COWLES FOUNDATION DISCUSSION PAPER NO. 56

Consumer Conformity in a Sample of College Students*

Donald D. Hester

August 13, 1958

* This project was developed and the data were collected in 1954-5 by the Yale University Graduate Seminar in Quantitative Economic Research conducted by Professors Tobin and Guthrie. Students in the seminar were R. B. Bristol, Jr., A. N. Katz, R. C. Porter, R. N. Rosett, and T. Y. Shen. The following results were obtained from these data. I am indebted to Professors Watts and Tobin for their many helpful suggestions and criticisms.

Consumer Conformity in a Sample of College Students

Donald D. Hester

The existence of tendencies towards group conformity in consumption in our society is obvious. This paper examines the nature of some of these tendencies in the sport coat consumption of a sample of Yale undergraduates. More explicitly, we assume the existence of reference groups to which an individual endeavors to conform. Employing certain characteristics of these groups as independent variables, we then attempt to explain the behavior of an undergraduate as represented by the price paid for a recent sport coat, the rate of purchase, and the change in number of sport coats owned while at Yale.

Interdependence among individuals in consumer behavior has been discussed in the literature of economics and social psychology. Veblen (7) dealt with this topic in his theory of conspicuous consumption. Liebenstein (5) classified various forms of consumer interdependence. The relative income hypothesis of Duesenberry (3) and Brady and Friedman (2) is a further example. Whyte (8) has attempted to analyze the retailing of consumer durables in terms of neighborhood effects. Prais and Houthakker (6) have developed a model of interdependence. In addition, "The Doctrine of Prestige, Suggestion, and Imitation in Social Psychology" (1) seems to be an analog of interdependence of consumer behavior in economics.

It is assumed in this study that identification with a reference group is not conditional upon the character of an individual's consumption. An individual identifies himself with a group because of his personality, attitude, interests, etc. Association with a group then causes him to want to emulate its members in certain visible items of consumption.

The problem of measuring the influence of reference groups is complex. The reference group of an individual is subjective and therefore difficult to identify. The individual himself probably cannot describe his reference group, that is, he cannot assign relative weights to those individuals or groups who influence him. The relevant reference group may be different for different kinds of behavior; for many items of consumption there may be no reference group. Statistically, the problem of estimating simultaneous relationships exists. For example, if A and B belong to a reference group and A buys a new sport coat, then B will want to buy one also. B's purchase can induce the same desire in A . Basically, we cannot identify and, therefore, estimate the structural relationships.

One way to approximate an individual's consumption reference groups in a sampling survey is to ascertain the formal or informal groups to which he belongs. The most eligible candidates would be those with which he has frequent contact. Another method would be to study those groups with which he has infrequent contact, but feels a strong desire to join. This class would include exclusive social sets, movie stars, athletes, etc. We shall employ the former method.

Two kinds of information concerning group characteristics that might influence an individual's behavior can be obtained. The first is objective information about the group, known from other sources or estimated from a sample of the members of the group. The second is the subjective estimate of the individual himself. The individual's perceptions of the group may be inaccurate, but the effects we are seeking are decidely subjective and objective accuracy is less important than relevance to the individual's decision. For this reason we shall rely on subjective or "perceived" estimates.

An objection to the use of this type of information has been made by S. E. Asch (1). Consider the question: "What would be your estimate of the price your friends usually pay for a sport coat?" Suppose that the respondent did not know the price of his friends' sport coats. What would be his response? Asch has suggested that because the question appears in the interview, the respondent feels that he is expected to know the answer. If he does not have the answer, he will make a guess at it. A likely guess would be the price of a recently purchased sport coat. If we were using the perceived price of sport coats to explain the price which a respondent pays for sport coats, the resulting correlation would be misleading. Only further research on respondent behavior in such situations can refute this suggestion. Here it is necessary to assume that the "Asch effect" is not relevant.

A consumer will be more group conscious the more conspicuous a commodity is.

> "For ties, nylon stockings, and television sets or, let us say ostentatious commodities in general, there will be some degree of interdependence which must be taken into account if the results of demand studies are to be properly interpreted." [Prais, S. J. and Houthakker, H.S.; The Analysis of Family Budgets, p. 19.]

The sample employed in this study was collected in February 1955 from students enrolled in undergraduate economics courses at Yale. The working sample is composed of 265 observations and includes no freshmen or transfer students. While the sample is neither random nor representative of college students in general, it is believed that this investigation will provide some insight into the relations which exist and the appropriate methodology for exposing them.

The questionnaire gathered information on a large variety of subjects relating to undergraduate life. It emphasized student ownership and purchases of sport coats and cars. The present report concerns only sport coats. A sport coat was defined to include ordinary sport coats and a suit coat which is worn with slacks.

Information concerning a number of potential reference groups, of varying size, was obtained in the questionnaire, including friends of the respondent and the entire body of undergraduates at Yale.*

---

* Sample questions concerning these reference groups are as follows: "Let's consider your friends now -- the men you frequently join for eating, talking, and other social occasions. What would be your estimate of the price your friends usually pay for a sport coat? On the average, approximately how much would you say a Yale undergraduate spends for a year at Yale?"

Of the two, friends of the respondent seems more meaningful and this study examined that grouping.

Frequently in analyzing consumer behavior, the dependent variable is defined to be the quantity of a good purchased per unit time. But conformity may occur in other dimensions of consumer behavior, - i.e. stock or perhaps price. An analysis explaining only rate of purchase would be incomplete.*

---

* The distinction between rate of purchase and consumption should be emphasized. Rate of purchase implies the number of units bought per unit time. Consumption involves the using up or expending of a good. In this paper it is not assumed that these rates are proportionate.

The questionnaire provides the following information on sport coats:

1) Number of sport coats owned.

2) Change in number of sport coats owned while at Yale.

3) Price paid for the most recent sport coat.

4) Satisfaction with most recent sport coat.*

5) Adequacy of stock of sport coats.*

6) Time since purchase of most recent sport coat.

---

* a)"Now think of the newest sport coat you have here at Yale -- Do you think now you should have chosen a different sport coat or are you perfectly satisfied with the one you chose? Please explain. b) Is the number of sport coats you now have less than adequate, adequate, or more than adequate? Please explain."

The satisfaction and adequacy variables are not quantifiable. Furthermore, as will become evident, these variables are not convenient for the analytical model employed. They are undesirable as dependent variables because of their discrete character. For these reasons they have been ignored in the present analysis.

The first two variables differ in that the first includes the number of sport coats owned by the respondent upon coming to Yale. This pre-Yale stock is better regarded as an explanatory variable for subsequent behavior at Yale than as a variable to be explained. The second is employed in this analysis as a dependent variable in order to examine interdependence in stock of sport coats owned.

A second dimension in which interdependence might be evident is the rate at which sport coats are purchased. This can be derived from the sixth variable, time T since purchase of most recent sport coat. Consider an individual having a constant rate of purchase of sport coats r per month. This rate implies that the individual purchases a sport coat every 1/r months. Assume that on a given date T is distributed rectangularly among individuals with a common rate of purchase r . If we had a sample of n such individuals, the mean of the observed T u_T would tend towards $\frac{1}{2r}$. Therefore

$$r = \frac{1}{2u_T}$$

The third dependent variable included in the analysis was P , price paid for a recent sport coat.

All variables subsequently used are defined as follows:

$Q-Q_{-1}$	Change in number of sport coats owned while at Yale.
T	Age of most recent sport coat - in months.
P	Price of most recent sport coat - in dollars.
Y	Total college year spending (net of transportation costs to and from home) - in dollars.
C	Yale class - freshmen = 1, sophomores = 2, etc.
Q_{-1}	Pre-Yale stock.
W	Graduate of public high school = 1, other = 0 .
F	Member of social fraternity = 1, other = 0 .
D	Number of dates during previous semester.

R Number of activities over past two years.

P' Perceived friends' average price of sport coats - in dollars.

Y' Perceived friends' average total expenditure - in dollars.

Q' Perceived friends' average stock of sport coats.

Y'' Perceived mean total spending of Yale undergraduates.

P'' Perceived Yale undergraduate average price of sport coats.

The hypotheses were that if the effects of certain variables (including total spending) were eliminated, then the change in number of sport coats ($Q - Q_{-1}$) would be larger and/or price (P) of most recent sport coat would be higher and/or time (T) since most recent sport coat purchase would be shorter if:

1) Perceived friends' number of sport coats owned (Q') were higher.

2) Perceived friends' average price (P') were higher.

3) Perceived friends' spending (Y') were higher.

4)-6) The first three effects would be larger if the respondent were in a higher Yale class. (CY', CQ', and CP' would have positive coefficients.)

7)-9) Similarly, they would be larger if the respondent were in a lower total spending bracket. (YY', YQ', and YP' would have negative coefficients.)

10)-12) And they would be larger if the respondent felt his reference group's mean income was above the Yale average. (Y'Y'/Y'', Y'Q'/Y'', and Y'P'/Y'' would have positive coefficients.)

The following were employed as additional variables; total spending (Y), Yale class (C), pre-Yale stock (Q_{-1}), fraternity membership (F), dating (D), number of extra-curricular activities (R), and whether or not the respondent was a graduate of a public high school (W).*

---

* It should be noted that a controversial adage has been dealt a withering blow in this formulation. We assume that dating and fraternity and activity membership determine the number of sport coats owned. Clothes do not make the man!

Since the dependent set is composed of three variables, the statistical testing of these hypotheses is possible with ordinary regression techniques only if we are willing to assign arbitrary weights to its three constituents. A better procedure is to estimate coefficients for both explaining and explained variables, by the method of canonical correlation developed by Professor Hotelling (4). Briefly, Hotelling defines two linear variates, $u = \sum_1^m \gamma W$ and $v = \sum_1^p \delta Y$, where the W's constitute a set of m independent or explaining variables and the Y's constitute a set of p dependent or explained variables. The coefficients, γ's and δ's, are estimated so that the correlation between the two variates is at a maximum. Thus the linear function of the W's calculated with the γ's is the best predictor of the linear function of the Y's calculated with the δ's. The correlation between u and v is defined to be the canonical correlation.

A basis for selecting between alternative hypotheses is provided by computing the corresponding canonical correlations. It is possible to measure the increase in the canonical correlation (ΔR_c^2) attributable to the addition of a new explaining variable. A ratio of this increase to the total unexplained variation $(1-R_c^2)$, divided by the appropriate degrees of freedom, is analogous to the F-ratio used for testing the influence of an additional independent variable in a regression.* The ratio thus computed was tested against the

* The decision as to how many degrees of freedom belong to the unexplained variation was complicated by the necessity to take into account the use of data to estimate the γ's and δ's. The number of observations minus the number of estimated parameters (sum of γ's and δ's) was considered the appropriate number of degrees of freedom.

F-distribution, although the applicability of the F-test to canonical correlation has not been proven.

What are the implications of using canonical correlation rather than regression? Consider the case where each of the dependent variables is regressed upon a set of independent variables.* In this situation

* This is a special case of the previously mentioned arbitrary set of weights. Here we assign weights of zero to two of the dependent variables and a weight of one to the third.

it is possible to appraise the effects of an independent variable on any one dependent variable. But it is not possible to examine the importance of an independent variable in explaining the dependent set. By employing canonical correlation, we argue that the various dependent variables can be explained by a set of independent variables whose coefficients differ only by a factor of proportionality. Thus a restriction is placed upon the coefficients of independent variables which does not exist when estimating regressions for each of the dependent variables. Unfortunately, it is not possible to examine the significance of relationships between an independent variable and a particular dependent variable when using canonical correlation.

What can canonical correlation suggest in terms of economic behavior? Suppose that there are two types of consumers in the population being examined.* One type will respond to increases in the price

* The supposition need not be this discrete. It is possible that there are gradations between stock responses and price responses where individuals respond in both dimensions. It is assumed that all points fall on a line.

of their friend's sport coats by owning a larger stock of sport coats. The second will pay a higher price for sport coats in response to the same stimulus, but will not change its stock of coats. While a simple regression of one of these variables (say stock) might have a significant multiple correlation coefficient, it is likely that a regression on a weighted sum of both price and stock would yield a larger multiple correlation. If the canonical correlation greatly exceeds the largest multiple correlation of a simple regression, then we can infer that the supposed population structure exists. If this is not the case, then either individuals increase both stock and price when perceived price rises or they uniformly increase either stock or price.

The procedure employed in this study may be described as follows. Canonical correlations were computed for the various previously specified hypotheses. Those explaining variables which were significant in the F-test were examined in a regression analysis in which each of the dependent variables was regressed on a set of independent variables. Our interpretations will depend upon the significance of relationships in the regressions and the relative magnitudes of the canonical and multiple correlations.

The iterative computations necessary for obtaining the canonical relations were performed on an IBM 650 computer. The principal results are recorded in Table 1.

Table 1

Results of Canonical Correlation Analysis

variable	one	two	three	four	five
			canonical equation number		
	coefficients in explained set				
$Q - Q_{-1}$	1.000	1.000	1.000	1.000	1.000
T	- .194	.012	.013	- .093	- .533
P	.733	- .061	.484	.701	-14.439
	canonical correlations				
	.235	.271	.492	.258	.508
	coefficients in explaining set				
Y	.010	.000	.003	.004	- .075
C	1.634	- .107	.955	1.890	-85.910
Q_{-1}	.450	- .500	- .104	.424	-17.063
W	-3.024	.495	- .623	-3.443	20.506
F	1.553	- .118	- .989	.945	27.173
D	- .035	- .003	.003	- .035	- .275
R	.423	.018	.147	.370	- 2.388
Q'		.467			
P'			.456		-13.207
Y'				.008	
CQ'					13.140

The results in Table 1 may be interpreted as follows. When each of the variables Q', P', and Y' was added to the set of explaining variables in equation one, the change in R_c^2 was significant at the one percent level. However, if P' were included, then the addition of Q' and Y' did not significantly change R_c^2. When P' was included, the addition of CQ' (class multiplied by perceived friends' stock of sport coats) was significant. The other previously specified hypotheses were rejected at the one percent level when P' was included.

Table 2*

	Regression equations			Canonical equation two
$Q-Q_{-1}$	1.000			1.000
T		1.000		.012
P			1.000	- .061
R^2 or R_c^2	.252	.040	.240	.271
Y	.001** (.000166)	- .002 (.00130)	.013** (.00225)	.000
C	.034 (.114)	.357 (.897)	2.380 (1.547)	- .107
Q_{-1}	- .407** (.0607)	.238 (.475)	1.560 (.820)	- .500
W	.245 (.171)	.262 (1.340)	-4.045 (2.311)	.495
F	.011 (.165)	.512 (1.296)	2.215 (2.236)	- .118
D	- .005 (.00701)	.003 (.0549)	.039 (.0948)	- .003
R	.047 (.0455)	- .280 (.357)	.424 (.615)	.018
Q'	.383** (.0598)	- .974 (.469)	-1.566 (.809)	.467

* Standard errors for regression coefficients are listed in parentheses. ** indicates significance at the .01 level.

By inspection of Table 2, it is evident that Q' was significant only in the first regression equation. Those individuals who have the higher estimates of their friends' stocks of sport coats tend to have larger increases in their own sport coat stock. Stock conformity provides an interesting variation on Leibenstein's "Bandwagon effect." The latter states that when an individual perceives aggregate demand rising, he shifts his demand curve to the right. Emphasis on conformity in stocks rather than flows may be a constructive result for future examinations of interdependence. With the appropriate assumptions, it may be argued that we have observed a "bandwagon effect."*

---

* If r_d = rate of discard, $\frac{\partial q}{\partial t}$ = rate of purchase, t_1 = number of years at Yale, t_2 = number of years that friends have been buying sport coats, and $\frac{\partial q'}{\partial t}$ = perceived average rate of purchase by friends, it can be shown that $\frac{\partial q}{\partial t} = g(\frac{\partial q'}{\partial t})$ with the following argument. Assume r_d is proportional to $\frac{\partial q}{\partial t}$ - i.e. $r_d = k\,\frac{\partial q}{\partial t}$.

$$Q - Q_{-1} = a\,Q'$$

$$(1-k)t_1(\frac{\partial q}{\partial t}) = a(1-k)f(t_2, \frac{\partial q'}{\partial t}) \quad \text{assuming } t_2 \text{ independent of } \frac{\partial q'}{\partial t}$$

$$\frac{\partial q}{\partial t} = a\,\frac{t_2}{t_1} \cdot f(\frac{\partial q'}{\partial t})$$

$$\frac{\partial q}{\partial t} = g(\frac{\partial q'}{\partial t})$$

Clearly the canonical correlation only slightly exceeds the largest multiple correlation.

Table 3 suggests that P' is significantly related to P in the third regression. When perceived price is higher, an individual will tend to pay a higher price for his most recent sport coat. This result is to be distinguished from the Veblen effect. The latter states that a consumer will buy more of a good if the price which he is generally believed to have paid for it is higher. R_c^2 does not greatly exceed R^2 .

Table 3

	Regression equations			Canonical equation three
$Q-Q_{-1}$	1.000			1.000
T		1.000		.013
P			1.000	.484
R^2 or R_c^2	.147	.025	.484	.492
Y	.001**	- .002	.005**	.003
	(.000187)	(.00138)	(.00195)	
C	.078	.235	1.805	.955
	(.122)	(.902)	(1.272)	
Q_{-1}	- .258**	- .156	.322	- .104
	(.0593)	(.438)	(.618)	
W	.458**	- .217	-2.227	- .623
	(.182)	(1.343)	(1.895)	
F	- .076	.627	-1.903	- .989
	(.180)	(1.332)	(1.879)	
D	- .004	.001	.015	.003
	(.0750)	(.0555)	(.0782)	
R	.036	- .256	.237	.147
	(.0486)	(.360)	(.507)	
P'	.017	- .191	9.077**	.456
	(.00773)	(.572)	(.0809)	

The variable Y' is positively related to P in the third regression in Table 4. The price paid for a recent sport coat will be higher if perceived friends' total spending is larger. If we associate quality with price, the implication is that ceteris paribus an individual will buy better quality sport coats if a reference group has a higher per capita total expenditure. He apparently will not buy or own more coats. Again the canonical correlation does not greatly exceed the largest multiple correlation coefficient.

Table 4

	Regression equations			Canonical equation four
$Q-Q_{-1}$	1.000			1.000
T		1.000		- .093
P			1.000	.701
R^2 or R_c^2	.131	.028	.255	.258
Y	.001** (.000293)	- .004 (.00214)	.004 (.00364)	.004
C	.089 (.124)	.309 (.904)	2.610 (1.536)	1.890
Q_{-1}	- .247** (.0596)	- .162 (.436)	.936 (.741)	.424
W	.405 (.183)	- .313 (1.342)	-5.530** (2.278)	-3.443
F	- .006 (.180)	.381 (1.315)	1.407 (2.233)	.945
D	- .005 (.00755)	.002 (.0553)	.043 (.0939)	- .035
R	.039 (.0490)	- .264 (.359)	.437 (.609)	.370
Y'	.000 (.000311)	.002 (.00228)	.011** (.00387)	.008

Table 5 suggests that CQ' (class multiplied by perceived quantity) is positively related to $Q-Q_{-1}$. This may be interpreted as further evidence of conformity in stock of sport coats. This effect is stronger when the respondent is in a higher Yale class. The negative relationship between C and $Q-Q_{-1}$ is probably attributable to the tradition at Yale whereby some sophomores purchase a college blazer. This is an addition to their sport coat stock which is not likely to be duplicated in their junior and senior years in college.*

---

* I am indebted to the members of the seminar in quantitative economic research (mentioned in the first footnote) for this suggestion.

The negative relationship between CQ' and P was not expected. No obvious explanation of this result seems to exist.

The possibility that an individual has more than one reference group must not be ignored. Only a superficial test of this suggestion can be made here. To those variables in canonical equation five, a new explaining variable, P'' (perceived average undergraduate price of sport coats), was added.* The hypothesis was analogous to hypothesis

---

* The question was, "What would be your estimate of the amount the average Yale student pays for a sport coat?"

3.) above. The change in the canonical correlation was not significant. Further examination of interdependence is necessary in order to establish if multiple reference groups exist.

Table 5

	Regression equations			Canonical equation five
$Q-Q_{-1}$	1.000			1.000
T		1.000		- .533
P			1.000	-14.439
R^2 or R_c^2	.266	.039	.507	.508
Y	.001** (.000173)	-.002 (.00138)	.005** (.00191)	- .075
C	- .606** (.155)	1.810 (1.229)	5.841** (1.706)	-85.910
Q_{-1}	- .398** (.0592)	.167 (.469)	1.148 (.650)	-17.063
W	.318 (.170)	.105 (1.348)	-1.402 (1.871)	20.506
F	- .074 (.167)	.624 (1.325)	-1.910 (1.840)	27.173
D	- .005 (.00697)	.003 (.0552)	.019 (.0766)	- .275
R	.045 (.0452)	- .279 (.358)	.179 (.497)	-2.388
P'	.015 (.00719)	- .016 (.0569)	.916** (.0790)	-13.207
CQ'	.150** (.0233)	- .346 (.184)	- .887** (.256)	13.140

Canonical correlation has been employed to examine the relations between an independent variable and a set of dependent variables. It provides a method for examining behavior where more than one type of response exists. It seems likely that canonical correlation may have been inappropriate for the present analysis. This follows from the fact that in no case do regressions of each dependent variable on the independent set yield coefficients differing solely by a factor of proportionality. Nevertheless, in every case the largest multiple correlation coefficient was only slightly smaller than the canonical correlation. This implies that the hypothesized social structure does not exist. In the case of an increase in P' an individual will either both pay a higher price for a recent coat and own more coats or solely pay more for a new sport coat. An analogous argument applies for an increase in Q' .

This experiment has yielded results which confirm the existence of conformity tendencies in a sample of Yale undergraduates. Both a price and a stock effect were found. In general, variations in Q' are effective in explaining differences in sport coat ownership, while variations in P' are effective in describing the price paid for a recent sport coat. Differences in Y' tend to be positively related to the price paid for a recent coat.

If these phenomena exist in the American population, many demand studies could be subject both to inconsistent estimates and to a rather serious autocorrelation of residuals. Gregariousness in consumer behavior may exist in small towns, professions, social organizations, etc. Further investigation is desirable in order to obtain greater knowledge of the nature and extent of conformity tendencies.

BIBLIOGRAPHY

1. Asch, S. E. "The Doctrine of Suggestion, Prestige, and Imitation in Social Psychology" Psychological Review, Vol. 55, Sept. 1948.

2. Brady, D. S. and Friedman, R. D. "Savings and the Income Distribution" Conference on Research in Income and Wealth, Vol. 10, N. B. E. R., 1947.

3. Duesenberry, J. S. Income, Saving, and the Theory of Consumer Behavior. Cambridge: Harvard University Press, 1952.

4. Hotelling, H. "Relations Between Two Sets of Variates," Biometrica. Cambridge: Cambridge University Press, Vol. 28, 1936.

5. Leibenstein, H. "Bandwagon, Snob, and Veblen Effects in the Theory of Consumer's Demand" The Quarterly Journal of Economics. Cambridge: Harvard University Press, 1950.

6. Prais, S. J. and Houthakker, H. S. The Analysis of Family Budgets. Cambridge: Cambridge University Press, 1950.

7. Veblen, T. The Theory of the Leisure Class. New York: The Vanguard Press, 1912.

8. Whyte, W. H., Jr. "The Web of Word of Mouth" Consumer Behavior. New York: New York University Press, Vol. 11, 1955.