The BCHOICE Procedure (Experimental)

Example 27.1 Alternative-Specific and Individual-Specific Effects

In many situations, a choice model that includes characteristics of both the alternatives and the individuals is needed for investigating consumer choice.

Consider an example of travel demand. People are asked to choose among travel by auto, plane, or public transit (bus or train). The following SAS statements create the data set *Travel*. The variables *AutoTime*, *PlanTime*, and *TranTime* represent the total travel time that is required to get to a destination by using auto, plane, or public transit, respectively. The variable *Age* represents the age of each individual who is surveyed, and the variable *Chosen* contains each individual's choice of travel mode.

```
data Travel;
   input AutoTime PlanTime TranTime Age Chosen $;
   AgeCtr=Age-34;
   datalines;
10.0 4.5 10.5 32 Plane
5.5 4.0 7.5 13 Auto
4.5 6.0 5.5 41 Transit
3.5 2.0 5.0 41 Transit
1.5 4.5 4.0 47 Auto
10.5 3.0 10.5 24 Plane
7.0 3.0 9.0 27 Auto
9.0 3.5 9.0 21 Plane
4.0 5.0 5.5 23 Auto
22.0 4.5 22.5 30 Plane
7.5 5.5 10.0 58 Plane
11.5 3.5 11.5 36 Transit
3.5 4.5 4.5 43 Auto
12.0 3.0 11.0 33 Plane
18.0 5.5 20.0 30 Plane
23.0 5.5 21.5 28 Plane
4.0 3.0 4.5 44 Plane
5.0 2.5 7.0 37 Transit
3.5 2.0 7.0 45 Auto
12.5 3.5 15.5 35 Plane
1.5 4.0 2.0 22 Auto
;
```

In this example, the *AutoTime*, *PlanTime*, and *TranTime* variables apply to the alternatives, whereas *Age* is a characteristic of the individuals. *AgeCtr*, a centered version of *Age*, is created by subtracting the sample's mean age from each individual's age. To study how the choice depends on both the travel time and age of the individuals, you need to incorporate both types of variables.

Before you invoke PROC BCHOICE to fit a choice logit model, you must arrange your data in such a way that there is one observation for each combination of individual and alternative. In this example, let *Subject* identify the individuals, let *TravTime* represent the travel time for each mode of transportation, and let *Choice* have the value 1 if the alternative is chosen and 0 otherwise. The following SAS statements rearrange the data set *Travel* into a new data set, *Travel2*, and display the first nine observations:

```
data Travel2(keep=Subject Mode TravTime Age AgeCtr Choice);
   array Times[3] AutoTime PlanTime TranTime;
   array Allmodes[3] $ _temporary_ ('Auto' 'Plane' 'Transit');
   set Travel;
   Subject = _n_;
   do i = 1 to 3;
      Mode = Allmodes[i];
      TravTime = Times[i];
      Choice = (Chosen eq Mode);
      output;
   end;
run;

proc print data=Travel2 (obs=20);
   by Subject;
   id Subject;
run;
```

The data for the first nine observations is shown in Output 27.1.1.

Output 27.1.1: Data for the First Nine Observations

Subject	Age	AgeCtr	Mode	TravTime	Choice
1	32	-2	Auto	10.0	0
	32	-2	Plane	4.5	1
	32	-2	Transit	10.5	0

Subject	Age	AgeCtr	Mode	TravTime	Choice
2	13	-21	Auto	5.5	1
	13	-21	Plane	4.0	0
	13	-21	Transit	7.5	0

Subject	Age	AgeCtr	Mode	TravTime	Choice
3	41	7	Auto	4.5	0
	41	7	Plane	6.0	0
	41	7	Transit	5.5	1

Notice that each subject in the data set *Travel* corresponds to a block of three observations in the data set *Travel2*, one for each travel alternative. The response variable *Choice* indicates the chosen alternative by the value 1 and the unchosen alternative by the value 0; exactly one alternative is chosen. The following SAS statements invoke PROC BCHOICE to fit the choice logit model:

```
proc bchoice data=Travel2 seed=124;
   class Mode Subject / param=ref order=data;
   model Choice = Mode TravTime / choiceset=(Subject);
run;
```

The *"Choice Sets Summary"* table shows that there are 21 choice sets and that each consists of three alternatives and one chosen alternative (each subject chooses one out of the three travel modes). It seems that the data are arrayed correctly.

Output 27.1.2: Choice Sets Summary

The BCHOICE Procedure

Choice Sets Summary				
Pattern	Choice Sets	Total Alternatives	Chosen Alternatives	Not Chosen
1	21	3	1	2

Summary statistics are shown in Output 27.1.3.

Output 27.1.3: PROC BCHOICE Posterior Summary Statistics

Posterior Summaries and Intervals					
Parameter	N	Mean	Standard Deviation	95% HPD Interval	
Mode Auto	5000	-0.1678	0.7440	-1.7017	1.2396
Mode Plane	5000	-1.8794	1.2683	-4.6055	0.3801
TravTime	5000	-0.5695	0.2047	-0.9943	-0.2328

When *Transit* is the reference mode (normalized to 0), the part-worth (posterior mean) of *Auto*, which is negative, might reflect that driving is more inconvenient than traveling by bus or train, and the negative part-worth of *Plane* might reflect that traveling by plane is more expensive than traveling by bus or train. However, neither is statistically significant, because both the 95% HPD intervals have 0 in them. The posterior mean of *TravTime* is negative, and the estimate is significant. These results make sense because having to spend more time en route is often unfavorable.

To study the relationship between the choice of transportation and the age of people who make the choice, you need to create an interaction between *AgeCtr* and *Mode*. *AgeCtr* is not estimable by itself, because it is the same throughout a choice set for an individual. The following statements request the interaction between *AgeCtr* and *Mode*:

```
proc bchoice data=Travel2 seed=124;
   class Mode Subject / param=ref order=data;
   model Choice = Mode Mode*AgeCtr TravTime / choiceset=(Subject);
run;
```

Output 27.1.4: PROC BCHOICE Posterior Summary Statistics

The BCHOICE Procedure

Posterior Summaries and Intervals

Parameter	N	Mean	Standard Deviation	95% HPD Interval	
Mode Auto	5000	-0.2634	0.7883	-1.8072	1.1395
Mode Plane	5000	-2.8210	1.5370	-5.9228	-0.0686
AgeCtr*Mode Auto	5000	-0.0986	0.0678	-0.2182	0.0350
AgeCtr*Mode Plane	5000	0.0251	0.0775	-0.1268	0.1618
TravTime	5000	-0.7608	0.2564	-1.2943	-0.3473

The parameter estimate for *Mode Auto* reflects the part-worth of *Auto* for an individual of mean age (34 years old), whereas the parameter estimate for *Mode Plane* is the part-worth of *Plane* for an individual of mean age. There are two interaction effects: the first corresponds to the effect of a one-unit change in age on the probability of choosing *Auto* over *Transit*, and the second corresponds to the effect of a one-unit change in age on the probability of choosing *Plane* over *Transit*.