# Week 10: Mixed Logit Model

ResEcon 703: Topics in Advanced Econometrics

Matt Woerman
University of Massachusetts Amherst

# Agenda

Last week

- Generalized extreme value models

This week

- Mixed logit model overview
- Mixed logit choice probabilities
- Random coefficients
- Substitution patterns
- Panel data
- Empirical considerations
- Mixed logit R example

This week's reading

- Train textbook, chapter 6

# Mixed Logit Model Overview

# Discrete Choice Models Recap

Logit model

- Strong assumption that unobserved components of utility are i.i.d.
- Simple closed-form expressions for choice probabilities
- Preference variation can only be represented by observed data
- Panel data applications are limited by the i.i.d. assumption

Nested logit model (and other generalized extreme value models)

- Correlation between unobserved components of utility can be modeled
- Choice probabilities are more complex but still closed-form
- Preference variation can only be represented by observed data
- Panel data applications are limited by an i.i.d. assumption

What if we want a richer representation of preference variation, more flexible substitution patterns, and an ability to use panel data?

- Mixed logit model!

# Mixed Logit Model

The mixed logit model overcomes three limitations of the logit model

- Unobserved (or random) preference variation
- Unrestricted substitution patterns
- Correlations in unobserved factors over time

What is the catch?

- Mixed logit choice probabilities are not closed-form
- Estimation requires numerical simulation

# Mixed Logit Coefficients

How does the mixed logit model achieves this level of flexibility?

- It does not use a set of fixed coefficients for the entire population, $\boldsymbol{\beta}$
- It assumes there is a distribution of coefficients throughout the population, $f(\boldsymbol{\beta} \mid \boldsymbol{\theta})$

Distributions of coefficients overcome the three main limitations of the logit model

- Model distributions of unobserved preferences among the sample of decision makers
- Impose correlations in unobserved utility among alternatives
- Represent individual preferences over time

# Mixed Logit Choice Probabilities

# Mixed Logit Choice Probabilities

Unlike the logit and nested logit models, the mixed logit model is not defined by an assumption about the joint density of unobserved utility, $\varepsilon_n$

A mixed logit model is any discrete choice model with choice probabilities of the form

$$P_{ni} = \int L_{ni}(\beta) f(\beta \mid \boldsymbol{\theta}) d\beta$$

where $L_{ni}(\beta)$ is the logit probability at a given set of coefficients, $\beta$

$$L_{ni}(\beta) = \frac{e^{V_{ni}(\beta)}}{\sum_{j=1}^{J} e^{V_{nj}(\beta)}}$$

and $f(\beta \mid \boldsymbol{\theta})$ is a density function of coefficients $\beta$, which depends on a vector of parameters, $\boldsymbol{\theta}$

- The mixed logit choice probability is not a closed-form expression

# Mixed Logit Choice Probabilities Intuition

$$P_{ni} = \int L_{ni}(\beta) f(\beta \mid \boldsymbol{\theta}) d\beta$$

$$L_{ni}(\beta) = \frac{e^{V_{ni}(\beta)}}{\sum_{j=1}^{J} e^{V_{nj}(\beta)}}$$

The mixed logit choice probability is a weighted average of logit choice probabilities

- The logit choice probabilities are evaluated at different values of $\beta$
- Each logit choice probability is weighted by the density $f(\beta \mid \boldsymbol{\theta})$

The mixed logit choice probability is a mixed function of logit choice probabilities, $L_{ni}(\beta)$, with the mixing distribution $f(\beta \mid \boldsymbol{\theta})$

# Distributions of Coefficients

We previously used the terms "parameter" and "coefficient" interchangeably, but we will use them to mean different things in the mixed logit model

- $\beta$: Coefficients that appear in the utility expression
  - We will not estimate these coefficients in the mixed logit model
  - These coefficients are integrated out of the choice probability
- $\boldsymbol{\theta}$: Parameters that define the density of random coefficients
  - We will estimate these parameters in the mixed logit model

Some examples of distributions of coefficients that we can model

- Normal: $\beta \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, we estimate $\boldsymbol{\theta} = \{\boldsymbol{\mu}, \boldsymbol{\Sigma}\}$
- Log-normal: $\ln \beta \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, we estimate $\boldsymbol{\theta} = \{\boldsymbol{\mu}, \boldsymbol{\Sigma}\}$
- Uniform: $\beta \sim \mathcal{U}(\boldsymbol{a}, \boldsymbol{b})$, we estimate $\boldsymbol{\theta} = \{\boldsymbol{a}, \boldsymbol{b}\}$
- Triangular: $\beta \sim Tri(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$, we estimate $\boldsymbol{\theta} = \{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$
- Many other distributions to choose from

# Random Coefficients

# Random Utility Model with Individual-Specific Coefficients

A common way to rationalize the mixed logit model is the random utility model with individual-specific coefficients

The utility that decision maker $n$ obtains from alternative $j$ is

$$U_{nj} = \boldsymbol{\beta}_n' \boldsymbol{x}_{nj} + \varepsilon_{nj}$$

- $\boldsymbol{x}_{nj}$: data for alternative $j$ and decision maker $n$
- $\boldsymbol{\beta}_n$: individual-specific vector of coefficients
- $\varepsilon_{nj}$: i.i.d. extreme value error term

Each decision maker knows their own $\boldsymbol{\beta}_n$ and $\varepsilon_{nj}$, so the choice is deterministic from their perspective

- $n$ chooses $i$ if and only if $U_{ni} > U_{nj} \; \forall j \neq i$

# Random Utility Model with Random Coefficients

The utility that decision maker $n$ obtains from alternative $j$ is

$$U_{nj} = \boldsymbol{\beta}_n' \mathbf{x}_{nj} + \varepsilon_{nj}$$

But we (the researchers) do not observe $\boldsymbol{\beta}_n$ for any individual

- We model $\boldsymbol{\beta}_n$ as a random variable with density $f(\beta \mid \boldsymbol{\theta})$

If we did know $\boldsymbol{\beta}_n$, then the model would be a standard logit with the *conditional* choice probability

$$L_{ni}(\boldsymbol{\beta}_n) = \frac{e^{\boldsymbol{\beta}_n' \mathbf{x}_{ni}}}{\sum_{j=1}^{J} e^{\boldsymbol{\beta}_n' \mathbf{x}_{nj}}}$$

But we do not know $\boldsymbol{\beta}_n$, so we have to integrate over the density of the random coefficients to obtain the *unconditional* choice probability

$$P_{ni} = \int \frac{e^{\beta' \mathbf{x}_{ni}}}{\sum_{j=1}^{J} e^{\beta' \mathbf{x}_{nj}}} f(\beta \mid \boldsymbol{\theta}) d\beta$$

# Substitution Patterns

# Random Utility Model with Flexible Random Utility

Another common way to rationalize the mixed logit model is the random utility model with fully flexible correlations in random utilities

The utility that decision maker $n$ obtains from alternative $j$ is

$$U_{nj} = \boldsymbol{\alpha}'\boldsymbol{x}_{nj} + \boldsymbol{\mu}_n'\boldsymbol{z}_{nj} + \varepsilon_{nj}$$

- $\boldsymbol{x}_{nj}$, $\boldsymbol{z}_{nj}$: data for alternative $j$ and decision maker $n$
- $\boldsymbol{\alpha}$: vector of fixed coefficients
- $\boldsymbol{\mu}_n$: vector of random coefficients with mean zero
- $\varepsilon_{nj}$: i.i.d. extreme value error term

We can combine the random components of utility into a single composite random utility term, $\eta_{nj}$

$$\eta_{nj} = \boldsymbol{\mu}_n'\boldsymbol{z}_{nj} + \varepsilon_{nj}$$

# Random Utility Model with Correlated Random Utility

The utility that decision maker $n$ obtains from alternative $j$ is

$$U_{nj} = \boldsymbol{\alpha}'\boldsymbol{x}_{nj} + \eta_{nj}$$

where the random utility term, $\eta_{nj}$, is given by

$$\eta_{nj} = \boldsymbol{\mu}_n'\boldsymbol{z}_{nj} + \varepsilon_{nj}$$

These flexible random utility terms can be correlated over alternatives

$$Cov(\eta_{ni}, \eta_{nj}) = \boldsymbol{z}_{ni}'\boldsymbol{\Sigma}\boldsymbol{z}_{nj}$$

where $\boldsymbol{\Sigma}$ is the variance-covariance matrix of $\boldsymbol{\mu}$

The flexible specification of random utility yields flexible correlations between alternatives, which we can estimate

# Special Cases of Correlated Random Utility

This representation of random utility correlations generalizes all of the previous discrete choice models we have discussed

$$Cov(\eta_{ni}, \eta_{nj}) = \boldsymbol{z}_{ni}' \Sigma \boldsymbol{z}_{nj}$$

Logit model

- When $z_{nj} = 0 \; \forall j$, the random utility of alternatives is uncorrelated

Nested logit model

- $\boldsymbol{z}_{nj}$ is a set of indicator variables for each nest
  - $K$ variables that equal 1 for the alternatives in that nest
- $\mu_k \sim \mathcal{N}(0, \sigma_k)$ and independent across nests
- Covariance between alternatives in nest $k$ is $\sigma_k$, which we estimate

# Equivalence of Random Utility Models

We have motivated the mixed logit model from two random utility models

- These random utility models are mathematically equivalent

Starting from the random coefficients expression

$$U_{nj} = \boldsymbol{\beta}_n' \boldsymbol{x}_{nj} + \varepsilon_{nj}$$

Express $\boldsymbol{\beta}_n$ as $\boldsymbol{\beta}_n = \boldsymbol{\alpha} + \boldsymbol{\mu}_n$ by decomposing $\boldsymbol{\beta}_n$ into the means, $\boldsymbol{\alpha}$, and the deviations from the means, $\boldsymbol{\mu}_n$

$$U_{nj} = \boldsymbol{\alpha}' \boldsymbol{x}_{nj} + \boldsymbol{\mu}_n' \boldsymbol{z}_{nj} + \varepsilon_{nj}$$

where $\boldsymbol{z}_{nj}$ is the subset of data, $\boldsymbol{x}_{nj}$, with random coefficients

- If all variables have random coefficients, then $\boldsymbol{z}_{nj} = \boldsymbol{x}_{nj}$

Your motivation will affect which coefficients you model as random and whether to allow correlations between the coefficients

# Mixed Logit Substitution Patterns

If representative utility is linear, $V_{nj} = \beta_n' \boldsymbol{x}_{nj}$, then the ratio of the choice probabilities of alternative $i$ and alternative $m$ is

$$\frac{P_{ni}}{P_{nm}} = \frac{\displaystyle\int \frac{e^{\beta' \boldsymbol{x}_{ni}}}{\sum_{j=1}^{J} e^{\beta' \boldsymbol{x}_{nj}}} f(\beta \mid \boldsymbol{\theta}) d\beta}{\displaystyle\int \frac{e^{\beta' \boldsymbol{x}_{nm}}}{\sum_{j=1}^{J} e^{\beta' \boldsymbol{x}_{nj}}} f(\beta \mid \boldsymbol{\theta}) d\beta}$$

- Choice probability denominators are inside the integrals, so they cannot cancel as they did in previous models

The ratio of any two choice probabilities depends on all alternatives

- The mixed logit model does not exhibit independence of irrelevant alternatives (IIA)

# Mixed Logit Elasticities

The own elasticity of alternative $i$ with respect to its attribute $z_{ni}$ (where $z_{ni}$ is any element of $\boldsymbol{x}_{ni}$) is

$$E_{iz_{ni}} = \frac{z_{ni}}{P_{ni}} \int \beta_z L_{ni}(\beta)[1 - L_{ni}(\beta)] f(\beta \mid \boldsymbol{\theta}) d\beta$$

The cross elasticity of alternative $i$ with respect to attribute $z_{nj}$ of alternative $j$ (where $z_{nj}$ is any element of $\boldsymbol{x}_{nj}$) is

$$E_{iz_{nj}} = -\frac{z_{nj}}{P_{ni}} \int \beta_z L_{ni}(\beta) L_{nj}(\beta) f(\beta \mid \boldsymbol{\theta}) d\beta$$

As with the mixed logit choice probabilities, mixed logit elasticities do not have closed-form expressions

These expressions depend on how $L_{ni}(\beta)$ varies and how $L_{ni}(\beta)$ and $L_{nj}(\beta)$ covary as you integrate over the density of $\beta$, which is determined by which parameters you specify as random and whether they covary

# Panel Data

# Mixed Logit and Panel Data

The structure of the mixed logit model allows for more flexibility in representing how a single decision maker makes multiple choices over time periods, so it provides a better model for most panel data settings

- In the logit and nested logit models, all unobserved factors that influence utility are represented by $\varepsilon_{nt}$, which is i.i.d., so those models cannot accommodate correlations over time
- But the mixed logit model allows for unobserved preference variation through random coefficients, which yields correlations in utility over time for the same decision maker

# Mixed Logit Model with Panel Data

We add a time index, $t$, to our random utility model but assume the random coefficients, $\beta_n$, are constant for an individual

$$U_{njt} = \beta_n' \boldsymbol{x}_{njt} + \varepsilon_{njt}$$

and we consider the vector of alternatives that decision maker $n$ chooses over the $T$ time periods

$$\boldsymbol{i} = (i_1, \ldots, i_T)$$

Then the conditional logit probability for a sequence of choices is

$$L_{n\boldsymbol{i}}(\beta) = \prod_{t=1}^{T} \frac{e^{\beta' \boldsymbol{x}_{ni_t t}}}{\sum_{j=1}^{J} e^{\beta' \boldsymbol{x}_{njt}}}$$

and integrating over the density of coefficients gives the choice probability

$$P_{n\boldsymbol{i}} = \int L_{n\boldsymbol{i}}(\beta) f(\beta \mid \boldsymbol{\theta}) d\beta$$

# Dynamics in a Mixed Logit Model

Some "dynamics" can be represented in a mixed logit model using panel data

- Past and future exogenous variables can be included to model lagged or anticipatory behavior
- Lagged dependent variables can be included to represent state dependence

This approach is a relatively naive way of incorporating dynamics into a discrete choice model

- We are essentially modeling a sequence of static choices
- A fully "dynamic discrete choice model" would model how every choice affects all subsequent choices
- We will talk about dynamic discrete choice models in the final week of the semester

# Empirical Considerations

# Mixed Logit Estimation

Mixed logit choice probabilities do not have a closed-form expression

$$P_{ni} = \int \frac{e^{\beta' \mathbf{x}_{ni}}}{\sum_{j=1}^{J} e^{\beta' \mathbf{x}_{nj}}} f(\beta \mid \boldsymbol{\theta}) d\beta$$

- We cannot estimate a mixed logit model using maximum likelihood because we cannot calculate the log-likelihood function

Instead, we can approximate choice probabilities through numerical simulation, calculate the simulated log-likelihood function, and estimate using maximum simulated likelihood

- More on other simulation-based estimation methods next week

We can use the `mlogit()` function in R to estimate mixed logit models

- We specify our random coefficients using the `rpar` argument in the `mlogit()` function

# Mixed Logit with Market-Level Data

The mixed logit model can also be estimated from market-level data

- You observe the price, market share, and characteristics of every cereal brand at the grocery store, and you want to estimate the structural parameters of consumer decision making that explain those purchases

When aggregated over many consumers, choice probabilities become market shares

$$S_i = \int \frac{e^{\beta' \mathbf{x}_i}}{\sum_{j=1}^{J} e^{\beta' \mathbf{x}_j}} f(\beta \mid \boldsymbol{\theta}) d\beta$$

- Because of the integral, mixed logit market shares do not reduce to a linear model as they did for logit and nested logit
- Demand estimation using random coefficients logit often uses the "BLP" method of Berry, Levinsohn, and Pakes (1995)
- We will talk about BLP in the final week of the semester

# Mixed Logit R Example

# Mixed Logit Model Example

We are again studying how consumers make choices about expensive and highly energy-consuming systems in their homes

- We have (real) data on 250 households in California and the type of HVAC (heating, ventilation, and air conditioning) system in their home. Each household has the following choice set, and we observe the following data

Choice set

- `ec`: electric central
- `ecc`: electric central with AC
- `er`: electric room
- `erc`: electric room with AC
- `gc`: gas central
- `gcc`: gas central with AC
- `hpc`: heat pump with AC

Alternative-specific data

- `ich`: installation cost for heat
- `icca`: installation cost for AC
- `och`: operating cost for heat
- `occa`: operating cost for AC

Household demographic data

- `income`: annual income

# Random Utility Model of HVAC System Choice

We model the utility to household $n$ of installing HVAC system $j$ as

$$U_{nj} = V_{nj} + \varepsilon_{nj}$$

where $V_{nj}$ depends on the data about alternative $j$ and household $n$

The probability that household $n$ installs HVAC system $i$ is

$$P_{ni} = \int_{\varepsilon} \mathbb{1}(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj} \; \forall j \neq i) f(\boldsymbol{\varepsilon}_n) d\boldsymbol{\varepsilon}_n$$

Under the logit assumption, these choice probabilities simplify to

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}}$$

# Representative Utility of HVAC System Choice

What might affect the utility of the different HVAC systems?

- Installation cost
- Annual operating cost
- HVAC system technology
  - Systems with cooling might be preferred to heating-only systems
  - Gas systems might be preferred to electric systems
  - Central systems might be preferred to room systems
- Anything else?

We model the representative utility of HVAC system $j$ to household $n$ as

$$V_{nj} = \alpha_j + \beta_1 IC_{nj} + \beta_2 OC_{nj}$$

# Load Dataset

```r
## Load tidyverse and mlogit
library(tidyverse)
library(mlogit)
## Load dataset from mlogit package
data('HC', package = 'mlogit')
```

# Dataset

```
## Look at dataset
tibble(HC)
## # A tibble: 250 x 18
##    depvar ich.gcc ich.ecc ich.erc ich.hpc ich.gc ich.ec ich.er  icca
##    <fct>    <dbl>   <dbl>   <dbl>   <dbl>  <dbl>  <dbl>  <dbl> <dbl>
##  1 erc       9.7     7.86    8.79   11.4    24.1   24.5   7.37  27.3
##  2 hpc       8.77    8.69    7.09    9.37   28     32.7   9.33  26.5
##  3 gcc       7.43    8.86    6.94   11.7    25.7   31.7   8.14  22.6
##  4 gcc       9.18    8.93    7.22   12.1    29.7   26.7   8.04  25.3
##  5 gcc       8.05    7.02    8.44   10.5    23.9   28.4   7.15  25.4
##  6 gcc       9.32    8.03    6.22   12.6    27.0   21.4   8.6   19.9
##  7 gc        7.11    8.78    7.36   12.4    22.9   28.6   6.41  27.0
##  8 hpc       9.38    7.48    6.72    8.93   26.2   27.9   7.3   18.1
##  9 gcc       8.08    7.39    8.79   11.2    23.0   22.6   7.85  22.6
## 10 gcc       6.24    4.88    7.46    8.28   19.8   27.5   6.88  25.8
## # ... with 240 more rows, and 9 more variables: och.gcc <dbl>,
## #   och.ecc <dbl>, och.erc <dbl>, och.hpc <dbl>, och.gc <dbl>,
## #   och.ec <dbl>, och.er <dbl>, occa <dbl>, income <dbl>
```

# Clean Dataset

```r
## Combine heating and cooling costs into one variable
hvac_clean <- HC %>%
  mutate(id = 1:n(),
         ic.gcc = ich.gcc + icca, ic.ecc = ich.ecc + icca,
         ic.erc = ich.erc + icca, ic.hpc = ich.hpc + icca,
         ic.gc = ich.gc, ic.ec = ich.ec, ic.er = ich.er,
         oc.gcc = och.gcc + occa, oc.ecc = och.ecc + occa,
         oc.erc = och.erc + occa, oc.hpc = och.hpc + occa,
         oc.gc = och.gc, oc.ec = och.ec, oc.er = och.er) %>%
  select(id, depvar, starts_with('ic.'), starts_with('oc.'), income)
```

# Cleaned Dataset

```
## Look at cleaned dataset
tibble(hvac_clean)
## # A tibble: 250 x 17
##       id depvar ic.gcc ic.ecc ic.erc ic.hpc ic.gc ic.ec ic.er oc.gcc
##    <int> <fct>   <dbl>  <dbl>  <dbl>  <dbl> <dbl> <dbl> <dbl>  <dbl>
##  1     1 erc      37.0   35.1   36.1   38.6  24.1  24.5  7.37   5.21
##  2     2 hpc      35.3   35.2   33.6   35.9  28    32.7  9.33   3.93
##  3     3 gcc      30.1   31.5   29.6   34.3  25.7  31.7  8.14   4.46
##  4     4 gcc      34.5   34.3   32.6   37.5  29.7  26.7  8.04   5.32
##  5     5 gcc      33.5   32.5   33.9   36.0  23.9  28.4  7.15   5.29
##  6     6 gcc      29.2   28.0   26.2   32.5  27.0  21.4  8.6    4.67
##  7     7 gc       34.2   35.8   34.4   39.4  22.9  28.6  6.41   4.18
##  8     8 hpc      27.5   25.6   24.8   27.0  26.2  27.9  7.3    5.37
##  9     9 gcc      30.6   30.0   31.4   33.7  23.0  22.6  7.85   4.74
## 10    10 gcc      32.0   30.6   33.2   34.0  19.8  27.5  6.88   4.32
## # ... with 240 more rows, and 7 more variables: oc.ecc <dbl>,
## #   oc.erc <dbl>, oc.hpc <dbl>, oc.gc <dbl>, oc.ec <dbl>,
## #   oc.er <dbl>, income <dbl>
```

# Convert Dataset to dfidx Format

```r
## Convert cleaned dataset to dfidx format
hvac_dfidx <- dfidx(hvac_clean, shape = 'wide',
                    choice = 'depvar', varying = 3:16)
```

# Dataset in dfidx Format

```
## Look at data in dfidx format
tibble(hvac_dfidx)
## # A tibble: 1,750 x 6
##        id depvar income    ic    oc idx$id1 $id2
##     <int> <lgl>   <dbl> <dbl> <dbl>   <int> <fct>
##  1      1 FALSE      20 24.5   4.09       1 ec
##  2      1 FALSE      20 35.1   7.04       1 ecc
##  3      1 FALSE      20  7.37  3.85       1 er
##  4      1 TRUE       20 36.1   6.8        1 erc
##  5      1 FALSE      20 24.1   2.26       1 gc
##  6      1 FALSE      20 37.0   5.21       1 gcc
##  7      1 FALSE      20 38.6   4.68       1 hpc
##  8      2 FALSE      50 32.7   2.69       2 ec
##  9      2 FALSE      50 35.2   4.32       2 ecc
## 10      2 FALSE      50  9.33  3.45       2 er
## # ... with 1,740 more rows
```

# Multinomial Logit Model

We model the representative utility of HVAC system $j$ to household $n$ as

$$V_{nj} = \alpha_j + \beta_1 IC_{nj} + \beta_2 OC_{nj}$$

```
## Model choice using alternative intercepts and cost data
model_1 <- mlogit(formula = depvar ~ ic + oc | 1 | 0,
                  data = hvac_dfidx,
                  reflevel = 'hpc')
```

# Multinomial Logit Model Summary

```
## Summarize model results
summary(model_1)
##
## Call:
## mlogit(formula = depvar ~ ic + oc | 1 | 0, data = hvac_dfidx,
##     reflevel = "hpc", method = "nr")
##
## Frequencies of alternatives:choice
##   hpc    ec   ecc    er   erc    gc   gcc
## 0.104 0.004 0.016 0.032 0.004 0.096 0.744
##
## nr method
## 7 iterations, 0h:0m:0s
## g'(-H)^-1g = 1.94E-05
## successive function values within tolerance limits
##
## Coefficients :
##                  Estimate Std. Error z-value  Pr(>|z|)
## (Intercept):ec  -6.305515   1.201159 -5.2495 1.525e-07 ***
## (Intercept):ecc  2.142493   0.944478  2.2684    0.0233 *
## (Intercept):er  -7.779634   1.369346 -5.6813 1.337e-08 ***
## (Intercept):erc  0.944165   1.337210  0.7061    0.4801
## (Intercept):gc  -6.120762   0.964956 -6.3430 2.253e-10 ***
## (Intercept):gcc  2.997228   0.395830  7.5720 3.664e-14 ***
## ic              -0.225519   0.043764 -5.1531 2.562e-07 ***
## oc              -1.937800   0.373606 -5.1867 2.140e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log-Likelihood: -198.57
## McFadden R^2:  0.11831
## Likelihood ratio test : chisq = 53.291 (p.value = 2.6789e-12)
```

# Mixed Logit Model of HVAC System Choice

We modeled the utility to household $n$ of installing HVAC system $j$ as

$$U_{nj} = \alpha_j + \beta_1 IC_{nj} + \beta_2 OC_{nj} + \varepsilon_{nj}$$

- But what if the marginal utilities of current and future costs (or income) vary throughout the population?

Using the mixed logit model, we can allow for unobserved variation in these coefficients

$$U_{nj} = \alpha_j + \beta_{1n} IC_{nj} + \beta_{2n} OC_{nj} + \varepsilon_{nj}$$

where the random coefficients are normally distributed

$$\beta_{1n} \sim \mathcal{N}(\mu_1, \sigma_1^2)$$
$$\beta_{2n} \sim \mathcal{N}(\mu_2, \sigma_2^2)$$

# Mixed Logit Choice Probabilities for HVAC System Choice

The mixed logit choice probabilities are integrals over the densities of our random coefficients

$$P_{ni} = \int \frac{e^{\alpha_i + \beta_1 IC_{ni} + \beta_2 OC_{ni}}}{\sum_{j=1}^{J} e^{\alpha_j + \beta_1 IC_{nj} + \beta_2 OC_{nj}}} f(\boldsymbol{\beta} \mid \boldsymbol{\theta}) d\boldsymbol{\beta}$$

We will use maximum simulated likelihood to find the set of parameters—$\alpha_j$, $\mu_1$, $\sigma_1^2$, $\mu_2$, and $\sigma_2^2$—that maximize the simulated likelihood of generating the choices that we observe

We can use the `mlogit` function in R to estimate mixed logit models

- We specify our random coefficients using the `rpar` argument in the `mlogit()` function

# Mixed Logit Models in R

```
## Help file for the mlogit function
?mlogit
## Arguments for mlogit mixed logit functionality
mlogit(formula, data, reflevel, rpar, correlation, R, seed, ...)
```

`mlogit()` arguments for mixed logit

1. `formula, data, reflevel`: same as a multinomial logit model
2. `rpar`: named vector of random coefficients and their distributions
3. `correlation`: `TRUE` models random coefficient correlations
4. `R`: number of random draws in numerical simulation
5. `seed`: seed for random draws in simulation

# Mixed Logit Model with Random Coefficients

We model the utility of HVAC system $j$ to household $n$ as

$$U_{nj} = \alpha_j + \beta_{1n} IC_{nj} + \beta_{2n} OC_{nj} + \varepsilon_{nj}$$

where the random coefficients are normally distributed

$$\beta_{1n} \sim \mathcal{N}(\mu_1, \sigma_1^2)$$
$$\beta_{2n} \sim \mathcal{N}(\mu_2, \sigma_2^2)$$

```r
## Model choice using alt intercepts and cost data with normal coefs
model_2 <- mlogit(formula = depvar ~ ic + oc | 1 | 0,
                  data = hvac_dfidx,
                  reflevel = 'hpc',
                  rpar = c(ic = 'n', oc = 'n'),
                  R = 1000, seed = 703)
```

# Mixed Logit Model Summary

```
## Summarize model results
summary(model_2)
##
## Call:
## mlogit(formula = depvar ~ ic + oc | 1 | 0, data = hvac_dfidx,
##     reflevel = "hpc", rpar = c(ic = "n", oc = "n"), R = 1000,
##     seed = 703)
##
## Frequencies of alternatives:choice
##   hpc    ec   ecc    er   erc    gc   gcc
## 0.104 0.004 0.016 0.032 0.004 0.096 0.744
##
## bfgs method
## 23 iterations, 0h:0m:18s
## g'(-H)^-1g = 8.04E-08
## gradient close to zero
##
## Coefficients :
##                   Estimate Std. Error z-value Pr(>|z|)
## (Intercept):ec  -14.42155    7.00416 -2.0590 0.039495 *
## (Intercept):ecc   4.75138    2.02256  2.3492 0.018814 *
## (Intercept):er  -28.53708   15.70081 -1.8176 0.069132 .
## (Intercept):erc   3.51896    2.45641  1.4326 0.151983
## (Intercept):gc  -16.46511    7.52932 -2.1868 0.028757 *
## (Intercept):gcc   4.74834    1.51929  3.1254 0.001776 **
## ic               -0.57118    0.25936 -2.2023 0.027647 *
## oc               -3.97472    1.50884 -2.6343 0.008431 **
## sd.ic             0.36460    0.25271  1.4427 0.149094
## sd.oc            -1.30715    0.83330 -1.5686 0.116731
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log-Likelihood: -192.14
## McFadden R^2:  0.14683
```

# Interpreting Parameters

```
## Display model coefficients
coef(model_2)
##  (Intercept):ec (Intercept):ecc  (Intercept):er (Intercept):erc
##     -14.4215520       4.7513809     -28.5370829       3.5189647
##  (Intercept):gc (Intercept):gcc              ic              oc
##     -16.4651095       4.7483421      -0.5711827      -3.9747232
##           sd.ic           sd.oc
##       0.3646017      -1.3071521
```

How do we interpret these coefficients?

- Alternative-specific intercepts
  - `ecc` and `gcc` provide more utility, *ceteris paribus*, than `hpc`
  - `erc` provides the same utility, *ceteris paribus*, as `hpc`
  - `ec`, `er`, and `gc` provide less utility, *ceteris paribus*, than `hpc`
- An additional \$100 of installation cost reduces utility by 0.57 on average with a standard deviation of 0.36
- An additional \$100 of annual operating cost reduces utility by 3.97 on average with a standard deviation of 1.31

# Test of Random Coefficient Variances

We can test if the random coefficient variances are statistically different from those implied by the multinomial logit model

- The multinomial logit model assumes all coefficients are fixed, or $\sigma_k^2 = 0 \; \forall k$

We will test the null hypothesis

$$H_0\colon \; \sigma_1^2 = \sigma_2^2 = 0$$

We are using maximum simulated likelihood to estimate this model, so we can use the same tests that apply to maximum likelihood

- We will use a (simulated) likelihood ratio test

# Likelihood Ratio Test

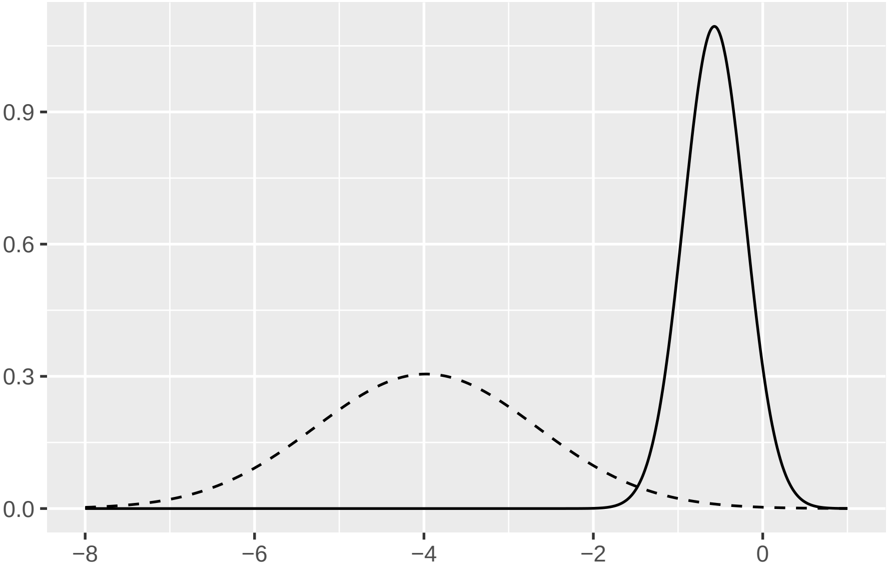
`lrtest()` conducts a (simulated) likelihood ratio test on the two models we specify as arguments

```
## Conduct likelihood ratio test of models 1 and 2
lrtest(model_1, model_2)
## Likelihood ratio test
##
## Model 1: depvar ~ ic + oc | 1 | 0
## Model 2: depvar ~ ic + oc | 1 | 0
##   #Df  LogLik Df  Chisq Pr(>Chisq)
## 1   8 -198.57
## 2  10 -192.14  2 12.847   0.001623 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

# Plot of Random Coefficients

```r
## Plot distributions of random coefficients
ggplot(data = data.frame(x = c(-8, 1)), aes(x)) +
  stat_function(fun = dnorm, n = 1001,
                args = list(mean = coef(model_2)[7],
                            sd = abs(coef(model_2)[9]))) +
  stat_function(fun = dnorm, n = 1001,
                args = list(mean =coef(model_2)[8],
                            sd = abs(coef(model_2)[10])),
                linetype = 'dashed') +
  xlab(NULL) +
  ylab(NULL)
```

# Plot of Random Coefficients



# Cost Trade-Offs

How do consumers trade off the installation cost and the annual operating cost when choosing an HVAC system?

- What reduction in installation cost offsets a \$1 increase in the annual operating cost?

$$
\begin{aligned}
U_{ni} &= \alpha_i + \beta_{1n} IC_{ni} + \beta_{2n} OC_{ni} + \varepsilon_{ni} \\
dU_{ni} &= \beta_{1n} dIC_{ni} + \beta_{2n} dOC_{ni} \\
dU_{ni} &= 0 \quad \Rightarrow \quad \frac{dIC_{ni}}{dOC_{ni}} = -\frac{\beta_{2n}}{\beta_{1n}}
\end{aligned}
$$

We do not know each decision maker's coefficients, but we do know the distribution of these individual-specific coefficients

- We could find the distribution that corresponds to this ratio of distributions, but the ratio of normal distributions is challenging

To make this question easier, we can estimate a model with a fixed $\beta_1$

# Mixed Logit Model with Random OC Coefficient

We model the utility of HVAC system $j$ to household $n$ as

$$U_{nj} = \alpha_j + \beta_1 IC_{nj} + \beta_{2n} OC_{nj} + \varepsilon_{nj}$$

where the random coefficient is normally distributed

$$\beta_{2n} \sim \mathcal{N}(\mu_2, \sigma_2^2)$$

```
## Model choice using alt intercepts and cost data with normal oc coef
model_3 <- mlogit(formula = depvar ~ ic + oc | 1 | 0,
                  data = hvac_dfidx,
                  reflevel = 'hpc',
                  rpar = c(oc = 'n'),
                  R = 1000, seed = 703)
```

# Mixed Logit Model Summary

```
## Summarize model results
summary(model_3)
##
## Call:
## mlogit(formula = depvar ~ ic + oc | 1 | 0, data = hvac_dfidx,
##     reflevel = "hpc", rpar = c(oc = "n"), R = 1000, seed = 703)
##
## Frequencies of alternatives:choice
##   hpc    ec   ecc    er   erc    gc   gcc
## 0.104 0.004 0.016 0.032 0.004 0.096 0.744
##
## bfgs method
## 9 iterations, 0h:0m:7s
## g'(-H)^-1g = 2.15E-07
## gradient close to zero
##
## Coefficients :
##                  Estimate Std. Error z-value  Pr(>|z|)
## (Intercept):ec  -6.988571   1.777744 -3.9311 8.454e-05 ***
## (Intercept):ecc  1.989181   0.963761  2.0640 0.0390199 *
## (Intercept):er  -9.033176   2.282496 -3.9576 7.571e-05 ***
## (Intercept):erc  0.719607   1.620491  0.4441 0.6569940
## (Intercept):gc  -7.304331   2.161066 -3.3800 0.0007249 ***
## (Intercept):gcc  3.036280   0.477932  6.3530 2.112e-10 ***
## ic              -0.267513   0.081086 -3.2991 0.0009699 ***
## oc              -2.061273   0.486978 -4.2328 2.308e-05 ***
## sd.oc            0.535617   0.526907  1.0165 0.3093769
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Log-Likelihood: -198.12
## McFadden R^2:  0.12029
## Likelihood ratio test : chisq = 54.18 (p.value = 1.0269e-11)
##
```

# Interpreting Parameters

```
## Display model coefficients
coef(model_3)
##  (Intercept):ec (Intercept):ecc  (Intercept):er (Intercept):erc
##      -6.9885711       1.9891815      -9.0331761       0.7196068
##  (Intercept):gc (Intercept):gcc              ic              oc
##      -7.3043311       3.0362800      -0.2675133      -2.0612730
##           sd.oc
##       0.5356169
```

How do we interpret these coefficients?

- Alternative-specific intercepts
  - `ecc` and `gcc` provide more utility, *ceteris paribus*, than `hpc`
  - `erc` provides the same utility, *ceteris paribus*, as `hpc`
  - `ec`, `er`, and `gc` provide less utility, *ceteris paribus*, than `hpc`
- An additional \$100 of installation cost reduces utility by 0.27
- An additional \$100 of annual operating cost reduces utility by 2.06 on average with a standard deviation of 0.54

# Test of Random Coefficient Variance

We just estimated a model that restricted $\beta_1$ to be fixed, which is equivalent to saying $\beta_1$ is random with zero variance, or $\sigma_1^2 = 0$

- We can test if this imposed restriction is correct

We will test the null hypothesis

$$H_0\colon\ \sigma_1^2 = 0$$

```
## Conduct likelihood ratio test of models 2 and 3
lrtest(model_2, model_3)
## Likelihood ratio test
##
## Model 1: depvar ~ ic + oc | 1 | 0
## Model 2: depvar ~ ic + oc | 1 | 0
##   #Df  LogLik Df  Chisq Pr(>Chisq)
## 1  10 -192.14
## 2   9 -198.12 -1 11.958  0.0005443 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

# Cost Trade-Offs

How do consumers trade off the installation cost and the annual operating cost when choosing an HVAC system?

- What reduction in installation cost offsets a $1 increase in the annual operating cost?

$$\frac{dIC_{ni}}{dOC_{ni}} = -\frac{\beta_{2n}}{\beta_1}$$

Because $\beta_{2n}$ is normally distributed and $\beta_1$ is fixed, the ratio is also a normal distribution

$$\frac{dIC_{ni}}{dOC_{ni}} \sim \mathcal{N}\left(\frac{\mu_2}{\beta_1}, \frac{\sigma_2^2}{\beta_1^2}\right)$$

- What are the mean and standard deviation of this distribution?

```
## Calculate ic equivalence of an increase in oc
c(-coef(model_3)[8] / coef(model_3)[7],
  abs(coef(model_3)[9] / coef(model_3)[7]))
##         oc     sd.oc
## -7.705311  2.002207
```

# Plot of Cost Trade-offs

```
## Plot distributions of random coefficients
ggplot(data = data.frame(x = c(-18, 2)), aes(x)) +
  stat_function(fun = dnorm, n = 1001,
              args = list(mean = -coef(model_3)[8] / coef(model_3)[7],
                          sd = abs(coef(model_3)[9] /
                                   coef(model_3)[7]))) +
  xlab(NULL) +
  ylab(NULL)
```

# Plot of Cost Trade-offs