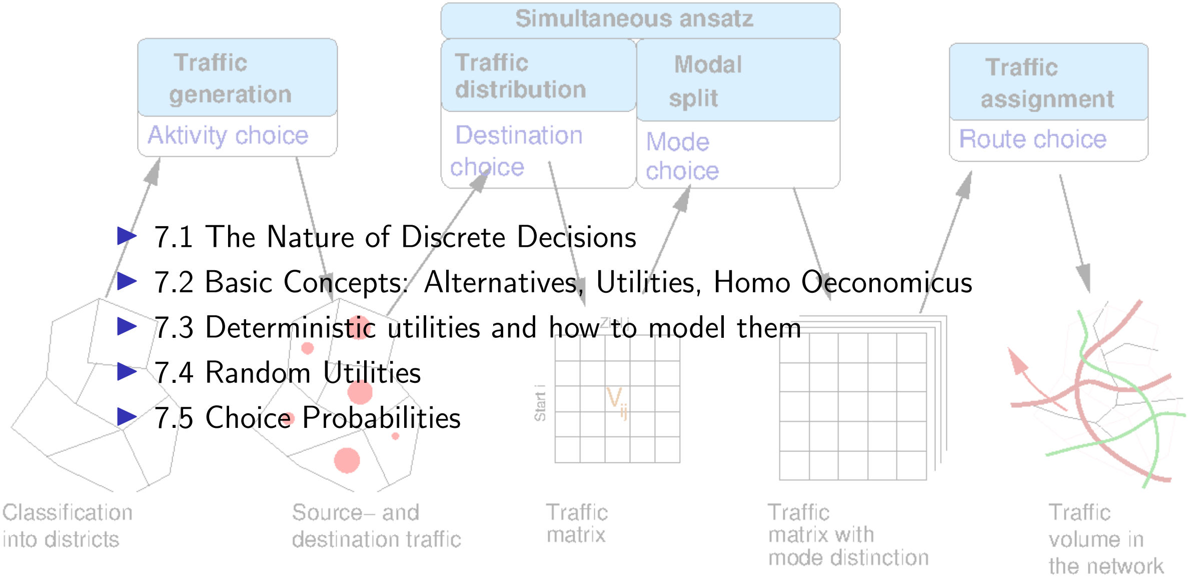
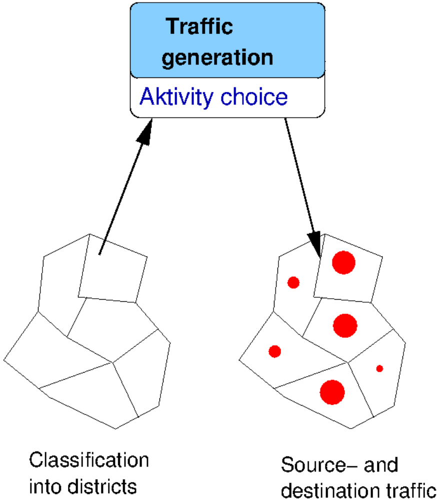
7. Discrete-Choice Theory: the Basics



- 7.1 The Nature of Discrete Decisions
- 7.2 Basic Concepts: Alternatives, Utilities, Homo Oeconomicus
- 7.3 Deterministic utilities and how to model them
- 7.4 Random Utilities
- 7.5 Choice Probabilities

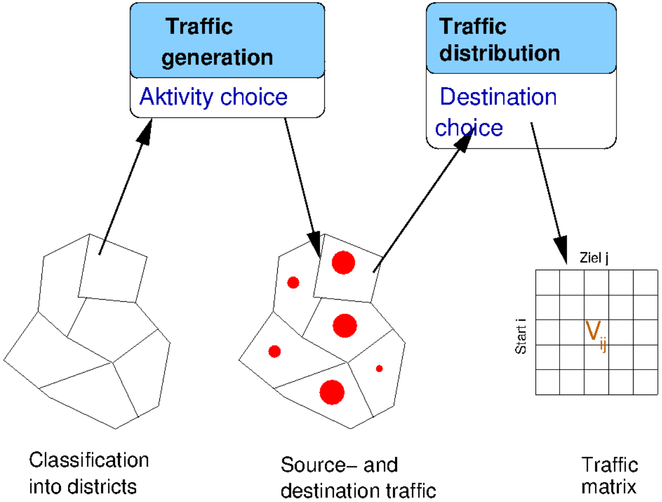
7.1 The Nature of Discrete Decisions

Example: the four-step model of transportation planning



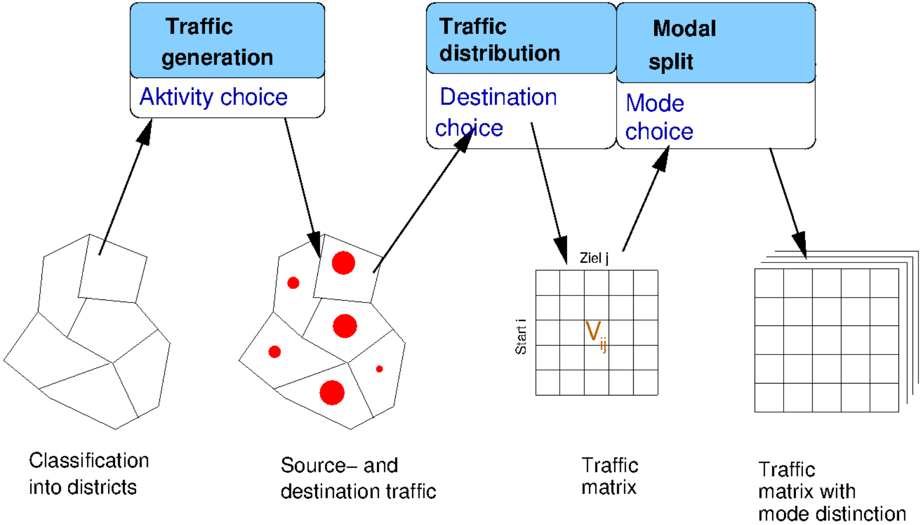
7.1 The Nature of Discrete Decisions

Example: the four-step model of transportation planning



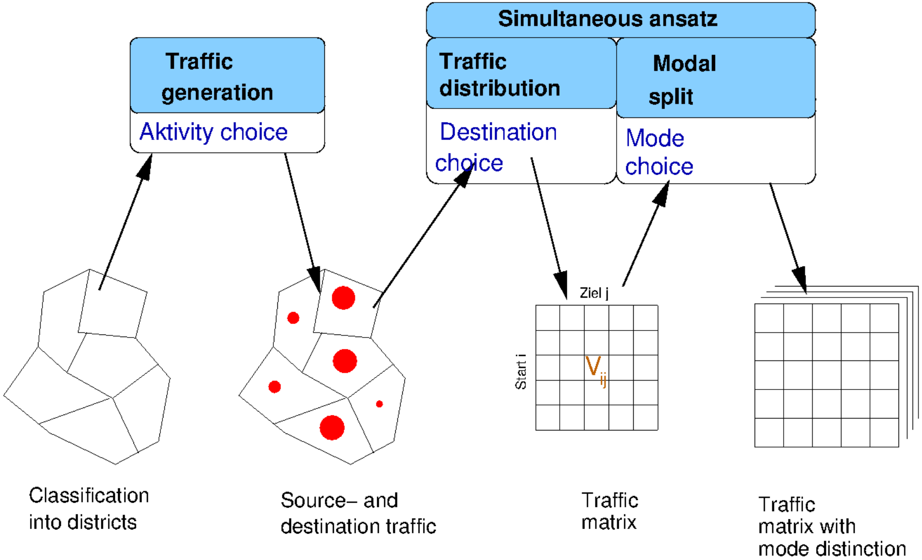
7.1 The Nature of Discrete Decisions

Example: the four-step model of transportation planning



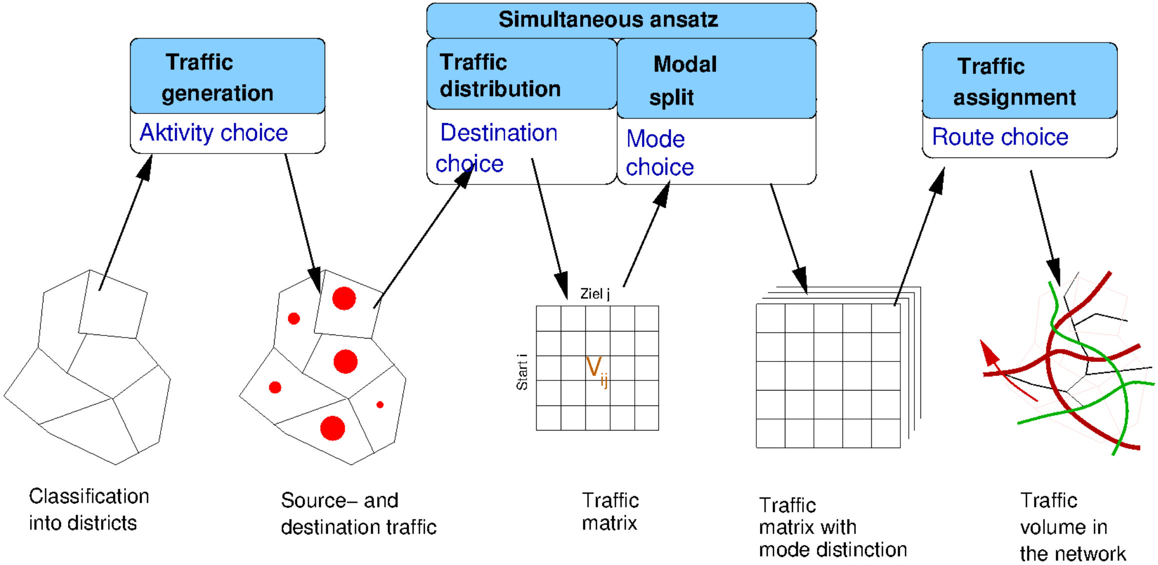
7.1 The Nature of Discrete Decisions

Example: the four-step model of transportation planning



7.1 The Nature of Discrete Decisions

Example: the four-step model of transportation planning



7.2 Basic Concepts: Alternatives

- Each person n has a certain set of discrete alternatives: $\mathcal{A}_n = \{a_{ni}\}$ containing alternatives $i = 1, ..., I_n$.
  Example: Person 1 can chose the modes pedestrian or public transport, Person 2 pedestrian, bike, and car
- The number I_n of alternatives is finite and should not be too large.
  Example: speeding or not speeding; counterexample: choosing the speed [km/h]
- The alternative set needs to be **exclusive** (non-cumulative), i.e., a person can chose at most one alternative.
  Example: you cannot live at two places simultaneously
- The set must be **complete**, i.e., at least one alternative must be chosen. Complete and exclusive $\Rightarrow$ exactly one alternative must be chosen.
  Counterexample: Till Eulenspiegel adopted several (fake) professions
- The alternatives need to be sufficiently different from each other.
  Counterexample: two routes differing by only a small fraction of links

7.2 Basic Concepts: Alternatives

- Each person n has a certain set of discrete alternatives: $\mathcal{A}_n = \{a_{ni}\}$ containing alternatives $i = 1, ..., I_n$.
  Example: Person 1 can chose the modes pedestrian or public transport, Person 2 pedestrian, bike, and car

- The number I_n of alternatives is finite and should not be too large.
  Example: speeding or not speeding; counterexample: choosing the speed [km/h]

- The alternative set needs to be **exclusive** (non-cumulative), i.e., a person can chose at most one alternative.
  Example: you cannot live at two places simultaneously

- The set must be **complete**, i.e., at least one alternative must be chosen. Complete and exclusive $\Rightarrow$ exactly one alternative must be chosen.
  Counterexample: Till Eulenspiegel adopted several (fake) professions

- The alternatives need to be sufficiently different from each other.
  Counterexample: two routes differing by only a small fraction of links

7.2 Basic Concepts: Alternatives

- Each person n has a certain set of discrete alternatives: $\mathcal{A}_n = \{a_{ni}\}$ containing alternatives $i = 1, ..., I_n$.
  Example: Person 1 can chose the modes pedestrian or public transport, Person 2 pedestrian, bike, and car

- The number I_n of alternatives is finite and should not be too large.
  Example: speeding or not speeding; counterexample: choosing the speed [km/h]

- The alternative set needs to be **exclusive** (non-cumulative), i.e., a person can chose at most one alternative.
  Example: you cannot live at two places simultaneously

- The set must be **complete**, i.e., at least one alternative must be chosen. Complete and exclusive $\Rightarrow$ exactly one alternative must be chosen.
  Counterexample: Till Eulenspiegel adopted several (fake) professions

- The alternatives need to be sufficiently different from each other.
  Counterexample: two routes differing by only a small fraction of links

7.2 Basic Concepts: Alternatives

- Each person n has a certain set of discrete alternatives: $\mathcal{A}_n = \{a_{ni}\}$ containing alternatives $i = 1, ..., I_n$.
  Example: Person 1 can chose the modes pedestrian or public transport, Person 2 pedestrian, bike, and car

- The number I_n of alternatives is finite and should not be too large.
  Example: speeding or not speeding; counterexample: choosing the speed [km/h]

- The alternative set needs to be **exclusive** (non-cumulative), i.e., a person can chose at most one alternative.
  Example: you cannot live at two places simultaneously

- The set must be **complete**, i.e., at least one alternative must be chosen. Complete and exclusive $\Rightarrow$ exactly one alternative must be chosen.
  Counterexample: Till Eulenspiegel adopted several (fake) professions

- The alternatives need to be sufficiently different from each other.
  Counterexample: two routes differing by only a small fraction of links

7.2 Basic Concepts: Alternatives

- Each person n has a certain set of discrete alternatives: $\mathcal{A}_n = \{a_{ni}\}$ containing alternatives $i = 1, ..., I_n$.
  Example: Person 1 can chose the modes pedestrian or public transport, Person 2 pedestrian, bike, and car
- The number I_n of alternatives is finite and should not be too large.
  Example: speeding or not speeding; counterexample: choosing the speed [km/h]
- The alternative set needs to be **exclusive** (non-cumulative), i.e., a person can chose at most one alternative.
  Example: you cannot live at two places simultaneously
- The set must be **complete**, i.e., at least one alternative must be chosen. Complete and exclusive $\Rightarrow$ exactly one alternative must be chosen.
  Counterexample: Till Eulenspiegel adopted several (fake) professions
- The alternatives need to be sufficiently different from each other.
  Counterexample: two routes differing by only a small fraction of links

Questions on specifying alternatives

? What to do if a person has the option to not select anything or possibly select something that is not on the list?

! Just include a "do-nothing" and/or an "other" alternative

? What to do if multi-modal trips (e.g., bike+tram) are possible?

! Just add a "multi-modal" alternative

? Assume that someone has no car or bike available. How to model the four alternatives ped, bike, car, PT for this person? Give two possible solutions

! 1. Exclude these alternatives for this person by reducing his/her choice set $\mathcal{A}_n$;
2. Give prohibitive penalties for the "forbidden" alternatives

? Explain "sufficiently different alternatives" by red and blue buses

! Red and blue buses are much less different than both buses from trams, and even buses and trams have many "public transport" commonalities (such as fixed schedules and stops), hence are very similar compared to, e.g., cars or bikes

? What to do with continuous alternatives such as desired speed?

! Define ordinal scales (speed classes) or nominal scales (speeding or not speeding)

Questions on specifying alternatives

? What to do if a person has the option to not select anything or possibly select something that is not on the list?

! Just include a "do-nothing" and/or an "other" alternative

? What to do if multi-modal trips (e.g., bike+tram) are possible?

! Just add a "multi-modal" alternative

? Assume that someone has no car or bike available. How to model the four alternatives ped, bike, car, PT for this person? Give two possible solutions

! 1. Exclude these alternatives for this person by reducing his/her choice set $\mathcal{A}_n$;
2. Give prohibitive penalties for the "forbidden" alternatives

? Explain "sufficiently different alternatives" by red and blue buses

! Red and blue buses are much less different than both buses from trams, and even buses and trams have many "public transport" commonalities (such as fixed schedules and stops), hence are very similar compared to, e.g., cars or bikes

? What to do with continuous alternatives such as desired speed?

! Define ordinal scales (speed classes) or nominal scales (speeding or not speeding)

Questions on specifying alternatives

? What to do if a person has the option to not select anything or possibly select something that is not on the list?

! Just include a "do-nothing" and/or an "other" alternative

? What to do if multi-modal trips (e.g., bike+tram) are possible?

! Just add a "multi-modal" alternative

? Assume that someone has no car or bike available. How to model the four alternatives ped, bike, car, PT for this person? Give two possible solutions

! 1. Exclude these alternatives for this person by reducing his/her choice set $\mathcal{A}_n$;
2. Give prohibitive penalties for the "forbidden" alternatives

? Explain "sufficiently different alternatives" by red and blue buses

! Red and blue buses are much less different than both buses from trams, and even buses and trams have many "public transport" commonalities (such as fixed schedules and stops), hence are very similar compared to, e.g., cars or bikes

? What to do with continuous alternatives such as desired speed?

! Define ordinal scales (speed classes) or nominal scales (speeding or not speeding)

Questions on specifying alternatives

? What to do if a person has the option to not select anything or possibly select something that is not on the list?

! Just include a "do-nothing" and/or an "other" alternative

? What to do if multi-modal trips (e.g., bike+tram) are possible?

! Just add a "multi-modal" alternative

? Assume that someone has no car or bike available. How to model the four alternatives ped, bike, car, PT for this person? Give two possible solutions

! 1. Exclude these alternatives for this person by reducing his/her choice set $\mathcal{A}_n$;
2. Give prohibitive penalties for the "forbidden" alternatives

? Explain "sufficiently different alternatives" by red and blue buses.

! Red and blue buses are much less different than both buses from trams, and even buses and trams have many "public transport" commonalities (such as fixed schedules and stops), hence are very similar compared to, e.g., cars or bikes

? What to do with continuous alternatives such as desired speed?

! Define ordinal scales (speed classes) or nominal scales (speeding or not speeding)

Questions on specifying alternatives

? What to do if a person has the option to not select anything or possibly select something that is not on the list?

! Just include a “do-nothing” and/or an “other” alternative

? What to do if multi-modal trips (e.g., bike+tram) are possible?

! Just add a “multi-modal” alternative

? Assume that someone has no car or bike available. How to model the four alternatives ped, bike, car, PT for this person? Give two possible solutions

! 1. Exclude these alternatives for this person by reducing his/her choice set $\mathcal{A}_n$;
2. Give prohibitive penalties for the “forbidden” alternatives

? Explain “sufficiently different alternatives” by red and blue buses.

! Red and blue buses are much less different than both buses from trams, and even buses and trams have many “public transport” commonalities (such as fixed schedules and stops), hence are very similar compared to, e.g., cars or bikes

? What to do with continuous alternatives such as desired speed?

! Define ordinal scales (speed classes) or nominal scales (speeding or not speeding)

Questions on specifying alternatives

? What to do if a person has the option to not select anything or possibly select something that is not on the list?

! Just include a “do-nothing” and/or an “other” alternative

? What to do if multi-modal trips (e.g., bike+tram) are possible?

! Just add a “multi-modal” alternative

? Assume that someone has no car or bike available. How to model the four alternatives ped, bike, car, PT for this person? Give two possible solutions

! 1. Exclude these alternatives for this person by reducing his/her choice set $\mathcal{A}_n$;
2. Give prohibitive penalties for the “forbidden” alternatives

? Explain “sufficiently different alternatives” by red and blue buses.

! Red and blue buses are much less different than both buses from trams, and even buses and trams have many “public transport” commonalities (such as fixed schedules and stops), hence are very similar compared to, e.g., cars or bikes

? What to do with continuous alternatives such as desired speed?

! Define ordinal scales (speed classes) or nominal scales (speeding or not speeding)

Questions on specifying alternatives

? What to do if a person has the option to not select anything or possibly select something that is not on the list?

! Just include a "do-nothing" and/or an "other" alternative

? What to do if multi-modal trips (e.g., bike+tram) are possible?

! Just add a "multi-modal" alternative

? Assume that someone has no car or bike available. How to model the four alternatives ped, bike, car, PT for this person? Give two possible solutions

! 1. Exclude these alternatives for this person by reducing his/her choice set $\mathcal{A}_n$;
2. Give prohibitive penalties for the "forbidden" alternatives

? Explain "sufficiently different alternatives" by red and blue buses.

! Red and blue buses are much less different than both buses from trams, and even buses and trams have many "public transport" commonalities (such as fixed schedules and stops), hence are very similar compared to, e.g., cars or bikes

? What to do with continuous alternatives such as desired speed?

! Define ordinal scales (speed classes) or nominal scales (speeding or not speeding)

Questions on specifying alternatives

? What to do if a person has the option to not select anything or possibly select something that is not on the list?

! Just include a "do-nothing" and/or an "other" alternative

? What to do if multi-modal trips (e.g., bike+tram) are possible?

! Just add a "multi-modal" alternative

? Assume that someone has no car or bike available. How to model the four alternatives ped, bike, car, PT for this person? Give two possible solutions

! 1. Exclude these alternatives for this person by reducing his/her choice set $\mathcal{A}_n$;
2. Give prohibitive penalties for the "forbidden" alternatives

? Explain "sufficiently different alternatives" by red and blue buses.

! Red and blue buses are much less different than both buses from trams, and even buses and trams have many "public transport" commonalities (such as fixed schedules and stops), hence are very similar compared to, e.g., cars or bikes

? What to do with continuous alternatives such as desired speed?

! Define ordinal scales (speed classes) or nominal scales (speeding or not speeding)

Questions on specifying alternatives

? What to do if a person has the option to not select anything or possibly select something that is not on the list?

! Just include a "do-nothing" and/or an "other" alternative

? What to do if multi-modal trips (e.g., bike+tram) are possible?

! Just add a "multi-modal" alternative

? Assume that someone has no car or bike available. How to model the four alternatives ped, bike, car, PT for this person? Give two possible solutions

! 1. Exclude these alternatives for this person by reducing his/her choice set $\mathcal{A}_n$;
2. Give prohibitive penalties for the "forbidden" alternatives

? Explain "sufficiently different alternatives" by red and blue buses.

! Red and blue buses are much less different than both buses from trams, and even buses and trams have many "public transport" commonalities (such as fixed schedules and stops), hence are very similar compared to, e.g., cars or bikes

? What to do with continuous alternatives such as desired speed?

! Define ordinal scales (speed classes) or nominal scales (speeding or not speeding)

Questions on specifying alternatives

? What to do if a person has the option to not select anything or possibly select something that is not on the list?

! Just include a "do-nothing" and/or an "other" alternative

? What to do if multi-modal trips (e.g., bike+tram) are possible?

! Just add a "multi-modal" alternative

? Assume that someone has no car or bike available. How to model the four alternatives ped, bike, car, PT for this person? Give two possible solutions

! 1. Exclude these alternatives for this person by reducing his/her choice set $\mathcal{A}_n$;
2. Give prohibitive penalties for the "forbidden" alternatives

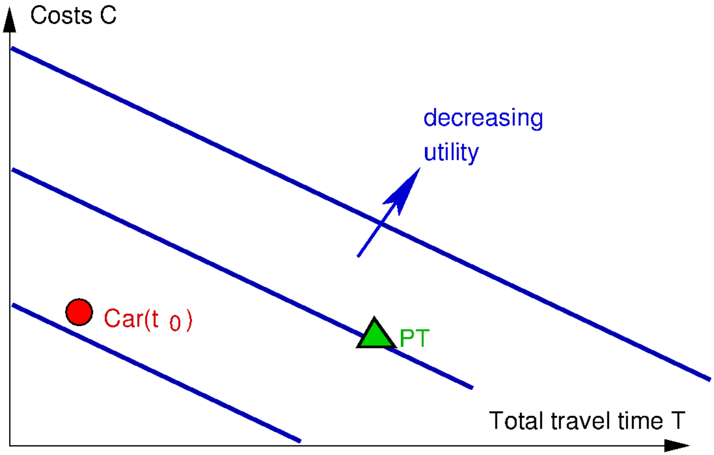
? Explain "sufficiently different alternatives" by red and blue buses.

! Red and blue buses are much less different than both buses from trams, and even buses and trams have many "public transport" commonalities (such as fixed schedules and stops), hence are very similar compared to, e.g., cars or bikes

? What to do with continuous alternatives such as desired speed?

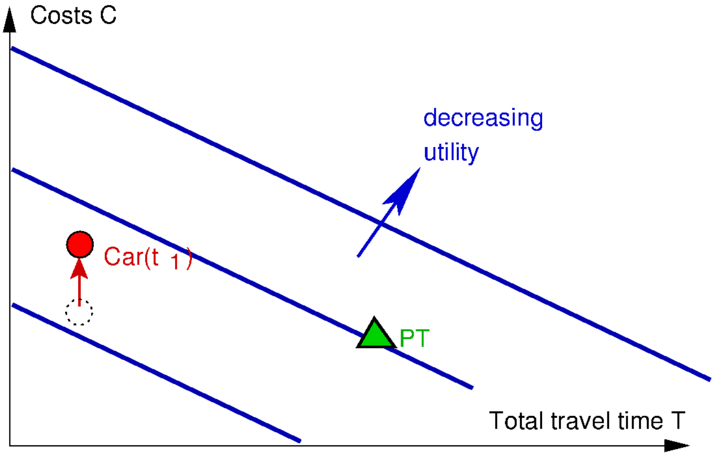
! Define ordinal scales (speed classes) or nominal scales (speeding or not speeding)

7.2 The essence of the Homo Oeconomicus: two alternatives



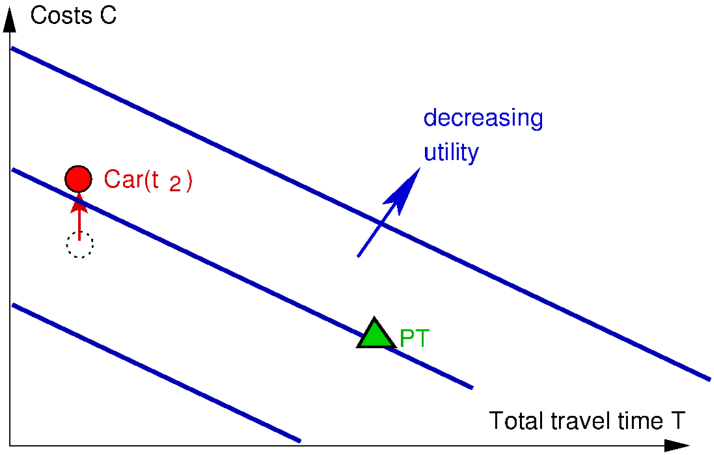
Time t_0: Chosen alternative $i_{\text{selected}} = \arg\max_i U_i = 1$ (public transport)

7.2 The essence of the Homo Oeconomicus: two alternatives



Time t_1: Chosen alternative $i_{\text{selected}} = \arg\max_i U_i = 1$ (public transport)

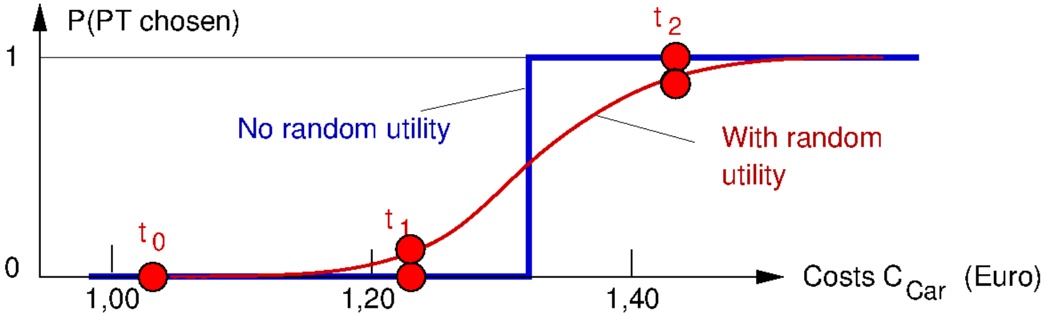
7.2 The essence of the Homo Oeconomicus: two alternatives



Time t_2: Chosen alternative $i_{\text{selected}} = \arg\max_i U_i = 2$ (car)

sudden change $\Rightarrow$ intrinsically nonlinear response!

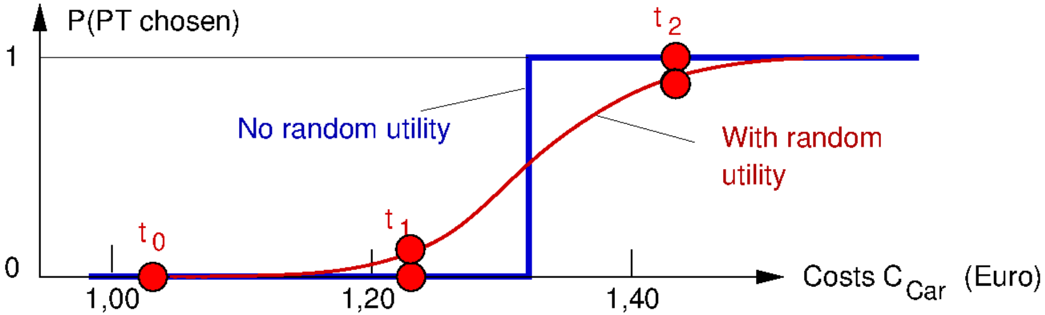
Decisions under uncertainty



$$\text{Alternative } i_{\text{selected}} = \arg\max_i U_i = \arg\max_i V_i + \epsilon_i$$

- On an *individual* level, the decision is "yes" (1) or "no" (0)
- On an *aggregated* level, we have a certain probability between 0 and 1.

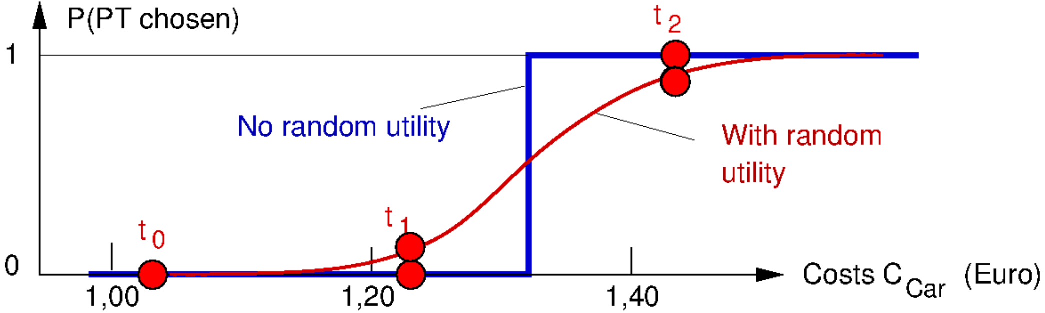
Decisions under uncertainty



$$\text{Alternative } i_{\text{selected}} = \text{arg}\max_i U_i = \text{arg}\max_i V_i + \epsilon_i$$

- On an *individual* level, the decision is "yes" (1) or "no" (0)
- On an *aggregated* level, we have a certain probability between 0 and 1.

Decisions under uncertainty



$$\text{Alternative } i_{\text{selected}} = \text{arg}\max_i U_i = \text{arg}\max_i V_i + \epsilon_i$$

- On an *individual* level, the decision is “yes” (1) or “no” (0)
- On an *aggregated* level, we have a certain probability between 0 and 1.

Structure and classification of dicrete-choice models

Exogenous variable ("characteristics")

Total travel times T_{ni} Costs C_{ni}
Gender g_i Age τ_i

Structure and classification of dicrete-choice models

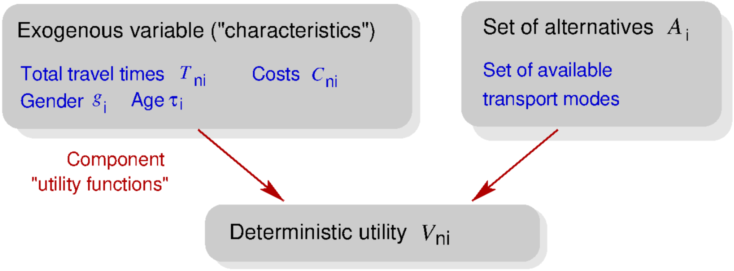
Exogenous variable ("characteristics")

Total travel times T_{ni} Costs C_{ni}
Gender g_i Age τ_i

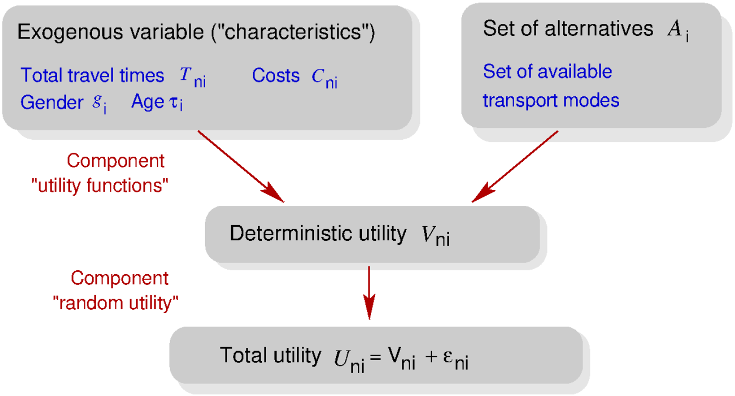
Set of alternatives A_i

Set of available transport modes

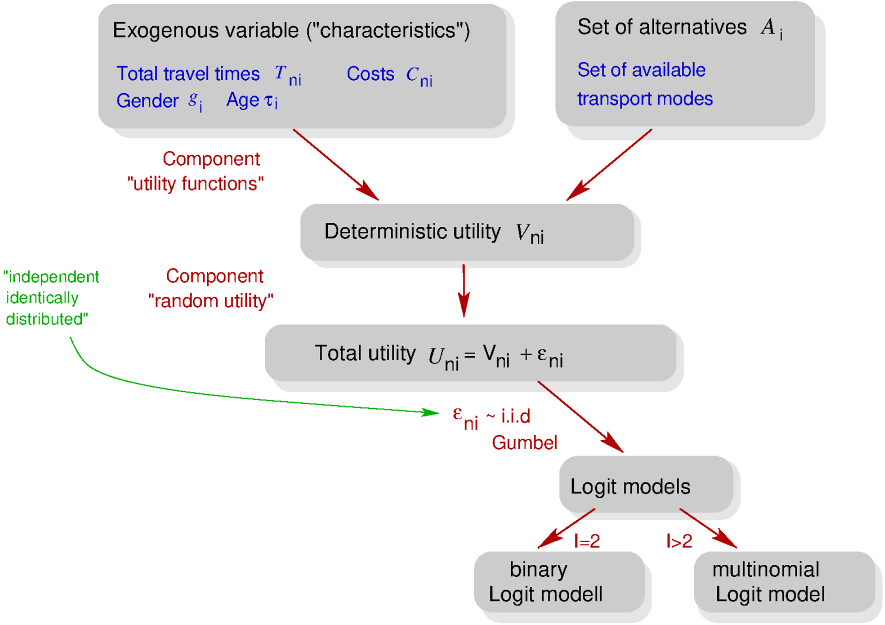
Structure and classification of dicrete-choice models



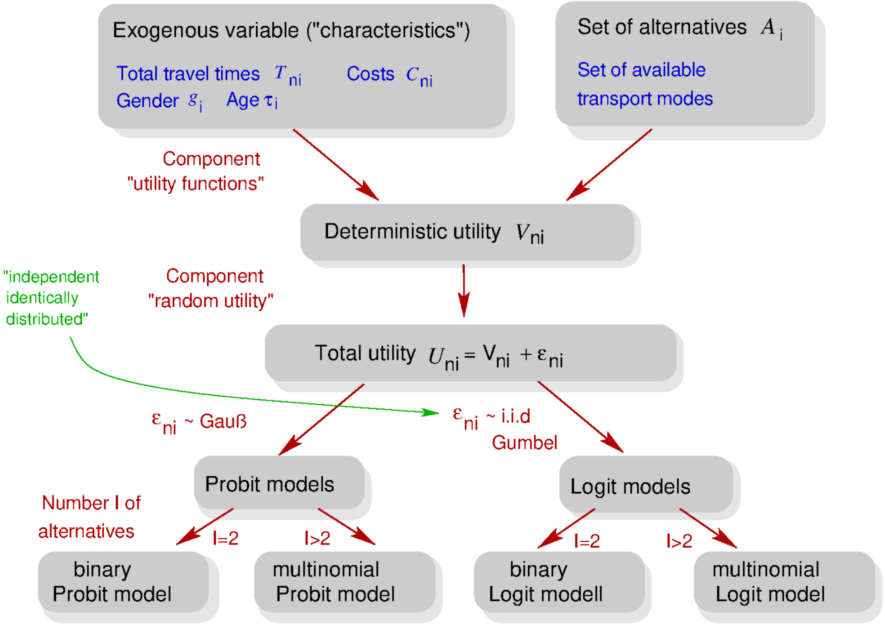
Structure and classification of dicrete-choice models



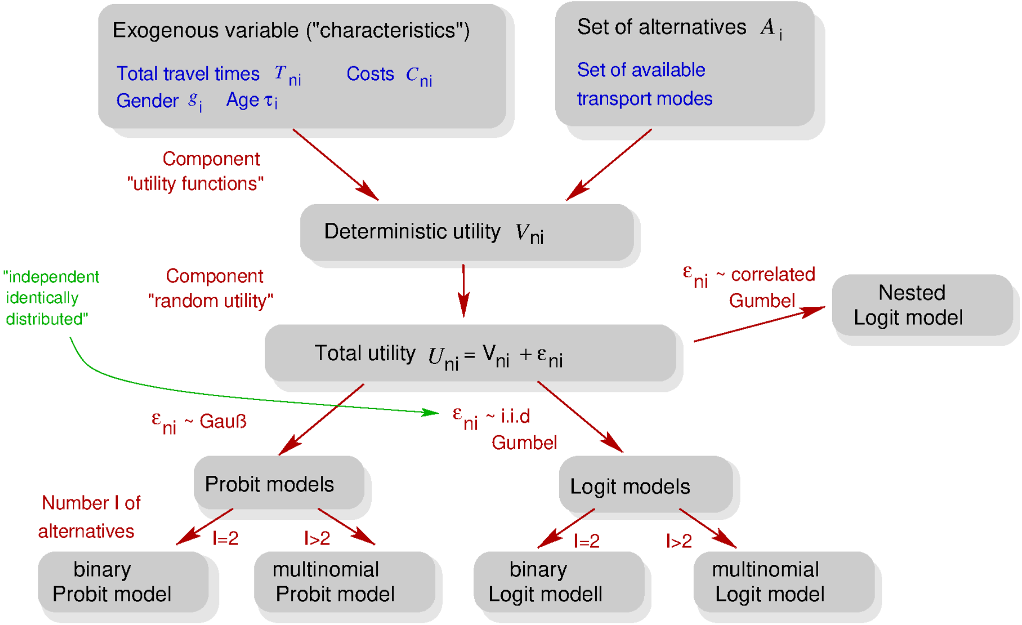
Structure and classification of dicrete-choice models



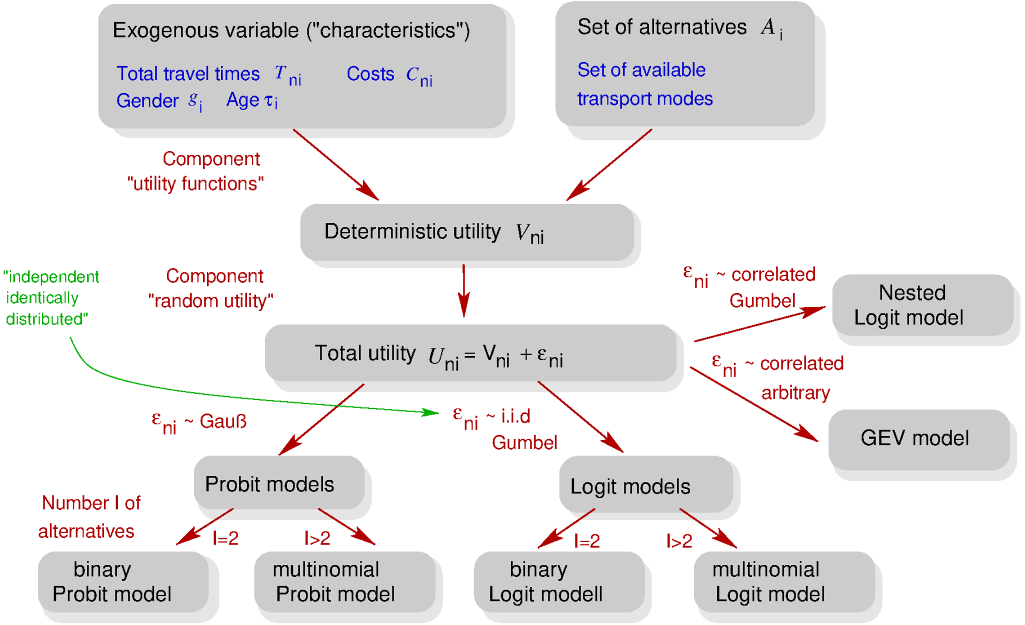
Structure and classification of dicrete-choice models



Structure and classification of dicrete-choice models



Structure and classification of dicrete-choice models



7.3 Deterministic Utilities

- The **deterministic utilities** V_{ni} of alternative i for person n are like the endogenous variables of regression models: continuous and made up of linear factors:

$$V_{ni} = \sum_m \beta_m X_{mni}$$

- The person index (or choice-set index) n plays the role of a data-point index i in regression models (the index naming is by convention) and X_{mni} corresponds to the system matrix.

- The factors may contain alternative-specific constants and three categories of exogenous variables:
  - **alternative-specific constants (ACs)** play the role of constants in regression models,
  - **characteristics** are attributes of the alternatives,
  - **socioeconomic factors** are attributes of the decision makers,
  - **external factors** depend neither on the alternatives nor the persons but influence the decisions.

7.3 Deterministic Utilities

- The **deterministic utilities** V_{ni} of alternative i for person n are like the endogenous variables of regression models: continuous and made up of linear factors:

$$V_{ni} = \sum_m \beta_m X_{mni}$$

- The person index (or choice-set index) n plays the role of a data-point index i in regression models (the index naming is by convention) and X_{mni} corresponds to the system matrix.
- The factors may contain alternative-specific constants and three categories of exogenous variables:
  - **alternative-specific constants (ACs)** play the role of constants in regression models,
  - **characteristics** are attributes of the alternatives,
  - **socioeconomic factors** are attributes of the decision makers,
  - **external factors** depend neither on the alternatives nor the persons but influence the decisions.

7.3 Deterministic Utilities

- The **deterministic utilities** V_{ni} of alternative i for person n are like the endogenous variables of regression models: continuous and made up of linear factors:

$$V_{ni} = \sum_m \beta_m X_{mni}$$

- The person index (or choice-set index) n plays the role of a data-point index i in regression models (the index naming is by convention) and X_{mni} corresponds to the system matrix.
- The factors may contain alternative-specific constants and three categories of exogenous variables:
  - **alternative-specific constants (ACs)** play the role of constants in regression models,
  - **characteristics** are attributes of the alternatives,
  - **socioeconomic factors** are attributes of the decision makers,
  - **external factors** depend neither on the alternatives nor the persons but influence the decisions.

Factors I: Alternative-specific constants (ACs)

- In most situations, there are systematic ("ad-hoc") preferences for certain alternatives not explained by the characteristics or socioeconomic variables.
- Since only utility *differences* are relevant for the choice, normal constants are of no use. We need **alternative-specific constants** or **ACs** which are essentially *selector dummy variables*:

$$V_{ni} = \beta_{01}\delta_{1i} + ... + \beta_{0,I-1}\delta_{I-1,i}, \quad \delta_{ji} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

- The parameter β_{0i} denotes the **ad-hoc preference** of Alternative i (in utility units) over the **reference alternative** I if all non-AC factors are zero.

? What would happen if using a normal constant, $V_{ni} = \beta_0$?
Nothing since the utility differences are unchanged

? Give the ad-hoc preference of alternative j over k, $j, k \neq I$
$\beta_{0j}\delta_{ji} - \beta_{0k}\delta_{ki}$ such that $V_{nj} - V_{nk} = \beta_{0j} - \beta_{0k} +$ other factors

? What changes when making Alternative 1 the reference?
$V_{ni} = \beta_{02}\delta_{2i} + ... + \beta_{0I}\delta_{Ii}$

Factors I: Alternative-specific constants (ACs)

- In most situations, there are systematic ("ad-hoc") preferences for certain alternatives not explained by the characteristics or socioeconomic variables.
- Since only utility *differences* are relevant for the choice, normal constants are of no use. We need **alternative-specific constants** or **ACs** which are essentially *selector dummy variables*:

$$V_{ni} = \beta_{01}\delta_{1i} + ... + \beta_{0,I-1}\delta_{I-1,i}, \quad \delta_{ji} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

- The parameter β_{0i} denotes the **ad-hoc preference** of Alternative i (in utility units) over the **reference alternative** I if all non-AC factors are zero.

? What would happen if using a normal constant, $V_{ni} = \beta_0$?

Nothing since the utility differences are unchanged

? Give the ad-hoc preference of alternative j over k, $j, k \neq I$

$\beta_{0j}\delta_{ji} - \beta_{0k}\delta_{ki}$ such that $V_{nj} - V_{nk} = \beta_{0j} - \beta_{0k} +$ other factors

? What changes when making Alternative 1 the reference?

$V_{ni} = \beta_{02}\delta_{2i} + ... + \beta_{0I}\delta_{Ii}$

Factors I: Alternative-specific constants (ACs)

- In most situations, there are systematic ("ad-hoc") preferences for certain alternatives not explained by the characteristics or socioeconomic variables.
- Since only utility *differences* are relevant for the choice, normal constants are of no use. We need **alternative-specific constants** or **ACs** which are essentially *selector dummy variables*:

$$V_{ni} = \beta_{01}\delta_{1i} + ... + \beta_{0,I-1}\delta_{I-1,i}, \quad \delta_{ji} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

- The parameter β_{0i} denotes the **ad-hoc preference** of Alternative i (in utility units) over the **reference alternative** I if all non-AC factors are zero.

? What would happen if using a normal constant, $V_{ni} = \beta_0$?

Nothing since the utility differences are unchanged

? Give the ad-hoc preference of alternative j over k, $j,k \neq I$

$\beta_{0j}\delta_{ji} - \beta_{0k}\delta_{ki}$ such that $V_{nj} - V_{nk} = \beta_{0j} - \beta_{0k}$ + other factors

? What changes when making Alternative 1 the reference?

$V_{ni} = \beta_{02}\delta_{2i} + ... + \beta_{0I}\delta_{Ii}$

Factors I: Alternative-specific constants (ACs)

- In most situations, there are systematic ("ad-hoc") preferences for certain alternatives not explained by the characteristics or socioeconomic variables.
- Since only utility *differences* are relevant for the choice, normal constants are of no use. We need **alternative-specific constants** or **ACs** which are essentially *selector dummy variables*:

$$V_{ni} = \beta_{01}\delta_{1i} + ... + \beta_{0,I-1}\delta_{I-1,i}, \quad \delta_{ji} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

- The parameter β_{0i} denotes the **ad-hoc preference** of Alternative i (in utility units) over the **reference alternative** I if all non-AC factors are zero.

? What would happen if using a normal constant, $V_{ni} = \beta_0$?

Nothing since the utility differences are unchanged

? Give the ad-hoc preference of alternative j over k, $j, k \neq I$

$\beta_{0j}\delta_{ji} - \beta_{0k}\delta_{ki}$ such that $V_{nj} - V_{nk} = \beta_{0j} - \beta_{0k}$ + other factors

? What changes when making Alternative 1 the reference?

$V_{ni} = \beta_{02}\delta_{2i} + ... + \beta_{0I}\delta_{Ii}$

Factors I: Alternative-specific constants (ACs)

- In most situations, there are systematic ("ad-hoc") preferences for certain alternatives not explained by the characteristics or socioeconomic variables.
- Since only utility *differences* are relevant for the choice, normal constants are of no use. We need **alternative-specific constants** or **ACs** which are essentially *selector dummy variables*:

$$V_{ni} = \beta_{01}\delta_{1i} + ... + \beta_{0,I-1}\delta_{I-1,i}, \quad \delta_{ji} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

- The parameter β_{0i} denotes the **ad-hoc preference** of Alternative i (in utility units) over the **reference alternative** I if all non-AC factors are zero.

? What would happen if using a normal constant, $V_{ni} = \beta_0$?
Nothing since the utility differences are unchanged

? Give the ad-hoc preference of alternative j over k, $j, k \neq I$
$\beta_{0j}\delta_{ji} - \beta_{0k}\delta_{ki}$ such that $V_{nj} - V_{nk} = \beta_{0j} - \beta_{0k} +$ other factors

? What changes when making Alternative 1 the reference?
$V_{ni} = \beta_{02}\delta_{2i} + ... + \beta_{0I}\delta_{Ii}$

Factors I: Alternative-specific constants (ACs)

- In most situations, there are systematic ("ad-hoc") preferences for certain alternatives not explained by the characteristics or socioeconomic variables.
- Since only utility *differences* are relevant for the choice, normal constants are of no use. We need **alternative-specific constants** or **ACs** which are essentially *selector dummy variables*:

$$V_{ni} = \beta_{01}\delta_{1i} + \ldots + \beta_{0,I-1}\delta_{I-1,i}, \quad \delta_{ji} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

- The parameter β_{0i} denotes the **ad-hoc preference** of Alternative i (in utility units) over the **reference alternative** I if all non-AC factors are zero.

? What would happen if using a normal constant, $V_{ni} = \beta_0$?
Nothing since the utility differences are unchanged

? Give the ad-hoc preference of alternative j over k, $j, k \neq I$
$\beta_{0j}\delta_{ji} - \beta_{0k}\delta_{ki}$ such that $V_{nj} - V_{nk} = \beta_{0j} - \beta_{0k}$ + other factors

? What changes when making Alternative 1 the reference?
$V_{ni} = \beta_{02}\delta_{2i} + \ldots + \beta_{0I}\delta_{Ii}$

Factors I: Alternative-specific constants (ACs)

- In most situations, there are systematic ("ad-hoc") preferences for certain alternatives not explained by the characteristics or socioeconomic variables.
- Since only utility *differences* are relevant for the choice, normal constants are of no use. We need **alternative-specific constants** or **ACs** which are essentially *selector dummy variables*:

$$V_{ni} = \beta_{01}\delta_{1i} + ... + \beta_{0,I-1}\delta_{I-1,i}, \quad \delta_{ji} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

- The parameter β_{0i} denotes the **ad-hoc preference** of Alternative i (in utility units) over the **reference alternative** I if all non-AC factors are zero.

? What would happen if using a normal constant, $V_{ni} = \beta_0$?
Nothing since the utility differences are unchanged

? Give the ad-hoc preference of alternative j over k, $j, k \neq I$
$\beta_{0j}\delta_{ji} - \beta_{0k}\delta_{ki}$ such that $V_{nj} - V_{nk} = \beta_{0j} - \beta_{0k} +$ other factors

? What changes when making Alternative 1 the reference?
$V_{ni} = \beta_{02}\delta_{2i} + ... + \beta_{0I}\delta_{Ii}$

Factors I: Alternative-specific constants (ACs)

- In most situations, there are systematic ("ad-hoc") preferences for certain alternatives not explained by the characteristics or socioeconomic variables.
- Since only utility *differences* are relevant for the choice, normal constants are of no use. We need **alternative-specific constants** or **ACs** which are essentially *selector dummy variables*:

$$V_{ni} = \beta_{01}\delta_{1i} + ... + \beta_{0,I-1}\delta_{I-1,i}, \quad \delta_{ji} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

- The parameter β_{0i} denotes the **ad-hoc preference** of Alternative i (in utility units) over the **reference alternative** I if all non-AC factors are zero.

? What would happen if using a normal constant, $V_{ni} = \beta_0$?
Nothing since the utility differences are unchanged

? Give the ad-hoc preference of alternative j over k, $j, k \neq I$
$\beta_{0j}\delta_{ji} - \beta_{0k}\delta_{ki}$ such that $V_{nj} - V_{nk} = \beta_{0j} - \beta_{0k}$ + other factors

? What changes when making Alternative 1 the reference?
$V_{ni} = \beta_{02}\delta_{2i} + ... + \beta_{0I}\delta_{Ii}$

Factors I: Alternative-specific constants (ACs)

- In most situations, there are systematic ("ad-hoc") preferences for certain alternatives not explained by the characteristics or socioeconomic variables.
- Since only utility *differences* are relevant for the choice, normal constants are of no use. We need **alternative-specific constants** or **ACs** which are essentially *selector dummy variables*:

$$V_{ni} = \beta_{01}\delta_{1i} + ... + \beta_{0,I-1}\delta_{I-1,i}, \quad \delta_{ji} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

- The parameter β_{0i} denotes the **ad-hoc preference** of Alternative i (in utility units) over the **reference alternative** I if all non-AC factors are zero.

? What would happen if using a normal constant, $V_{ni} = \beta_0$?
Nothing since the utility differences are unchanged

? Give the ad-hoc preference of alternative j over k, $j, k \neq I$
$\beta_{0j}\delta_{ji} - \beta_{0k}\delta_{ki}$ such that $V_{nj} - V_{nk} = \beta_{0j} - \beta_{0k} +$ other factors

? What changes when making Alternative 1 the reference?
$V_{ni} = \beta_{02}\delta_{2i} + ... + \beta_{0I}\delta_{Ii}$

Factors II: Characteristics

- The characteristic C_{mni} is the m^{th} attribute of alternative i for person (choice set) n.
- Examples:
  - Complex travel time $C_{1in} = T_{ni}$ for person n when travelling by transport mode i,
  - Ad-hoc costs $C_{2in} = C_{ni}$ for person n when using mode i
- Characteristics can be identified directly with factors, $X_{mni} = C_{mni}$. Modelling variants:
  - **generic** ansatz $V_{ni} = ... + \beta_m C_{mni}$ (the same parameter β_m for all alternatives i),
  - **alternative-specific** ansatz $V_{ni} = ... + \beta_{mi} C_{mni}$ with alternative-specific parameters β_{mi}.

? Which way is to be preferred when modelling (i) travel times and (ii) costs? (justify!)

One minute of public transport is weighted differently to a bike minute ⇒ model travel times alternative-specifically. However, people typically have only one *mental account* for small spendings ⇒ model costs generically

? Give an example of a characteristic not depending on the person.

Tricky since both time and costs generally depend on the person (even for the same OD relation). *Reliability* would be a good candidate

Factors II: Characteristics

- The characteristic C_{mni} is the m^{th} attribute of alternative i for person (choice set) n.
- Examples:
  - Complex travel time $C_{1in} = T_{ni}$ for person n when travelling by transport mode i,
  - Ad-hoc costs $C_{2in} = C_{ni}$ for person n when using mode i
- Characteristics can be identified directly with factors, $X_{mni} = C_{mni}$. Modelling variants:
  - **generic** ansatz $V_{ni} = ... + \beta_m C_{mni}$ (the same parameter β_m for all alternatives i),
  - **alternative-specific** ansatz $V_{ni} = ... + \beta_{mi} C_{mni}$ with alternative-specific parameters β_{mi}.

? Which way is to be preferred when modelling (i) travel times and (ii) costs? (justify!)
One minute of public transport is weighted differently to a bike minute $\Rightarrow$ model travel times alternative-specifically. However, people typically have only one *mental account* for small spendings $\Rightarrow$ model costs generically

? Give an example of a characteristic not depending on the person.
Tricky since both time and costs generally depend on the person (even for the same OD relation). *Reliability* would be a good candidate

Factors II: Characteristics

- The characteristic C_{mni} is the m^{th} attribute of alternative i for person (choice set) n.
- Examples:
  - Complex travel time $C_{1in} = T_{ni}$ for person n when travelling by transport mode i,
  - Ad-hoc costs $C_{2in} = C_{ni}$ for person n when using mode i
- Characteristics can be identified directly with factors, $X_{mni} = C_{mni}$. Modelling variants:
  - **generic** ansatz $V_{ni} = ... + \beta_m C_{mni}$ (the same parameter β_m for all alternatives i),
  - **alternative-specific** ansatz $V_{ni} = ... + \beta_{mi} C_{mni}$ with alternative-specific parameters β_{mi}.

? Which way is to be preferred when modelling (i) travel times and (ii) costs? (justify!)

One minute of public transport is weighted differently to a bike minute $\Rightarrow$ model travel times alternative-specifically. However, people typically have only one *mental account* for small spendings $\Rightarrow$ model costs generically

? Give an example of a characteristic not depending on the person.

Tricky since both time and costs generally depend on the person (even for the same OD relation). *Reliability* would be a good candidate

Factors II: Characteristics

- The characteristic C_{mni} is the m^{th} attribute of alternative i for person (choice set) n.
- Examples:
  - Complex travel time $C_{1in} = T_{ni}$ for person n when travelling by transport mode i,
  - Ad-hoc costs $C_{2in} = C_{ni}$ for person n when using mode i
- Characteristics can be identified directly with factors, $X_{mni} = C_{mni}$. Modelling variants:
  - **generic** ansatz $V_{ni} = ... + \beta_m C_{mni}$ (the same parameter β_m for all alternatives i),
  - **alternative-specific** ansatz $V_{ni} = ... + \beta_{mi} C_{mni}$ with alternative-specific parameters β_{mi}.

? Which way is to be preferred when modelling (i) travel times and (ii) costs? (justify!)

One minute of public transport is weighted differently to a bike minute $\Rightarrow$ model travel times alternative-specifically. However, people typically have only one *mental account* for small spendings $\Rightarrow$ model costs generically

? Give an example of a characteristic not depending on the person.

Tricky since both time and costs generally depend on the person (even for the same OD relation). *Reliability* would be a good candidate

Factors II: Characteristics

- The characteristic C_{mni} is the m^{th} attribute of alternative i for person (choice set) n.
- Examples:
  - Complex travel time $C_{1in} = T_{ni}$ for person n when travelling by transport mode i,
  - Ad-hoc costs $C_{2in} = C_{ni}$ for person n when using mode i
- Characteristics can be identified directly with factors, $X_{mni} = C_{mni}$. Modelling variants:
  - **generic** ansatz $V_{ni} = ... + \beta_m C_{mni}$ (the same parameter β_m for all alternatives i),
  - **alternative-specific** ansatz $V_{ni} = ... + \beta_{mi} C_{mni}$ with alternative-specific parameters β_{mi}.

? Which way is to be preferred when modelling (i) travel times and (ii) costs? (justify!)
One minute of public transport is weighted differently to a bike minute $\Rightarrow$ model travel times alternative-specifically. However, people typically have only one *mental account* for small spendings $\Rightarrow$ model costs generically

? Give an example of a characteristic not depending on the person.
Tricky since both time and costs generally depend on the person (even for the same OD relation). *Reliability* would be a good candidate

Factors II: Characteristics

- The characteristic C_{mni} is the m^{th} attribute of alternative i for person (choice set) n.
- Examples:
  - Complex travel time $C_{1in} = T_{ni}$ for person n when travelling by transport mode i,
  - Ad-hoc costs $C_{2in} = C_{ni}$ for person n when using mode i
- Characteristics can be identified directly with factors, $X_{mni} = C_{mni}$. Modelling variants:
  - **generic** ansatz $V_{ni} = ... + \beta_m C_{mni}$ (the same parameter β_m for all alternatives i),
  - **alternative-specific** ansatz $V_{ni} = ... + \beta_{mi} C_{mni}$ with alternative-specific parameters β_{mi}.

? Which way is to be preferred when modelling (i) travel times and (ii) costs? (justify!)

One minute of public transport is weighted differently to a bike minute $\Rightarrow$ model travel times alternative-specifically. However, people typically have only one *mental account* for small spendings $\Rightarrow$ model costs generically

? Give an example of a characteristic not depending on the person.

Tricky since both time and costs generally depend on the person (even for the same OD relation). *Reliability* would be a good candidate

Factors II: Characteristics

- The characteristic C_{mni} is the m^{th} attribute of alternative i for person (choice set) n.
- Examples:
  - Complex travel time $C_{1in} = T_{ni}$ for person n when travelling by transport mode i,
  - Ad-hoc costs $C_{2in} = C_{ni}$ for person n when using mode i
- Characteristics can be identified directly with factors, $X_{mni} = C_{mni}$. Modelling variants:
  - **generic** ansatz $V_{ni} = ... + \beta_m C_{mni}$ (the same parameter β_m for all alternatives i),
  - **alternative-specific** ansatz $V_{ni} = ... + \beta_{mi} C_{mni}$ with alternative-specific parameters β_{mi}.

? Which way is to be preferred when modelling (i) travel times and (ii) costs? (justify!)
One minute of public transport is weighted differently to a bike minute $\Rightarrow$ model travel times alternative-specifically. However, people typically have only one *mental account* for small spendings $\Rightarrow$ model costs generically

? Give an example of a characteristic not depending on the person.
Tricky since both time and costs generally depend on the person (even for the same OD relation). *Reliability* would be a good candidate

Factors III: Socio-economic variables

- The **socioeconomic variable** S_{mn} denotes the m^{th} attribute of person n, e.g., age, gender, income, or the distance to the next PT station.
- Sometimes, socio-economic variables have a *nominal* scaling, e.g. having a season ticket or not, or the gender g_n with the values *male* and *female*. As in regression, the way to deal with this are **dummy variables**, e.g.

$$g_n = \begin{cases} 1 \\ 0 \end{cases}$$

- Since, by definition, socio-economic variables do *not* depend on the alternative and the choice depend on utility *differences*, they cannot identified directly as a factor.
- Instead, there are two ways to specify valid linear factors:
  1. formulate it in an alternative-specific way with ACs:

$$V_{ni} = ... + \beta_{m1} S_{mn} \delta_{1i} + ... + \beta_{m,I-1} S_{mn} \delta_{I-1,i}, \quad \delta_{ji} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

  2. let it **interact** with a characteristic m', e.g., $X_{mni} = S_{mn} C_{m'ni}$ and model it in an alternative-specific way $V_{ni} = ... + \beta_{mi} S_{mn} C_{m'ni}$

Factors III: Socio-economic variables

- The **socioeconomic variable** S_{mn} denotes the m^{th} attribute of person n, e.g., age, gender, income, or the distance to the next PT station.
- Sometimes, socio-economic variables have a *nominal* scaling, e.g. having a season ticket or not, or the gender g_n with the values *male* and *female*. As in regression, the way to deal with this are **dummy variables**, e.g.

$$g_n = \begin{cases} 1 & \female \\ 0 & \male \end{cases}$$

- Since, by definition, socio-economic variables do *not* depend on the alternative and the choice depend on utility *differences*, they cannot identified directly as a factor.
- Instead, there are two ways to specify valid linear factors:
  1. formulate it in an alternative-specific way with ACs:

$$V_{ni} = ... + \beta_{m1} S_{mn} \delta_{1i} + ... + \beta_{m,I-1} S_{mn} \delta_{I-1,i}, \quad \delta_{ji} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

  2. let it **interact** with a characteristic m', e.g., $X_{mni} = S_{mn} C_{m'ni}$ and model it in an alternative-specific way $V_{ni} = ... + \beta_{mi} S_{mn} C_{m'ni}$

Factors III: Socio-economic variables

- The **socioeconomic variable** S_{mn} denotes the m^{th} attribute of person n, e.g., age, gender, income, or the distance to the next PT station.
- Sometimes, socio-economic variables have a *nominal* scaling, e.g. having a season ticket or not, or the gender g_n with the values *male* and *female*. As in regression, the way to deal with this are **dummy variables**, e.g.

$$g_n = \begin{cases} 1 & \female \\ 0 & \male \end{cases}$$

- Since, by definition, socio-economic variables do *not* depend on the alternative and the choice depend on utility *differences*, they cannot identified directly as a factor.
- Instead, there are two ways to specify valid linear factors:
  1. formulate it in an alternative-specific way with ACs:

$$V_{ni} = \ldots + \beta_{m1} S_{mn} \delta_{1i} + \ldots + \beta_{m,I-1} S_{mn} \delta_{I-1,i}, \quad \delta_{ji} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

  2. let it **interact** with a characteristic m', e.g., $X_{mni} = S_{mn} C_{m'ni}$ and model it in an alternative-specific way $V_{ni} = \ldots + \beta_{mi} S_{mn} C_{m'ni}$

Factors III: Socio-economic variables

- The **socioeconomic variable** S_{mn} denotes the m^{th} attribute of person n, e.g., age, gender, income, or the distance to the next PT station.
- Sometimes, socio-economic variables have a *nominal* scaling, e.g. having a season ticket or not, or the gender g_n with the values *male* and *female*. As in regression, the way to deal with this are **dummy variables**, e.g.

$$g_n = \begin{cases} 1 & ♀ \\ 0 & ♂ \end{cases}$$

- Since, by definition, socio-economic variables do *not* depend on the alternative and the choice depend on utility *differences*, they cannot identified directly as a factor.
- Instead, there are two ways to specify valid linear factors:
  1. formulate it in an alternative-specific way with ACs:

$$V_{ni} = ... + \beta_{m1} S_{mn} \delta_{1i} + ... + \beta_{m,I-1} S_{mn} \delta_{I-1,i}, \quad \delta_{ji} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

  2. let it **interact** with a characteristic m', e.g., $X_{mni} = S_{mn} C_{m'ni}$ and model it in an alternative-specific way $V_{ni} = ... + \beta_{mi} S_{mn} C_{m'ni}$

Factors III: Socio-economic variables

- The **socioeconomic variable** S_{mn} denotes the m^{th} attribute of person n, e.g., age, gender, income, or the distance to the next PT station.
- Sometimes, socio-economic variables have a *nominal* scaling, e.g. having a season ticket or not, or the gender g_n with the values *male* and *female*. As in regression, the way to deal with this are **dummy variables**, e.g.

$$g_n = \begin{cases} 1 & ♀ \\ 0 & ♂ \end{cases}$$

- Since, by definition, socio-economic variables do *not* depend on the alternative and the choice depend on utility *differences*, they cannot identified directly as a factor.
- Instead, there are two ways to specify valid linear factors:
  1. formulate it in an alternative-specific way with ACs:

$$V_{ni} = ... + \beta_{m1} S_{mn} \delta_{1i} + ... + \beta_{m,I-1} S_{mn} \delta_{I-1,i}, \quad \delta_{ji} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

  2. let it **interact** with a characteristic m', e.g., $X_{mni} = S_{mn} C_{m'ni}$ and model it in an alternative-specific way $V_{ni} = ... + \beta_{mi} S_{mn} C_{m'ni}$

Questions on socio-economic variables

? What would happen when setting $V_{ni} = \ldots + \beta_m S_{mn}$ for all i?
nothing

? How to model that woman prefer PT over other modes more than men?
Assume $i = 3$ is the PT and $S_{mn} = g_n = 1$ for woman and =0 for men.

- If the PT alternative $i = 3$ is not the reference alternative, add a term $+\beta_g g_n \delta_{i3}$ with β_g expected to be positive
- If the PT alternative $i = 3$ is the reference, then add a term $-\beta_g g_n \sum_{j \neq 3} \delta_{ji}$ where β_g has *exactly* the same value (why?)

? How to model an income-dependent time sensitivity?
Let the income I_n interact with the travel time by a factor $+\beta_I I_n T_{ni}$ with β_I expected to be negative

? Why it is not possible to use the age directly?
Since very young and very old persons often act more similar than middle-aged persons, age (as a socio-economic variable) needs not only be coupled to ACs or characteristics but also be formulated in terms of age-group dummies, e.g., $S_{yn} = 1$ if young, zero otherwise, $S_{on} = 1$ if old, zero otherwise (the reference is the "middle age")

Questions on socio-economic variables

? What would happen when setting $V_{ni} = ... + \beta_m S_{mn}$ for all i?
nothing

? How to model that woman prefer PT over other modes more than men?
Assume $i = 3$ is the PT and $S_{mn} = g_n = 1$ for woman and =0 for men.

- If the PT alternative $i = 3$ is not the reference alternative, add a term $+\beta_g g_n \delta_{i3}$ with β_g expected to be positive
- If the PT alternative $i = 3$ is the reference, then add a term $-\beta_g g_n \sum_{j \neq 3} \delta_{ji}$ where β_g has *exactly* the same value (why?)

? How to model an income-dependent time sensitivity?
Let the income I_n interact with the travel time by a factor $+\beta_I I_n T_{ni}$ with β_I expected to be negative

? Why it is not possible to use the age directly?
Since very young and very old persons often act more similar than middle-aged persons, age (as a socio-economic variable) needs not only be coupled to ACs or characteristics but also be formulated in terms of age-group dummies, e.g., $S_{yn} = 1$ if young, zero otherwise, $S_{on} = 1$ if old, zero otherwise (the reference is the "middle age")

Questions on socio-economic variables

? What would happen when setting $V_{ni} = ... + \beta_m S_{mn}$ for all i?
nothing

? How to model that woman prefer PT over other modes more than men?
Assume $i = 3$ is the PT and $S_{mn} = g_n = 1$ for woman and =0 for men.

- If the PT alternative $i = 3$ is not the reference alternative, add a term $+\beta_g g_n \delta_{i3}$ with β_g expected to be positive
- If the PT alternative $i = 3$ is the reference, then add a term $-\beta_g g_n \sum_{j \neq 3} \delta_{ji}$ where β_g has *exactly* the same value (why?)

? How to model an income-dependent time sensitivity?
Let the income I_n interact with the travel time by a factor $+\beta_I I_n T_{ni}$ with β_I expected to be negative

? Why it is not possible to use the age directly?
Since very young and very old persons often act more similar than middle-aged persons, age (as a socio-economic variable) needs not only be coupled to ACs or characteristics but also be formulated in terms of age-group dummies, e.g., $S_{yn} = 1$ if young, zero otherwise, $S_{on} = 1$ if old, zero otherwise (the reference is the "middle age")

Questions on socio-economic variables

? What would happen when setting $V_{ni} = ... + \beta_m S_{mn}$ for all i?
nothing

? How to model that woman prefer PT over other modes more than men?
Assume $i = 3$ is the PT and $S_{mn} = g_n = 1$ for woman and =0 for men.

- If the PT alternative $i = 3$ is not the reference alternative, add a term $+\beta_g g_n \delta_{i3}$ with β_g expected to be positive
- If the PT alternative $i = 3$ is the reference, then add a term $-\beta_g g_n \sum_{j \neq 3} \delta_{ji}$ where β_g has *exactly* the same value (why?)

? How to model an income-dependent time sensitivity?
Let the income I_n interact with the travel time by a factor $+\beta_I I_n T_{ni}$ with β_I expected to be negative

? Why it is not possible to use the age directly?
Since very young and very old persons often act more similar than middle-aged persons, age (as a socio-economic variable) needs not only be coupled to ACs or characteristics but also be formulated in terms of age-group dummies, e.g., $S_{yn} = 1$ if young, zero otherwise, $S_{on} = 1$ if old, zero otherwise (the reference is the "middle age")

Questions on socio-economic variables

? What would happen when setting $V_{ni} = ... + \beta_m S_{mn}$ for all i?
nothing

? How to model that woman prefer PT over other modes more than men?
Assume $i = 3$ is the PT and $S_{mn} = g_n = 1$ for woman and =0 for men.

- If the PT alternative $i = 3$ is not the reference alternative, add a term $+\beta_g g_n \delta_{i3}$ with β_g expected to be positive
- If the PT alternative $i = 3$ is the reference, then add a term $-\beta_g g_n \sum_{j \neq 3} \delta_{ji}$ where β_g has *exactly* the same value (why?)

? How to model an income-dependent time sensitivity?
Let the income I_n interact with the travel time by a factor $+\beta_I I_n T_{ni}$ with β_I expected to be negative

? Why it is not possible to use the age directly?
Since very young and very old persons often act more similar than middle-aged persons, age (as a socio-economic variable) needs not only be coupled to ACs or characteristics but also be formulated in terms of age-group dummies, e.g., $S_{yn} = 1$ if young, zero otherwise, $S_{on} = 1$ if old, zero otherwise (the reference is the "middle age")

Questions on socio-economic variables

? What would happen when setting $V_{ni} = ... + \beta_m S_{mn}$ for all i?
nothing

? How to model that woman prefer PT over other modes more than men?
Assume $i = 3$ is the PT and $S_{mn} = g_n = 1$ for woman and =0 for men.

- If the PT alternative $i = 3$ is not the reference alternative, add a term $+\beta_g g_n \delta_{i3}$ with β_g expected to be positive
- If the PT alternative $i = 3$ is the reference, then add a term $-\beta_g g_n \sum_{j \neq 3} \delta_{ji}$ where β_g has *exactly* the same value (why?)

? How to model an income-dependent time sensitivity?
Let the income I_n interact with the travel time by a factor $+\beta_I I_n T_{ni}$ with β_I expected to be negative

? Why it is not possible to use the age directly?
Since very young and very old persons often act more similar than middle-aged persons, age (as a socio-economic variable) needs not only be coupled to ACs or characteristics but also be formulated in terms of age-group dummies, e.g., $S_{yn} = 1$ if young, zero otherwise, $S_{on} = 1$ if old, zero otherwise (the reference is the "middle age")

Questions on socio-economic variables

? What would happen when setting $V_{ni} = ... + \beta_m S_{mn}$ for all i?

nothing

? How to model that woman prefer PT over other modes more than men?

Assume $i = 3$ is the PT and $S_{mn} = g_n = 1$ for woman and =0 for men.

- If the PT alternative $i = 3$ is not the reference alternative, add a term $+\beta_g g_n \delta_{i3}$ with β_g expected to be positive
- If the PT alternative $i = 3$ is the reference, then add a term $-\beta_g g_n \sum_{j \neq 3} \delta_{ji}$ where β_g has *exactly* the same value (why?)

? How to model an income-dependent time sensitivity?

Let the income I_n interact with the travel time by a factor $+\beta_I I_n T_{ni}$ with β_I expected to be negative

? Why it is not possible to use the age directly?

Since very young and very old persons often act more similar than middle-aged persons, age (as a socio-economic variable) needs not only be coupled to ACs or characteristics but also be formulated in terms of age-group dummies, e.g., $S_{yn} = 1$ if young, zero otherwise, $S_{on} = 1$ if old, zero otherwise (the reference is the "middle age")

Questions on socio-economic variables

? What would happen when setting $V_{ni} = ... + \beta_m S_{mn}$ for all i?

nothing

? How to model that woman prefer PT over other modes more than men?

Assume $i = 3$ is the PT and $S_{mn} = g_n = 1$ for woman and =0 for men.

- If the PT alternative $i = 3$ is not the reference alternative, add a term $+\beta_g g_n \delta_{i3}$ with β_g expected to be positive
- If the PT alternative $i = 3$ is the reference, then add a term $-\beta_g g_n \sum_{j \neq 3} \delta_{ji}$ where β_g has *exactly* the same value (why?)

? How to model an income-dependent time sensitivity?

Let the income I_n interact with the travel time by a factor $+\beta_I I_n T_{ni}$ with β_I expected to be negative

? Why it is not possible to use the age directly?

Since very young and very old persons often act more similar than middle-aged persons, age (as a socio-economic variable) needs not only be coupled to ACs or characteristics but also be formulated in terms of age-group dummies, e.g., $S_{yn} = 1$ if young, zero otherwise, $S_{on} = 1$ if old, zero otherwise (the reference is the "middle age")

Factors IV: External variables

- **External variables** influence the decisions although they neither depend on the alternatives nor the persons, e.g., the weather with the dummy $W=0$ (no rain) or $W=1$ (rain).
- The specification depends on the way they influence decisions:
  - Affecting the alternatives directly: $X_{mni}=W\delta_{mi}$,
  - affecting the *alternatives* indirectly via a characteristic m', e.g. travel time: $X_{mni}=WC_{m'ni}$,
  - affecting the *decision maker* via a socioeconomic attribute m', e.g., the gender: $X_{mni}=WS_{m'n}\delta_{mi}$,
  - combinations thereof.

Assume there are three modes bike, PT and car and define $W=0$ (not snowing) and $W=1$ (snowing). Model the effect of snow on the decision making in following situations:

? The car becomes less attractive because I need to shovel off the snow On-off disutility ⇒ coupling W directly with the AC for cars

? The bike is protected in a shed but driving becomes less attractive because the road is slippery and it's cold Disutility increases with time ⇒ let the weather dummy interact with the bike travel time

? Woman dislike biking even more than men when it's snowing Make a douple interaction gender dummy-weather dummy-bike travel time (in addition to the previous factor)

Factors IV: External variables

- **External variables** influence the decisions although they neither depend on the alternatives nor the persons, e.g., the weather with the dummy $W=0$ (no rain) or $W=1$ (rain).
- The specification depends on the way they influence decisions:
  - Affecting the alternatives directly: $X_{mni} = W\delta_{mi}$,
  - affecting the *alternatives* indirectly via a characteristic m', e.g. travel time: $X_{mni} = WC_{m'ni}$,
  - affecting the *decision maker* via a socioeconomic attribute m', e.g., the gender: $X_{mni} = WS_{m'n}\delta_{mi}$,
  - combinations thereof.

Assume there are three modes bike, PT and car and define $W=0$ (not snowing) and $W=1$ (snowing). Model the effect of snow on the decision making in following situations:

? The car becomes less attractive because I need to shovel off the snow On-off disutility ⇒ coupling W directly with the AC for cars

? The bike is protected in a shed but driving becomes less attractive because the road is slippery and it's cold Disutility increases with time ⇒ let the weather dummy interact with the bike travel time

? Woman dislike biking even more than men when it's snowing Make a double interaction gender dummy-weather dummy-bike travel time (in addition to the previous factor).

Factors IV: External variables

- **External variables** influence the decisions although they neither depend on the alternatives nor the persons, e.g., the weather with the dummy $W = 0$ (no rain) or $W = 1$ (rain).
- The specification depends on the way they influence decisions:
  - Affecting the alternatives directly: $X_{mni} = W\delta_{mi}$,
  - affecting the *alternatives* indirectly via a characteristic m', e.g. travel time: $X_{mni} = WC_{m'ni}$,
  - affecting the *decision maker* via a socioeconomic attribute m', e.g., the gender: $X_{mni} = WS_{m'n}\delta_{mi}$,
  - combinations thereof.

Assume there are three modes bike, PT and car and define $W = 0$ (not snowing) and $W = 1$ (snowing). Model the effect of snow on the decision making in following situations:

? The car becomes less attractive because I need to shovel off the snow On-off disutility $\Rightarrow$ coupling W directly with the AC for cars

? The bike is protected in a shed but driving becomes less attractive because the road is slippery and it's cold Disutility increases with time $\Rightarrow$ let the weather dummy interact with the bike travel time

? Woman dislike biking even more than men when it's snowing Make a douple interaction gender dummy-weather dummy-bike travel time (in addition to the previous factor)

Factors IV: External variables

- **External variables** influence the decisions although they neither depend on the alternatives nor the persons, e.g., the weather with the dummy $W = 0$ (no rain) or $W = 1$ (rain).
- The specification depends on the way they influence decisions:
  - Affecting the alternatives directly: $X_{mni} = W\delta_{mi}$,
  - affecting the *alternatives* indirectly via a characteristic m', e.g. travel time: $X_{mni} = WC_{m'ni}$,
  - affecting the *decision maker* via a socioeconomic attribute m', e.g., the gender: $X_{mni} = WS_{m'n}\delta_{mi}$,
  - combinations thereof.

Assume there are three modes bike, PT and car and define $W = 0$ (not snowing) and $W = 1$ (snowing). Model the effect of snow on the decision making in following situations:

? The car becomes less attractive because I need to shovel off the snow On-off disutility ⇒ coupling W directly with the AC for cars

? The bike is protected in a shed but driving becomes less attractive because the road is slippery and it's cold Disutility increases with time ⇒ let the weather dummy interact with the bike travel time

? Woman dislike biking even more than men when it's snowing Make a double interaction gender dummy-weather dummy-bike travel time (in addition to the previous factor)

Factors IV: External variables

- **External variables** influence the decisions although they neither depend on the alternatives nor the persons, e.g., the weather with the dummy $W = 0$ (no rain) or $W = 1$ (rain).
- The specification depends on the way they influence decisions:
  - Affecting the alternatives directly: $X_{mni} = W\delta_{mi}$,
  - affecting the *alternatives* indirectly via a characteristic m', e.g. travel time: $X_{mni} = WC_{m'ni}$,
  - affecting the *decision maker* via a socioeconomic attribute m', e.g., the gender: $X_{mni} = WS_{m'n}\delta_{mi}$,
  - combinations thereof.

Assume there are three modes bike, PT and car and define $W = 0$ (not snowing) and $W = 1$ (snowing). Model the effect of snow on the decision making in following situations:

? The car becomes less attractive because I need to shovel off the snow On-off disutility $\Rightarrow$ coupling W directly with the AC for cars

? The bike is protected in a shed but driving becomes less attractive because the road is slippery and it's cold Disutility increases with time $\Rightarrow$ let the weather dummy interact with the bike travel time

? Woman dislike biking even more than men when it's snowing Make a douple interaction gender dummy-weather dummy-bike travel time (in addition to the previous factor)

Factors IV: External variables

- **External variables** influence the decisions although they neither depend on the alternatives nor the persons, e.g., the weather with the dummy $W = 0$ (no rain) or $W = 1$ (rain).
- The specification depends on the way they influence decisions:
  - Affecting the alternatives directly: $X_{mni} = W\delta_{mi}$,
  - affecting the *alternatives* indirectly via a characteristic m', e.g. travel time: $X_{mni} = WC_{m'ni}$,
  - affecting the *decision maker* via a socioeconomic attribute m', e.g., the gender: $X_{mni} = WS_{m'n}\delta_{mi}$,
  - combinations thereof.

Assume there are three modes bike, PT and car and define $W = 0$ (not snowing) and $W = 1$ (snowing). Model the effect of snow on the decision making in following situations:

? The car becomes less attractive because I need to shovel off the snow On-off disutility ⇒ coupling W directly with the AC for cars

? The bike is protected in a shed but driving becomes less attractive because the road is slippery and it's cold Disutility increases with time ⇒ let the weather dummy interact with the bike travel time

? Woman dislike biking even more than men when it's snowing Make a douple interaction gender dummy-weather dummy-bike travel time (in addition to the previous factor)

Factors IV: External variables

- **External variables** influence the decisions although they neither depend on the alternatives nor the persons, e.g., the weather with the dummy $W = 0$ (no rain) or $W = 1$ (rain).
- The specification depends on the way they influence decisions:
  - Affecting the alternatives directly: $X_{mni} = W\delta_{mi}$,
  - affecting the *alternatives* indirectly via a characteristic m', e.g. travel time: $X_{mni} = WC_{m'ni}$,
  - affecting the *decision maker* via a socioeconomic attribute m', e.g., the gender: $X_{mni} = WS_{m'n}\delta_{mi}$,
  - combinations thereof.

Assume there are three modes bike, PT and car and define $W = 0$ (not snowing) and $W = 1$ (snowing). Model the effect of snow on the decision making in following situations:

? The car becomes less attractive because I need to shovel off the snow On-off disutility ⇒ coupling W directly with the AC for cars

? The bike is protected in a shed but driving becomes less attractive because the road is slippery and it's cold Disutility increases with time ⇒ let the weather dummy interact with the bike travel time

? Woman dislike biking even more than men when it's snowing Make a douple interaction gender dummy-weather dummy-bike travel time (in addition to the previous factor)

Factors IV: External variables

- **External variables** influence the decisions although they neither depend on the alternatives nor the persons, e.g., the weather with the dummy $W = 0$ (no rain) or $W = 1$ (rain).
- The specification depends on the way they influence decisions:
  - Affecting the alternatives directly: $X_{mni} = W\delta_{mi}$,
  - affecting the *alternatives* indirectly via a characteristic m', e.g. travel time: $X_{mni} = WC_{m'ni}$,
  - affecting the *decision maker* via a socioeconomic attribute m', e.g., the gender: $X_{mni} = WS_{m'n}\delta_{mi}$,
  - combinations thereof.

Assume there are three modes bike, PT and car and define $W = 0$ (not snowing) and $W = 1$ (snowing). Model the effect of snow on the decision making in following situations:

? The car becomes less attractive because I need to shovel off the snow On-off disutility ⇒ coupling W directly with the AC for cars

? The bike is protected in a shed but driving becomes less attractive because the road is slippery and it's cold Disutility increases with time ⇒ let the weather dummy interact with the bike travel time

? Woman dislike biking even more than men when it's snowing Make a douple interaction gender dummy-weather dummy-bike travel time (in addition to the previous factor)

Wrap up: modelling a certain choice situation

Given is a SP survey for mode choice with three alternatives

- $i = 1$: pedestrian mode: door-to-door travel time T_1
- $i = 2$: bicycle: door-to-door travel time T_2
- $i = 3$: motorized: door-to-door travel time T_3, ad-hoc costs C_3

Furthermore, we distinguish the gender of the deciding person ($g = 0$: male; $g = 1$: female) and the weather ($W = 0$: good; $W = 1$: bad).

? Specify a model for generic time and cost sensitivities making $i = 3$ the reference. Give the meaning and expected signs of the parameters.

? Now, formulate the travel time dependence alternative-specifically.

? Include the weather dependence assuming that the motorized mode is favoured in bad weather proportionally to the travel time.

? Now assume that women give the motorized modes a fixed "bonus" compared to men and that they also are more sensitive to the weather.

? Which parameters would change in which way (i) if $W = 1$ stands for good instead of bad weather; (ii) if $g = 1$ stands for men instead of woman; (iii) if the reference alternative is $i = 1$ instead of $i = 3$?

Wrap up: modelling a certain choice situation

Given is a SP survey for mode choice with three alternatives

- $i = 1$: pedestrian mode: door-to-door travel time T_1
- $i = 2$: bicycle: door-to-door travel time T_2
- $i = 3$: motorized: door-to-door travel time T_3, ad-hoc costs C_3

Furthermore, we distinguish the gender of the deciding person ($g = 0$: male; $g = 1$: female) and the weather ($W = 0$: good; $W = 1$: bad).

? Specify a model for generic time and cost sensitivities making $i = 3$ the reference. Give the meaning and expected signs of the parameters.

? Now, formulate the travel time dependence alternative-specifically.

? Include the weather dependence assuming that the motorized mode is favoured in bad weather proportionally to the travel time.

? Now assume that women give the motorized modes a fixed "bonus" compared to men and that they also are more sensitive to the weather.

? Which parameters would change in which way (i) if $W = 1$ stands for good instead of bad weather; (ii) if $g = 1$ stands for men instead of woman; (iii) if the reference alternative is $i = 1$ instead of $i = 3$?

Wrap up: modelling a certain choice situation

Given is a SP survey for mode choice with three alternatives

- $i = 1$: pedestrian mode: door-to-door travel time T_1
- $i = 2$: bicycle: door-to-door travel time T_2
- $i = 3$: motorized: door-to-door travel time T_3, ad-hoc costs C_3

Furthermore, we distinguish the gender of the deciding person ($g = 0$: male; $g = 1$: female) and the weather ($W = 0$: good; $W = 1$: bad).

? Specify a model for generic time and cost sensitivities making $i = 3$ the reference. Give the meaning and expected signs of the parameters.

? Now, formulate the travel time dependence alternative-specifically.

? Include the weather dependence assuming that the motorized mode is favoured in bad weather proportionally to the travel time.

? Now assume that women give the motorized modes a fixed "bonus" compared to men and that they also are more sensitive to the weather.

? Which parameters would change in which way (i) if $W = 1$ stands for good instead of bad weather; (ii) if $g = 1$ stands for men instead of woman; (iii) if the reference alternative is $i = 1$ instead of $i = 3$?

Wrap up: modelling a certain choice situation

Given is a SP survey for mode choice with three alternatives

- $i = 1$: pedestrian mode: door-to-door travel time T_1
- $i = 2$: bicycle: door-to-door travel time T_2
- $i = 3$: motorized: door-to-door travel time T_3, ad-hoc costs C_3

Furthermore, we distinguish the gender of the deciding person ($g = 0$: male; $g = 1$: female) and the weather ($W = 0$: good; $W = 1$: bad).

? Specify a model for generic time and cost sensitivities making $i = 3$ the reference. Give the meaning and expected signs of the parameters.

? Now, formulate the travel time dependence alternative-specifically.

? Include the weather dependence assuming that the motorized mode is favoured in bad weather proportionally to the travel time.

? Now assume that women give the motorized modes a fixed "bonus" compared to men and that they also are more sensitive to the weather.

? Which parameters would change in which way (i) if $W = 1$ stands for good instead of bad weather; (ii) if $g = 1$ stands for men instead of woman; (iii) if the reference alternative is $i = 1$ instead of $i = 3$?

Wrap up: modelling a certain choice situation

Given is a SP survey for mode choice with three alternatives

- $i=1$: pedestrian mode: door-to-door travel time T_1
- $i=2$: bicycle: door-to-door travel time T_2
- $i=3$: motorized: door-to-door travel time T_3, ad-hoc costs C_3

Furthermore, we distinguish the gender of the deciding person ($g=0$: male; $g=1$: female) and the weather ($W=0$: good; $W=1$: bad).

? Specify a model for generic time and cost sensitivities making $i=3$ the reference. Give the meaning and expected signs of the parameters.

? Now, formulate the travel time dependence alternative-specifically.

? Include the weather dependence assuming that the motorized mode is favoured in bad weather proportionally to the travel time.

? Now assume that women give the motorized modes a fixed "bonus" compared to men and that they also are more sensitive to the weather.

? Which parameters would change in which way (i) if $W=1$ stands for good instead of bad weather; (ii) if $g=1$ stands for men instead of woman; (iii) if the reference alternative is $i=1$ instead of $i=3$?

Example: SP In-class survey WS18/19 (red: bad weather)

Choice Set	Alt. 1: Ped	Alt. 2: Bike	Alt. 3: PT/Car	Alt 1	Alt 2	Alt 3
1	30 min	20 min	20 min+0€	1	3	7
2	30 min	20 min	20 min+2€	2	9	2
3	30 min	20 min	20 min+1€	1	5	7
4	30 min	20 min	30 min+0€	2	9	3
5	50 min	20 min	30 min+0€	0	9	4
6	50 min	30 min	30 min+0€	0	3	9
7	50 min	40 min	30 min+0€	0	2	10
8	180 min	60 min	60 min+2€	0	4	11
9	180 min	40 min	60 min+2€	0	9	6
10	180 min	40 min	60 min+2€	0	1	14
11	12 min	8 min	10 min+0€	3	5	6
12	12 min	8 min	10 min+1€	5	7	2

$$
\begin{aligned}
V_i &= \beta_0\delta_{i1} + \beta_1\delta_{i2} \\
&+ \beta_2 K_i \\
&+ \beta_{31}T_1\delta_{i1} + \beta_{32}T_2\delta_{i2} \\
&+ \beta_{33}T_3\delta_{i3} + \beta_4 W\delta_{i3}
\end{aligned}
$$

or

$$
\begin{aligned}
V_1 &= \beta_0 + \beta_2 K_1 + \beta_{31}K_1, \\
V_2 &= \beta_1 + \beta_2 K_2 + \beta_{32}K_2, \\
V_3 &= \beta_2 K_3 + \beta_{33}K_3 + \beta_4 W
\end{aligned}
$$

7.4 Random Utilities
where do randum utilities come from?

- Not all relevant characteristics $\boldsymbol{C}$ and socioeconomic variables $\boldsymbol{S}$ are included:

$$U_i = U(\underbrace{\boldsymbol{C}_i, \boldsymbol{S}}_{\text{known}}, \underbrace{\boldsymbol{C}'_i, \boldsymbol{S}'}_{\text{unknown}}) = V(\boldsymbol{C}_i, \boldsymbol{S}) + \epsilon_i^{(1)}.$$

  whatch out for neglected systematic influences leading to a bias

- Measuring/observation errors

$$U_i = U(\boldsymbol{C}_i \underbrace{+\boldsymbol{\epsilon}_i}_{\text{measuring error}}, \boldsymbol{S} \underbrace{+\boldsymbol{\epsilon}}_{\text{measuring error}}) = V(\boldsymbol{C}_i, \boldsymbol{S}) + \epsilon_i^{(2)}.$$

- Relevant variables are only indirectly observable via **instrument variables** such as the address S' of one's home for the income S:

$$U_i = U(\boldsymbol{C}_i, \boldsymbol{S}) = V(\boldsymbol{C'}_i, \boldsymbol{S}') + \epsilon_i^{(3)}.$$

- True irrationality $\epsilon_i^{(4)}$.

7.4 Random Utilities
where do randum utilities come from?

- Not all relevant characteristics $\boldsymbol{C}$ and socioeconomic variables $\boldsymbol{S}$ are included:

$$U_i = U(\underbrace{\boldsymbol{C}_i, \boldsymbol{S}}_{\text{known}}, \underbrace{\boldsymbol{C}'_i, \boldsymbol{S}'}_{\text{unknown}}) = V(\boldsymbol{C}_i, \boldsymbol{S}) + \epsilon_i^{(1)}.$$

whatch out for neglected systematic influences leading to a bias

- Measuring/observation errors

$$U_i = U(\boldsymbol{C}_i \underbrace{+\boldsymbol{\epsilon}_i}_{\text{measuring error}}, \boldsymbol{S} \underbrace{+\boldsymbol{\epsilon}}_{\text{measuring error}}) = V(\boldsymbol{C}_i, \boldsymbol{S}) + \epsilon_i^{(2)}.$$

- Relevant variables are only indirectly observable via **instrument variables** such as the address S' of one's home for the income S:

$$U_i = U(\boldsymbol{C}_i, \boldsymbol{S}) = V(\boldsymbol{C'}_i, \boldsymbol{S}') + \epsilon_i^{(3)}.$$

- True irrationality $\epsilon_i^{(4)}$.

7.4 Random Utilities
where do randum utilities come from?

- Not all relevant characteristics $\boldsymbol{C}$ and socioeconomic variables $\boldsymbol{S}$ are included:

$$U_i = U(\underbrace{\boldsymbol{C}_i, \boldsymbol{S}}_{\text{known}}, \underbrace{\boldsymbol{C}'_i, \boldsymbol{S}'}_{\text{unknown}}) = V(\boldsymbol{C}_i, \boldsymbol{S}) + \epsilon_i^{(1)}.$$

  whatch out for neglected systematic influences leading to a bias

- Measuring/observation errors

$$U_i = U(\boldsymbol{C}_i \underbrace{+\boldsymbol{\epsilon}_i}_{\text{measuring error}}, \boldsymbol{S} \underbrace{+\boldsymbol{\epsilon}}_{\text{measuring error}}) = V(\boldsymbol{C}_i, \boldsymbol{S}) + \epsilon_i^{(2)}.$$

- Relevant variables are only indirectly observable via **instrument variables** such as the address S' of one's home for the income S:

$$U_i = U(\boldsymbol{C}_i, \boldsymbol{S}) = V(\boldsymbol{C'}_i, \boldsymbol{S}') + \epsilon_i^{(3)}.$$

- True irrationality $\epsilon_i^{(4)}$.

7.4 Random Utilities
where do randum utilities come from?

- Not all relevant characteristics $\boldsymbol{C}$ and socioeconomic variables $\boldsymbol{S}$ are included:

$$U_i = U(\underbrace{\boldsymbol{C}_i, \boldsymbol{S}}_{\text{known}}, \underbrace{\boldsymbol{C}'_i, \boldsymbol{S}'}_{\text{unknown}}) = V(\boldsymbol{C}_i, \boldsymbol{S}) + \epsilon_i^{(1)}.$$

whatch out for neglected systematic influences leading to a bias

- Measuring/observation errors

$$U_i = U(\boldsymbol{C}_i \underbrace{+\boldsymbol{\epsilon}_i}_{\text{measuring error}}, \boldsymbol{S} \underbrace{+\boldsymbol{\epsilon}}_{\text{measuring error}}) = V(\boldsymbol{C}_i, \boldsymbol{S}) + \epsilon_i^{(2)}.$$

- Relevant variables are only indirectly observable via **instrument variables** such as the address S' of one's home for the income S:

$$U_i = U(\boldsymbol{C}_i, \boldsymbol{S}) = V(\boldsymbol{C'}_i, \boldsymbol{S}') + \epsilon_i^{(3)}.$$

- True irrationality $\epsilon_i^{(4)}$.

7.5 Choice probabilities

- The three basic components of discrete-choice models are **derterminstic (explained) utilities** V_i, **random utilities (RUs)** ϵ_i, and a decision rule based on the **Homo Oeconomicus**:

$$i_{\text{selected}} = \text{arg}\max_i U_i = \text{arg}\max_i (V_i + \epsilon_i)$$

- By virtue of the RUs, this microscopically fixed decision rule leads to **choice probabilities** P_i when aggregating over many decisions.

7.5 Choice probabilities

- The three basic components of discrete-choice models are **determinstic (explained) utilities** V_i, **random utilities (RUs)** ϵ_i, and a decision rule based on the **Homo Oeconomicus**:

$$i_{\text{selected}} = \text{arg}\max_i U_i = \text{arg}\max_i (V_i + \epsilon_i)$$

- By virtue of the RUs, this microscopically fixed decision rule leads to **choice probabilities** P_i when aggregating over many decisions.

7.5 Choice probabilities

- The three basic components of discrete-choice models are **determinstic (explained) utilities** V_i, **random utilities (RUs)** ϵ_i, and a decision rule based on the **Homo Oeconomicus**:

$$i_{\text{selected}} = \text{arg}\max_i U_i = \text{arg}\max_i (V_i + \epsilon_i)$$

- By virtue of the RUs, this microscopically fixed decision rule leads to **choice probabilities** P_i when aggregating over many decisions.

Binomial case: $P_1 = \text{Prob}(U_1 \geq U_2)$

7.5 Choice probabilities

- The three basic components of discrete-choice models are **determinstic (explained) utilities** V_i, **random utilities (RUs)** ϵ_i, and a decision rule based on the **Homo Oeconomicus**:

$$i_{\text{selected}} = \text{arg}\max_i U_i = \text{arg}\max_i (V_i + \epsilon_i)$$

- By virtue of the RUs, this microscopically fixed decision rule leads to **choice probabilities** P_i when aggregating over many decisions.

Binomial case:

$$\begin{aligned} P_1 &= \text{Prob}(U_1 \geq U_2) \\ &= \text{Prob}(V_1 + \epsilon_1 \geq V_2 + \epsilon_2) \end{aligned}$$

7.5 Choice probabilities

- The three basic components of discrete-choice models are **determinstic (explained) utilities** V_i, **random utilities (RUs)** ϵ_i, and a decision rule based on the **Homo Oeconomicus**:

$$i_{\text{selected}} = \arg\max_i U_i = \arg\max_i (V_i + \epsilon_i)$$

- By virtue of the RUs, this microscopically fixed decision rule leads to **choice probabilities** P_i when aggregating over many decisions.

Binomial case:

$$\begin{aligned} P_1 &= \text{Prob}(U_1 \ge U_2) \\ &= \text{Prob}(V_1 + \epsilon_1 \ge V_2 + \epsilon_2) \\ &= \text{Prob}(\epsilon_2 - \epsilon_1 \le V_1 - V_2) \end{aligned}$$

7.5 Choice probabilities

- The three basic components of discrete-choice models are **determinstic (explained) utilities** V_i, **random utilities (RUs)** ϵ_i, and a decision rule based on the **Homo Oeconomicus**:

$$i_{\text{selected}} = \text{arg}\max_i U_i = \text{arg}\max_i (V_i + \epsilon_i)$$

- By virtue of the RUs, this microscopically fixed decision rule leads to **choice probabilities** P_i when aggregating over many decisions.

Binomial case:

$$\begin{aligned} P_1 &= \text{Prob}(U_1 \geq U_2) \\ &= \text{Prob}(V_1 + \epsilon_1 \geq V_2 + \epsilon_2) \\ &= \text{Prob}(\epsilon_2 - \epsilon_1 \leq V_1 - V_2) \\ &= F_{\epsilon_2 - \epsilon_1}(V_1 - V_2) \end{aligned}$$

Questions on choice probabilities

? What does this general result $P_1 = F_{\epsilon_2-\epsilon_1}(V_1 - V_2)$ mean in words?

The probability of chosing alternative 1 ist given by the distribution function of the RU difference at the deterministic utility difference

? Is the result plausible? Check the change if $V_1 - V_2$ changes and the limits of $V_1 - V_2 \to \pm\infty$

- The probability depends only on the deterministic utility difference $V_1 - V_2$ OK
- The sensitivity $\partial P_1/\partial \Delta V = f_{\epsilon_2-\epsilon_1}(\Delta V)$ is given by the probability density of the RU diference, i.e. nonnegative OK
- For $\Delta V \to -\infty$, we obtain $P_1 \to 0$ because of the s-shaped form of the distribution function OK
- For the same reason, $P_1 \to 1$ for $\Delta V \to +\infty$ OK

? Derive explicit formulas for $\epsilon_i \sim$ i.i.d $U(-1,1)$ and $\epsilon_i \sim$ i.i.d $N(0,1)$

The probability density of a sum of independent random variables is equal to the convolution of the individual densities. For symmetric densities this also applies for the difference.

- The convolution of two i.i.d. uniform densities is the triangular density $f^{(1)}_{\Delta\epsilon} = 1/2(1 - 1/2|\Delta\epsilon|)$ and P_1 is the integral of it at $V_1 - V_2$
- The convolution of two i.i.d $N(0,1)$ distributed RUs is $\sim N(0,2)$, hence $P_1(V_1 - V_2) = \Phi((V_1 - V_2)/\sqrt{2})$ with $\Phi(.)$ the standardnormal distribution

? What would happen for 100 % correlated RUs? The RU difference $\Delta\epsilon = 0$. Hence, we have a deterministic situation $P_1 = 1$ if $V_1 - V_2 \geq 0$ and $P_1 = 0$, otherwise.

Questions on choice probabilities

? What does this general result $P_1 = F_{\epsilon_2-\epsilon_1}(V_1 - V_2)$ mean in words?

The probability of chosing alternative 1 ist given by the distribution function of the RU difference at the deterministic utility difference

? Is the result plausible? Check the change if $V_1 - V_2$ changes and the limits of $V_1 - V_2 \to \pm\infty$

- The probability depends only on the deterministic utility difference $V_1 - V_2$ OK
- The sensitivity $\partial P_1/\partial \Delta V = f_{\epsilon_2-\epsilon_1}(\Delta V)$ is given by the probability density of the RU diference, i.e. nonnegative OK
- For $\Delta V \to -\infty$, we obtain $P_1 \to 0$ because of the s-shaped form of the distribution function OK
- For the same reason, $P_1 \to 1$ for $\Delta V \to +\infty$ OK

? Derive explicit formulas for $\epsilon_i \sim$ i.i.d $U(-1,1)$ and $\epsilon_i \sim$ i.i.d $N(0,1)$

The probability density of a sum of independent random variables is equal to the convolution of the individual densities. For symmetric densities this also applies for the difference.

- The convolution of two i.i.d. uniform densities is the triangular density $f^{(1)}_{\Delta\epsilon} = 1/2(1 - 1/2|\Delta\epsilon|)$ and P_1 is the integral of it at $V_1 - V_2$
- The convolution of two i.i.d $N(0,1)$ distributed RUs is $\sim N(0,2)$, hence $P_1(V_1 - V_2) = \Phi((V_1 - V_2)/\sqrt{2})$ with $\Phi(.)$ the standardnormal distribution

? What would happen for 100 % correlated RUs? The RU difference $\Delta\epsilon = 0$. Hence, we have a deterministic situation $P_1 = 1$ if $V_1 - V_2 \geq 0$ and $P_1 = 0$, otherwise.

Questions on choice probabilities

? What does this general result $P_1 = F_{\epsilon_2-\epsilon_1}(V_1 - V_2)$ mean in words?

The probability of chosing alternative 1 ist given by the distribution function of the RU difference at the deterministic utility difference

? Is the result plausible? Check the change if $V_1 - V_2$ changes and the limits of $V_1 - V_2 \to \pm\infty$

- The probability depends only on the deterministic utility difference $V_1 - V_2$ OK
- The sensitivity $\partial P_1/\partial \Delta V = f_{\epsilon_2-\epsilon_1}(\Delta V)$ is given by the probability density of the RU diference, i.e. nonnegative OK
- For $\Delta V \to -\infty$, we obtain $P_1 \to 0$ because of the s-shaped form of the distribution function OK
- For the same reason, $P_1 \to 1$ for $\Delta V \to +\infty$ OK

? Derive explicit formulas for $\epsilon_i \sim$ i.i.d $U(-1, 1)$ and $\epsilon_i \sim$ i.i.d $N(0, 1)$

The probability density of a sum of independent random variables is equal to the convolution of the individual densities. For symmetric densities this also applies for the difference.

- The convolution of two i.i.d. uniform densities is the triangular density $f_{\Delta\epsilon}^{(1)} = 1/2(1 - 1/2|\Delta\epsilon|)$ and P_1 is the integral of it at $V_1 - V_2$
- The convolution of two i.i.d $N(0, 1)$ distributed RUs is $\sim N(0, 2)$, hence $P_1(V_1 - V_2) = \Phi((V_1 - V_2)/\sqrt{2})$ with $\Phi(.)$ the standardnormal distribution

? What would happen for 100 % correlated RUs? The RU difference $\Delta\epsilon = 0$. Hence, we have a deterministic situation $P_1 = 1$ if $V_1 - V_2 \geq 0$ and $P_1 = 0$, otherwise.

Questions on choice probabilities

? What does this general result $P_1 = F_{\epsilon_2-\epsilon_1}(V_1 - V_2)$ mean in words?

The probability of chosing alternative 1 ist given by the distribution function of the RU difference at the deterministic utility difference

? Is the result plausible? Check the change if $V_1 - V_2$ changes and the limits of $V_1 - V_2 \to \pm\infty$

- The probability depends only on the deterministic utility difference $V_1 - V_2$ OK
- The sensitivity $\partial P_1/\partial \Delta V = f_{\epsilon_2-\epsilon_1}(\Delta V)$ is given by the probability density of the RU diference, i.e. nonnegative OK
- For $\Delta V \to -\infty$, we obtain $P_1 \to 0$ because of the s-shaped form of the distribution function OK
- For the same reason, $P_1 \to 1$ for $\Delta V \to +\infty$ OK

? Derive explicit formulas for $\epsilon_i \sim$ i.i.d $U(-1,1)$ and $\epsilon_i \sim$ i.i.d $N(0,1)$

The probability density of a sum of independent random variables is equal to the convolution of the individual densities. For symmetric densities this also applies for the difference.

- The convolution of two i.i.d. uniform densities is the triangular density $f^{(1)}_{\Delta\epsilon} = 1/2(1 - 1/2|\Delta\epsilon|)$ and P_1 is the integral of it at $V_1 - V_2$
- The convolution of two i.i.d $N(0,1)$ distributed RUs is $\sim N(0,2)$, hence $P_1(V_1 - V_2) = \Phi((V_1 - V_2)/\sqrt{2})$ with $\Phi(.)$ the standardnormal distribution

? What would happen for 100 % correlated RUs? The RU difference $\Delta\epsilon = 0$. Hence, we have a deterministic situation $P_1 = 1$ if $V_1 - V_2 \geq 0$ and $P_1 = 0$, otherwise.

Questions on choice probabilities

? What does this general result $P_1 = F_{\epsilon_2-\epsilon_1}(V_1 - V_2)$ mean in words?

The probability of chosing alternative 1 ist given by the distribution function of the RU difference at the deterministic utility difference

? Is the result plausible? Check the change if $V_1 - V_2$ changes and the limits of $V_1 - V_2 \to \pm\infty$

- The probability depends only on the deterministic utility difference $V_1 - V_2$ OK
- The sensitivity $\partial P_1/\partial \Delta V = f_{\epsilon_2-\epsilon_1}(\Delta V)$ is given by the probability density of the RU diference, i.e. nonnegative OK
- For $\Delta V \to -\infty$, we obtain $P_1 \to 0$ because of the s-shaped form of the distribution function OK
- For the same reason, $P_1 \to 1$ for $\Delta V \to +\infty$ OK

? Derive explicit formulas for $\epsilon_i \sim$ i.i.d $U(-1,1)$ and $\epsilon_i \sim$ i.i.d $N(0,1)$

The probability density of a sum of independent random variables is equal to the convolution of the individual densities. For symmetric densities this also applies for the difference.

- The convolution of two i.i.d. uniform densities is the triangular density $f^{(1)}_{\Delta\epsilon} = 1/2(1 - 1/2|\Delta\epsilon|)$ and P_1 is the integral of it at $V_1 - V_2$
- The convolution of two i.i.d $N(0,1)$ distributed RUs is $\sim N(0,2)$, hence $P_1(V_1 - V_2) = \Phi((V_1 - V_2)/\sqrt{2})$ with $\Phi(.)$ the standardnormal distribution

? What would happen for 100 % correlated RUs? The RU difference $\Delta\epsilon = 0$. Hence, we have a deterministic situation $P_1 = 1$ if $V_1 - V_2 \geq 0$ and $P_1 = 0$, otherwise.

Questions on choice probabilities

? What does this general result $P_1 = F_{\epsilon_2 - \epsilon_1}(V_1 - V_2)$ mean in words?

The probability of chosing alternative 1 ist given by the distribution function of the RU difference at the deterministic utility difference

? Is the result plausible? Check the change if $V_1 - V_2$ changes and the limits of $V_1 - V_2 \to \pm\infty$

- The probability depends only on the deterministic utility difference $V_1 - V_2$ OK
- The sensitivity $\partial P_1 / \partial \Delta V = f_{\epsilon_2 - \epsilon_1}(\Delta V)$ is given by the probability density of the RU diference, i.e. nonnegative OK
- For $\Delta V \to -\infty$, we obtain $P_1 \to 0$ because of the s-shaped form of the distribution function OK
- For the same reason, $P_1 \to 1$ for $\Delta V \to +\infty$ OK

? Derive explicit formulas for $\epsilon_i \sim$ i.i.d $U(-1,1)$ and $\epsilon_i \sim$ i.i.d $N(0,1)$

The probability density of a sum of independent random variables is equal to the convolution of the individual densities. For symmetric densities this also applies for the difference.

- The convolution of two i.i.d. uniform densities is the triangular density $f_{\Delta\epsilon}^{(1)} = 1/2(1 - 1/2|\Delta\epsilon|)$ and P_1 is the integral of it at $V_1 - V_2$
- The convolution of two i.i.d $N(0,1)$ distributed RUs is $\sim N(0,2)$, hence $P_1(V_1 - V_2) = \Phi((V_1 - V_2)/\sqrt{2})$ with $\Phi(.)$ the standardnormal distribution

? What would happen for 100 % correlated RUs? The RU difference $\Delta\epsilon = 0$. Hence, we have a deterministic situation $P_1 = 1$ if $V_1 - V_2 \geq 0$ and $P_1 = 0$, otherwise.

Questions on choice probabilities

? What does this general result $P_1 = F_{\epsilon_2 - \epsilon_1}(V_1 - V_2)$ mean in words?

The probability of chosing alternative 1 ist given by the distribution function of the RU difference at the deterministic utility difference

? Is the result plausible? Check the change if $V_1 - V_2$ changes and the limits of $V_1 - V_2 \to \pm\infty$

- The probability depends only on the deterministic utility difference $V_1 - V_2$ OK
- The sensitivity $\partial P_1 / \partial \Delta V = f_{\epsilon_2 - \epsilon_1}(\Delta V)$ is given by the probability density of the RU diference, i.e. nonnegative OK
- For $\Delta V \to -\infty$, we obtain $P_1 \to 0$ because of the s-shaped form of the distribution function OK
- For the same reason, $P_1 \to 1$ for $\Delta V \to +\infty$ OK

? Derive explicit formulas for $\epsilon_i \sim$ i.i.d $U(-1, 1)$ and $\epsilon_i \sim$ i.i.d $N(0, 1)$

The probability density of a sum of independent random variables is equal to the convolution of the individual densities. For symmetric densities this also applies for the difference.

- The convolution of two i.i.d. uniform densities is the triangular density $f_{\Delta\epsilon}^{(1)} = 1/2(1 - 1/2|\Delta\epsilon|)$ and P_1 is the integral of it at $V_1 - V_2$
- The convolution of two i.i.d $N(0, 1)$ distributed RUs is $\sim N(0, 2)$, hence $P_1(V_1 - V_2) = \Phi((V_1 - V_2)/\sqrt{2})$ with $\Phi(.)$ the standardnormal distribution

? What would happen for 100 % correlated RUs? The RU difference $\Delta\epsilon = 0$. Hence, we have a deterministic situation $P_1 = 1$ if $V_1 - V_2 \geq 0$ and $P_1 = 0$, otherwise.

Questions on choice probabilities

? What does this general result $P_1 = F_{\epsilon_2 - \epsilon_1}(V_1 - V_2)$ mean in words?

The probability of chosing alternative 1 ist given by the distribution function of the RU difference at the deterministic utility difference

? Is the result plausible? Check the change if $V_1 - V_2$ changes and the limits of $V_1 - V_2 \to \pm\infty$

- The probability depends only on the deterministic utility difference $V_1 - V_2$ OK
- The sensitivity $\partial P_1 / \partial \Delta V = f_{\epsilon_2 - \epsilon_1}(\Delta V)$ is given by the probability density of the RU diference, i.e. nonnegative OK
- For $\Delta V \to -\infty$, we obtain $P_1 \to 0$ because of the s-shaped form of the distribution function OK
- For the same reason, $P_1 \to 1$ for $\Delta V \to +\infty$ OK

? Derive explicit formulas for $\epsilon_i \sim$ i.i.d $U(-1, 1)$ and $\epsilon_i \sim$ i.i.d $N(0, 1)$

The probability density of a sum of independent random variables is equal to the convolution of the individual densities. For symmetric densities this also applies for the difference.

- The convolution of two i.i.d. uniform densities is the triangular density $f^{(1)}_{\Delta\epsilon} = 1/2(1 - 1/2|\Delta\epsilon|)$ and P_1 is the integral of it at $V_1 - V_2$
- The convolution of two i.i.d $N(0, 1)$ distributed RUs is $\sim N(0, 2)$, hence $P_1(V_1 - V_2) = \Phi((V_1 - V_2)/\sqrt{2})$ with $\Phi(.)$ the standardnormal distribution

? What would happen for 100 % correlated RUs? The RU difference $\Delta\epsilon = 0$. Hence, we have a deterministic situation $P_1 = 1$ if $V_1 - V_2 \geq 0$ and $P_1 = 0$, otherwise.

Questions on choice probabilities

? What does this general result $P_1 = F_{\epsilon_2 - \epsilon_1}(V_1 - V_2)$ mean in words?

The probability of chosing alternative 1 ist given by the distribution function of the RU difference at the deterministic utility difference

? Is the result plausible? Check the change if $V_1 - V_2$ changes and the limits of $V_1 - V_2 \to \pm\infty$

- The probability depends only on the deterministic utility difference $V_1 - V_2$ OK
- The sensitivity $\partial P_1 / \partial \Delta V = f_{\epsilon_2 - \epsilon_1}(\Delta V)$ is given by the probability density of the RU diference, i.e. nonnegative OK
- For $\Delta V \to -\infty$, we obtain $P_1 \to 0$ because of the s-shaped form of the distribution function OK
- For the same reason, $P_1 \to 1$ for $\Delta V \to +\infty$ OK

? Derive explicit formulas for $\epsilon_i \sim$ i.i.d $U(-1, 1)$ and $\epsilon_i \sim$ i.i.d $N(0, 1)$

The probability density of a sum of independent random variables is equal to the convolution of the individual densities. For symmetric densities this also applies for the difference.

- The convolution of two i.i.d. uniform densities is the triangular density $f_{\Delta\epsilon}^{(1)} = 1/2(1 - 1/2|\Delta\epsilon|)$ and P_1 is the integral of it at $V_1 - V_2$
- The convolution of two i.i.d $N(0, 1)$ distributed RUs is $\sim N(0, 2)$, hence $P_1(V_1 - V_2) = \Phi((V_1 - V_2)/\sqrt{2})$ with $\Phi(.)$ the standardnormal distribution

? What would happen for 100 % correlated RUs? The RU difference $\Delta\epsilon = 0$. Hence, we have a deterministic situation $P_1 = 1$ if $V_1 - V_2 \geq 0$ and $P_1 = 0$, otherwise.

Questions on choice probabilities

? What does this general result $P_1 = F_{\epsilon_2-\epsilon_1}(V_1 - V_2)$ mean in words?

The probability of chosing alternative 1 ist given by the distribution function of the RU difference at the deterministic utility difference

? Is the result plausible? Check the change if $V_1 - V_2$ changes and the limits of $V_1 - V_2 \to \pm\infty$

- The probability depends only on the deterministic utility difference $V_1 - V_2$ OK
- The sensitivity $\partial P_1/\partial \Delta V = f_{\epsilon_2-\epsilon_1}(\Delta V)$ is given by the probability density of the RU diference, i.e. nonnegative OK
- For $\Delta V \to -\infty$, we obtain $P_1 \to 0$ because of the s-shaped form of the distribution function OK
- For the same reason, $P_1 \to 1$ for $\Delta V \to +\infty$ OK

? Derive explicit formulas for $\epsilon_i \sim$ i.i.d $U(-1, 1)$ and $\epsilon_i \sim$ i.i.d $N(0, 1)$

The probability density of a sum of independent random variables is equal to the convolution of the individual densities. For symmetric densities this also applies for the difference.

- The convolution of two i.i.d. uniform densities is the triangular density $f^{(1)}_{\Delta\epsilon} = 1/2(1 - 1/2|\Delta\epsilon|)$ and P_1 is the integral of it at $V_1 - V_2$
- The convolution of two i.i.d $N(0, 1)$ distributed RUs is $\sim N(0, 2)$, hence $P_1(V_1 - V_2) = \Phi((V_1 - V_2)/\sqrt{2})$ with $\Phi(.)$ the standardnormal distribution

? What would happen for 100 % correlated RUs? The RU difference $\Delta\epsilon = 0$. Hence, we have a deterministic situation $P_1 = 1$ if $V_1 - V_2 \geq 0$ and $P_1 = 0$, otherwise.

Translation and scale invariance

The general choice formula $i_{\text{selected}} = \arg\max_i (V_i + \epsilon_i)$ leads to two far-reaching consequences:

- **Translation invariance**:

$$V_1 + \epsilon_1 > V_2 + \epsilon_2 \Longleftrightarrow V_1 + \epsilon_1 + c > V_2 + \epsilon_2 + c, \quad c \in I\!R$$

  - Only differences matter. In particular, one can set one deterministic utility $V_{i'} = 0$ making i' the reference alternative.
  - Nonzero expectation values $E(\epsilon_i)$ (as in Logit models) are irrelevant.

- **Scale invariance**:

$$V_1 + \epsilon_1 > V_2 + \epsilon_2 \Longleftrightarrow \lambda(V_1 + \epsilon_1) > \lambda(V_2 + \epsilon_2), \quad \lambda > 0$$

  - A *simultaneous* scaling of all V_i *and* ϵ leaves the choice unchanged (one could even apply a strictly monotonously increasing function defined in $I\!R$ such as exp(.) to all $V_i + \epsilon_i$)
  - One can normalize ϵ_i without loss of generality. In any case, V_i is given in multiples of the standard deviation of ϵ_i

Translation and scale invariance

The general choice formula $i_{\text{selected}} = \arg\max_i(V_i + \epsilon_i)$ leads to two far-reaching consequences:

- **Translation invariance**:

$$V_1 + \epsilon_1 > V_2 + \epsilon_2 \Longleftrightarrow V_1 + \epsilon_1 + c > V_2 + \epsilon_2 + c, \quad c \in \mathbb{R}$$

  - Only differences matter. In particular, one can set one deterministic utility $V_{i'} = 0$ making i' the reference alternative.
  - Nonzero expectation values $E(\epsilon_i)$ (as in Logit models) are irrelevant.

- **Scale invariance**:

$$V_1 + \epsilon_1 > V_2 + \epsilon_2 \Longleftrightarrow \lambda(V_1 + \epsilon_1) > \lambda(V_2 + \epsilon_2), \quad \lambda > 0$$

  - A *simultaneous* scaling of all V_i *and* ϵ leaves the choice unchanged (one could even apply a strictly monotonously increasing function defined in $\mathbb{R}$ such as $\exp(.)$ to all $V_i + \epsilon_i$)
  - One can normalize ϵ_i without loss of generality. In any case, V_i is given in multiples of the standard deviation of ϵ_i

Translation and scale invariance

The general choice formula $i_{\text{selected}} = \text{arg}\max_i(V_i + \epsilon_i)$ leads to two far-reaching consequences:

- **Translation invariance**:

$$V_1 + \epsilon_1 > V_2 + \epsilon_2 \iff V_1 + \epsilon_1 + c > V_2 + \epsilon_2 + c, \quad c \in I\!R$$

  - Only differences matter. In particular, one can set one deterministic utility $V_{i'} = 0$ making i' the reference alternative.
  - Nonzero expectation values $E(\epsilon_i)$ (as in Logit models) are irrelevant.

- **Scale invariance**:

$$V_1 + \epsilon_1 > V_2 + \epsilon_2 \iff \lambda(V_1 + \epsilon_1) > \lambda(V_2 + \epsilon_2), \quad \lambda > 0$$

  - A *simultaneous* scaling of all V_i *and* ϵ leaves the choice unchanged (one could even apply a strictly monotonously increasing function defined in $I\!R$ such as $\exp(.)$ to all $V_i + \epsilon_i$)
  - One can normalize ϵ_i without loss of generality. In any case, V_i is given in multiples of the standard deviation of ϵ_i

Translation and scale invariance

The general choice formula $i_{\text{selected}} = \arg\max_i(V_i + \epsilon_i)$ leads to two far-reaching consequences:

- **Translation invariance**:

$$V_1 + \epsilon_1 > V_2 + \epsilon_2 \iff V_1 + \epsilon_1 + c > V_2 + \epsilon_2 + c, \quad c \in \mathbb{R}$$

  - Only differences matter. In particular, one can set one deterministic utility $V_{i'} = 0$ making i' the reference alternative.
  - Nonzero expectation values $E(\epsilon_i)$ (as in Logit models) are irrelevant.

- **Scale invariance**:

$$V_1 + \epsilon_1 > V_2 + \epsilon_2 \iff \lambda(V_1 + \epsilon_1) > \lambda(V_2 + \epsilon_2), \quad \lambda > 0$$

  - A *simultaneous* scaling of all V_i *and* ϵ leaves the choice unchanged (one could even apply a strictly monotonously increasing function defined in $\mathbb{R}$ such as $\exp(.)$ to all $V_i + \epsilon_i$)
  - One can normalize ϵ_i without loss of generality. In any case, V_i is given in multiples of the standard deviation of ϵ_i

Translation and scale invariance

The general choice formula $i_{\text{selected}} = \arg\max_i (V_i + \epsilon_i)$ leads to two far-reaching consequences:

- **Translation invariance**:

$$V_1 + \epsilon_1 > V_2 + \epsilon_2 \Longleftrightarrow V_1 + \epsilon_1 + c > V_2 + \epsilon_2 + c, \quad c \in \mathbb{R}$$

  - Only differences matter. In particular, one can set one deterministic utility $V_{i'} = 0$ making i' the reference alternative.
  - Nonzero expectation values $E(\epsilon_i)$ (as in Logit models) are irrelevant.

- **Scale invariance**:

$$V_1 + \epsilon_1 > V_2 + \epsilon_2 \Longleftrightarrow \lambda(V_1 + \epsilon_1) > \lambda(V_2 + \epsilon_2), \quad \lambda > 0$$

  - A *simultaneous* scaling of all V_i *and* ϵ leaves the choice unchanged (one could even apply a strictly monotonously increasing function defined in $\mathbb{R}$ such as $\exp(.)$ to all $V_i + \epsilon_i$)
  - One can normalize ϵ_i without loss of generality. In any case, V_i is given in multiples of the standard deviation of ϵ_i

Translation and scale invariance

The general choice formula $i_{\text{selected}} = \text{arg}\max_i(V_i + \epsilon_i)$ leads to two far-reaching consequences:

- **Translation invariance**:

$$V_1 + \epsilon_1 > V_2 + \epsilon_2 \Longleftrightarrow V_1 + \epsilon_1 + c > V_2 + \epsilon_2 + c, \quad c \in I\!R$$

  - Only differences matter. In particular, one can set one deterministic utility $V_{i'} = 0$ making i' the reference alternative.
  - Nonzero expectation values $E(\epsilon_i)$ (as in Logit models) are irrelevant.

- **Scale invariance**:

$$V_1 + \epsilon_1 > V_2 + \epsilon_2 \Longleftrightarrow \lambda(V_1 + \epsilon_1) > \lambda(V_2 + \epsilon_2), \quad \lambda > 0$$

  - A *simultaneous* scaling of all V_i *and* ϵ leaves the choice unchanged (one could even apply a strictly monotonously increasing function defined in $I\!R$ such as $\exp(.)$ to all $V_i + \epsilon_i$)
  - One can normalize ϵ_i without loss of generality. In any case, V_i is given in multiples of the standard deviation of ϵ_i