TN 28: A COMPARISON OF THE REGRESSION RESULTS OF APPLYING TWO ATTRACTIVENESS FACTORS TO ESTIMATE PARK USE

By H.K. Cheung

ABSTRACT

In this paper, the results of using two attractiveness factors in regressions designed to estimate park use are compared. The two attractiveness factors thus compared were derived by two independent approaches. Cheung's measure was derived by applying a formula that defines attractiveness using characteristics of a park and popularity of activities that can be participated in within the park. Cesario's measure, in contrast, is defined by visitor flows from various origins to parks. The explanatory power associated with each attractiveness measure was found to depend on the particular form of the estimating equation used. This finding led to the conclusion that neither the Cheung, nor the Cesario, attractiveness factor performs better than the other.

INTRODUCTION

Cheung, in an attempt to explain the variance found in a set of Saskatchewan main destination day-user data (see TN 1) developed the following model using multiple regression analysis:

(1) V(i, j) = 1.33 + (120P(i) - 36.6P(i)A(i,j) + 1.25Tj - 104)/g(D(i,j))

WHERE for origin, i, and destination, j, and

V(i,j) = hundreds of day-use visitor vehicles going to park j from origin i per season;

P(i) = population, in thousands;

D(i,j) = road distance in miles, from i to j;

A(i,j) = alternative factor = $\Sigma_k D(i,k)^{-1/2}$ for, D(i,k), distances to alternatives with $k \leq 100$ miles and for $k \neq j$ (the sum is for parks within 100 miles but not including j)

Tj = attractiveness of park j: and

g(D(i,j))	= D(i,j)**1/2	for	0	<	D(i,j)	< 20
g(D(i,j))	= D(i,j)	for	20	≤	D(i,j)	< 55
g(D(i,j))	= D(i,j)**3/2	for	55	≤	D(i,j)	

How Cheung defined and measured Tj , the attractiveness of park j, is found in TN 1. It should be noted that Tj explained less than one percent of the total variance in V(i,j), the dependent variable. Doubts were thus raised about the soundness of the deductive ("ad hoc") way Tj was formulated (e.g. see TN 9 and reviews of Ch. 2 and 3).

RESEARCH PURPOSE

Given different approaches to measuring attractiveness (e.g., see TN 9) a question seemed reasonable to the author. That question is: Given the Saskatchewan data and the same functional form, would Cesario's approach of measuring park attractiveness explain more variance than Cheung's?

The purpose of this paper is to consider answers to that question. To that end a comparison is made of the results obtained by using Cheung's (TN 1) and the Cesario's (TN 4) attractiveness measures.

THE CHEUNG AND THE CESARIO ATTRACTIVENESS MEASURES

Logically, there are two very different approaches to defining park attractiveness. These are the inductive the deductive approaches. In the inductive approach specific components (elements) of attractiveness are designated, often subjectively. They may consist of the number of picnic tables, the length of a swimming beach, water quality and so on, depending on the type of users under study. Each component is given a score, based on some quantity or quality criteria and a numerical value. The "attractiveness of a park" is calculated by algebraically combining the scores. The Cheung attractiveness measure used in the Saskatchewan day-use model study (TN 1) is an example of an inductively defined measure. Briefly, Cheung took the attractiveness of a park with regard to a particular use (e.g., main destination day-use) to be measured by a weighted sum of scores for the day-use facilities offered at that park. Weights were defined on the basis of popularity of activities among the other considerations, but information on the day use of the park was not used in defining its attractivity. In contrast, the deductive approach to defining a measure of park attractiveness involves computing a numerical value reflecting attractiveness of a park based on analysis of something reflecting actual behaviour such as visitor flows (e.g., day-use traffic volume). Both Cesario (TN 4) and Ross (TN 2) have defined deductive ways of estimating attractiveness measures.

COMPARISON OF ATTRACTIVENESS MEASURES

Given that the purpose of this paper is to consider explaining variance in visitor flows to destinations, attractiveness measures must have certain properties. Because Ross's attractiveness measure is ordinal, it should not be used in a regression equation. Therefore, only Cesario and Cheung measures are compared in this analysis.

The attractiveness measures as developed by Cheung and Cesario represent two theoretical constructs designed to determine the relative attractiveness of parks. Although the measures were developed using different approaches, they both were intended to produce a "formula" for calculating park attractiveness which a park planner or model builder could use to advantage.

It was recognized that with the Cesario attractiveness measure available for Saskatchewan, it would be possible to use the data to compare the measures' "performance" with that of Cheung. So Cesario's method was used to derive attractiveness estimates for the twelve Saskatchewan parks for which Cheung had made estimates (see TN1). To do so, it was necessary to apply Cesario's model (Equation 4 in its logarithmic form) to the Saskatchewan main destination day-use data used previously by Cheung. The Cheung and the Cesario attractiveness values of the parks under study are presented in Table 1.

It should be noted that the absolute values of the Cheung or Cesario park attractiveness factors are not important. Rather, it is the relative values that are meaningful. For example, the relative attractiveness of Good Spirit provincial park to that of Moose Mountain provincial park is 0.68 (= 76.56/113.11) according to the Cheung attractiveness scale and 0.62 (= .93/1.50) according to the Cesario attractiveness scale. (The Cheung and the Cesario attractiveness scales both have the properties of direction and magnitude and are said to be interval scales.)

After the two sets of attractiveness values were obtained, they were used in separate regression runs to explain the number of visits from different origin areas to the twelve Saskatchewan parks considered in this analysis. Actually, three sets of regression runs using the attractiveness values and other regressors were made. There are, of course, an infinite number of regressions that could be made containing the attractiveness factor. The three following functional forms (Equations 2, 3, and 4) chosen in this study are some of the commonly used functional forms.

(2) $E(V(i,j))=C_0+(C_1P()+C_2P(i)A(i)+C_3T_k+C_4)/g(D(i,j))$

(3) $E(V(j)/P(i))=C_5+(C_6A(i)+C_7T_k)/g(D(i,j))$

(4) $E(log((V(j)+1)/P(i,j))=C_8+C_9log(A(i))+C_{10}log(D(i,j))+C_{11}log(T_k)$

WHERE

$T_k = T_j$ = the Cheung attractiveness factor,

T_f = the Cesario attractiveness factor,

CO, •••., C11 = parameters to be estimated, and

E( ) = the expected value of the quantity in ( )'se The equations derived were:

(5) $V(i,j) = 1.33 + (120.31P(i) = 36.60P(i)A(i)+1.25T_j- 104.56)/g(D(i,j))$(Cheung)

(5a) $V(i, j) = 1.32 + (116.54P(i) - 34.60P(i)A(1) + 106.19T_f- 118.08)/g(D(i,j))$ (Cesario)

(6) $V(i,j)/P(i) = 0.07+(0.97T_j-7.56A(i))/g(D(i,j))$(Cheung)

(6a) $V(i,j)/P(i) = 0.11 + (70.38T_j- 6.33A(i))/g(D(i,j))$ (Cesario)

TABLE 1: THE CHEUNG AND THE CESARIO ATTRACTIVENESS FACTORS OF TWELVE SASKATCHEWAN PARKS*

Park	Cheung	Cesario
Buffalo Pound		0.55
Cypress Hill		2.02
Duck Mountain	1	1.52
Echo Valley	1	0.63
Good Spirit		0.93
Green Water		1.06
Prince Albert		1.46
Moose Mountain	1	1.50
Pike Lake		0.83
Rowan's Ravine		0.65
Battleford's	1	1.59
Besant		0.51

***The Pearson product moment correlation coefficient is 0.28.**

(7) $log((V(i,j) + 1)/P(i)) = 2.65 + 0.19 log(T_j)- 1.83 logD(i,j)- 0.581 log(A(i))$ (Cheung)

(7a) $log((V(i,j) + 1)/P(i)) = 3.10 + 1.21 log(T_f)-1.88 logD(i,j) - 0.33 log(A(i))$ (Cesario)

In the first set of regressions (see Tables 2 and 3) both the Cheung and Cesario attractiveness factors, when combined with the explanatory variable defined by the distance function g(D(i,j)), did almost equally poorly in explaining the variance found in the Saskatchewan day-use data. The reason may be that the regressors P(i)/g((D(i,j)) and P(i)A(i)/g(D(i,j)) explained so much variance (about 90 percent) in the day-use data that there was not much variance left for the attractiveness regressor to explain.

In the second set of regressions (see Tables 4 and 5) the Cheung attractiveness factor, when combined with the distance function g(D(i,j)), explained 69 percent of the total variance in the data. On the other hand, the Cesario attractiveness factor, when combined with the same distance function, explained 66 percent of the total variance. The reason why the attractiveness regressors had so much explanatory power when the dependent variable was defined as participation rate is not clear. The author's conjecture is that it may have something to do with the constant variance assumption implicit in ordinary least squares (OLS) regression that was used in this study. Certainly, in the Saskatchewan main destination day-use data, large flows were associated with large population centers. Thus, when the flows were "weighted" by the corresponding populations, the homogeneous variance assumption of OLS was more nearly met than when they were not weighted. (On this problem, see TN 19).

TABLE 2: STATISTICS ON THE REGRESSION COEFFICIENTS OF EQUATION 5*

Regression Coefficient of	Coefficient Value	Error	F-Value	R^2
constant	-1.33			
P(i)/g(D(i,j))	120.31	5.80	429.80	0.8416
P(i)A(i)/g(D(i,j))	- 36.60	3.12	137.81	0.9029
1/g(D(i,j))	-104.56	27.30	14.63	0.9088
T_j /g(D(i,j))	1.25	0.40	9.85	0.9048

***The F-value and the standard error of estimate of this equation are 562.93 and 7.59 respectively, with 226 error degrees of freedom. Also all regression coefficients are significant at the one per cent probability level and all have the expected signs.**

TABLE 3: STATISTICS ON THE REGRESSION COEFFICIENTS OF EQUATION 5a*

Regression Coefficient	Standard Value	Error	F-Value	R^2
constant	1.32			
P(i)/g(D(i,j))	116.54	5.61	430.99	.8416
P(i)A(i)/g(D(i,j))	- 34.06	3.02	127.62	.9031
1/g(D(i,j))	-118.08	31.42	14.13	.9049
T_f/g(D(i,j))	106.19	34.13	9.68	.9089

***The F-value and the standard error of estimate of this equation are 563.37 and 7.59 respectively, with 226 error degrees of freedom. Also all regression coefficients are significant at the one per cent probability level and all have the expected signs.**

TABLE 4: STATISTICS ON THE REGRESSION COEFFICIENTS OF EQUATION 6*

Regression Coefficient	Standard Value	Error	F-Value	R^2
constant	0.07			
T_j/g(D(i,j))	0.97	0.06	297.22	.6862
A(i)/g(D(i,j))	-7.56	2.98	6.41	.6947

***The F-value and the standard error of estimate for this equation are 259.46 and 1.32 respectively, with 228 error degrees of freedom. Also all regression coefficients are significant at the one per cent probability level and all have the expected signs.**

It is seen from Tables 4 and 5 that the Cheung attractiveness regressor, Tj /g(D(i,j)), had a better overall performance than the Cesario attractiveness regressor, T_f/g(D(i,j)), in terms of a larger ratio of the regression coefficient to its standard error (.97/.06 = 16.17 versus 70.38/4.42 = 15.92), and a higher increase in the R^2 (.6862 versus .6610). It is also seen that Equation 6 incorporating the Cheung attractiveness regressor has a smaller standard error of estimate than the Equation 6a incorporating the Cesario attractiveness regressor.

TABLE 5: STATISTICS ON THE REGRESSION COEFFICIENTS OF EQUATION 6a*

Regression Coefficient	Value	Standard Error	F-Value	R2
constant	0.11			
$T_f/g(D(i,j))$	70.38	4.42	253.71	.6610
$A(i)/g(D(i,j))$	-6.33	3.12	4.11	.6670

***The F-value and the standard error of estimate of this equation are 228.53 and 1.38 respectively, with 228 error degrees of freedom. Also all regression coefficients are significant at the one per cent probability level and all have the expected signs.**

In the third set of regressions the functional form of the estimating equation was double-logarithmic. This time, as seen from Tables 6 and 7, there was a decrease in explanatory power of the regressor containing the Cheung attractiveness factor and the Cheung attractiveness regressor was not found to be significant. There was, however, a marked improvement in the explanatory power of the regressor incorporating the Cesario attractiveness factor. This was to be expected since the Cesario attractiveness factor was derived using an Equation that is very similar to Equation 7a, the estimating equation.

TABLE 6: STATISTICS ON THE REGRESSION COEFFICIENTS OF EQUATION 7*

Regression Coefficient	Value	Standard Error	F-Value	R^2
constant	2.65			
Log(D(i,j))	-1.82	0.09	411.78	0.6385
Log(A(i))	-0.58	0.11	27.30	0.6791
$Log(T_j)$	0.19	0.14	1.84	0.6817

***The F-value and the standard error of estimate of this equation are 162.06 and 0.39 respectively, with 227 error degrees of freedom. Also all regression coefficients are significant at the one per cent probability level except that of log T_j which is not significant, and all have the expected signs.**

TABLE 7: STATISTICS ON THE REGRESSION COEFFICIENTS OF EQUATION 7a*

Regression Coefficient	Value	Standard Error	F-Value	R^2
constant	3.10			
Log(D(i,j))	-1.88	0.08	562.18	.6385
Log(A(i))	-0.33	0.10	10.05	.7389
$Log(T_f)$	1.02	0.13	64.30	.7500

***The F-value and the standard error of estimate of this equation are 226.95 and 0.34 respectively, with 227 error degrees of freedom. Also all regression coefficients are significant at the one per cent probability level and all have the expected signs.**

SUMMARY AND CONCLUSIONS

In this paper two methodologies designed to measure the attractiveness of a park have been evaluated. The Cheung attractiveness measure was based on an ad hoc definition procedure developed by an inductive approach so that an overall rating of a site could be arrived at by considering a set of site characteristics and services offered. The Cesario attractiveness measure was a component of a trip-making model, the number of trips made from an origin to a destination considering spatial operation. It is estimated, based on a function of the Characteristics of the origin, the characteristics of the destination, and the spatial separation of the origin and the destination. Cesario's measure of attractivity was thus described as being defined deductively.

Using Equation 2, 3, and 4 to compare the effectiveness of the two attractiveness factors, using the regression results, and based on using the increase in the R^2 value as the main criterion for judging the performance of the Cheung and Cesario attractiveness factors, it is difficult to say whether one performs better than the other. The regression results presented showed that the efficiency of the attractiveness factors depended on the particular form of the estimating equation used. One can see the problem by noting that when the dependent variable was defined as participation rate, the Cheung attractiveness regressor, defined as T(c)/g(D(I,j)), slightly outperformed the Cesario attractiveness regressor, defined as Tf /g(D(i,j)). However, when the dependent variable used was log((V(i,j)1)/P(i)), the Cesario attractiveness regressor, defined as log Tf , greatly outperformed the Cheung attractiveness regressor, defined as log T(c), in the sense that the former explained more than ten percent of the total variance whereas the latter explained less than one percent.