Multiattribute shopping models and ridge regression analysis

Citation for published version (APA):
Timmermans, H. J. P. (1981). Multiattribute shopping models and ridge regression analysis. *Environment and Planning A*, *13*(1), 43-56. https://doi.org/10.1068/a130043

DOI:
10.1068/a130043

Document status and date:
Published: 01/01/1981

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Environment and Planning A, 1981, volume 13, pages 43-56

Multiattribute shopping models and ridge regression analysis

H J P Timmermans
Department of Architecture, Building and Planning, University of Technology, 5600 MB Eindhoven, The Netherlands
Received 27 November 1979, in revised form 22 April 1980

Abstract. Policy decisions regarding retailing facilities essentially involve multiple attributes of shopping centres. If mathematical shopping models are to contribute to these decision processes, their structure should reflect the multiattribute character of retailing planning. Examination of existing models shows that most operational shopping models include only two policy variables. A serious problem in the calibration of the existing multiattribute shopping models is that of multicollinearity arising from the fact that strong linear relationships among policy variables frequently occur in real world situations. This paper points at the technique of ridge regression analysis to overcome the problem of multicollinearity in the development of multiattribute shopping models. The use of ridge regression analysis is illustrated in an application of the multiplicative competitive interaction model to spatial shopping behaviour.

1 Introduction: problem setting

Spatial shopping behaviour must be considered as a very complex phenomenon resulting from an interplay of environmental, psychological, cultural, and socioeconomic factors. In order to grasp the main characteristics of this phenomenon scholars have attempted to build models which relate aggregate shopping patterns to a preselected set of independent or predictor variables. These models not only have a theoretical meaning, they have also been used in applied planning research to predict the effects of alternative policy decisions regarding retailing facilities on spatial consumer behaviour as well as spending between various competing shopping centres. Basically, these policy decisions concern manipulating a number of physical attributes of shopping centres such as floor space, parking facilities, type of shops, and relative location. In other words, these policy decisions are multiattribute decisions and it is the author's contention that, as a necessary although not sufficient condition, the structure of shopping models should reflect the multiattribute character of policy decisions regarding retailing facilities if mathematical model building is to contribute to the solving of planning problems. Evidently then, there is a strong need for multiattribute shopping models.

Curiously enough, however, most operational shopping models are not multiattribute models. Most attempts at modelling spatial shopping behaviour have remained direct modifications of the model developed by Huff (1963) in the early sixties. This model, essentially a production-constrained spatial interaction model, may be expressed as

$$p_{ij} = \frac{A_j}{d_{ij}^{\beta}} \Bigg/ \sum_j \frac{A_j}{d_{ij}^{\beta}} , \qquad (1)$$

where

p_{ij} is the probability that a consumer located at i will patronize shopping alternative j,
A_j is an attractiveness term for shopping alternative j,
d_{ij} is the distance-separation between residence zone i and shopping alternative j, and
β is a distance parameter to be estimated.

More recent versions of the Huff model differ only from the original model in terms of the measurement of the attractiveness term and the specification of the distance function as well as the method used for the calibration of the model (see for example Lakshmanan and Hansen, 1965; Lewis and Traill, 1968; Batty, 1971; Batty and Mackie, 1972; Gibson and Pullen, 1972; Openshaw, 1975; Stetzer, 1976; Smith et al, 1977; Pankhurst and Roe, 1978). Evidently, the production-constrained spatial interaction model includes only two policy variables, namely the attractiveness variable and relative location, implying that at best it can be used to assess the effects of policy decisions with respect to shifts in the locational pattern of the shopping centres vis-à-vis the residence zones and changes in, for example, the floor space of one or more of the shopping centres in the study area.

This limitation is shared by the studies designed in the tradition of Rushton's preference scaling approach (for example, Rushton, 1969a; 1969b; 1969c; 1971a; 1971b; 1976; Lentnek et al, 1975; Girt, 1976; Timmermans, 1979; see also, Pirie, 1976; MacLennan and Williams, 1979; Timmermans and Rushton, 1979), a major competing approach for the examination of aggregate spatial shopping patterns. An important step in this approach is the definition of locational types, combinations of two environmental stimuli serving as surrogate variables for, respectively, the attractiveness and the distance-separation of shopping alternatives. If the approach is used for applied research, preference scale values, describing the positioning of the locational types on a preference function, are related to these two stimuli. This again shows that at best the effects of only two policy variables can be determined.

Only recently, a number of operational multiattribute models have been developed, mostly for application outside the field of spatial shopping behaviour (for example, Cesario, 1973; 1974; 1975; 1976; Hudson, 1976; Ewing, 1976; 1978; Recker and Kostyniuk, 1978; Baxter, 1979a; 1979b; Baxter and Ewing, 1979; Tobler, 1979). Most of these models have attempted to derive an attractiveness parameter from observed spatial interaction patterns which may subsequently be explained in terms of a number of physical attributes of destinations by means of regression procedures. Other models have specified a priori a functional relationship between some measure of spatial behaviour and a set of predictor variables. Again, the effect of the predictor variables on the measure of spatial behaviour is estimated by means of standard regression procedures. Both types of models, however, are calibrated on data, usually collected in nonexperimental settings in which the relationships among the predictor variables cannot be controlled. Consequently, it frequently occurs that there are strong linear relationships among the predictor variables; multicollinearity appears to be a serious problem in the calibration of these types of multiattribute models.

The effects of multicollinearity on the least squares estimates of the regression coefficients are well-known. Multicollinearity may result in estimates of the regression coefficients with high variances and which, consequently, may be far removed from the true population values. In addition, the least squares estimates may be too large in absolute value and it is possible that some estimates will even be of the wrong sign. The resulting regression equation may thus be quite unreliable. Furthermore, it is well-known that multicollinearity may result in an unstable least squares solution. The estimates of the regression coefficients may change dramatically with slight alternations of the data. As a result, it is almost impossible to untangle statistically the effects of the individual predictor variables.

Spatial analysts have traditionally used principal components regression or stepwise regression analysis to overcome the problem of multicollinearity, although alternative procedures such as computation of all possible regressions and optimal subset selection using tree search algorithms exist (see, for example, Beale et al, 1967; Boyce et al, 1974). For example, in a recent article published in this journal Hubbard (1979) performed a principal components analysis among three image variables 'to minimise the possibility of markedly imprecise OLS (ordinary least squares) estimates resulting from a high degree of multicollinearity among these image variables'. Although these procedures may be useful in theory-oriented research, both the principal components regression and the stepwise regression approaches are, however, unsatisfactory from the point of view of applied planning research. In the stepwise regression approach variables are deleted from the model through some statistical rule. The decision regarding which variables to include is therefore rather arbitrary and, in addition, it is well-known that stepwise regression analysis does not always succeed in selecting the best subset of predictor variables in terms of maximizing explained variance. Still more important, however, is the fact that a number of variables are dropped from the model which is clearly at variance with the principal purpose of a multiattribute model in applied planning research, namely to predict the effects of all variables which are considered to be directly changeable through policy decisions. Principal components regression involves reducing the multicollinearity among the predictor variables by extracting principal components which subsequently serve as new predictor variables in a regression format. The interpretation of the principal components may however be very difficult and generally the components cannot be directly related to physical attributes of the destinations. If therefore the intent of the researcher is to forecast the impact of new proposed retailing developments, the usefulness of principal components regression analysis is limited. The usefulness of the procedure is further restricted by the fact that it cannot be employed in situations where the researcher for some reason has to perform a logarithmic transformation of the predictor variables before obtaining least squares estimates since the logarithmic transformation of negative components scores is undefined. This point is again clearly illustrated in Hubbard's recent paper (Hubbard, 1979), in which some 'hybrid' multiattribute model is developed, necessarily consisting both of logarithmic and of nonlogarithmic variables.

Given these general considerations, the starting point for the present paper is the existing need for the development of multiattribute shopping models, together with the problem of multicollinearity hindering such a development. Specifically, the present paper points at the technique of ridge regression analysis to overcome the problem of multicollinearity in the calibration of multiattribute shopping models based upon observed spatial interaction patterns in real world situations. In addition, the use of ridge regression analysis will be illustrated in a study of spatial shopping behaviour in Southeast Brabant in the Netherlands, that is, it will be shown how ridge regression analysis may be used to calibrate a multiplicative competitive interaction model.

2 Ridge regression analysis

2.1 *General outline*

Consider the standard multiple linear regression model

$$y = \mathbf{X}\beta + e\,, \qquad (2)$$

where

$\mathbf{X}$ is a fixed $(n \times m)$ matrix of full rank of observations on m predictor variables,

y is an $(n \times 1)$ column vector of observations on a dependent variable,

β is an $(m \times 1)$ column vector of parameters,

e is an $(n \times 1)$ column vector of error terms.

It is assumed that the elements of the matrices X^TX and X^Ty represent correlation coefficients. Further, it is assumed that

$$E(e) = 0 , \tag{3}$$

and

$$E(ee^T) = \sigma^2 I_n . \tag{4}$$

Under these conditions the ordinary least squares (OLS) estimators $\hat{\beta}$ of the parameters are given by

$$\hat{\beta} = (X^TX)^{-1}X^Ty . \tag{5}$$

Furthermore, it can be shown (Hoerl and Kennard, 1970a) that the total mean square error, $E(L^2)$, a measure of the square of the distance between an estimate and the true value of the parameter is

$$E[(\hat{\beta}-\beta)^T(\hat{\beta}-\beta)] = E(L^2) = \sigma^2 \operatorname{tr}(X^TX)^{-1} , \tag{6}$$

which is equivalent to

$$E(L^2) = \sigma^2 \sum_{i=1}^{m} \frac{1}{\lambda_i} , \tag{7}$$

where λ_i is the ith eigenvalue of the (X^TX) matrix, and

$$\lambda_{max} = \lambda_1 \geqslant ... \geqslant \lambda_m = \lambda_{min} > 0 .$$

In the absence of multicollinearity X^TX is an identity matrix and consequently the eigenvalues λ_i will all be equal to 1·0.

However, when near-multicollinearity occurs one or more of the eigenvalues will be small, resulting in a large value for $1/\lambda_{min}$ and hence $E(L^2)$. Under such circumstances the least squares estimates, although unbiased, may be far removed from the true population values. In addition, the least squares estimates may become very sensitive to the structure of the data and the addition of only a few more observations can lead to pronounced shifts in some of the regression coefficients. As a result it becomes hard and often impossible to obtain stable and precise estimates of all the individual parameters. Evidently then, this situation is a serious limitation when a researcher wishes to use the estimated coefficients for forecasting.

To overcome this problem of multicollinearity, Hoerl and Kennard (1970a) have introduced the technique of ridge regression analysis. Ridge regression analysis is a technique based upon

$$\hat{\beta}(k) = (X^TX+kI)^{-1}X^Ty \tag{8}$$

where k is a small nonnegative constant ($k > 0$).

The total mean square error of the ridge estimates is

$$E[L^2(k)] = E\{[\hat{\beta}(k)-\beta]^T[\hat{\beta}(k)-\beta]\} \tag{9}$$

$$= \sigma^2 \sum_{i=1}^{m} \left[\frac{\lambda_i}{(\lambda_i+k)^2}\right] + k^2\beta^T(X^TX+kI)^{-2}\beta \tag{10}$$

$$= \operatorname{tr}\{V[\hat{\beta}(k)]\}+b^2 \tag{11}$$

$$= \gamma_1(k)+\gamma_2(k) , \tag{12}$$

where b is the bias and $V[\hat{\beta}(k)]$ is the variance–covariance matrix of the ridge regression estimate. If $k = 0$, the ridge regression estimator reduces to the ordinary least squares estimator; if k increases, $\beta(k)$ tends to zero.

If $k > 0$, $\hat{\boldsymbol{\beta}}(k)$ is a biased estimator and the extent of the bias $[\gamma_2(k)]$ increases with increasing values of k, $\gamma_1(k)$—the variance of $\hat{\boldsymbol{\beta}}(k)$—however is a monotonic decreasing function of k. Hoerl and Kennard (1970a) and Theobald (1974) have shown the existence property that, if $\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\beta}$ is bounded, there exists a $k > 0$ such that

$$E[\mathrm{L}^2(k > 0)] < E[\mathrm{L}^2(k = 0)] \,. \tag{13}$$

In other words, ridge estimators are biased but they have better mean square error properties than ordinary least square estimators for some values of k.

2.2 *The selection of k*

The idea of ridge regression analysis is to select a value of k for which the reduction in total variance is not exceeded by the increase in bias. Hoerl and Kennard (1970a; 1970b) have suggested than an appropriate value of k may be determined from an inspection of a 'ridge trace', a simultaneous plot of the components of $\hat{\boldsymbol{\beta}}(k)$ versus k, together with some complementary statistics for $\hat{\boldsymbol{\beta}}(k)$. The ridge trace serves to portray the interrelationships between the predictor variables including the effects of these interrelationships on the estimation of $\boldsymbol{\beta}$.

In the presence of multicollinearity, the ridge trace may portray dramatic changes of the ridge estimators as k is slowly increased from zero. Large negative values may tend to zero, large positive values may decrease, some ridge estimators may even experience a change in sign. Eventually, the ridge estimators will stabilize, and Hoerl and Kennard (1970a; 1970b) recommended the selection of the smallest value of k for which $\hat{\boldsymbol{\beta}}(k)$ is stable. In addition, they suggested that one should check whether the variance-covariance matrix of $\hat{\boldsymbol{\beta}}(k)$ has "the general characteristics of an orthogonal system", the coefficients have reasonable absolute values with respect to the factors for which they represent rates of change, the coefficients with apparently incorrect signs at $k = 0$ have changed to have the proper sign, and the residual sum of squares has remained close to its minimum value.

The inspection of the ridge trace is an inherently subjective procedure which may be considered as a disadvantage. On the other hand, the ridge trace provides the researcher with a compact visual picture of the effects of nonorthogonality of $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ on the estimation of $\boldsymbol{\beta}$. Anyway, several authors have provided and evaluated alternative schemes for the automatic selection of k (see, for example, Marquardt, 1970; Goldstein and Smith, 1974; Farebrother, 1975; Guilkey and Murphy, 1975; Hemmerle, 1975; Hoerl et al, 1975; Marquardt and Snee, 1975; McDonald and Galarneau, 1975; Vinod, 1976; Swindel, 1976; Hoerl and Kennard, 1976; Dempster et al, 1977; Hemmerle and Brantle, 1978; Wichern and Churchill, 1978; Golub et al, 1979). Marquardt (1970), for example, suggested using a value of k for which the maximum variance inflation factor (VIF) is "between one and ten and closer to one". The VIF associated with each coefficient measures the degree by which the variance of that coefficient is inflated by the intercorrelations between the predictor variables.

Specifically, the VIFs are the diagonal elements of the inverse of the correlation matrix describing the correlations between the predictor variables. The VIFs will increase as these intercorrelations increase, and reduce markedly as a function of increasing k, the degree of reduction being dependent upon the strength of the intercorrelations.

2.3 *Generalized ridge regression analysis*

A natural extension of the ridge estimator is to consider a diagonal matrix $\mathbf{k}$ rather than the scalar k in $k\mathbf{I}$. In general, it is possible to reduce the general linear regression problem to a canonical form by applying an orthogonal transformation $\mathbf{P}$

such that

$$\mathbf{P}^{\mathrm{T}}(\mathbf{X}^{\mathrm{T}}\mathbf{X})\mathbf{P} = \Lambda \,, \tag{14}$$

where

Λ is a diagonal matrix of eigenvalues, λ_i, of $\mathbf{X}^{\mathrm{T}}\mathbf{X}$, and

$\mathbf{P}$ is a matrix whose columns are the corresponding eigenvectors.

The linear model may be written as

$$y = \mathbf{X}^{*}\alpha + e \,, \tag{15}$$

where

$$\mathbf{X}^{*} = \mathbf{XP} \,, \qquad \alpha = \mathbf{P}^{\mathrm{T}}\beta \,, \quad \text{and} \quad \mathbf{X}^{*\mathrm{T}}\mathbf{X}^{*} = \Lambda \,. \tag{16}$$

Then the general ridge estimation procedure is defined as

$$\hat{\alpha}(k) = (\Lambda + k)^{-1}\mathbf{X}^{*\mathrm{T}}y \,. \tag{17}$$

The total mean square error now becomes

$$E\{[\hat{\beta}(k) - \beta]^{\mathrm{T}}[\hat{\beta}(k) - \beta]\} = \sigma^2 \sum_{i=1}^{m} \frac{\lambda_i}{(\lambda_i + k_i)^2} + \sum_{i=1}^{m} \frac{k_i^2\alpha_i^2}{(\lambda_i + k_i)^2} \,. \tag{18}$$

Hoerl and Kennard (1970a) suggested an iterative procedure for minimizing the total mean square error by taking $k_i = \sigma^2/\alpha_i^2$.

At the first iteration, k_i is obtained by using ordinary least squares estimates for σ^2 and α_i. Then, a new value of $\hat{\alpha}(k)$ is computed by using equation (17). This process is repeated until no further significant changes in the elements of $\hat{\alpha}(k)$ occur. The generalized ridge estimators for β in the linear model are obtained as

$$\hat{\beta}(k) = \mathbf{P}\hat{\alpha}(k) \,. \tag{19}$$

Allen (1974), on the other hand, suggested a computational procedure utilizing nonlinear regression techniques for minimizing with respect to k the function he labelled Press:

$$\mathrm{Press}(k) = \sum_{i=1}^{n} [y_i - \hat{y}_{(i)}]^2 \,, \tag{20}$$

where $\hat{y}_{(i)}$ is the estimator of $E(y_i)$ obtained from equations (17) and (19) and excluding the ith observation.

Still other authors have suggested alternative procedures for estimating the elements of k. For further details, the reader is referred to the papers mentioned in section 2.2.

3 Structure and calibration of the multiplicative competitive interaction model

As has been noted in the introduction, the use of ridge regression analysis will be illustrated by means of the multiplicative competitive interaction model in a study of spatial shopping behaviour. The multiplicative competitive interaction model may be mathematically expressed as

$$p_{ij} = X_{1ij}^{\alpha_1} \dots X_{rij}^{\alpha_r} \bigg/ \sum_{j=1}^{t} (X_{1ij}^{\alpha_1} \dots X_{rij}^{\alpha_r}) \,, \tag{21}$$

$$p_{ij} = \prod_{k=1}^{r} X_{kij}^{\alpha_k} \bigg/ \sum_{j=1}^{t} \prod_{k=1}^{r} X_{kij}^{\alpha_k} \,, \tag{22}$$

where

p_{ij} is the probability that an individual located at i will choose alternative j,

X_{kij} is the value of the kth attribute of alternative j for individuals located at i,

t is the total number of alternatives,
r is the total number of attributes considered, and
α_k is a parameter for the kth attribute.

As will be evident from equation (22) the multiplicative competitive interaction model may be considered as a multiattribute extension of the production-constrained spatial-interaction model. Its structure corresponds with theories of spatial consumer behaviour and therefore the model might prove useful in applied research of spatial consumer behaviour, at least from a descriptive point of view. The structure of the multiplicative competitive interaction model is a simple one making a direct specification of the variables influencing spatial shopping behaviour possible. In addition, the model presupposes that spatial shopping behaviour essentially is a choice process between alternative shopping opportunities—the competitive component of the model—and that low values on one attribute of a shopping alternative cannot be compensated by high values on another attribute. If any of the attributes of a shopping opportunity contributing to its attractiveness is close to zero, the model will predict a very low proportion of consumers patronizing this shopping opportunity. This assumption appears to be a fruitful one from a theoretical point of view.

The multiplicative competitive interaction model has its roots in the marketing literature where it has been used initially for the description of competitive market behaviour (Kotler, 1965), the description of retail trading areas (Haines et al, 1972; Mahajan et al, 1977), determining brand share (Urban, 1969; Pessemier et al, 1971; Lambin, 1972), the examination of voting behaviour (Nakanishi et al, 1974) and the measurement of advertising and promotion effectiveness (Kuehn et al, 1966; Nakanishi, 1972). To the author's knowledge the model has rarely been used for the examination of aggregate spatial shopping patterns in the context of physical planning.

The multiplicative competitive interaction model may be calibrated in several different ways. Some authors have used a nonlinear least squares method which minimized the sum of squared residuals by a direct search technique. Others have tried to derive maximum likelihood estimates by Newton's method or by a direct search technique. More recently, Nakanishi and Cooper (1974) have questioned these calibration methods since none of them guarantees that the global maxima or minima will be found, they are costly in terms of computing time and, most importantly, in most cases the statistical properties of the estimates are unknown. Consequently, they have developed least squares estimation techniques for the multiplicative competitive interaction model. They have shown that the estimation of the parameter coefficients of the model can be obtained through transformation of equation (22) into the linear form:

$$\log\frac{p_{ij}}{p_i^*} = \sum_{k=1}^{r} \alpha_k \log\frac{X_{kij}}{X_{ki}^*} , \qquad (23)$$

where

$$p_i^* = \left(\prod_{j=1}^{t} p_{ij}\right)^{1/t} , \quad \text{and} \quad X_{ki}^* = \left(\prod_{j=1}^{t} X_{kij}\right)^{1/t} . \qquad (24)$$

Depending upon the specification of the error terms, ordinary least squares (OLS), generalized least squares (GLS) or iterative generalized least squares (IGLS) estimates may be obtained, mostly requiring little more than standard multiple regression analysis programs. Evidently, if the attributes, that is the predictor variables are highly intercorrelated, the effect of one particular predictor variable cannot be separated statistically from the effects of the remaining predictor variables. Ridge regression analysis represents one possible way out of this problem of multicollinearity.

4 The application

4.1 *Data base*

In order to illustrate the use of ridge regression analysis and assess the potentials of the multiplicative competitive interaction model in the study of spatial shopping behaviour, data on shopping behaviour were collected for 1795 respondents located in Southeast Brabant, the Netherlands. The matrix of shopping trips was constructed by asking respondents to mention the shopping centres they usually patronize as well as the frequency of their visits during a one month time period for a number of shopping goods and a number of convenience goods. In the present study only the data for the shopping goods were used. Next, the individual data on consumer shopping behaviour were aggregated to yield 33 residence zones and 37 shopping zones within the study area. The resulting 33 x 37 matrix of shopping trips was then converted into proportions which were subsequently interpreted as choice probabilities.

Apart from data on shopping trip distributions the multiplicative competitive interaction model requires input data on the factors contributing to the attractiveness of the shopping zones. On the basis of a small literature search seven variables were selected. All of these variables can be considered as policy variables. The variables selected are given in table 1. They represent physical manifestations of the size, convenience, and choice dimensions of the attractiveness of the shopping zones.

The measurement of the distance-separation between the residence zones and the shopping zones was in travel-time units. First, a transportation network covering the study area was constructed. Each residence zone was joined to the nearest node of the network. Distances were then measured from these nodes. Next, the 33 x 37 matrix of travel-times between the residence zones and the shopping zones was computed by taking into account the spatial distribution of the respondents within the residence zones, the speed of the various transport nodes on the links of the network and the distribution of the transport nodes for each residence zone.

An important aspect of calibrating the model on the basis of real travel-time distances is that the parameters of the model are strongly influenced by the geometry of the study area. In addition, a serious problem is the fact that a particular distance does not always represent a comparable choice situation. For example, consumers located in the central parts of the study area will probably have an attractive shopping centre at a relatively short distance whereas consumers located at the periphery of the study area will have to travel farther to reach the nearest shopping zone. In other words, consumers located at the periphery of the study area will probably discount the distance to the nearest shopping opportunity and then view the distances to the other more distant shopping opportunities. To include this notion in the model, the travel-time distance from a residence zone to the nearest shopping opportunity was set to 1·0, and the travel-time distances to the other shopping zones were measured accordingly.

Table 1. The selected attributes of the shopping zones.

Variable	Description
X_1	number of parking facilities
X_2	number of shops
X_3	variety in shops (functional complexity)
X_4	number of employees
X_5	number of superstores
X_6	distance to shopping zone

4.2 *Calibration and results*

The multiplicative competitive interaction model was calibrated using the ridge regression approach. Obviously, in applying ridge regression techniques it is recommendable to use different procedures to select reasonable values of k since in this way one can have an idea of the reliability of the obtained values and of the properties of the different procedures.

In addition, it is possible to examine the stability of the estimated regression coefficients through the derivation of the variance and estimation components and proportions (for details, see Belsley, 1976; and Moulaert, 1979). For the moment, however, we use the single parameter ridge trace approach since this approach ascertains that the effects of multicollinearity on the estimated regression coefficients may be visually examined, a clear advantage given the primarily illustrative purpose of this application, which is not shared by the more sophisticated procedures.

Table 2 gives the correlation coefficients between the predictor variables. Table 2 clearly shows that there are some strong intercorrelations between variables X_2 (number of shops), X_3 (variety of shops) and X_4 (number of employees). These strong intercorrelations point to the existence of multicollinearity. This is also illustrated by the variance inflation factors which are respectively 3·37, 584·12, 147·25, 338·99, 2·87 and 1·24.

The results of the ridge regression analysis are presented in table 3 and figure 1. Examination of table 3 gives rise to the following interpretations and observations. First, table 3 and figure 1 indicate that the coefficients for $k = 0$ are very likely to be overestimated; collectively the regression coefficients are not stable. Table 3 shows that the initial values of the regression coefficients alter rapidly as k increases. Especially variable X_4 (number of employees) drops dramatically in absolute value. The absolute value of the coefficients of variables X_1 (number of parking facilities) and X_6 (distance to shopping zone) also decreases as k slowly increases.

Variables X_2 (number of shops), X_3 (variety in shops), and X_5 (number of superstores) even experience a change in sign.

Figure 1 shows that the regression coefficients tend to stabilize between $k = 0{\cdot}2$ and $k = 0{\cdot}4$, whereas the performance of variable X_3 is hard to interpret throughout the whole range of k. This finding is further emphasized by the variance inflation factors (table 4). Table 4 shows that the variance inflation factors are strongly reduced for increasing values of k. At $k = 0{\cdot}4$ the variance inflation factors are approximately equal to 1·0, a value characteristic for an orthogonal system. Between $k = 0{\cdot}4$ and $k = 0{\cdot}9$ the regression coefficients of variables X_1 to X_5 are positive, indicating that the probability of selecting a particular shopping opportunity is positively related to the number of parking facilities, the number of shops, the variety in shops, the number of employees and the number of superstores. The coefficient of variable X_6 (the distance to the shopping zone) remains negative. This result is in correspondence with usual theoretical assumptions about spatial consumer behaviour.

Table 2. The correlations among the predictor variables.

	X_1	X_2	X_3	X_4	X_5	X_6
X_1	1·000	0·475	0·380	0·539	0·628	0·381
X_2	0·475	1·000	0·989	0·995	0·342	0·236
X_3	0·380	0·989	1·000	0·973	0·234	0·191
X_4	0·539	0·995	0·973	1·000	0·405	0·270
X_5	0·628	0·342	0·234	0·405	1·000	0·369
X_6	0·381	0·236	0·191	0·270	0·369	1·000

The probability of selecting a particular shopping opportunity decreases with the additional distance, as compared to the distance to the nearest shopping opportunity, a consumer has to travel in order to visit this shopping opportunity.

Second, table 3 provides information about the relative importance of the predictor variables. Since the model is expressed in terms of standardized variables, the size of the regression coefficients indicates the relative importance of the corresponding predictor variables.

Table 3 reveals that variables X_3 and X_4 lose their position as the most important predictor variables. At $k = 0 \cdot 2$, the relative importance of variables X_1 (the number of parking facilities) and X_6 (the distance to the shopping zone) significantly exceeds the relative importance of the remaining predictor variables in explaining the probability of selecting a particular shopping opportunity. This might indicate, although a causal interpretation might be misleading, that consumers organize their

Table 3. Ridge regression results.

k	Regression coefficients						Coefficient of determination
	X_1	X_2	X_3	X_4	X_5	X_6	
0·0	0·19665	−0·05161	−2·24460	2·73075	−0·05943	−0·71200	0·722
0·1	0·41397	0·10424	−0·11962	0·23125	0·20232	−0·60016	0·583
0·2	0·36964	0·09748	−0·04131	0·16528	0·19778	−0·52966	0·526
0·3	0·33276	0·08597	−0·01200	0·14102	0·18801	−0·47374	0·481
0·4	0·30299	0·08318	0·00369	0·12767	0·17784	−0·42822	0·446
0·5	0·27859	0·08108	0·01348	0·11877	0·16825	−0·39042	0·415
0·6	0·25823	0·07927	0·02009	0·11216	0·15946	−0·35855	0·390
0·7	0·24906	0·07763	0·02479	0·10690	0·15148	−0·33132	0·367
0·8	0·22611	0·07607	0·02821	0·10525	0·14425	−0·30780	0·348
0·9	0·21320	0·07590	0·03077	0·09874	0·13769	−0·28726	0·330

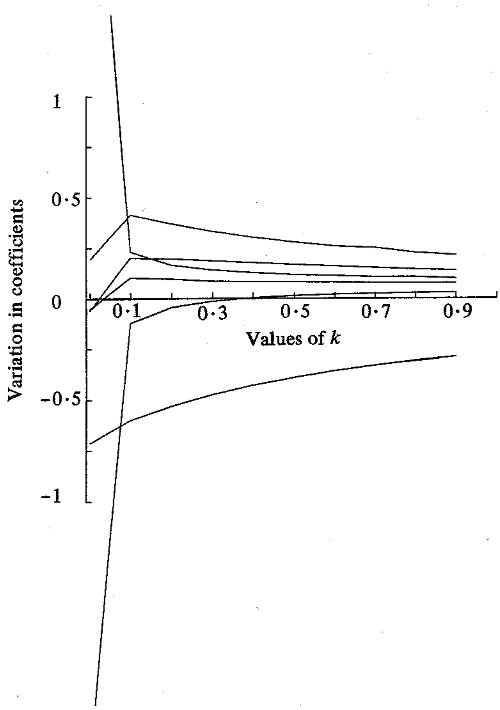

Figure 1. Ridge trace.

shopping trips mainly on the basis of the availability of parking space and distance separation between their residence zone and the shopping opportunities. Clearly, this result might have strong implications in the context of multiattribute planning since it implies that the strength of the effect of new retailing developments on existing shopping centres is to a relatively high degree dependent upon the projected parking facilities. In terms of modelling spatial shopping behaviour this result might indicate that the conceptual basis of the production constrained spatial interaction model relying heavily upon the size component of the attractiveness of shopping centres might not be the most suitable one for describing consumer spatial choice behaviour. At least this result clearly illustrates the usefulness of a multiattribute shopping model in the study of spatial shopping behaviour.

Finally, table 3 gives the predictive power of the multiplicative competitive interaction model in the context of shopping behaviour for durable goods in South-east Brabant, the Netherlands. If k equals 0 the model accounts for 72·2% of the variance in the trip probabilities between the 33 residence zones and the 37 shopping zones. However, this result is misleading because of the presence of multicollinearity among the predictor variables. This point is also shown in table 3. The percentage of explained variance in the shopping trip probabilities drops strongly with an increasing value of k.

At a value of 0·4 for k with relatively stable and interpretable regression coefficients the model only accounts for about 44·6% of the variance in the trip probabilities. Considering the multiattribute property of the model, this result might seem disappointing. On the other hand, the present analysis has been carried out primarily for illustrative purposes. Introduction of the transport mode used and identifying a more relevant zoning system might improve the empirical performance of the model. Further research will be needed to elaborate these points.

Table 4. The variance inflation factors.

k	X_1	X_2	X_3	X_4	X_5	X_6
0·0	3·367	584·117	147·253	338·999	2·873	1·240
0·1	1·624	6·721	6·079	6·761	1·460	1·073
0·2	1·339	3·442	3·202	3·494	1·219	0·962
0·3	1·149	2·336	2·200	2·374	1·058	0·872
0·4	1·011	1·779	1·689	1·807	0·939	0·798
0·5	0·904	1·442	1·378	1·464	0·846	0·736
0·6	0·819	1·216	1·168	1·234	0·771	0·683
0·7	0·749	1·054	1·016	1·068	0·710	0·638
0·8	0·691	0·932	0·901	0·944	0·657	0·598
0·9	0·642	0·836	0·811	0·846	0·613	0·563

5 Concluding remarks

The central contention of the paper has been that the development of multiattribute shopping models is a necessary condition for a better understanding of spatial shopping behaviour and for the assessment of the impact of new retailing developments on spatial shopping behaviour and the functioning of existing retailing facilities in the study area. The development of such models is however hindered by the existence of multicollinearity among the predictor variables of the model arising from the fact that such models are calibrated on the basis of real world data related to aggregate spatial shopping patterns and objective attractiveness variables which are hard to control statistically. The present paper has pointed at ridge regression analysis to portray the sensitivity of the regression estimates to problems of multicollinearity.

The application of ridge regression analysis may lead to models which are more effective in prediction as a result of an increase in precision of the regression coefficients in terms of mean square error, although at the cost of introducing some bias into the estimates. However, if a multiattribute shopping model is used for forecasting it is evident that the calibration of the model should be focused on smaller mean square errors rather than the least squares criterion. In addition ridge regression analysis represents a way of retaining all predictor variables in the model which, again, may be considered as a necessary condition if mathematical model building is to contribute to the evaluation of alternative retailing planning programs.

The application of the ridge regression analysis technique to the calibration of the multiplicative competitive interaction model has shown that this model might be a useful model in this respect. Ridge regression analysis can be used to estimate the effects of the predictor policy variables which are all retained in the model. On the other hand, the application has shown that only two variables, the number of parking facilities and the distance to the shopping zones, seem to be significant in explaining shopping trip probabilities. This result points at still another potential use of ridge regression methods, that is, ridge regression methods may be useful in identifying the best variables contributing to the attractiveness of shopping opportunities, for example in the context of the traditional production constrained spatial interaction models. Although ridge regression analysis is not developed explicitly for the purpose of variable selection, the technique has the property of an inherent deletion of variables, namely those whose regression coefficients go to zero with increasing values of k. Hoerl and Kennard (1970a) suggested that such variables should be eliminated from the model since they cannot hold their predicting power.

Although the present paper has illustrated the use of ridge regression analysis to shopping model development, it must be evident that ridge regression methods might be useful for parameter estimation in other fields of applied planning research where multicollinearity presents serious problems. It is clearly a good alternative to the commonly used stepwise regression methods and with respect to computational considerations ridge regression is quite efficient since reasonably good estimates can be obtained by using only a few values of k, mostly requiring little more than minor modifications of standard multiple regression analysis programs.

Acknowledgements. The author would like to thank Professor David E Boyce and an anonymous referee for particularly constructive comments.

References

Allen D M, 1974 "The relationship between variable selection and data augmentation and a method for prediction" *Technometrics* **16** 125-127

Batty M, 1971 "Exploratory calibration of a retail location model using search by golden section" *Environment and Planning* **3** 411-432

Batty M, Mackie S, 1972 "The calibration of gravity, entropy and related models of spatial interaction" *Environment and Planning* **4** 206-233

Baxter M J, 1979a "The application of logit regression analysis to production-constrained gravity models" *Journal of Regional Science* **10** 171-177

Baxter M J, 1979b "The interpretation of the distance and attractiveness components in models of recreational trips" *Geographical Analysis* **11** 311-315

Baxter M J, Ewing G O, 1979 "Calibration of production constrained trip distribution models and the effect of intervening opportunities" *Journal of Regional Science* **10** 319-330

Beale E M L, Kendall M G, Mann D W, 1967 "The discarding of variables in multivariate analysis" *Biometrika* **54** 357-366

Belsley D A, 1976 "Multicollinearity: diagnosing its presence and assessing the potential damage it causes least-squares estimation" WP-154, National Bureau of Economic Research, Cambridge, Mass

Boyce D E, Farhi A, Weischedel R, 1974 *Optimal Subset Selection: Multiple Regression, Interdependence and Optimal Network Algorithms* (Springer, Berlin)

Cesario F J, 1973 "A generalized trip distribution model" *Journal of Regional Science* **13** 233-248
Cesario F J, 1974 "More on the generalized trip distribution model" *Journal of Regional Science* **14** 389-397
Cesario F J, 1975 "A new method for analyzing outdoor recreation trip data" *Journal of Leisure Research* 7 200-215
Cesario F J, 1976 "Alternative models of spatial choice" *Economic Geography* **52** 363-373
Dempster A P, Schatzoff M, Wermuth N, 1977 "A simulation study of alternatives to ordinary least squares" *Journal of the American Statistical Association* **72** 77-90
Ewing G O, 1976 "Environmental and spatial preferences of interstate migrants in the United States" in *Spatial Choice and Spatial Behavior* Eds R G Golledge, G Rushton (Ohio State University Press, Columbus, Ohio) pp 249-270
Ewing G O, 1978 "The interpretation and estimation of parameters in constrained and unconstrained trip distribution models" *Economic Geography* **54** 264-273
Farebrother R W, 1975 "The minimum mean square error linear estimator and ridge regression" *Technometrics* **17** 127-128
Gibson M, Pullen M, 1972 "Retail turnover in the East Midlands: a regional application of a gravity model" *Regional Studies* **6** 183-196
Girt J L, 1976 "Some extensions to Rushton's spatial preference scaling model" *Geographical Analysis* **8** 137-156
Goldstein M, Smith A F M, 1974 "Ridge type estimators for regression analysis" *Journal of the Royal Statistical Society* **36B** 284-291
Golub G H, Heath H, Wahba G, 1979 "Generalized cross-validation as a method for choosing a good ridge parameter" *Technometrics* **21** 215-223
Guilkey D K, Murphy J L, 1975 "Directed ridge regression techniques in cases of multicollinearity" *Journal of the American Statistical Association* **70** 769-775
Haines G H, Simon L S, Alexis M, 1972 "Maximum likelihood estimation of central-city food trading areas" *Journal of Marketing Research* **9** 154-159
Hemmerle W J, 1975 "An explicit solution for generalized ridge regression" *Technometrics* **17** 309-314
Hemmerle W J, Brantle T F, 1978 "Explicit and constrained generalized ridge estimation" *Technometrics* **20** 109-120
Hoerl A E, Kennard R W, 1970a "Ridge regression: biased estimation for nonorthogonal problems" *Technometrics* **12** 56-68
Hoerl A E, Kennard R W, 1970b "Ridge regression: application to nonorthogonal problems" *Technometrics* **12** 69-92
Hoerl A E, Kennard R W, 1976 "Ridge regression: iterative estimation of the biasing parameter" *Communications in Statistics* **5** 77-88
Hoerl A E, Kennard R W, Baldwin K F, 1975 "Ridge regression: some simulations" *Communications in Statistics* **4** 105-123
Hubbard R, 1979 "Parameter stability in cross-sectional models of ethnic shopping behaviour" *Environment and Planning A* **11** 977-992
Hudson R, 1976 "Linking studies of the individual with models of aggregate behaviour: an empirical example" *Transactions of the Institute of British Geographers* **1** 159-174
Huff D, 1963 "A probability analysis of shopping center trading areas" *Land Economics* **53** 81-90
Kotler P, 1965 "Competitive strategies for new product marketing over the life cycle" *Management Science* **12** 104-119
Kuehn A A, McGuire T W, Weiss D L, 1966 "Measuring the effectiveness of advertising" in *Proceedings Fall Conference American Marketing Association* (American Marketing Association, Chicago, Ill.) pp 185-194
Lakshmanan T R, Hansen W G, 1965 "A retail market potential model" *Journal of the American Institute of Planners* **31** 134-143
Lambin J J, 1972 "A computer-on-line marketing mix model" *Journal of Marketing Research* **9** 119-126
Lentnek B, Lieber S R, Sheskin I, 1975 "Consumer behavior in different areas" *Annals of the Association of American Geographers* **65** 538-545
Lewis J P, Traill A C, 1968 "An assessment of shopping potential and the demand for shops" *Town Planning Review* **38** 317-326
MacLennan D, Williams N J, 1979 "Revealed space preference theory: a cautionary note" *Tijdschrift voor Economische en Sociale Geografie* **70** 307-309
Mahajan V, Jain A K, Bergier M, 1977 "Parameter estimation in marketing models in the presence of multicollinearity: an application of ridge regression" *Journal of Marketing Research* **14** 586-591

Marquardt D W, 1970 "Generalized inverses, ridge regression, biased linear and nonlinear estimation" *Technometrics* **12** 591-612
Marquardt D W, Snee R D, 1975 "Ridge regression in practice" *The American Statistician* **29** 3-19
McDonald G C, Galarneau D I, 1975 "A Monte Carlo evaluation of some ridge-type estimators" *Journal of the American Statistical Association* **70** 407-416
Moulaert F, 1979 "Estimability problems in the univariate homoscedastic linear regression model" WP-12, Department of Regional Science, University of Pennsylvania, Philadelphia, Pa
Nakanishi M, 1972 "Measurement of sales promotion effect at the retail level—a new approach" in *Proceedings Fall Conference American Marketing Association* (American Marketing Association, Chicago, Ill.) pp 338-343
Nakanishi M, Cooper L G, 1974 "Parameter estimation of a multiplicative competitive interaction model—least squares approach" *Journal of Marketing Research* **11** 303-311
Nakanishi M, Cooper L G, Kassarjian H M, 1974 "Voting for a political candidate under conditions of minimal information" *Journal of Consumer Research* **1** 36-43
Openshaw S, 1975 *Concepts and Techniques in Modern Geography 4. Some Theoretical and Applied Aspects of Spatial Interaction Shopping Models* (Geo Abstracts, Norwich)
Pankhurst I C, Roe P E, 1978 "An empirical study of two shopping models" *Regional Studies* **12** 727-748
Pessemier E, Burger P, Teach R, Tigert D, 1971 "Using laboratory brand preference scales to predict consumer brand preferences" *Management Science* **17** 371-385
Pirie G H, 1976 "Thoughts on revealed preference and spatial behaviour" *Environment and Planning A* **8** 947-955
Recker W W, Kostyniuk L P, 1978 "Factors influencing destination choice for the urban grocery shopping trip" *Transportation* **7** 19-33
Rushton G, 1969a "Analysis of spatial behavior by revealed space preference" *Annals of the Association of American Geographers* **59** 391-400
Rushton G, 1969b "Temporal changes in space preference structures" *Proceedings of the Association of American Geographers* **1** 129-132
Rushton G, 1969c "The scaling of locational preferences" in *Department of Geography Studies in Geography 17: Behavioral Problems in Geography* Eds K Cox, R G Golledge (Northwestern University, Evanston, Ill.) pp 197-227
Rushton G, 1971a "Behavioral correlates of urban spatial structure" *Economic Geography* **47** 49-58
Rushton G, 1971b "Preference and choice in different environments" *Proceedings of the Association of American Geographers* **3** 146-150
Rushton G, 1976 "Decomposition of space preference functions" in *Spatial Choice and Spatial Behavior* Eds R G Golledge, G Rushton (Ohio State University Press, Columbus, Ohio) pp 119-133
Smith A P, Whitehead P J, Mackett R L, 1977 "The utilization of services" in *Models of Cities and Regions* Eds A G Wilson, P Rees, C M Leigh (John Wiley, New York) pp 323-403
Stetzer F, 1976 "Parameter estimation for the constrained gravity model: a comparison of six methods" *Environment and Planning A* **8** 673-683
Swindel B F, 1976 "Good ridge estimators on prior information" *Communications in Statistics* **5** 985-997
Theobald C M, 1974 "Generalisation of mean square error applied to ridge regression" *Journal of the Royal Statistical Society* **36B** 103-106
Timmermans H J P, 1979 "A spatial preference model of regional shopping behaviour" *Tijdschrift voor Economische en Sociale Geografie* **70** 45-48
Timmermans H J P, Rushton G, 1979 "Revealed space preference theory: a rejoinder" *Tijdschrift voor Economische en Sociale Geografie* **70** 309-312
Tobler W R, 1979 "Estimation of attractivities from interactions" *Environment and Planning A* **11** 121-127
Urban G L, 1969 "Mathematical modeling approach to product line decisions" *Journal of Marketing Research* **6** 40-47
Vinod H D, 1976 "Application of ridge regression methods to a study of Bell system scale economies" *Journal of the American Statistical Association* **71** 835-841
Wichern D W, Churchill G A, 1978 "A comparison of ridge estimators" *Technometrics* **20** 301-311