITY(I)
1 CONTINUE
DO 2J=1,NMX
READ(5,51) NMXCOD(J) ,XC0NS(J), YCONS(J),NACT(J)
2 CONTINUE
DO 3 J=1, NMX
BIG(J) =0.0
DO 4 I = 1,N
XDIST =XSTORE(I) - XCONS(J)
YDIST = YSTORE (I) -YCONS (J)
AIRDST(I) =SORT (XDIST** 2+YDIST**2)
XMILES(I) = AIRDST (I)/97.
VALUE (I) =UTILITY(I)/XMILES(I) **1.5
IF(VALUE(I) .LE.BIG(J) )G0T04
BIG(J) = VALUE(I)
NPRED(J) =NCODE(I)
4 CONTINUE
IF(NPRED(J) .NE.NACT(J) ) GOTO5
SAME =SAME +1.0
5 WRITE(6,60) NMXCOD(J), NACT(J), NPRED(J)
3 CONTINUE
CORPER =SAME/NMX *100
WRITE(6,61) CORPER
STOP
END
```
The first statement in the programme is: DIMENSION NCODE (14),
XSTORE (14), YSTORE (14), UTILITY (14), NMXCOD (60), XCONS (60), YCONS (60),
NACT (60), BIG (60), XMILES (14), VALUE (14), NPRED (60).
This statement instructs the computer to set aside storage space for various
items of information which will be read into the computer on the store and
consumer data cards and for the results of calculations which will be made
by the computer as it works through this programme. NCODE is the name of the
array in which store code numbers will be stored and NMXCOD is the name of
the array in which consumer code numbers will be stored. XSTORE and YSTORE
name the arrays in which the X and Y co-ordinates of the various stores are
held, while XCONS and YCONS do the same for the X and Y co-ordinates of the
consumers. The store utilities read in on the store data cards are stored
in the array named UTILITY. NACT is the name of the array in which the
code numbers of the stores actually chosen by the consumer are stored and
NPRED is the name for the array in which the code numbers of the stores
predicted by the model are stored. Finally, XMILES stores, for a given
consumer, the distance between that consumer and each of the fourteen stores,
while VALUE stores the attractiveness indices computed for that consumer
in relation to the fourteen stores. These values are retained only until
they are printed in the outer DO-loop and then the XMILES and VALUE arrays
are used to store distance and attractiveness indices for the next consumer.

Following the DIMENSION statement, there are four FORMAT statements.
These statements are used to give the computer information about the form
and size of numbers to be read into storage or to be printed. They tell the
computer if a number is an integer or a real number and specify where a
number will be located on a data card and where it should be printed. They
also tell the computer how large a number is to be and how many decimal places a real number is taken to. These FORMAT statements will be discussed when considering the READ, WRITE or PRINT statements to which they refer.

The two statements \( N = 14 \) and \( \text{NMX} = 60 \) inform the computer that the number of store data cards to be read into storage will be 14 and the number of consumer data cards to be read in will be 60. The reason that the number of consumer data cards is 60 instead of 240 is that the programme was designed so that it could be run in four segments on the Debug terminal at the University of Waterloo rather than the Main terminal. There were two reasons for this design. First, programmes that run on the Debug terminal do so free of charge. This meant that more variations of the model could be tested in an attempt to reach the highest level of prediction. Secondly, this design was a time-saving device. Programmes can usually be run on the Debug terminal in a few minutes, while there is usually a half day wait on the Main terminal because of the computer's priority system.

The statement \( \text{SAME} = 0.0 \) sets to zero a counter which will be used later in the programme to store the number of stores which the computer model is able to allocate correctly.

With all the preliminary statements having been made, the next step in the programme involves two DO-loops which cause the computer to read and store the information on the store and consumer data cards. The first DO-loop consists of three statements:

\[
\text{DO 1 I}=1,N \text{ READ (5,50) NCODE(I), XSTORE (I), YSTORE(I), UTILITY (I) CONTINUE}
\]

The DO 1 I=1, N statement established a CO-loop which causes the computer to move from that statement to the CONTINUE statement, carrying out the
instructions of all statements in between. The CONTINUE statement causes the computer to return to the beginning of the DO-loop. This loop is repeated N times, with the value of I incrementing by one on each run through the DO-loop. The READ statement instructs the computer to read a data card which contains the store identification code (NCODE(I)), the store's X and Y co-ordinates (XSTORE(I) and YSTORE(I)) and the store's utility value (UTILITY(I)). Each run through the DO-loop causes one store data card to be read and stored. When N is reached, all 14 store data cards are read and the programme moves on to the next statement.

The second DO-loop causes the computer to read the 60 consumer data cards that are run in each of the four segments of the programme. The three statements are:

```
DO 2 J=1, NMX
   READ (5,51) NMXCOD(J), XCONS(J), YCONS(J), NACT(J)
2 CONTINUE
```

On each run through this DO-loop, the computer reads one of the consumer data cards. Each of these cards has a consumer identification code (NMXCOD(J)), the consumer's locational co-ordinates (XCONS(J) and YCONS(J)), and the code number of the store to which the consumer actually went to buy his groceries (NACT(J)).

Once the computer has read and stored the data on the store and consumer data cards, the programme goes on to use this data to allocate consumers to various stores.

The statement DO 3 J=1, NMX begins an outer DO-loop which runs to the CONTINUE statement numbered 3. It successively calls from the computer's memory the data which was read and stored for each consumer in the second
DO-loop. This data is then used in the fourth DO-loop to calculate distances between a consumer and each of the 14 stores and an attractiveness index for each store in relationship to the same consumer.

BIG=0.0 sets to zero a storage location that is used to hold the largest attractiveness index that is computed for a particular consumer. It is used in the consumer allocation process of the inner DO-loop and is reset to zero for each consumer.

The fourth DO-loop of the programme is nested within the third DO-loop. It causes the computer to perform a series of calculations 14 times each time a new consumer is brought forward by the outer DO-loop. In other words, this series of calculations is carried out once for each consumer in relation to each store. This inner DO-loop begins with the statement DO 4 I=1,N and ends with the CONTINUE statement numbered 4. Most of the important calculations of the programme are carried out within this DO-loop.

The first two statements within the inner DO-loop calculate the distance between the store and the consumer along the X and Y axes of the location grid. These statements are:

\[ \text{XDIST} = \frac{(X\text{STORE}(I) - X\text{CONS}(J))}{97} \]  
\[ \text{XDIST} = \frac{(Y\text{STORE}(I) - Y\text{CONS}(J))}{97} \]

Each of these computed values is divided by 97, in order to convert the distances measured on the arbitrary location grid to miles; there are 97 grid units to the mile.

The X and Y distances are then utilized in the formula for finding the hypotenuse of a right-angled triangle. This formula is \( A^2 = B^2 + C^2 \). Therefore, the straight line distance between a consumer and a store is
computed by the statement: $XMILES(I) = SQRT(XDIST^2 + YDIST^2)$.

Once the airline distance between a consumer and a store has been calculated, this distance is utilized along with the store's utility value in a gravity formula to calculate an attractiveness index for that store in relation to the consumer being handled in the outer Do-loop. The required calculation is generated by the following statement:

$VALUE(I) = UTILITY(I)/XMILES(I)^2$. It will be noted that the value of the distance exponent was varied in different runs of the programme in order to find the value of the exponent which resulted in the highest level of prediction.

Once the attractiveness index of a particular store has been calculated it is compared to the value of BIG in an IF statement. IF($VALUE(I)$. LE.$BIG(J)$ ) GOTO 4 is the statement involved. If the attractiveness index does not exceed the value of BIG($J$), the programme is directed to the end of the fourth or inner DO-loop and the calculations are all repeated for the next store. However, if the value of the attractiveness index does exceed the value of BIG($J$), then the programme is directed to the statements BIG($J$)=VALUE($I$) and NPRED($J$)=NCODE($I$). BIG($J$)=VALUE($I$) causes BIG($J$) to be replaced by larger attractiveness indices until it contains the largest attractiveness index when the inner DO-loop is fulfilled. NPRED($J$)=NCODE($I$) causes the code number of the store with the largest attractiveness index to be stored at NPRED($J$). This represents the store to which the model allocates the consumer.

Returning to the outer DO-loop, the store to which the model allocates the consumer is compared to the store at which the consumer actually
shops for his groceries. This comparison is made by the following IF statement:

IF(NPRED(J).NE.NACT(J)) GOTO 5. If the store predicted is not the same as the store actually chosen, an incorrect allocation has been made and the programme goes to the WRITE statement numbered 5 and then to the end of the outer DO-loop to begin the allocation process for another consumer. But if the two code numbers are identical, the value of SAME is incremented by one in the statement SAME=SAME+1.0. SAME is thus used to store the number of consumers whose choice of store is correctly predicted by the model.

WRITE(6,60) NMXCOD(J),NACT(J), NPRED(J) is the final active statement in the outer DO-loop. It causes the computer to print the consumer's code number (NMXCOD(J) ), the code number of the store at which he actually shops (NACT(J) ), and the code number of the store to which the model allocated him (NPRED(J) ). The outer DO-loop repeats itself until all the consumers have been allocated.

When all consumers have been allocated, the percentage of consumers correctly predicted is calculated by the statement CORPER=SAME/NMX*100. The statement WRITE(6,61) CORPER instructs the computer to print this percentage and when it has been printed, the programme comes to an end.
APPENDIX II

Results of the Best Version of the Model

<table>
<thead>
<tr>
<th>Consumer Code #</th>
<th>Actual Choice</th>
<th>Predicted Store</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>11</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>12</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>13</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>14</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>15</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>16</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>17</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>18</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>19</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>20</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>21</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>22</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>23</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>24</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>26</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>27</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>28</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>29</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>30</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>31</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>32</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>33</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>34</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>35</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>36</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>37</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>38</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>39</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>40</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>41</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>42</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>43</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>44</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>45</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>58</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>59</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>73</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>82</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>83</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>84</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>86</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>87</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>88</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>89</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>91</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>92</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>93</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>96</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>97</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>98</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>99</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>100</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>101</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>102</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>103</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>104</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>105</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>106</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>107</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>108</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>109</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>110</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>111</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>112</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>113</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>114</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>115</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>116</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>117</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>118</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>119</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>120</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>121</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>122</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>123</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>124</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>125</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>126</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>127</td>
<td>16</td>
<td>9</td>
</tr>
<tr>
<td>128</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>129</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>130</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>131</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>132</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>133</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>134</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>135</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>136</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>137</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>138</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>139</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>140</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>141</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>142</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>143</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>144</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>145</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>146</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>147</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>148</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>149</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>150</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
1. By what means of transportation do you usually go to the store where you do most of your weekly grocery shopping?
- by car
- by taxi
- by bus
- on foot (with delivery)
- on foot (carrying groceries)
- other

2. How far do you travel to reach the store where you do most of your weekly grocery shopping? ____ miles ____ blocks.

3. How long does it take you to reach the store where you do most of your weekly grocery shopping?

4. Would you please list the stores where you buy your groceries and try to indicate the percentage of your weekly grocery shopping that is done at each store?

<table>
<thead>
<tr>
<th>Store Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. On what day or days of the week do you usually shop for your groceries and at what time of the day do you usually shop? Please try to estimate the percentage of your weekly shopping that is done on different days.

<table>
<thead>
<tr>
<th>Day</th>
<th>% of Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday</td>
<td></td>
</tr>
<tr>
<td>Tuesday</td>
<td></td>
</tr>
<tr>
<td>Wednesday</td>
<td></td>
</tr>
<tr>
<td>Thursday</td>
<td></td>
</tr>
<tr>
<td>Friday</td>
<td></td>
</tr>
<tr>
<td>Saturday</td>
<td></td>
</tr>
<tr>
<td>Sunday</td>
<td></td>
</tr>
</tbody>
</table>

6. Which, in your opinion, are the three best supermarkets in the Kitchener-Waterloo area? (1) (2) (3)

7. Do you usually shop for groceries on your way to or from work?
- Yes
- No

8. If your answer to question 7 was yes, how far is that store from your normal route to work?

9. Do you usually visit other stores or services while on your major weekly trip for groceries?
- Yes
- No

10. If you usually visit other stores or services while on your major weekly grocery shopping trip, what types of stores or services do you usually visit?

11. What is the size of your weekly grocery bill?
- 0-$5.00
- $5.00-$10.00
- $10.00-$15.00
- $15.00-$20.00
- $20.00-$25.00
- $25.00-$30.00
- $30.00-$35.00
- $35.00-$40.00
- $40.00-$45.00
- $45.00-$50.00
- $50.00-$55.00
- over $50.00

12. How many times did you shop for groceries last week?
Would you please examine the list below and number the five which you feel influence your decision to shop at any particular store the most. If, for example, you feel that ease of parking is most important to you, it should be numbered 1.

For the five items that you have numbered as important, please indicate which store you feel is best in that feature.

<table>
<thead>
<tr>
<th>low prices</th>
<th>advertising</th>
<th>near to your home</th>
<th>good service</th>
<th>near to your work</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Store Number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>good selection of brands</th>
<th>near to other stores</th>
<th>good delivery service</th>
<th>easy parking</th>
<th>good quality of meat</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>friendly employees</th>
<th>good quality of produce</th>
<th>lack of crowding</th>
<th>can shop quickly</th>
<th>gives trading stamps</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

14. If you have changed stores for your major grocery purchases in the past year, why did you do so?

________________________________________________________________________

15. What is the occupation of the head of the household? _______________________

16. How long have you lived at this location? _________________________________

17. How many are there in your family? _________________________________________

18. How old is the head of the household?

<table>
<thead>
<tr>
<th>10-20</th>
<th>20-30</th>
<th>30-40</th>
<th>40-50</th>
<th>50-60</th>
<th>60-70</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

19. What is your average annual income?

<table>
<thead>
<tr>
<th>0- $2,500</th>
<th>$2,500- $5,000</th>
<th>$5,000- $7,500</th>
<th>$7,500- $10,000</th>
<th>$10,000- $12,500</th>
<th>$12,500- $15,000</th>
<th>Over $27,500</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Blount, S.F., "The Spatial Distribution of Shoppers Around the Valley Shopping Center, St. Charles, Ill.," Professional Geographer, 16(5), 1964, pp. 8-17.


