Estimation of Price Response Functions

Outline

◆ Regression
  – Linear Demand
  – Constant Elasticity
  – Logit
    » Given D
    » Joint D
◆ Maximum Likelihood Estimation
  – Logit
  – Multinomial Logit
Linear Demand and Linear Regression

- Constant willingness to pay $d(p)=D-mp$
- Estimate market size $D$ and slope $m$

- Use standard linear regression
  - Excel – Office Button – Excel Options – Add-Ins – Analysis ToolPak – Go
  - Wait for installation
  - Excel Toolbar Data – Click on Data Analysis – Pick Regression

- Given data $(d_i,p_i)$
  $$\text{Min}_{D,m} \sum_i (D - mp_i - d_i)^2$$

  Use standard Solver function
  - Excel – Office Button – Excel Options – Add-Ins – Solver – Go
  - Wait for installation
  - Excel Toolbar Data – Click on Solver

<see linear.xls>
Constant Elasticity

- Constant elasticity: \( d(p) = C p^{-\varepsilon} \).
- Estimate parameters \( C \) and elasticity \( \varepsilon \).

- Make log transformation: \( \log d(p) = \log C - \varepsilon \log p \).
- Make change of variable: \( y = \log d(p) \), \( x = \log p \), intercept = \( \log C \), slope = \( \varepsilon \) so that \( y = \text{intercept} - \text{slope} \times x \).

- Use standard **Linear Regression** to find intercept and slope.
- Given data \((y_i, x_i)\) use **Solver** to find

\[
\text{Min} \sum_{\text{Intercept, Slope}} (\text{intercept} - \text{slope} \times x_i - y_i)^2
\]

<see constant_elasticity.xls>
Logit Price Response
Given Market Size $D$

- Logit Price Response function.

$$d(p) = D \left(1 - \frac{1}{1 + e^{-(a+bp)}}\right)$$

- For given market size $D$, estimate parameters $a$ and $b$.

- Make ln transformation: $\ln \left((D-d(p))/d(p)\right) = a+bp$.

- Make change of variable: $y=\ln \left(D/d(p)-1\right)$, $x=p$, intercept=$a$, slope = $b$ so that $y=\text{intercept}+\text{slope}*x$.

- Use standard **Linear Regression** to find intercept and slope.

- Given data $(y_i,x_i)$ use **Solver** to find

$$\text{Min}_{\text{Intercept, Slope}} \sum_{i} (\text{intercept} + \text{slope} * x_i - y_i)^2$$

<see logit_givenD.xls>
Logit Price Response
Unknown Market Size $D$

- Logit Price Response function.

$$d(p) = D \left( 1 - \frac{1}{1 + e^{-(a+bp)}} \right)$$

- Estimate parameters $D$, $a$ and $b$.

- No log or ln transformation possible to cast the estimation as a linear regression.

- Given data $(d_i, p_i)$, solve to estimate $a$, $b$, $D$

$$\text{Min}_{D, a, b} \sum_i \left( D \left( 1 - \frac{1}{1 + e^{-(a+bp_i)}} \right) - d_i \right)^2$$

- Solver does not (convincingly) solve this problem.
Logit Price Response Unknown Market Size $D$

Sequential Approach

- Suppose that $D$ is given.
- Estimate $a, b$.
- Update $D$ such that it solves
  \[
  \frac{\partial}{\partial D} \sum_i \left( \left( 1 - \frac{1}{1 + e^{-a-bp_i}} \right) D - d_i \right)^2 = 0
  \]
  \[
  \sum_i 2 \left( \left( 1 - \frac{1}{1 + e^{-a-bp_i}} \right) D - d_i \right) \left(1 - \frac{1}{1 + e^{-a-bp_i}} \right) = 0
  \]
- So
  \[
  D(d_i, p_i; a, b) = \frac{\sum_i d_i \left( 1 - \frac{1}{1 + e^{-a-bp_i}} \right)}{\sum_i \left( 1 - \frac{1}{1 + e^{-a-bp_i}} \right)^2}
  \]

Sequential Algorithm

0. Start with an initial guess for $D$
1. Use current $D$ to estimate $a$ and $b$ by using the procedure of Logit with given $D$.
2. Update $D$ so that it sets the derivative of the sum of squares of errors to zero
3. Stop or go to 1.
Logit Price Response
Sequential Algorithm

- Sequential algorithm is implemented in logit_jointD.xls.
- Unfortunately, $D - D(d_i, p_i; a, b)$ is small:
  - In other words, updates do not change the market size much.
- Hypothesis: $a$ and $b$ are sufficient to represent market size.
- Graph actual demand vs. estimated demand for

<table>
<thead>
<tr>
<th>Actual</th>
<th>Est-1</th>
<th>Est-2</th>
<th>Est-3</th>
<th>LSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>D=100</td>
<td>37</td>
<td>70</td>
<td>150</td>
<td>200</td>
</tr>
<tr>
<td>a=0.1</td>
<td>-3.29</td>
<td>-0.41</td>
<td>0.76</td>
<td>0.19</td>
</tr>
<tr>
<td>b=0.3</td>
<td>0.64</td>
<td>0.31</td>
<td>0.26</td>
<td>0.33</td>
</tr>
</tbody>
</table>
Logit Price Response
Sequential Algorithm Graphs

Based on these computations

Conclusion: LSE estimates are significantly worse than the Sequential ones.
Reason: Excel Solver fails to minimize sum of squares, a highly nonlinear objective.
Suggestion: Select 2-3 values of D and use Sequential algorithm.
Change of Perspective

From curve fitting
Minimize the difference between data and curve
Min \( d_i - \text{curve}(p_i; \text{parameters}) \)

To probabilistic model fitting
Maximize the probability of observing the data
Curve \( \leftarrow \) A probabilistic model
Max Probability(Observing \( d_i \) with parameters)
Maximum Likelihood Estimation (MLE)  
Maximizing to Match Model and Data

- Recall: Probability of selling: \(1 - W(p; a, b)\) with the price \(p\) and the WTP function \(W\), parameters \(a, b\).

- Consider 1 customer with WTP function \(W\) and conduct the price experiment:
  - Offer price \(p\) to this customer and record a sale as \(y = 1\) and no-sale as \(y = 0\)
  - Collect (scanner) data \(\{(y, p)\}\)

- When \(y = 1\) in the data, we ideally want the consistent outcome from the model \(1 - W(p; a, b) = 1\)
  - However, \(1 - W(p; a, b) = 1\) hardly happens except when \(p\) is equal to the lowest WTP.

- Give up idealism and pragmatically match
  - [Sales in the model] with [Sales in the data]
    - [Model’s \(1 - W(p; a, b)\) to be high] when [data \(y = 1\)] \(\Rightarrow\) \(\max y(1 - W(p; a, b))\)
  - [No-sales in the model] with [No-sales in the data]
    - [Model’s \(W(p; a, b)\) to be high] when [data \(y = 0\)] \(\Rightarrow\) \(\max (1 - y)W(p; a, b)\)
    - Since \(y \in \{0, 1\}\), combine two objectives to obtain \(\Rightarrow\) \(\max y(1 - W(p; a, b)) + (1 - y)W(p; a, b)\)

- Consider sales to the customer as a binary random variable \(Y \in \{0, 1\}\) with probabilities
  - \(P(Y = 1) = (1 - W(p; a, b))\) and \(P(Y = 0) = W(p; a, b)\)
  - \(P(Y = y) = y(1 - W(p; a, b)) + (1 - y)W(p; a, b)\) for \(y \in \{0, 1\}\).
  - The objective we arrived at through consistency argument is \(\max P(Y = y)\)

- Maximize the likelihood (probability) of the realized event \([Y = y]\) (sales)
MLE with $N \geq 2$ Customers

- For $N$ independent customers with identical WTP functions: the same $W, a, b$.
  - Repeat price experiment twice and collect data $\text{Data} = \{(y_1, p_1), (y_2, p_2)\}$
  - $\max P(Y_1 = y_1, Y_2 = y_2) = \max P(Y_1 = y_1)P(Y_2 = y_2)

- Maximize the likelihood function $L$ whose
  $$L(a, b; \{(y_1, p_1), (y_2, p_2)\}) = \left[y_1 \left(1 - W(p_1; a, b)\right) + (1 - y_1)W(p_1; a, b)\right] \times \left[y_2 \left(1 - W(p_2; a, b)\right) + (1 - y_2)W(p_2; a, b)\right]$$

- Variables $(a, b)$ are the parameters of WTP and
- parameters $(y_n, p_n)$ are the variables of WTP & its consequence $1 - W(\cdot; a, b)$.

- For $N$ independent customers with continuous WTP,
  $$L(a, b; \text{Data}) = \prod_{n=1}^{N} \left[y_n P(\text{Sale at price } p_n) + (1 - y_n)P(\text{No sale at } p_n)\right]$$
  $$= \prod_{n=1}^{N} \left[y_n \left(1 - W(p_n; a, b)\right) + (1 - y_n)W(p_n; a, b)\right].$$

- For $N$ independent customers with discrete WTP,
  $$L(a, b; \text{Data}) = \prod_{n=1}^{N} \left[y_n \left(1 - W(p_n^-; a, b)\right) + (1 - y_n)W(p_n^-; a, b)\right],$$
  where $p_n^- :=$ the largest WTP value strictly less than $p_n$.

Likelihoods are multiplications of probability $\Rightarrow L(a, b; \text{Data}) \leq 1$ and $L(a, b; \text{Data}) \downarrow N.$
Examples of MLE

Ex: Suppose that the WTP for a shirt is uniformly distributed between unknown parameters $a \leq 40$ and $b \geq 60$. As the retailer you dealt with two customers, one bought at the price $40$ and the other did not at the price $60$. What are the associated scanner data? Set up the likelihood function to estimate $a, b$.

- $D = \{(y_1 = 1, p_1 = 40), (y_2 = 0, p_2 = 60)\}$
- $L(a, b; \{(y_1, p_1), (y_2, p_2)\})$

\[
= \left[ y_1 (1 - W(p_1; a, b)) + (1 - y_1) W(p_1; a, b) \right] \left[ y_2 (1 - W(p_2; a, b)) + (1 - y_2) W(p_2; a, b) \right]
\]

\[
= \left[ 1 \frac{b-p_1}{b-a} + 0 \right] \left[ 0 + 1 \frac{p_2-a}{b-a} \right] \frac{(b-40)(60-a)}{(b-a)^2} \leq 1
\]

- Maximum value of $1$ obtained at $a = 40$ and $b = 60$.

Ex: Suppose that the WTP for a shirt is uniformly distributed between unknown parameters $a \leq 40$ and $b \geq 60$. As the retailer you dealt with three customers, one bought at the price $40$, the other bought at the price $50$ but the last did not at the price $60$. What are the associated scanner data? Set up the likelihood function to estimate $a, b$.

- $D = \{(y_1 = 1, p_1 = 40), (y_2 = 1, p_2 = 50), (y_3 = 0, p_3 = 60)\}$
- $L(a, b; \{(y_1, p_1), (y_2, p_2), (y_3, p_3)\})$

\[
= \left[ 1 \frac{b-p_1}{b-a} + 0 \right] \left[ 1 \frac{b-p_2}{b-a} + 0 \right] \left[ 0 + 1 \frac{p_3-a}{b-a} \right] \frac{(b-40)(b-50)(60-a)}{(b-a)^3} \leq 1
\]

- For each fixed $b$, we want $a$ as large as possible. Because denominator is cubic, decreasing in $a$.
- We set $a = 40$.
- The objective then is $(b-40)(b-50)/(b-40)^3$.
- This is larger when $b$ is smaller. We set $b = 60$.
- The maximum likelihood value is 0.5
- In comparison with the last example, likelihood dropped with the third data point $(y_3, p_3)$.

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
b \backslash a & 40 & 38 & 36 & 34 & 32 \\
\hline
60 & 0.500 & 0.413 & 0.347 & 0.296 & 0.255 \\
62 & 0.496 & 0.420 & 0.361 & 0.313 & 0.274 \\
64 & 0.486 & 0.421 & 0.367 & 0.324 & 0.287 \\
66 & 0.473 & 0.417 & 0.370 & 0.330 & 0.296 \\
68 & 0.459 & 0.411 & 0.369 & 0.333 & 0.302 \\
\hline
\end{array}
\]

<see maxLikelihood.xls>
Multinomial Logit Choice Model

The most common consumer choice model is multinomial logit (MNL) for \( K \) products

\[
\text{Probability of buying product } k = f_k(p = [p_1, p_2, \ldots, p_K]) = \frac{e^{-b_k p_k}}{\sum_{j=1}^{K} e^{-b_j p_j}}.
\]

Parameters are \( b_1, \ldots, b_K \). Parameter \( b_k \) is large when product \( k \) is price sensitive.

Ex: Suppose \( K = 3 \) products A, B, C, each with the probability of choice \( f_k \), \( k \in \{A, B, C\} \). \( N = 9 \) customers bought products as follows. 2 bought A; 3 bought B and 4 bought C. What is the probability of this event?

\[
P(2A; 3B; 4C) = \frac{9!}{2!3!4!} f_A^2 f_B^3 f_C^4.\]

This is also known as multinomial distribution.

For customer 1, let \( y_{1A} = 1, y_{1B} = 0, y_{1C} = 0 \). Similarly define \( y_2 = [y_{2A}, y_{2B}, y_{2C}] \) for customer 2 such that \( y_2 = y_1 \).

For customer 3, let \( y_{3A} = 0, y_{3B} = 1, y_{3C} = 0 \). Similarly define \( y_4 = y_5 := y_3 \) for customers 4 and 5.

For customer 6, let \( y_{6A} = 0, y_{6B} = 0, y_{6C} = 1 \). Define \( y_7 = y_8 = y_9 := y_6 \) for customers 7, 8 and 9.

We can rewrite \( P(y's) = [\text{Constant}] f_A^{y_{1A}+y_{2A}} f_B^{y_{3B}+y_{4B}+y_{5B}} f_C^{y_{6C}+y_{7C}+y_{8C}+y_{9C}} \)

\[
= [\text{Constant}] f_A^{y_{1A}+y_{2A}} f_B^{y_{3B}+y_{4B}+y_{5B}} f_C^{y_{6C}+y_{7C}+y_{8C}+y_{9C}}
\]

\[
= [\text{Constant}] \Pi_{k \in \{A,B,C\}} \Pi_{n=1}^{9} f_k^{y_{nk}}
\]

Multinomial Logit maximum likelihood estimators for a given sample of 1, 2, \ldots, \( N \) individuals

- Each individual \( n \) made a choice \( y_{nk} \) in response to prices \([p_{n1}, p_{n2}, \ldots, p_{nK}]\).
  Indicator variable \( y_{nk} = 1 \) if individual \( n \) chooses product \( k \); otherwise zero.
- Data for individual \( n \): \([y_{n1}, y_{n2}, \ldots, y_{nK}, p_{n1}, p_{n2}, \ldots, p_{nK}]\).
- What are the most likely values of parameters \([b_1, b_2, \ldots, b_K]\), i.e., what parameter values maximize the probability of choices?
MLE of Multinomial Logit Choice Model

- From the multinomial probability mass function, the likelihood is

\[ L([b_1, b_2, \ldots, b_K]) = \text{[Const.]} \prod_{n,k} f_k^{y_{nk}} = \text{[Const.]} \prod_{n,k} \left( \frac{e^{-b_k p_{nk}}}{\sum_{j=1}^{K} e^{-b_j p_{nj}}} \right)^{y_{nk}} \]

- Instead maximize the logarithm of the likelihood without the constant

\[ \log L([b_1, b_2, \ldots, b_K]) = \sum_{n,k} y_{nk} \log \left( \frac{e^{-b_k p_{nk}}}{\sum_{j=1}^{K} e^{-b_j p_{nj}}} \right) \]

- Use a solver to find parameters \([b_1, b_2, \ldots, b_n]\) that maximize \(\log L([b_1, b_2, \ldots, b_n])\).
Consider the Fishing mode and Income example on Page 494 of Microeconometrics: Methods and Applications (2005) by Cameron, Adrian Colin.; Trivedi, P. K. Book available as an ebook from UTD Library.

Fishing modes are: Beach, Pier, Private Boat, Charter (boat).
Individuals choose one of these.
Each individual reports own income and choice.
Data: For individual n: \([y_{n,\text{Beach}}, y_{n,\text{Pier}}, y_{n,\text{Private}}, y_{n,\text{Charter}}; \text{income}_n]\)
Statistical model with parameters \(a\) and \(b\) for an individual:

\[
P(y_{\text{Beach}} = 1) = \frac{e^{a_{\text{Beach}} + b_{\text{Beach}} \cdot \text{Income}}}{e^{a_{\text{Beach}} + b_{\text{Beach}} \cdot \text{Income}} + e^{a_{\text{Pier}} + b_{\text{Pier}} \cdot \text{Income}} + e^{a_{\text{Private}} + b_{\text{Private}} \cdot \text{Income}} + e^{a_{\text{Charter}} + b_{\text{Charter}} \cdot \text{Income}}}\]

\[
P(y_{\text{Pier}} = 1) = \frac{e^{a_{\text{Pier}} + b_{\text{Pier}} \cdot \text{Income}}}{e^{a_{\text{Beach}} + b_{\text{Beach}} \cdot \text{Income}} + e^{a_{\text{Pier}} + b_{\text{Pier}} \cdot \text{Income}} + e^{a_{\text{Private}} + b_{\text{Private}} \cdot \text{Income}} + e^{a_{\text{Charter}} + b_{\text{Charter}} \cdot \text{Income}}}\]
Downloading R

1. Download standard ≥ R-2.9.2 version from www.r-project.org
2. This should create R-2.9.2 directory and, underneath it, directories
   bin; doc; etc; include; library; modules; share; src; Tcl
3. Among those directories, library is important to us as we shall add estimation specific and other packages to
   this library. Library directory should have 27 subdirectories.

Downloading Useful Packages

Download the following packages from http://cran.r-project.org/web/packages/
   xlsReadWrite.zip, maxLik.zip, mlogit.zip, Ecdat.zip; nls2.zip
   into new directories that you create under R-2.9.2\library with names xlsReadWrite, maxLik, mlogit, Ecdat.

The roles of these packages are:

xlsReadWrite;  ## Required for im(ex)porting excel files
maxLik;        ## Required for Maximum Likelihood Estimation
mlogit;        ## Required for Multinomial Logit model
Ecdat;         ## Interesting Econometric data files
nls2;          ## Nonlinear least squares

Check the directories in R-2.9.2\library

You should have directories 27 standard R directories, plus 5 that you have manually added above. At this point the
entire R-2.9.2 directory takes 86,392,832 bytes on my hard disk.
Starting R

1. Start R (click on the icon on your desktop, on the quick start button on your Start menu, or click on R-2.9.2\bin\Rgui.exe). This will start R with a home directory of R-2.9.2.
2. I suggest that you keep your data in a different directory, say C:\Demreman\R\.
3. You have to tell R which directory you want to work in. Go to File menu and then click on to Change dir(ectory)

Make R read your packages
   In R, issue commands
   > library(xlsReadWrite);
   > library(maxLik);
   > library(mlogit);
   > library(Ecdat);
   > library(nls2);
Multinomial Example: Fishing Mode

🔹 To read the Fishing data, issue command

  > data("Fishing",package="mlogit");

🔹 Briefly Fishing data are about 1182 individuals' fishing mode choices. Data come from a survey conducted by Thomson and Crooke (1991).

🔹 Issue

  > fix(Fishing);

  to see what is inside the Fishing dataframe structure. It has 1 row for each individual and 1182 rows in total.

🔹 It has 12 columns = 3 columns + 4 columns + 4 columns + 1 column.
  – The first 3 columns have the chosen mode of fishing. Its price and the probability of catching a fish.
  – The next four columns have the price for each mode of fishing. These prices change from one individual to another as they can depend on the location and access of the individual.
  – The next four columns have the catch probability for each mode of fishing.
  – The last column contains the monthly income of the individual.
Multinomial Example: Fishing Mode
Data Manipulation

◆ Issue command
  
  ```R
  > Fish <- mlogit.data(Fishing, varying = c(4:11), shape = "wide", choice = "mode")
  ```
  
  to prepare Fishing data for multinomial logit regression.

  - Immediately the number of rows become 4728 (=1182*4), so one row for each individual and each alternative. See the Fish dataframe by issuing
    
    ```R
    > fix(Fish);
    ```
  
  - The first four rows are now for the first individual whose identity is in the column named chid. The next column is the alternative. By putting individual id and alternative together we obtain the first column, named row.names.

◆ The important columns for our purpose are named mode and income.

◆ Manipulation on the model

\[
\frac{P(y_{Pier} = 1)}{P(y_{Beach} = 1)} = \frac{e^{a_{Pier} + b_{Pier} \cdot \text{Income}}}{e^{a_{Beach} + b_{Beach} \cdot \text{Income}}} = \exp(a_{Pier} - a_{Beach} + (b_{Pier} - b_{Beach}) \cdot \text{Income}),
\]

\[
\log \left( \frac{P(y_{Pier} = 1)}{P(y_{Beach} = 1)} \right) = a_{Pier} - a_{Beach} + (b_{Pier} - b_{Beach}) \cdot \text{Income}
\]

◆ So \(a_{Beach}\) and \(b_{Beach}\) can be assumed to be zero to estimate the other parameters with respect to these two.
### Multinomial Example: Fishing Mode Logit Regression

- **Issue command**
  ```r
  > summary(mlogit(mode~1|income, data=Fish));
  ```
  to estimate \(a, b\) parameters. R outputs
  - Call: mlogit(formula = mode ~ 1 | income, data = Fish)
  - Frequencies of alternatives:
    - beach: 0.11337
    - boat: 0.35364
    - charter: 0.38240
    - pier: 0.15059
  - Newton-Raphson maximisation gradient close to zero. May be a solution 5 iterations, 0h:0m:0s \(g'(H)^{-1}g = 9.47E-30\)
  - Coefficients:

| Parameter   | Estimate     | Std. Error   | t-value | Pr(>|t|)       |
|-------------|--------------|--------------|---------|--------------|
| altboat     | 7.3892e-01   | 1.9673e-01   | 3.7560  | 0.0001727 ***|
| altcharter  | 1.3413e+00   | 1.9452e-01   | 6.8955  | 5.367e-12 ***|
| altpier     | 8.1415e-01   | 2.2863e-01   | 3.5610  | 0.0003695 ***|
| altboat:income | 9.1906e-05 | 4.0664e-05   | 2.2602  | 0.0238116 *   |
| altcharter:income | -3.1640e-05 | 4.1846e-05  | -0.7561 | 0.4495908      |
| altpier:income | -1.4340e-04 | 5.3288e-05   | -2.6911 | 0.0071223 **   |

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

##Note that numbers in red are the coefficients in column MNL of Table 15.2 of Cameron and Trivedi (2005).
- Log-Likelihood: -1477.2; McFadden R^2: 0.77463; Likelihood ratio test: chisq = 41.145 (p.value=6.0931e-09)
Multinomial Example: Fishing Mode
Interpretation of Logit Regression

- Inserting the coefficients

\[
\log \left( \frac{P(y_{\text{Boat}} = 1)}{P(y_{\text{Beach}} = 1)} \right) = 0.739 + 0.092 \cdot \frac{\text{Income}}{1000}
\]

\[
\log \left( \frac{P(y_{\text{Charter}} = 1)}{P(y_{\text{Beach}} = 1)} \right) = 1.341 - 0.032 \cdot \frac{\text{Income}}{1000}
\]

\[
\log \left( \frac{P(y_{\text{Pier}} = 1)}{P(y_{\text{Beach}} = 1)} \right) = 0.814 - 0.143 \cdot \frac{\text{Income}}{1000}
\]

- As the income increases, one is much less likely to fish on the pier (-0.143), less likely to fish on a charter boat (-0.032) and more likely to fish on a private boat (0.092).
Summary

◆ Regression
  – Linear Demand
  – Constant Elasticity
  – Logit
    » Given D
    » Joint D

◆ Maximum Likelihood Estimation
  – Logit
  – Multinomial Logit

Another Example in R: Transportation Mode

Issue command

```r
> data("Mode",package="mlogit");
```
to read the transportation Mode data.

Transportation modes are Car, Carpool, Bus, Rail.

Data are from 453 individuals.

Issue

```r
> fix(Mode)
```
to see what is inside the Mode dataframe. It has 1 row for each individual and 453 rows in total.

It has 12 columns = 1 column + 4 columns + 4 columns.
- The first column is the chosen mode of transportation.
- The next four columns are the price of each mode.
- The next four columns are the duration of each mode.

Issue command

```r
> market <- mlogit.data(Mode, alt.levels=c("car", "carpool", "bus", "rail"), shape="wide", choice="choice");
```
to prepare Mode data for multinomial logit regression.
Transportation Mode Results

◆ Issue command

```r
> summary(mlogit(choice~1|cost.car+cost.carpool+cost.bus+cost.rail, data=market));
```

to estimate parameters. R outputs

◆ Coefficients:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>altcarpool</td>
<td>-5.26903</td>
</tr>
<tr>
<td>altbus</td>
<td>-2.07793</td>
</tr>
<tr>
<td>altrail</td>
<td>-3.65221</td>
</tr>
<tr>
<td>altcarpool:cost.car</td>
<td>0.82269</td>
</tr>
<tr>
<td>altbus:cost.car</td>
<td>0.95693</td>
</tr>
<tr>
<td>altrail:cost.car</td>
<td>0.98898</td>
</tr>
<tr>
<td>altcarpool:cost.carpool</td>
<td>-0.41234</td>
</tr>
<tr>
<td>altbus:cost.carpool</td>
<td>0.17596</td>
</tr>
<tr>
<td>altrail:cost.carpool</td>
<td>0.13842</td>
</tr>
<tr>
<td>altcarpool:cost.bus</td>
<td>0.54418</td>
</tr>
<tr>
<td>altbus:cost.bus</td>
<td>-1.74465</td>
</tr>
<tr>
<td>altrail:cost.bus</td>
<td>0.21560</td>
</tr>
<tr>
<td>altcarpool:cost.rail</td>
<td>-0.47507</td>
</tr>
<tr>
<td>altbus:cost.rail</td>
<td>-0.19281</td>
</tr>
<tr>
<td>altrail:cost.rail</td>
<td>-1.11535</td>
</tr>
</tbody>
</table>

◆ Estimations are with respect to Car probability.

◆ If cost.car increases, the probability of choosing another mode is higher.

◆ If cost.carpool increases, probability of carpool drops, those of bus and rail increase.

◆ If cost.bus increases, probability of bus drops, those of carpool and rail increase.

◆ If cost.rail increases, probability of carpool, bus both rail all drop with respect to car. That is the probability of car increases.