



0191-2615(94)E0015-Q

AN OVERLAPPING CHOICE SET MODEL OF AUTOMOTIVE PRICE ELASTICITIES

ROBERT BORDLEY

Operating Sciences Department, General Motors Research & Development Center,
Warren, MI 48090-9055, U.S.A.

(Received 8 January 1991; in revised form 16 October 1993)

Abstract—This paper focusses on choice models in which individuals (a) determine which of the many available products are worthy of detailed consideration. We refer to the resulting smaller set of products as the individual's choice set; (b) choose among products in the choice set using a fairly simple logit model. The nested logit model (McFadden, 1978; 1983) is one common example of this model in which choice sets are mutually exclusive and collectively exhaustive. Unfortunately, there is strong empirical evidence suggesting that automobile buyers have overlapping choice sets. Thus, some buyers will consider both small and medium-sized cars whereas other buyers will consider both large and medium-sized cars. Hence, the nested logit model appears to unrealistically limit the allowable patterns of interproduct similarity. To avoid these problems, I allow choice sets to overlap but will restrict all choice sets to include the same number of products. As I describe, the resulting model is estimable with data on individual first and second choice preferences. We illustrate the model's utility by deducing demand-price elasticities for the automotive market.

INTRODUCTION

Making pricing changes in the automotive industry (as well as many other key decisions (Hagerty, Carman, & Russell, 1988; Hauser, 1988; Tellis, 1988) required some knowledge of how pricing changes impact the sales of all products. But estimating demand-price elasticities is complicated by the following:

1. *Data limitations.* Because vehicle prices tend to move together and because vehicle attributes tend to change from year to year, the available time-series data is generally insufficient for estimating demand-price elasticities for all vehicles.
2. *Complicated patterns of interproduct similarity.* These complicated patterns generally make the simpler nested logit models (Kamakura & Russell, 1989; Russell & Bolton, 1988) inapplicable.

Models capable of capturing complicated patterns of interproduct similarity (e.g., the multinomial probit (Daganzo, 1979) become computationally impractical when the number of products is not small.

This paper develops a new approach to estimating elasticities based on (a) Information on individual first and second choice preferences; (b) A flexible model of overlapping choice sets which can accommodate any pattern of demand-price cross-elasticity. I use the model to deduce more than 40,000 automotive demand-price elasticities and examine the sensitivity of our predictions to changes in modelling assumptions.

THE CHOICE MODEL

Assumption 1: With probability P_{Ω} , an individual will only consider buying a product in choice set Ω . Given Ω , the probability of buying a product in Ω (or not buying any product) is logit. Thus, the overall probability of buying product i is:

$$P_i = \sum_{\Omega} P_{\Omega} \frac{\exp(V_{i\Omega})}{\sum_{k \in \Omega} \exp(V_{k\Omega}) + \exp(V_{0\Omega})}$$

where $V_{i\Omega}$ is the value that an individual with choice set Ω attaches to buying product i (with $i = 0$ corresponding to the option of not buying any product in the market of interest.)

As the prices of all the products in one's choice set increases, an individual becomes more likely to drop out of the market. There is no possible price change that will cause the individual to switch a product outside his choice set. This model is consistent with a customer using a coarse screening process, not involving price, to select his/her choice set followed by a more focussed examination of the options in the choice set. This is a key difference between our model and the nested logit model in which the probability of having a given choice set does depend on prices. By making this restriction, our model, unlike nested logit, can allow choice sets to be overlapping.

If X_j is the price of product j ,

$$\frac{\partial P_i}{\partial X_j} = \sum_{\Omega \supset j} P_{\Omega} \frac{\partial V_{j\Omega}}{\partial X_j} \left[\delta_{i=j} \frac{\exp(V_{i\Omega})}{\sum_{k \in \Omega} \exp(V_{k\Omega}) + \exp(V_{0\Omega})} - \frac{\exp(V_{i\Omega})}{\sum_{k \in \Omega} \exp(V_{k\Omega}) + \exp(V_{0\Omega})} \frac{\exp(V_{j\Omega})}{\sum_{k \in \Omega} \exp(V_{k\Omega}) + \exp(V_{0\Omega})} \right]$$

where $\delta_{i=j}$ equals 1 for $i = j$ and equals 0 else.

Assumption 2: Suppose that $\frac{\partial V_{j\Omega}}{\partial X_j} = V'_j \delta_{j \in \Omega}$, i.e., the marginal utility of product j is independent of the other products in the choice set.

Suppose we also define

1. $P_{i|\Omega} = \frac{\exp(V_{i\Omega})}{\sum_{k \in \Omega} \exp(V_{k\Omega})}$ as the probability of buying i from choice set Ω , given one buys a product in the market.
2. $P_{c|\Omega} = 1 - \frac{\exp(V_{0\Omega})}{\sum_{k \in \Omega} \exp(V_{k\Omega}) + \exp(V_{0\Omega})}$ as the probability of buying some vehicle in the choice set (versus not buying at all).
3. $B_{ij} = \sum_{\Omega \supset j} P_{\Omega} [\delta_{ij} P_{c|\Omega} P_{i|\Omega} - P_{c|\Omega}^2 P_{i|\Omega} P_{j|\Omega}]$.

Then

$$\begin{aligned} \frac{\partial P_i}{\partial X_j} &= B_{ij} V'_j \\ \eta_{ij} &= \frac{X_j}{P_i} \frac{\partial P_i}{\partial X_j} = \frac{X_j}{P_i} B_{ij} V'_j \end{aligned} \quad (1)$$

ESTIMATING THE MODEL SPECIFYING B_{ij} USING FIRST CHOICE/SECOND CHOICE DATA

In the automotive context, we have information on the vehicle individuals buy (their first choice) and the vehicle they would have bought if their first choice were unavailable (their second choice)—conditioned on their second choice being a car. We let N_{ij} be the fraction of the population specifying i as their first choice and j as their second choice. We now assume:

Assumption 3(a): All choice sets contain the same number of vehicles.

Thus if each choice set only contains two vehicles, $N_{ij} + N_{ji}$ would be the fraction of the population considering both products i and j and $P_{i|ij} = \frac{N_{ij}}{N_{ij} + N_{ji}}$ would be the

fraction of those only considering products i and j who buy product i . If we let $P_{c|ij}$ be the fraction of individuals with products i and j as their choice set who buy a product in the market, then

$$B_{ik} = -P_{c|ik}^2 P_{ik} P_{i|ik} P_{k|ik} = -P_{c|ik}^2 P_{ik} \frac{N_{ik}}{N_{ik} + N_{ki}} \frac{N_{ki}}{N_{ik} + N_{ki}} = -P_{c|ik}^2 \frac{N_{ik} N_{ki}}{N_{ik} + N_{ki}}$$

To extrapolate our first choice/second choice data to choice sets with $r \geq 2$ vehicles, let $N(i_1, \dots, i_r)$ be the unobserved fraction of individuals with $i_1, \dots, i_r$ as their first through r th choices. Then suppose we estimate $N(i_1, \dots, i_r)$ by

$$N(i_1, i_2, i_3, \dots, i_r) = N(i_1, i_2) \frac{N(i_1, i_3)}{N(i_1) - N(i_1, i_2)} \dots \frac{N(i_1, i_r)}{N(i_1) - \sum_{k < r} N(i_1, i_k)}$$

From this, $P_{i_1|i_2, \dots, i_r}$ is easily estimated. If all choice sets contain r elements, we then compute

$$B_{i_1, i_2} = -P_{c|\Omega}^2 \sum_{\Omega \supset i_1, i_2} P_{\Omega} P_{i_1|\Omega} P_{i_2|\Omega}, i_1 \neq i_2 \quad (2)$$

For concreteness, we will assume:

Assumption 3(b). All choice sets contain three vehicles.

This paper will additionally assume that buyers are similar to nonbuyers, i.e.,

Assumption 4. $P_{c|\Omega}$ is a constant.

SPECIFYING V'_k USING SEGMENT ELASTICITY INFORMATION

While estimating the sales of specific vehicles as a function of prices is difficult, we can frequently group vehicles into segments and use time-series to estimate the sales in each segment I as a function of the average price, X_I , of vehicles within that segment. If m_i is the product i 's marketshare and P_I is the share of sales in segment I , then the segment own-elasticity is given by

$$\eta_{II} = \frac{1}{P_I} \sum_{i \in I} \frac{\partial P_i}{\partial \log X_I} = \sum_{i \in I} m_i \sum_{j \in I} \eta_{ij} \frac{d \log X_j}{d \log X_I} \quad (3)$$

We now assume:

Assumption 5(a) A percentage increase in the prices of all products in the same segment will not change the relative marketshares of products in that segment, i.e., $\sum_j \eta_{ij}$ is a constant for all $i \in I$.

Assumption 5(b): A percentage increase in the average price of all products in the same segment is historically associated with a percentage increase in the prices of all products in that segment, i.e., $\frac{d \log X_j}{d \log X_I} = 1$

Assumption 5 and equation 4 imply that $\eta_{II} = \sum_{j \in I} \eta_{ij}$, $i \in I$ and

$$P_i \eta_{II} = \sum_{j \in I} P_i \eta_{ij} = \sum_{j \in I} B_{ij} X_j V'_j$$

Defining $D_{ij} = B_{ij}$ if products i and j are in the same segment (and zero else) and inverting D gives

$$X_j V'_j = \eta_{II} \sum_{k \in I} D_{jk}^{-1} q_k, j \in I$$

Thus, we can estimate V'_j from our segment own-elasticities. Substituting into (1) gives

$$P_i \eta_{ij} = \eta_{II} B_{ij} \sum_{k \in I} D_{jk}^{-1} P_k, j \in I \quad (4)$$

SPECIFYING $P_{c| \Omega}$ USING MARKET OWN-ELASTICITY DATA

If $\theta'_i = \frac{d \log X_i}{d \log X_1}$, the market own-elasticity, η , is given by

$$\eta = \frac{\partial \log P}{\partial \log X} = \frac{\partial \log P / \partial \log X_i}{\partial \log X / \partial \log X_i} = \frac{\sum_j \theta'_j \partial \log P / \partial \log X_j}{\sum_j \theta'_j \partial \log X / \partial \log X_j}$$

where P is total sales and X is the average price of products in the market. Then

$$\begin{aligned} \frac{\partial \log P}{\partial \log X_j} &= \sum_i m_i \frac{\partial \log P_i}{\partial \log X_j} \\ \frac{\partial \log X}{\partial \log X_j} &= \frac{1}{X} \left(m_j X_j + \sum_i X_i \frac{\partial m_i}{\partial \log X_j} \right) \\ \frac{\partial m_i}{\partial \log X_j} &= m_i \left(\eta_{ij} - \sum_k m_k \eta_{kj} \right) \end{aligned}$$

Defining $w_j = \frac{m_j X_j}{X}$ and $\eta = \frac{\partial \log P}{\partial \log X}$ gives

$$\eta = \frac{\sum_{ij} \theta'_j m_i \eta_{ij}}{\sum_j \theta'_j (w_j + \sum_k (w_k - m_k) \eta_{kj})}$$

Note that η is invariant to which product is chosen as our reference product for θ'_j . We will specify θ'_j by assuming that

Assumption 6: θ'_j is a constant, i.e., all prices had the same percentage price changes over the period during which elasticities were estimated. Thus our products form a Hick-
sian composite (Deaton & Muellbauer, 1988).

Given this assumption

$$\eta = \frac{\sum_{ij} m_i \eta_{ij}}{1 + \sum_{kj} (w_k - m_k) \eta_{kj}} \quad (5)$$

Hence, a uniform 1% price increase causes individuals to switch to cheaper goods, leading to a less than one percent change in the average price of all products.

We can now specify $P_{c| \Omega}$ from η using the following algorithm:

1. Assume $P_{c| \Omega} = 0.5$
2. Compute B_{ij} from (2) and η_{ij} from (3)
3. Compute η' from (5)
4. Adjust $P_{c| \Omega}$ and iterate until η' equals the prespecified value of η .

APPLICATION TO THE AUTOMOTIVE MARKET

Results

We divided the automotive market into seven segments with segment elasticities estimated from a time-series model relating segment sales to average prices within each segment. The average segment elasticity was about -2 with a confidence interval ranging

from -1.5 to -3.0 . The market own-elasticity was assumed to equal -1 . Our first choice/second choice data was from a quarterly survey of more than 40,000 new car buyers.

Table 1 presents results for a single segment. Note that higher cross-elasticities are associated with products with higher first choice/second choice frequencies: e.g., Chevrolet's Celebrity's cross-elasticity with 6000 is high while its cross-elasticity with Pontiac's Bonneville is small. Cross-elasticities are asymmetric (e.g., Celebrity's cross-elasticity with 6000 is much smaller than 6000's cross-elasticity with Celebrity because Celebrity's sales of 231,000 units exceed 6000's sales of 117,000 units.)

The average productline own-elasticity is about -5 . I also find that products whose second choices lay in the same segment seemed to be more elastic than products whose second choices lay outside the segment. These own-elasticity estimates were considerably lower than those generally used by General Motor's (GM) marketing staff. This result, reinforced by other empirical studies, caused GM to reduce the extent to which it used price rebates to attract customers. Nested logit, when applied to this problem, tends to lead to even lower elasticity estimates because nested logit artificially restricts the degree of competition between product lines.

Sensitivity Analysis

I now explore the sensitivity of our elasticity predictions to various perturbations in the model assumptions and input data. As the accompanying Table 2 indicates, I first compared the maximum, mean and minimum own-elasticity and cross-elasticity inferred

Table 1. Mid-sized car segment demand/price elasticities for 1986

	Celebrity	Monte Carlo	6000	Bonneville	Grand Prix	Ciera	Supreme 4DR	Supreme 2DR	Century	Regal 2DR
Celebrity	-3.85	0.08	0.47	0.02	0.00	0.55	0.00	0.02	0.46	0.01
Monte Carlo	0.26	-3.96	0.05	0.03	0.34	0.06	0.00	0.28	0.03	0.41
6000	1.09	0.03	-6.41	0.14	0.14	1.18	0.00	0.04	1.17	0.06
Bonneville	0.26	0.13	0.88	-5.36	0.57	0.42	0.00	0.32	0.62	0.00
Grand Prix	0.03	1.02	0.59	0.38	-6.09	0.12	0.00	0.79	0.13	0.40
Ciera	0.53	0.02	0.49	0.03	0.01	-4.69	0.00	0.06	1.22	0.00
Supreme 4DR	0.00	0.00	0.00	0.00	0.00	0.00	-2.04	0.00	0.04	0.00
Supreme 2DR	0.01	0.05	0.01	0.01	0.04	0.03	0.00	-2.28	0.01	0.10
Century	0.60	0.01	0.65	0.05	0.02	1.65	0.28	0.02	-5.67	0.06
Regal 2DR	0.01	0.13	0.03	0.00	0.04	0.00	0.00	0.19	0.05	-2.53
Thunderbird	0.07	0.22	0.06	0.00	0.05	0.12	0.00	0.05	0.05	0.08
Taurus	0.23	0.02	0.23	0.01	0.00	0.27	0.00	0.04	0.24	0.00
Cougar	0.08	0.17	0.08	0.01	0.10	0.11	0.00	0.00	0.05	0.12
Sable	0.20	0.02	0.36	0.01	0.03	0.42	0.00	0.00	0.24	0.00
Lancer	0.23	0.00	0.08	0.00	0.00	0.04	0.00	0.00	0.09	0.00
Lebaron	0.15	0.04	0.13	0.00	0.04	0.41	0.00	0.00	0.33	0.05
Lebaron GTS	0.04	0.02	0.14	0.00	0.00	0.14	0.00	0.00	0.02	0.00
	Thunderbird	Taurus	Cougar	Sable	Lancer	Lebaron	Lebaron GTS			
Celebrity	0.03	0.09	0.03	0.03	0.02	0.03	0.01			
Monte Carlo	0.27	0.02	0.19	0.01	0.00	0.02	0.01			
6000	0.05	0.21	0.07	0.12	0.02	0.06	0.05			
Bonneville	0.00	0.06	0.06	0.03	0.00	0.00	0.00			
Grand Prix	0.19	0.00	0.33	0.04	0.00	0.08	0.00			
Ciera	0.04	0.10	0.04	0.06	0.00	0.08	0.02			
Supreme 4DR	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Supreme 2DR	0.01	0.01	0.00	0.00	0.00	0.00	0.00			
Century	0.03	0.12	0.02	0.05	0.01	0.08	0.01			
Regal 2DR	0.03	0.00	0.04	0.00	0.00	0.01	0.00			
Thunderbird	-4.11	0.34	0.91	0.08	0.02	0.04	0.01			
Taurus	0.35	-3.92	0.09	0.36	0.01	0.04	0.02			
Cougar	0.99	0.10	-4.05	0.19	0.00	0.03	0.02			
Sable	0.21	0.94	0.46	-5.01	0.00	0.05	0.08			
Lancer	0.09	0.06	0.00	0.00	-3.37	0.02	0.76			
Lebaron	0.08	0.08	0.06	0.04	0.01	-3.50	0.06			
Lebaron GTS	0.02	0.04	0.05	0.06	0.38	0.07	-2.98			

Table 2. Sensitivity analysis of elasticity estimates

	Own-Elasticity					Cross-Elasticity				
	Min	Mean	Max	Cor	σ	Min	Mean	Max	Cor	σ
Baseline model	-13.4	-4.74	-1.8	1.0	0	0	.03	3.9	1.0	0
1987 Version	-11.7	-4.76	-2.8	.91	.9	0	.03	4.8	.88	.07
Sensitivity to Assumptions										
Assumption 3										
$r = 2$	-10.9	-4.13	-1.8	.99	.5	0	.02	3.0	.98	.05
$r = 4$	-15.5	-5.22	-1.8	.99	.5	0	.03	4.9	.98	.05
Assumption 4	-16.2	-5.14	-1.8	.97	.6	0	.03	6.3	.96	.05
Assumption 5										
Equal own-elast	-7.8	-4.85	-2.2	.70	1.6	0	.03	4.1	.98	.03
Assumption 6										
$\theta'_i = P_i$	-16.7	-5.25	-1.8	.99	.5	0	.03	5.0	.99	.03
$\theta'_i = 1/P_i$	-11.8	-4.48	-1.8	.99	.3	0	.027	3.9	.99	.02
Sensitivity to Input Data										
Min segment Elasticities	-17.1	-5.84	-2.3	1.0	.5	0	.04	5.0	1.00	.04
Max segment Elasticities	-9.7	-3.62	-1.4	.99	.5	0	.02	3.0	1.00	.04
Market own										
$n = -1.5$	-8.1	-3.76	-1.8	.99	.5	0	.02	2.3	0.99	.03
$n = -0.5$	-21.1	-5.87	-1.8	.99	1.2	0	.04	6.6	0.99	.07
1st choice/ 2nd choice	-12.0	-4.69	-1.8	.88	1.1	0	.03	1.7	0.64	.11

from 1986 data with the estimates inferred using 1987 first choice/second choice data. I also computed the correlation of the 1987 estimates with my 1986 estimates as well as the square root of the mean squared differences (which I call σ) between the 1987 estimates and the 1986 estimates. I found little difference between my estimates.

I then varied assumption 3 (which specified a choice set of three elements) by estimating the model with the size of the choice set equal to 2 or 4. Adding one more element to the choice set caused our own-elasticity estimates to become 10% more negative. But the correlation between all three sets of estimates remained high.

Assumption 4 had assumed that nonbuyers had the same choice sets as buyers. We now alternatively assume that nonbuyers have the same choice set as buyers with comparable incomes, Y . Then

$$P_{c|\Omega} \propto \int_Y f_{c|Y} f_{Y|\Omega}$$

where $f_{c|Y}$ is the fraction of vehicle buyers with incomes Y and $f_{Y|\Omega}$ is the fraction of buyers with choice set Ω having income Y . This assumption causes our luxury car estimates to become 20% more negative though it only changed the average elasticity by about five percent.

Assumption 5 indicated that an increase in the average product price within a segment would not cause product shares within the segment to change. If we remove this assumption, define $w_{i|I} = \frac{P_i X_i}{\sum_{i \in I} P_i X_i}$ and $m_{i|I} = \frac{P_i}{\sum_{i \in I} P_i}$ and use the analogue of the arguments leading to (5), we get

$$\eta_{II} = \frac{\sum_{ij \in I} m_{i|I} \eta_{ij}}{1 + \sum_{ij \in I} (w_{i|I} - m_{i|I}) \eta_{ij}}$$

I now suppose that all own-elasticities within the same segment are equal. Because (1) gives $\frac{\eta_{ij}}{\eta_{jj}} = \frac{P_j B_{ij}}{P_i B_{ij}}$, this implies

$$\eta_{ij} = \frac{\eta_{II}}{\sum_{ij \in I} (P_j B_{ij} / P_i B_{jj}) [m_{iI} + \eta_{II} (m_{i|I} - w_{i|I})]}$$

Assumption 6 indicated that all products regardless of their price tended to have the same percentage price change over time. To vary that assumption, I first assume that products that are more expensive tend to have much higher percentage price increases. A second alternative is to assume that products that are more expensive have much lower percentage price increases. This does not appear to change our elasticity estimates significantly.

Assumption 1 specifies the structure of the model. Although I did not conduct a sensitivity analysis of our results to Assumption 1, Bordley (in press) constructed a very different kind of model that appeared to give roughly the same kinds of results. Hence, my elasticity predictions do, in fact, seem insensitive to model specification.

Thus, my elasticity predictions seem robust to my modelling assumptions. To explore my model's sensitivity to input data, I first examine how the elasticity estimates change when I use the maximum segment elasticity values (instead of the midpoints) and when I use the minimum segment elasticity values. I then vary my market own-elasticity. Finally, I construct a crude approximation to the first choice/second choice as follows:

1. Let n_{ij} be the number of car lines for which N_{ij} exceeds $\frac{m_i m_j}{1 - m_i}$
2. Set $N_{ij} = 0$ if $N_{ij} \leq \frac{m_i m_j}{1 - m_i}$
3. Set $N_{ij} = \frac{m_i}{n_{ij}}$ if $N_{ij} \geq \frac{m_i m_j}{1 - m_i}$

As Table 2 shows, varying the input data does change my elasticity estimates significantly. Varying the first choice/second choice data appears to change my cross-elasticities more than my own-elasticities.

Thus, my results seem robust to my modelling assumptions while being quite sensitive to my input data.

CONCLUSIONS

Nested logit attempts to model individual choice by dividing up a market of n products into, say, $\frac{n}{5}$ disjoint nests or segments. In the automotive market, this assumption of a small number of disjoint segments seems unrealistic. I developed an alternative model that divides the market up into $\frac{n!}{3!(n-3)!}$ overlapping segments, thus allowing for arbitrarily complicated patterns of interproduct similarity. I demonstrate the feasibility of the method by applying it to 1986 automotive data; I then demonstrated the robustness of our elasticity estimates to permutations in our assumptions.

REFERENCES

- Bordley R. F. (1989). Estimating cross-elasticities from first choice/second choice data. *Journal of Business & Economic Statistics*, 11, 141-146.
- Bordley R. F. (in press). Estimating automotive elasticities from segment elasticities & first choice/second choice data. *Review of Economics & Statistics*.
- Daganzo C. (1979). *Multinomial probit*. New York: Academic Press.
- Deaton A., & Muellbauer J. (1980). *Economics and consumer behavior*. London: Cambridge University Press.
- Hagerty M., Carman J., & Russell G. (1988). Estimating elasticities with PIMS data: Methodological issues & substantive implications. *Journal of Marketing Research*, 25, p. 1.
- Hauser J. (1988). Competitive price & positioning strategies. *Marketing Science*, 7, p. 1.
- Kamakura W., & Russell G. (1989). A probabilistic choice model for market segmentation & elasticity structure. *Journal of Marketing Research*, 26, pp. 379-390.

- Lave C., & Train K. (1979). A disaggregate model of auto-type choice. *Transportation Research A*, 13, pp. 1-10.
- McFadden D. (1978) Modelling the choice of residential location. Karlqvist A., Lundqvist L., Snickars F., Weibull (eds.) *Spatial interaction theory and planning models* (pp. 75-96). New York: North-Holland.
- McFadden D. (1983). Econometric models of probabilistic Choice. Manski, C., & McFadden, D. *Structural analysis of discrete data with econometric applications* (pp. 100-120). Cambridge, MA: MIT Press.
- Russell G., & Bolton R. (1988). Implications of market structure for elasticity structure. *Journal of Marketing Research*, 25, 199, 229-241.
- Tellis G. (1988). The price elasticity of selective demand: a meta-analysis of econometric models of sales. *Journal of Marketing Research*, 25, 331-341.