

LECTURE / DISCUSSION

Specification Issues

Alternative-Specific Constants

1. Adding the same constant to each alternative has no effect on choices or probabilities:

$$U_c = k + \alpha C_c + \beta T_c + \varepsilon_c$$

$$U_b = k + \alpha C_b + \beta T_b + \varepsilon_b$$

is the same as

$$U_c^* = \alpha C_c + \beta T_c + \varepsilon_c$$

$$U_b^* = \alpha C_b + \beta T_b + \varepsilon_b$$

That is, U_c exceeds U_b only if U_c^* exceeds U_b^* and vice versa.

2. Add a different constant to each alternative but one.

Example:

$$U_c = k + \alpha C_c + \beta T_c + \varepsilon_c$$

$$k = 3$$

$$U_b = a + \alpha C_b + \beta T_b + \varepsilon_b$$

$$a = 5$$

is the same as

$$U_c^* = \alpha C_c + \beta T_c + \varepsilon_c$$

$$U_b^* = a^* + \alpha C_b + \beta T_b + \varepsilon_b$$

$$a^* = 2$$

Properties of the Estimated Alternative-Specific Constants

1. Set average estimated probabilities equal to the actual shares in the sample.
2. Capture the average impact of omitted variables.
3. Correct for deviations from logit specification (e.g., failures of IIA, discussed tomorrow).

How to Add Constants in SST

```
set zero = 0
```

```
set one = 1
```

```
mnl dep[depvar] \  
  ivalt[ic:   ic1  ic2  ic3  ic4  ic5 \  
           oc:   oc1  oc2  oc3  oc4  oc5 \  
           c1:   one  zero zero zero zero \  
           c2:   zero one  zero zero zero \  
           c3:   zero zero one  zero zero \  
           c4:   zero zero zero one  zero] \  
  
  coef[beta]
```

Differences in Utility

Only differences in representative utility matter.

Proof I:

$$\begin{aligned} P_{in} &= \frac{e^{V_{in}}}{\sum e^{V_{jn}}} = \left(\frac{e^{-V_{kn}}}{e^{-V_{kn}}} \right) \frac{e^{V_{in}}}{\sum e^{V_{jn}}} \\ &= \frac{e^{V_{in} - V_{kn}}}{\sum e^{V_{jn} - V_{kn}}} \end{aligned}$$

Proof II:

$$\frac{P_{in}}{P_{kn}} = \frac{\frac{e^{V_{in}}}{\sum e^{V_{jn}}}}{\frac{e^{V_{kn}}}{\sum e^{V_{jn}}}} = \frac{e^{V_{in}}}{e^{V_{kn}}} = e^{V_{in} - V_{kn}}$$

Implications for Socioeconomic Variables

1. Socioeconomic variable cannot enter with same coefficient in each alternative:

$$V_c = \alpha C_c + \beta T_c + \lambda I$$

$$V_b = \alpha C_b + \beta T_b + \lambda I$$

This is equivalent to:

$$V_c^* = \alpha C_c + \beta T_c$$

$$V_b^* = \alpha C_b + \beta T_b$$

2. Socioeconomic variables can enter with different coefficients in each alternative. **But:** one coefficient is normalized to zero.

$$V_c = \alpha C_c + \beta T_c + \lambda I$$

$$V_b = \alpha C_b + \beta T_b + \tau I$$

This is equivalent to

$$V_c^* = \alpha C_c + \beta T_c$$

$$V_b^* = \alpha C_b + \beta T_b + \tau^* I$$

where $\tau^* = \tau - \lambda$.

$\tau^* > 0$: If I increases, P_b increases relative to P_c .

$\tau^* < 0$: If I increases, P_b decreases relative to P_c .

Multinomial Example

$$V_c = \alpha C_c + \beta T_c$$

$$V_b = \alpha C_b + \beta T_b - .04 I$$

$$V_r = \alpha C_r + \beta T_r - .02 I$$

If income increases, V_b decreases the most, V_r next most, and V_c not at all.

So: P_c increases relative to P_r and P_b .

P_r increases relative to P_b .

Socioeconomic Variables Interacting with Other Variables

$$V_c = \alpha C_c + \beta T_c$$

$$V_b = \alpha C_b + \beta T_b$$

But: $\beta = \delta \cdot \text{INCOME}$.

That is, the importance of time is proportional to a person's income.

$$V_c = \alpha C_c + \delta(\text{INCOME} \cdot T_c)$$

$$V_b = \alpha C_b + \delta(\text{INCOME} \cdot T_b)$$

How to Add Socioeconomic Variables in SST

Income entering all alternatives but last:

```

set zero = 0
mnl dep[depvar] \
  ivalt[ic:   ic1   ic2   ic3   ic4   ic5 \
    oc:  oc1  oc2   oc3   oc4   oc5 \
    inc1: income   zero   zero   zero   zero \
    inc2: zero  income zero   zero   zero \
    inc3: zero  zero  income zero   zero \
    inc4: zero  zero   zero  income zero] \
prob[p2 p3 p4 p5] coef[beta]
set p1 = 1 - p2 - p3 - p4 - p5

```

Operating cost times income:

```

set oci1 = oc1 * income
set oci2 = oc2 * income
set oci3 = oc3 * income
set oci4 = oc4 * income
set oci5 = oc5 * income
mnl dep[depvar] ivalt[ic: ic1 ic2 ic3 ic4 ic5 \
  oci: oci1 oci2 oci3 oci4 oci5]

```

A useful SST command:

```
array nalt = 5
foreach (i;{1 - $nalt}) {
  set oci$i = oc$i * income
}
```

This is equivalent to the five set statements on the previous page.

You can put as many statements as you want between the brackets.

Socioeconomic Variables, Alternative-Specific Constants, and Taste Variation

Model:

$$V_c = \alpha C_c + \beta T_c$$

$$V_b = a + \alpha C_b + \beta T_b$$

$$a = c + d \cdot \text{Income}$$

$$\beta = \gamma + \delta \cdot \text{Income}$$

Substituting in:

$$V_b = c + d \cdot \text{Income} + \alpha C_b + \gamma T_b + \delta (\text{Income} \cdot T_b)$$

WORKSHOP ON SPECIFICATION

Run models with the following explanatory variables, using **sim.cmd**.

MODEL 1

Installation cost, operating cost, and an alternative-specific constant in each of alternatives 2, 3, 4, and 5 (not 1). That is,

$$\begin{aligned}V_1 &= \beta_1 \cdot IC_1 + \beta_2 \cdot OC_1 \\V_2 &= \beta_1 \cdot IC_2 + \beta_2 \cdot OC_2 + \beta_3 \\V_3 &= \beta_1 \cdot IC_3 + \beta_2 \cdot OC_3 + \beta_4 \\V_4 &= \beta_1 \cdot IC_4 + \beta_2 \cdot OC_4 + \beta_5 \\V_5 &= \beta_1 \cdot IC_5 + \beta_2 \cdot OC_5 + \beta_6\end{aligned}$$

Verify that the predicted shares equal the actual shares.

Now consider a model with installation cost, operating cost, and an alternative-specific constant in each of alternatives 1, 3, 4, and 5 (not 2). What are the estimated coefficients of these variables? Figure them out logically from the results for model 1, rather than run the new model.

MODEL 2

Installation cost divided by income, operating cost, and an alternative-specific constant in alternatives 2, 3, 4, and 5.

What is the reason, or explanation, for the variable "installation cost divided by income?"

MODEL 3

Installation cost, operating cost, alternative-specific income effects entering alternatives 2, 3, 4, and 5, and alternative-specific constants entering alternatives 2, 3, 4, and 5.

Do you think model 2 or model 3 more accurately captures the effect of income?

MODEL 4

Try various specifications and find the model that you consider best.

DISCUSSION OF WORKSHOP RESULTS