WORKING PAPERS

THE RISK SEGMENTATION HYPOTHESIS:

A REINTERPRETATION OF THE BOCZAR STUDY

David M. Barton

WORKING PAPER NO. 56

March 1982

BUREAU OF ECONOMICS
FEDERAL TRADE COMMISSION
WASHINGTON, DC 20580

The Risk Segmentation Hypothesis:

A Reinterpretation of the Boczar Study

by

David M. Barton
Staff Economist
Bureau of Economics
Federal Trade Commission

The author wishes to thank Mr. Chaicho C. Wang for supplying data, and John Calfee, George Hay, and especially Douglas Greer for comments on earlier drafts.

I. Introduction

In a recent issue of this Journal, Gregory Boczar presented an empirical test of market segmentation in consumer credit, specifically "whether competition between banks and finance companies is limited by market segmentation on the basis of risk" (2, p. 245). According to Boczar's interpretation of the results, the risk segmentation hypothesis is not supported.

The study has considerable practical interest, since it bears directly on the question of how the relevant market should be defined when appraising the competitive consequences of mergers between lending institutions. The specific question addressed by the Boczar study--whether or not banks and finance companies compete in the same market--arose in a 1979 Section 7 (Clayton Act) suit brought by the Department of Justice to block the merger of Household Finance Company and American Investment Corporation (HFC/AIC merger). To win under Section 7, the DOJ was required to show that the merger would substantially lessen competition--a requirement which under guiding precedents would be met by showing that the merger would substantially increase concentration in the "relevant market." The DOJ argued for a narrow market definition (the business of making personal cash loans by finance companies), while the defense argued for a broader market definition (consumer installment credit) which would include consumer installment lending by S&L's, credit unions, retail stores, and most importantly, banks. Since the decision in this case so clearly depended on which market definition was accepted by the

court, the litigation was confined to this single issue.[1] Because of the conclusions of the Boczar study and its direct relevance to the question at trial, it was used as evidence by the defense.

In this note, I show that Bozcar's results support a conclusion opposite to the one he drew. While risk segmentation may be of limited relevance and thus an unfruitful working hypothesis, the purpose of this note is to clarify the appropriate use of Boczar's empirical methodology.

To test the risk segmentation hypothesis, Boczar employed probit analysis to discriminate between individuals in a 1970 sample of debtors having personal loans at either banks or finance companies. A number of socioeconomic variables plausibly related to riskiness and frequently included in credit scoring models were used as predictor variables. The sense of this approach is straightforward. If the discriminating ability of the model is "good," the risk segmentation hypothesis is supported. Conversely, if the discriminating ability is "poor," the hypothesis is "disconfirmed." Boczar's conclusion that "the evidence on borrower characteristics does not support the risk segmentation hypothesis" (2, p. 254) rests on the results reproduced in table 1, which show the proportion of the total sample and each borrower category separately, correctly and incorrectly classified.

(Table 1 goes here)

The link between these results and the conclusion drawn by Boczar should give pause for two reasons. First, it is hardly obvious that an overall error rate of 25 percent is "poor." Boczar never confronts this issue, though given the approach adopted, it would appear to be central. The second puzzle concerns the result which Boczar finds conclusive, namely, that "two out of every three finance company customers have 'risk-related' characteristics which we determined were highly similar to those of bank customers" (2, p. 254). Note, however, from table 1 that 93 percent of bank borrowers, which comprise 69 percent of the sample, were correctly classified. In view of this, Boczar's conclusion would appear to be that apples are indistinguishable from oranges, despite the fact that oranges are readily distinguished from apples.

In the following section, I show that the central result of Boczar's study--the high proportion of finance company borrowers erroneously classified as bank borrowers--is a red herring. In section III, I provide a statistical test of the classification results. The null hypothesis that the risk-related variables do not discriminate between borrowers is rejected at a confidence level exceeding .999999. Section IV repeats the analysis using more recent data.

II. The Effects of Sampling Proportions on Classification Error.

This section examines the relationship between prior probabilities (sample proportions) of two dichotomous categories and classification error. An example is provided using pseudo data. Following that, Boczar's results are revised using estimates of the population proportions of bank and finance company borrowers.

A. An Example

Hypothetical data on type B and type F individuals is given in table 2.

(Table 2 goes here)

Let the values 1,0 represent the presence or absence of risk-related characteristics X_1 and X_2. The frequency distribution of X_1, X_2 for the two individual categories is given in columns N_B and N_F, where 10 observations are assumed for each group.[2] While it is apparent that type B individuals are on average less risky, discrimination between the two groups using X_1 and X_2 will clearly be imperfect.

Using these data, two samples were constructed. For sample #1, the observations on the F's were combined with two replications of the observations on the B's to give a sample of size 30 with sample proportions $\pi' = \{\pi_B' = 2/3, \pi_F' = 1/3\}$. Following Boczar's methodology, probit analysis was used to estimate the conditional

probability that an individual belonged to the B group, and individuals were then classified according to: B if $P_B > .5$; F if $P_B < .5 (1-P_B = P_F > .5)$. The results are shown in the top portion of table 3. By design, the data were such that the results using sample #1 would have the same pattern as those of Boczar's study.

(Table 3 goes here)

For the second experiment, the same procedure was followed except that sample #2 was constructed using the observations on the B's and two replications of observations on the F's to give sample of 30 with proportions $\pi' = \{\pi_B' = 1/3, \pi_F' = 2/3\}$.

Before turning to the results, consider Boczar's speculation regarding the way his results would have been affected had his sample contained a greater proportion of finance company borrowers. After pointing out that finance company borrowers were probably underrepresented in his sample, he offers the following conjecture (2, p. 254):

> "For the sake of argument, suppose the sample contained an even split of borrowers from the two populations. Assuming that the percentages of misclassification from the holdout sample do not change (9.01 percent for bank borrowers and 68.06 percent for finance company borrowers), then the percentage of total misclassification would rise from 27.15 percent to 38.54 percent."[3]

From this it is clear that Boczar believed the effect of altered sample proportions can be inferred by simply reweighting the percentage of misclassifications for each category.

Our second experiment tests this conjecture. If Boczar is correct, the within-category error rates will be unaffected and the the total error rate will rise to 1/3(10%) + 2/3(60%) = 43 1/3%. Instead, we see from the bottom half of table 2 that the within-category error rates are substantially altered, while the overall error rate is only slightly affected.

B. Revised Results

Information provided by Boczar in his study under consideration here and in a previously published paper can be used to derive estimates of the population proportions of debtors having personal loans at banks and finance companies. Using these estimates his results are then revised.

In support of his suspicion that bank debtors appear overrepresented in the sample, Boczar says, ". . . Federal Reserve data for the same time, 1970, show banks hold about 45 percent of the dollar value of personal loans outstanding at the two institutions" (2, p. 254). Let L_B and L_F denote, respectively, the dollar volume of personal loans held by bank and finance companies. Then,

(1) $L_B = .45(L_B+L_F)$ or $L_B/L_F = .45/.55$

In an earlier paper, Boczar provides the following information on the average size of personal loans at the two institutions: "For the second quarter of 1971, the mean amount extended per contract by finance companies was $900. The comparable figure for commercial banks was $1,127" (3, p. 151).

Assume that the ratio of the average loan amount extended by banks and finance companies in the first quarter of 1971 is the same as ratio of the average loan amount in their 1970 portfolios. Then the estimate of the relative proportions of bank and finance company borrowers in the population given by

$$(2) \quad \hat{\pi}_B/\hat{\pi}_F = N_B/N_F = (900/1127)(L_B/L_F)$$

using (1). This yields $\hat{\pi}=\{\hat{\pi}_B=.395, \hat{\pi}_F=.605\}$ as the estimates of the population proportions.[4,5] Note that the these are approximately the reverse of the proportions in Boczar's sample.

We now use these estimates to revise Boczar's classification results. By analogy, our task is to determine the results of experiment #2 from those of experiment #1. The maximum likelihood estimates of conditional probabilities under two regimes where the population mixture varies are related by a simple transformation. Particularizing to the present case, let P_i' be the maximum likelihood estimate of the (conditional) probability of being a bank borrower using a sample containing bank and finance company borrowers with proportions $\pi' = \{\pi'_B , \pi'_F\}$. The maximum likelihood estimate of $\hat{P}i$, the conditional probability given proportions $\hat{\pi}=\{\hat{\pi}_B, \hat{\pi}_F\}$, is related to P_i' by the following formula:[6]

$$3)\quad \frac{\hat{P}_i}{1-\hat{P}_i}\cdot\frac{\hat{\pi}_F}{\hat{\pi}_B} = \frac{P_i'}{1-P_i'}\cdot\frac{\pi_F'}{\pi_B'}$$

Using this formula we determine the value of P_i' corresponding to $\hat{P}_i = .5$. This value then becomes the critical value for reclassifying individuals.[7] Reclassification is possible because Boczar provides a graph of the relative frequency distributions of P_i' for bank and finance company borrowers separately by .05 intervals.

To find the critical value, substitute the following values in (3) and solve for P_i':

$\hat{P}_i = .5$

$\hat{\pi}_B = .395$

$\hat{\pi}_F = .605$

$\pi_B' = (224/325) = .689$

$\pi_F' = (101/325) = .311$

The critical value of P_i' is found to be .773. The approximate relative frequency distributions of P_i' for bank and finance company borrowers, determined from Boczar's figure 1 (2, p. 251), are given in table 4.[8]

(Table 4 goes here)

Using $P_i^* \gtrless .773$ as the decision rule yields the revised prediction results shown in table 5.

(Table 5 goes here)

As could be anticipated from the experimental results, the error rate for finance comapny borrowers has fallen, that for bank borrowers has risen, while the total error rate is only slightly changed.

In this section we have shown that Boczar's central result (the 2/3 error rate in predicting finance company borrowers) was an artifact of the overrepresentation of bank borrowers in his sample. The fundamental lesson, however, is that the within-category error rates depend on prior probabilities. Since the question of interest is the contribution of the risk-related variables in discriminating between borrower categories, a test of the risk segmentation hypothesis, given the methodology under consideration, must in some sense normalize or correct for the influence of prior probabilities.

III. A Statistical Test

This section provides a test of the discriminating power of the risk-related variables. The test given is an alternative to the likelihood ratio test, but has the heuristic advantage of using the classification results directly.

Consider a chance mechanism which designates individuals as B's or F's with probabilities P_B, P_F (= $1-P_B$), respectively. Since a successful experiment is one where either a B or a F is correctly assigned, the probability of a success is

$$P_S = P_B\pi_B + P_F\pi_F, = P_B\pi_B + (1-P_B)\pi_F.$$

Thus, under the chance mechanism the number of successful experiments is distributed binomially, $B(X; n, P_B\pi_B + (1 - P_B)\pi_F)$. With π_B, π_F given, it remains to determine P_B in order to parameterize the distribution.

Recall that individual i was predicted to be a B or a F according as $P_i \gtrless .5$. Making explicit the fact that P_i is a conditional probability, we may write

$$P_i = P(B|R_i) = \frac{P(R_i|B)\cdot\pi_B}{P(R_i|B)\cdot\pi_B + P(R_i|F)\cdot\pi_F},$$

where R_i denotes a realization on the vector of risk-related variables. Since the null hypothesis of interest is that the risk-related variables have no information value, this means that the null hypothesis assumes $P(R_i|B) = P(R_i|F)$. (In other words, that the joint distribution of R is the same for both types of borrowers).

Making the substitution in the above expression yields $P_i = \pi_B$. Thus, under the null hypothesis that the risk-related variables do not discriminate between B's and F's, the probability that an individual will be assigned to category B is $P_B = \pi_B$. Hence the probability of a correct assignment is $P_S = \pi_B^2 + (1 - \pi_B)\pi_F = \pi_B^2 + \pi_F^2$. The number of correct predictions, therefore, is distributed $B(X; n, \pi_B^2 + \pi_F^2)$. In the present case, $n = 325$, $\pi_B = 224/325$, and $\pi_F = 101/325$. Although an exact test is possible, the sample size clearly justifies the use of the normal approximation. Thus we calculate the value of

$$Z = \frac{x - \mu}{\sigma},$$

where: $x = 244$ (the number of borrowers correctly identified, 325-81)

$$\mu = P_S n = (\pi_B^2 + \pi_F^2)n = 185.775$$

$$\sigma = n \cdot P_S \cdot (1 - P_S) = 8.921$$

This gives $Z = 6.527$; $P(Z > 6.527) < .0000003$.[9] The null hypothesis that the risk-related variables do not discriminate between F's and B's is rejected at the level of confidence greater than .999999. A researcher inclined to believe in the risk segmentation hypothesis would, of course, be pleased to impress readers with this result.

IV. Results Using More Recent Data

In deciding for the broader market definition urged by the defense (consumer installment credit), the District Court in the HFC/AIC merger case was mainly persuaded by the evidence showing the rapid growth since 1960 of competition from banks and other financial institutions in the consumer loan market (4, pp. 3-4).

As summarized by the Appellate Court, "The thrust of the District court's conclusion . . . is that although concedely some customers are uniquely served by finance companies, the number of such customers is insignificant" (5, pp. 13-14). The Appellate Court did not dispute the evidence of a trend, but found this insufficient to support the District Court's decision because, under applicable case law, it had also to show that the customer class uniquely served by finance companies was "insignificant." The Appellate Court then proceeded with a detailed review of the evidence bearing on the overlap equation. In contrast to the District Court, which if conceding the existence of a distinct class of finance company customers did so only by not specifically denying their existence, the Appellate Court took pains to show that the studies provided positive confirmation, and that "all the evidence suggested that these customers constituted anywhere from 15 to 50 percent of the finance company clientele."[10] In its view, "even if these customers constituted a far lesser percent . . . we would still find that group significant. Thus, the finding of the District Court that finance companies do not have a significant unique clientele was erroneous" (5, p. 18).

In the previous section we have shown that Boczar's study is at best "unsupportive" of defendants' position since the null hypothesis of no risk segmentation is decisively rejected. Since that study was cross-sectional using 1970 data, a remaining question of interest is whether or not the same test applied to more recent data would give the same result or show risk segmentation to have diminished in the sense that the overlap between bank and finance companies has increased (classification error has increased), perhaps to the point where the null hypothesis could no longer be rejected.

Using data from the Federal Reserve System's 1977 Consumer Credit Survey, Boczar's study was repeated using a similar set of explanatory variables.[11] Again the null hypotheis may be decisively rejected ($Z = 5.99$).[12] In table 6 below, we compare results normalized by assuming $\pi_B = \pi_F = .5$. That is, the results

(Table 6 goes here)

shown are those that would be expected had both samples contained a 50/50 split between bank and finance company borrowers.[13] This comparison indicates a slight reduction in the discriminating power of the risk-related variables.[14] The comparative results are thus at least consistent with the evidence introduced by the defense regarding the competitive inroads made by banks into consumer loan market. This is unlikely to have had any impact on the Appellate Court's decision, since it indicates no substantial

reduction in the proportion of finance company customers unlikely to qualify for bank credit (the standard of significance used by the Appellate Court). However, it is probable that measured as a proportion of all loan customers, the "significance" of the high-risk borrowers "uniquely served by finance companies" has declined substantially.[15]

V. Conclusion

In his study of market segmentation, Boczar addressed an issue of substantive policy interest, but drew a conclusion which is not supported by the results. Specifically, the fact that 2/3 of the finance company borrowers in his sample were misclassified as bank borrowers was an artifact of the sample proportion and hence does not "disconfirm" the risk segmentation hypothesis. The null hypothesis that the risk-related variables do not discriminate between borrower categories is decisively rejected using both 1970 and 1977 data. Despite his failure to draw the appropriate inference, the methodology does appear, as Boczar claims (2, p. 247), to offer several advantages over that used in earlier research on the same topic. For this reason, a clarification and warning regarding the appropriate use of that methodology, in particular the kinds of inferences which can be properly drawn, is warranted.

References

1. J. A. Anderson. "Separate Sample Logistic Discrimination," Biometrika, Volume 59, no. 1 (1972):19-35.

2. G. E. Boczar. "Competition Between Banks and Finance Companies: A Cross Section Study of Personal Loan Debtors," Journal of Finance, Volume 33 (March 1978):245-258.

3. G. E. Boczar. "The Evidence on Competition Between Commercial Banks and Finance Companies," Journal of Bank Research, Volume 6 (Summer 1975):150-154.

4. United States v. Household Finance Corporation, HFC American, Inc., and American Investment Company, No. 79C80 (N.D. Ill., March 22, 1979).

5. United States of America v. Household Finance Corporation, HFC American, Inc., and American Investment Company, No. 79-1313 (7th Cir., August 10, 1979).

Footnotes

[1] If the DOJ market definition had been accepted, the merger would have resulted in concentration ratios of 100 percent in many cities, while if the market had been taken to be consumer installment credit, the change in concentration would have been trivial.

The suit was filed in January of 1979, and the trial took place in February and March. The Department of Justice lost in the District Court decision delivered in May 1979, but won on appeal in August 1979. Not only was the case unusual in being confined to a narrow issue; it may also have set a record for shortness of total elapsed time.

[2] The data in table 2 is to be interpreted as follows. Taking the first row for example: 1 (10 percent) of the F types and 5 (50 percent) of the B types have low risk combination $X_1 = 0$, $X_2 = 0$.

[3] The holdout sample results came from using 250 randomly drawn observations to estimate the model, which is then used to predict the borrower category for the remaining 75 observations. This procedure was replicated 25 times. The percentages cited in the quoted passage refer to the average percentage of misclassifications for the 25 replications (2, p. 253), which are virtually identical to the results using the entire sample (given in table 1). This similarity is not concidence, since the expected values of the holdout sample proportions are equal to those of the data base from which the holdout samples were drawn.

Footnotes (continued)

4

$$\hat{\pi}_B/\hat{\pi}_F = \frac{.45}{.55} \times \frac{900}{1127} = .65338388.$$

Since $\hat{\pi}_B = 1 - \hat{\pi}_F$, we have $\hat{\pi}_B = .65338388(1 - \hat{\pi}_B)$, which gives $\hat{\pi}_B = .39517975$ and $1 - \hat{\pi}_B = \hat{\pi}_F = .60482030$.

5 The crucial assumption in deriving the estimates of the population proportions is that the relative size of new loans extended in 1971 can be used as an estimate of the average loan size in 1970. This assumption will produce an upward bias in the estimate of π_F if the average size of new personal loans extended by banks was increasing more slowly than for finance companies and the turnover (average maturity) of bank loans was shorter than of finance companies. Additional information given by Boczar on these two points provides a reasonable basis for believing the estimate of π_F is conservative. "With regard to maturity of personal loans, the average for the finance company industry for November 1973 was 34.2 months, while the most common maximum maturity for banks in 1973 was 24 months" (3, p. 151). Regarding the rate of increase in the size of new loans, the average size of new loans by banks and finance companies during 1973 was $1,645 and $1,121, respectively. Comparing these 1973 figures with those for 1971 given in the text, the average loan size increased by 46 percent for banks as compared to 24.5 percent for finance companies. (Substituting 1121/1645 for 900/1127 into (2) would give $\pi_F = .642$, $\pi_B = .358$).

Footnotes (continued)

[6] See Anderson, (1, p. 22). The transformation given by (3) is an application of Bayes formula. The general problem examined by Anderson was that of estimating a discriminant function using stratified sampling. This corresponds to our problem in that we may regard Boczar's data as a stratified sample, using sample proportions $\pi' = \{\pi_B' = 224/325, \pi_F' = 101/325\}$, from a population in which the proportions are $\hat{\pi} = \{\hat{\pi}_B = .395. \hat{\pi}_F = .605\}$. In the case where the function giving the posterior (conditional) probabilities is assumed to take the logistic form (the case examined by Anderson), a simple transformation exists relating the coefficients of the logit functions under different sampling regimes, as well as the posterior probabilities.

[7] In other words, an individual having $\hat{P}_i = .5$ under sampling regime $\hat{\pi}$ would have $P_i' = \alpha$ under sampling regime π', where α is found using the formula.

[8] A substantially larger copy of the graph, available from an earlier draft of Boczar's paper, was used to determine the numbers in table 4.

[9] The largest tabulated "Z" value found in readily available reference tables is 5.0; P(Z > .5) = 000000287.

Footnotes (continued)

[10] For example, the interpretation given to Boczar's study was as follows: "Certainly, the Boczar study demonstrates that a large number of finance customers, specifically two-thirds, have risks characteristics which appear to make them theoretically eligible for bank loans. However, such a conclusion necessarily entails the related conclusion that one-third of those customers have characteristics which suggest that they would probably not be bank customers. Thus, this study does not negate the existence of a group of customers whose credit needs are exclusively met by finance companies" (5, p. 11).

[11] I wish to thank Mr. Chaicho C. Wang of the Department of Justice (Economic Policy Office, Antitrust Division) for providing this data. The 1977 sample contained 360 bank borrowers and 93 finance company borrowers. All the variables used by Boczar were used here. The only difference was that age, education, and income were broken into fewer categories. For example, Boczar used dummy variables for age categories 18-24 (excluded category), 25-34, 35-49, 50 and over. For the probit regression using the 1977 data, the categories were 34 or less (excluded category), 35-50, and over 50. The results are available on request.

[12] The classification results using the 1977 sample were B|B = 351, B|F = 9, F|F = 14, and F|B = 79 (the notation X|Y means X's classified as Y's).

Footnotes (continued)

[13] A critical value of $P_i' = .70$ was used in reclassifying Boczar's sample here. This corresponds to a .05 interval given in Table 4 and is very close to the critical value of .689 found using (3) with $\hat{\pi}_B = \hat{\pi}_F = .5$.

[14] A test of the difference in proportions correctly classified with $\pi_F = \pi_B = .5$ gives $Z = 1.890$, which is statistically significant at the .97 level of confidence. However, in view of the approximations used here, this test of decline in discriminating power must be viewed cautiously.

[15] Not only would the trend evidence indicate this but it is also supported by the small proportion of finance company customers in the 1977 sample. Boczar's 1970 sample contained 31 percent finance company borrowers; in the 1977 sample this had declined to 20.5 percent. Although finance company borrowers were underrepresented in the 1970 sample, and likely to also have been underrepresented in 1977, the decline in the sample proportion is at least tentative evidence that the population proportion had also declined.

Table 1: Prediction Results--Boczar's Sample

Actual Group		Predicted Group: Bank	Predicted Group: Finance Company
Bank Borrower	224 (100.00)	208 (92.86)	16 (7.14)
Finance Company Borrower	101 (100.00)	65 (64.36)	36 (35.64)
Observations Misclassified:		81 (24.92)	

Table 2: Experimental Data

N_F	N_B	X_1	X_2
1	5	0	0
3	2	0	1
3	2	1	0
3	1	1	1

Table 3: Prediction Results Using Experimental Data

		Actual	Predicted Group	
			B	F
Sample #1	B	20 (100.00)	18 (90.00)	2 (10.00)
	F	10 (100.00)	6 (60.00)	4 (40.00)
Observations Misclassified: 8 (26.67)				
Sample #2	B	10 (100.00)	5 (50.00)	5 (50.00)
	F	20 (100.00)	2 (10.00)	18 (90.00)
Observations Misclassified: 7 (23.33)				

Table 4: Distribution of Bank and Finance Company Borrowers by P_i--Boczar's Sample

Value of P_i	Bank Borrowers (%)	Finance Company Borrowers* (%)
.70-.75	14.5	7.0
.75-.80	15.0	7.0
.80-.85	6.5	2.0
.85-.90	8.0	3.0
.90-.95	20.0	5.0
.95-1.00	9.5	1.0

* Note: These finance company borrowers were among those misclassified as bank borrowers in Table 1.

Table 5: Revised Prediction Results
Using Estimated Population Proportions

	Predicted Group (as percent of Actual Group)*	
Actual Group	Bank	Finance Company
Bank Borrower	44-59	56-41
Finance Company Borrower	11-18	89-82
Total Classification Error (percent):**	27.09-28.78	

*/ Figures given as a range using $P_i = .75$, and $P_i = .80$ respectively, as approximateions for $P_i = .773$

**/ This figure is the expected value of the total classification error with $\hat{\pi} = \{\hat{\pi}_B = .395, \hat{\pi}_F = .605\}$ using the indicated range for the within category error rates; i.e., $27.09 = .395(41) + .605(18)$, $28.78 = .395(56) + .605(11)$.

Table 6: Comparison of Prediction Results Using 1970 and 1977 Samples Standardized by Assuming $\pi_B = \pi_F = .5$

	Predicted Group (% of actual group)	
Actual Group	Bank	Finance Company
1970 Sample		
Bank Borrower	73.5	26.5
Finance Co. Borrower	25.0	75.0
Total Classification Error (%):		25.75
1977 Sample		
Bank Borrower	69.2	30.8
Finance Co. Borrower	34.4	65.6
Total Classification Error (%):		32.60