# Transportation Economics and Decision Making

Lecture-9

# Measurement Error

- Let us assume that different individuals have different travel times for the automobile modes.
- Specifically assume that the drive alone and carpool travel times for individuals are distributed in the following relative frequencies

| Percentage of individuals | 20% | 50% | 20% | 10% |
|---|---|---|---|---|
| DA Time | 0.4 | 0.5 | 0.60 | 0.70 |
| Carpool Time | 0.65 | 0.75 | 0.85 | 0.95 |

| Percentage Individuals | Zero Cars | | | | No Auto Ownership |
|---|---|---|---|---|---|
| | 20% | 50% | 20% | 10% | -1.17 |
| Drive Alone | -1.47 | -1.57 | -1.67 | -1.77 | -1.08 |
| Carpool | -1.18 | -1.28 | -1.38 | -1.48 | -1.25 |
| Bus | -1.25 | -1.25 | -1.25 | -1.25 | Carpool |
| Chosen Mode | Carpool | Bus | Bus | Bus | Carpool |

# Measurement Error

| Percentage Individuals | One Cars | | | | |
|---|---|---|---|---|---|
| | 20% | 50% | 20% | 10% | No Auto Ownership |
| Drive Alone | -1.07 | -1.17 | -1.27 | -1.37 | -1.17 |
| Carpool | -0.98 | -1.08 | -1.18 | -1.28 | -1.08 |
| Bus | -1.25 | -1.25 | -1.25 | -1.25 | -1.25 |
| Chosen Mode | Carpool | Carpool | Carpool | Bus | Carpool |

| Percentage Individuals | Two Cars | | | | |
|---|---|---|---|---|---|
| | 20% | 50% | 20% | 10% | No Auto Ownership |
| Drive Alone | -0.67 | -0.77 | -0.87 | -0.97 | -1.17 |
| Carpool | -0.78 | -0.88 | -0.98 | -1.08 | -1.08 |
| Bus | -1.25 | -1.25 | -1.25 | -1.25 | -1.25 |
| Chosen Mode | Drive Alone | Drive Alone | Drive Alone | Drive Alone | Carpool |

# Measurement Error (2)

- Ignoring distribution of travel times of zero and one car households result in predictions that do not reflect the true variations in mode choice.
- In other words, actual choices vary in ways not explained by the model used to make predictions.

# Differences of Preference Among Individuals

- It is possible that, for reasons not known to the analyst, different individuals have different preferences among the same sets of alternatives.
- For example a population consists of two groups in which time is valued differently.

# Differences of Preference Among Individuals (2)

- Group-1
  - $U_{DA}$ (1) = $-0.75T_{DA}-5C_{DA}/Y+0.4(A-1)$
  - $U_{CP}$ (1)= $-0.75T_{DA}-5C_{CP}/Y+0.2(A-1)$
  - $U_{B}$ (1) = $-0.75T_{B}-5C_{B}/Y$
- Group-2
  - $U_{DA}$ (2) = $-1.5T_{DA}-5C_{DA}/Y+0.4(A-1)$
  - $U_{CP}$ (2)= $-1.5T_{DA}-5C_{CP}/Y+0.2(A-1)$
  - $U_{B}$ (2) = $-1.5T_{B}-5C_{B}/Y$

# Differences of Preference Among Individuals (3)

| Mode | Time (T), Hours | Cost ©, $ |
|---|---|---|
| Drive Alone | 0.5 | 2 |
| Carpool | 0.75 | 1 |
| Bus | 1 | 0.75 |

| Mode | Zero Cars | | One Car | | Two Cars | |
|---|---|---|---|---|---|---|
| | Group-1 | Group-2 | Group-1 | Group-2 | Group-1 | Group-2 |
| Drive Alone | -1.44 | -1.82 | -1.04 | -1.42 | -0.64 | -1.02 |
| Carpool | -1.10 | -1.66 | -0.90 | -1.46 | -0.70 | -1.26 |
| Bus | -1.00 | -1.75 | -1.00 | -1.75 | -1.00 | -1.75 |
| Chosen Mode | Bus | Carpool | Carpool | Drive Alone | Drive Alone | Drive Alone |

# Differences of Preference Among Individuals (4)

- Notice that for zero and one car households, the mode chosen by members of group 1 is different from group 2.
- If the analyst does not know that individuals belong to different preference groups, it will appear that "identical" individuals (who have same income and auto ownership) make different choices when faced with identical alternatives.

# Multiple Sources of Unexplained Variation in Choices

- In the first example we considered all individuals will choose carpool

| Mode | Time (T), Hours | Cost ©, $ | One Car |
|---|---|---|---|
| Drive Alone | 0.5 | 2 | -1.17 |
| Carpool | 0.75 | 1 | -1.08 |
| Bus | 1 | 0.75 | -1.25 |

- However the actual choices of a specific individual will depend on
  - Which preference group he/she is
  - Number of automobiles owned by the household
  - True travel time
- Modes chosen will vary among individuals if they are not incorporated in the utility function

# Basic Formulation of Probabilistic Models

- The correct utility function can be written as the sum of utility function specified by the analyst and an error term
- U = V + e

| Mode | Specified Utility | Error Term |
|---|---|---|
| Drive Alone | $-T_{DA}-5C_{DA}/Y$ | 0.4(A-1) |
| Carpool | $-T_{CP}-5C_{CP}/Y$ | 0.2(A-1) |
| Bus | $-T_{B}-5C_{B}/Y$ | 0 |

# Basic Formulation of Probabilistic Models (2)

- Alternatively

| Mode | Specified Utility | Error Term |
|---|---|---|
| Drive Alone | $-T^*_{DA}-5C_{DA}/Y$ | $T_{DA}-T^*_{DA}$ |
| Carpool | $-T^*_{CP}-5C_{CP}/Y$ | $T_{CP}-T^*_{CP}$ |
| Bus | $-T_B-5C_B/Y$ | 0 |

- T* is the measured travel time and T is the true travel time

# Basic Formulation of Probabilistic Models (3)

- Utilities of drive alone and carpool include an error term because some individuals are measured with errors with these modes.
- There is no error term for bus because bus travel time is measured without error

| Mode | Specified Utility | Error Term |
|---|---|---|
| Drive Alone | $-T_{DA}-5C_{DA}/Y+0.4(A-1)$ | $0.25T_{DA}$ (group-1)<br>$-0.50T_{DA}$ (group-2) |
| Carpool | $-T_{CP}-5C_{CP}/Y+0.2(A-1)$ | $0.25T_{CP}$ (group-1<br>-0.50 (group-2) |
| Bus | $-T_B-5C_B/Y$ | |

# Basic Formulation of Probabilistic Models (4)

- If the true utility function is known then there is no need to specify an error term.
- But in real world we always have unknown factors
- Therefore, an error term of unknown size is always present in the analyst's specification of error term

# Dependence of Choice on Deterministic Component of Utility

- Let the population be distributed over automobile ownership classes and preference groups according to the following percentages

| Preference Group | Automobiles Owned | | |
|---|---|---|---|
| | 0 | 1 | 2 |
| Preference Group-1 | 15% | 25% | 5% |
| Preference Group-2 | 5% | 30% | 20% |

# Dependence of Choice on Deterministic Component of Utility (2)

Taking result from the previous example and merging with auto ownership

| Mode | Zero Cars | | One Car | | Two Cars | |
|---|---|---|---|---|---|---|
| | Group-1 | Group-2 | Group-1 | Group-2 | Group-1 | Group-2 |
| Drive Alone | -1.44 | -1.82 | -1.04 | -1.42 | -0.64 | -1.02 |
| Carpool | -1.10 | -1.66 | -0.90 | -1.46 | -0.70 | -1.26 |
| Bus | -1.00 | -1.75 | -1.00 | -1.75 | -1.00 | -1.75 |
| Chosen Mode | Bus | Carpool | Carpool | Drive Alone | Drive Alone | Drive Alone |

| Autos Owned | Group | % Population | Mode Chosen |
|---|---|---|---|
| 0 | 1 | 15 | Bus |
| | 2 | 5 | Carpool |
| 1 | 1 | 25 | Carpool |
| | 2 | 30 | Drive Alone |
| 2 | 1 | 5 | Drive Alone |
| | 2 | 20 | Drive Alone |

# Dependence of Choice on Deterministic Component of Utility (3)

- Now let us consider a parking tax is imposed on single-occupant vehicles, so automobile cost becomes $2.5
- The revised choices will be following

| Mode | Zero Cars | | One Car | | Two Cars | |
|---|---|---|---|---|---|---|
| | Group-1 | Group-2 | Group-1 | Group-2 | Group-1 | Group-2 |
| Drive Alone | -1.61 | -1.98 | -1.21 | -1.58 | -0.81 | -1.18 |
| Carpool | -1.10 | -1.66 | -0.90 | -1.46 | -0.70 | -1.26 |
| Bus | -1.00 | -1.75 | -1.00 | -1.75 | -1.00 | -1.75 |
| Chosen Mode | Bus | Carpool | Carpool | Carpool | Carpool | Drive Alone |

# Deterministic Component of Utility

- If we take the deterministic component of the utility only with variation in parking tax will result in the following

| Tax | Deterministic Component of Utility | % of Population Choosing | | |
|---|---|---|---|---|
| | | Drive Alone | Carpool | Bus |
| 0 | -1.17 | 55 | 30 | 15 |
| 0.25 | -1.25 | 20 | 65 | 15 |
| 0.50 | -1.34 | 20 | 65 | 15 |

- Note, with deterministic component of utility the outcome does not change with increase in tax.

# Deterministic Component of Utility (2)

- Suppose the income distribution is as following

| Income (1000) | % Zero Car Households | % One Car Households | % Two Car Households | Total % |
|---|---|---|---|---|
| 17.5 | 40 | 50 | 10 | 100 |
| 22.5 | 20 | 60 | 20 | 100 |
| 27.5 | 15 | 60 | 25 | 100 |
| 32.5 | 10 | 60 | 30 | 100 |
| 37.5 | 5 | 55 | 40 | 100 |
| 42.5 | 0 | 50 | 50 | 100 |

- The resulting choices will be as follows

| Income (1000) | % Zero Car Households | % One Car Households | % Two Car Households |
|---|---|---|---|
| 17.5 | Bus | Carpool | Drive Alone |
| 22.5 | Carpool | Drive Alone | Drive Alone |
| 27.5 | Carpool | Drive Alone | Drive Alone |
| 32.5 | Carpool | Drive Alone | Drive Alone |
| 37.5 | Carpool | Drive Alone | Drive Alone |
| 42.5 | Carpool | Drive Alone | Drive Alone |

# Population distribution and choices

- Percent choosing are

| Income | Drive Alone | Carpool | Bus |
|---|---|---|---|
| 17.5 | 10 | 50 | 40 |
| 22.5 | 80 | 20 | 0 |
| 27.5 | 85 | 15 | 0 |
| 32.5 | 90 | 10 | 0 |
| 37.5 | 95 | 5 | 0 |
| 42.5 | 100 | 0 | 0 |

- If population distribution is given then number of individuals choosing different modes are as follows

| Income | Number of Individuals | Drive Alone | Carpool | Bus |
|---|---|---|---|---|
| 17.5 | 20 | 10 | 50 | 40 |
| 22.5 | 60 | 80 | 20 | 0 |
| 27.5 | 100 | 85 | 15 | 0 |
| 32.5 | 100 | 90 | 10 | 0 |
| 37.5 | 80 | 95 | 5 | 0 |
| 42.5 | 40 | 100 | 0 | 0 |
| Total | 400 | 341 | 51 | 8 |

# The Logit Choice Model

- It was shown in the previous examples that
  - The analyst's limitation of knowledge make it difficult to model individual's mode choice
  - Hence it is necessary to predict these choices in terms of probabilities
- The probability that an alternative is chosen
  - should increase when the deterministic component of its utility increases
  - should decrease when deterministic component of any other alternative increases

# Binomial Logit Model

- The most frequently used model of probabilistic among two alternatives is binomial logit model
- In this model, the probability that alternative 1 is chosen when the choice set consists of two alternatives is given by

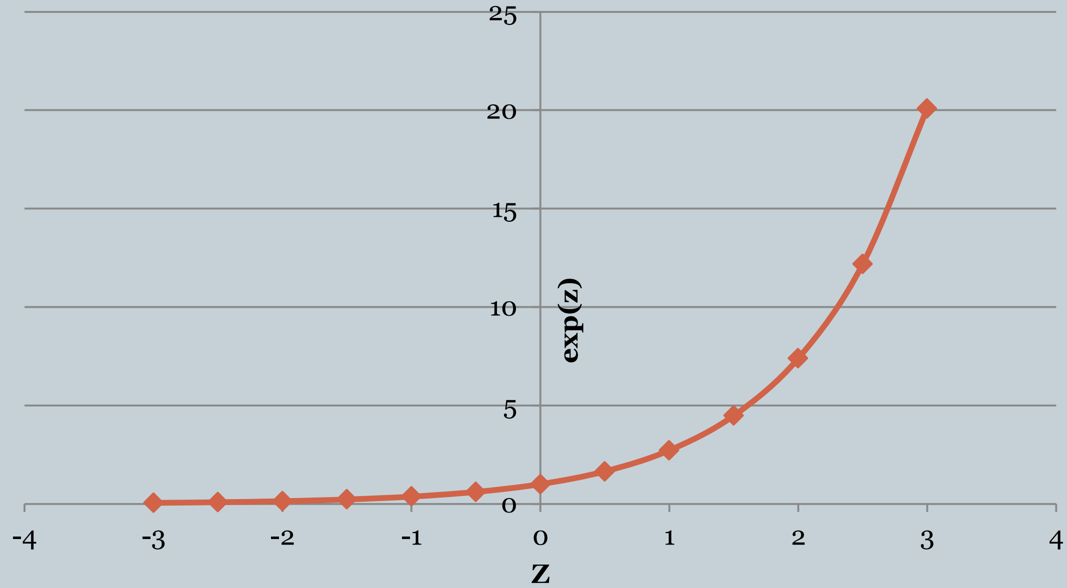
$$P(1) = \frac{e^{v_1}}{e^{v_1} + e^{v_2}}$$

- Where P(1) is the probability that an individual choses alternative 1; $v_1$ and $v_2$ are deterministic utility components of two alternatives

# Binomial Logit Model

- The exponential function is monotonic
  - i.e. Its value increases when the utility of the alternative increases

| v | Exp(v) |
|---|---|
| -3 | 0.049787 |
| -2.5 | 0.082085 |
| -2 | 0.135335 |
| -1.5 | 0.22313 |
| -1 | 0.367879 |
| -0.5 | 0.606531 |
| 0 | 1 |
| 0.5 | 1.648721 |
| 1 | 2.718282 |
| 1.5 | 4.481689 |
| 2 | 7.389056 |
| 2.5 | 12.18249 |
| 3 | 20.08554 |



# Binomial Logit Model (2)

- Further, P(1) is the probability that a randomly selected individual with deterministic components of utility $v_1$ and $v_2$ choses alternative 1.
- Since there are only two alternatives, probability that alternative 2 is chosen is given by

$$P(2) = 1 - P(1)$$

$$=1-\frac{e^{v_1}}{e^{v_1}+e^{v_2}}$$

$$=\frac{e^{v_2}}{e^{v_1}+e^{v_2}}$$

# Binomial Logit Model (3)

- Desirable properties
  - Property-1: Logit choice probabilities depend on the deterministic components of utilities of all alternatives
  - Property-2: Probability decreases when the deterministic component of the utility of other alternative increases
  - Property-3: Can not treat choice more than two alternatives

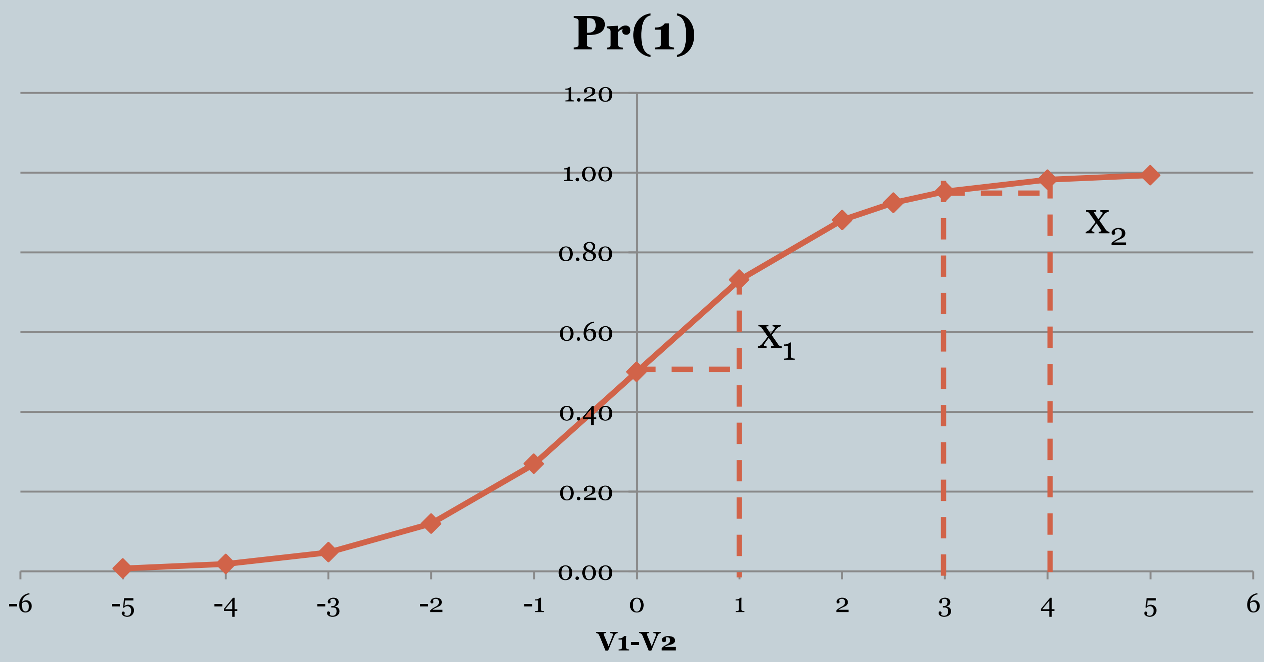
# Relative Utilities (1)

- The choice probabilities are most sensitive to changes in the deterministic component of the utilities when these components are approximately equal and the choice probabilities are close to 0.5

$$P(2) = \frac{e^{v_2}}{e^{v_1} + e^{v_2}}$$

$$= \frac{1}{1+e^{-(v_1 - v_2)}}$$

# Relative Utilities (2)



# Relative Utilities (3)

- The probabilities of choosing alternatives 1 and 2 are both equal to 0.5 when deterministic components of utilities are equal
- The probability of choosing alternative 1 is more sensitive to changes in the deterministic component of either alternative when P(1) is close to 0.5 than when P(1) is close to 1 (the same statement applies to P(2))
- When v1-v2 changes from 0 to 1 the probability difference is 0.23 (i.e. 0.73-0.5)
- But when v1-v2 changes from 3 to 4 the probability difference is 0.03 (i.e. 0.98-0.95)

# Relative Utilities (4)

- Another property of binomial logit model is that P(1) is affected equally by increase in the value of $v_1$ and decrease in value of $v_2$.

| Case | V1 | V2 | V1-V2 | Pr(1) |
|---|---|---|---|---|
| 1 | -5 | 0 | -5 | 0.01 |
| 2 | -4 | 0 | -4 | 0.02 |
| 3 | -3 | 0 | -3 | 0.05 |
| 4 | -2 | 0 | -2 | 0.12 |
| 5 | -1 | 0 | -1 | 0.27 |
| 6 | 0 | 0 | 0 | 0.50 |
| 7 | 1 | 0 | 1 | 0.73 |
| 8 | 2 | 0 | 2 | 0.88 |
| 9 | 2.5 | 0 | 2.5 | 0.92 |
| 10 | 2 | -0.5 | 2.5 | 0.92 |
| 11 | 3 | 0 | 3 | 0.95 |
| 12 | 4 | 0 | 4 | 0.98 |
| 13 | 5 | 0 | 5 | 0.99 |

# Multinomial Logit Model

- The binomial logit model can be easily extended to accommodate choices among more than two alternatives
- Let us consider three alternatives in the choice set
- Probability that alternative 1 is chosen

$$P(1) = \frac{e^{v_1}}{e^{v_1} + e^{v_2} + e^{v_3}}$$

# Multinomial Logit Model (2)

- If there are more alternatives than three then the probabilities can be expresses as follows

$$P(i) = \frac{e^{v_i}}{\sum_{i=1}^{I} e^{v_i}}$$

$$P(1) = \frac{e^{v_1}}{\sum_{i=1}^{I} e^{v_i}}$$

$$P(2) = \frac{e^{v_2}}{\sum_{i=1}^{I} e^{v_i}}$$

# Multinomial Logit Model (3)

- The multinomial logit model has all the desirable properties of the binomial logit model
- In addition, it can be applied to any number of alternatives
- The probability of choosing an alternative depends on the relative utilities with all other alternatives

$$P(1) = \frac{e^{v_1}}{e^{v_1} + e^{v_2} + e^{v_3}}$$

$$= \frac{1}{1+e^{-(v_1-v_2)}+e^{-(v_1-v_3)}}$$

# Example-1

- Consider travel to work and let there be three modes of choice set

| Mode | V |
|---|---|
| Drive Alone | 2.5 |
| Carpool | 2 |
| Bus | 1 |

| Mode | V | exp(v) |
|---|---|---|
| Drive Alone | 2.5 | 12.18249 |
| Carpool | 2 | 7.389056 |
| Bus | 1 | 2.718282 |
| Total | | 22.28983 |

| Mode | V | exp(v) | Probability |
|---|---|---|---|
| Drive Alone | 2.5 | 12.18249 | 0.546549 |
| Carpool | 2 | 7.389056 | 0.331499 |
| Bus | 1 | 2.718282 | 0.121952 |
| | | 22.28983 | 1 |

- As expected the mode with highest deterministic component of utility has the highest probability of being chosen