# CIE4801 Transportation and spatial modelling
## Trip generation

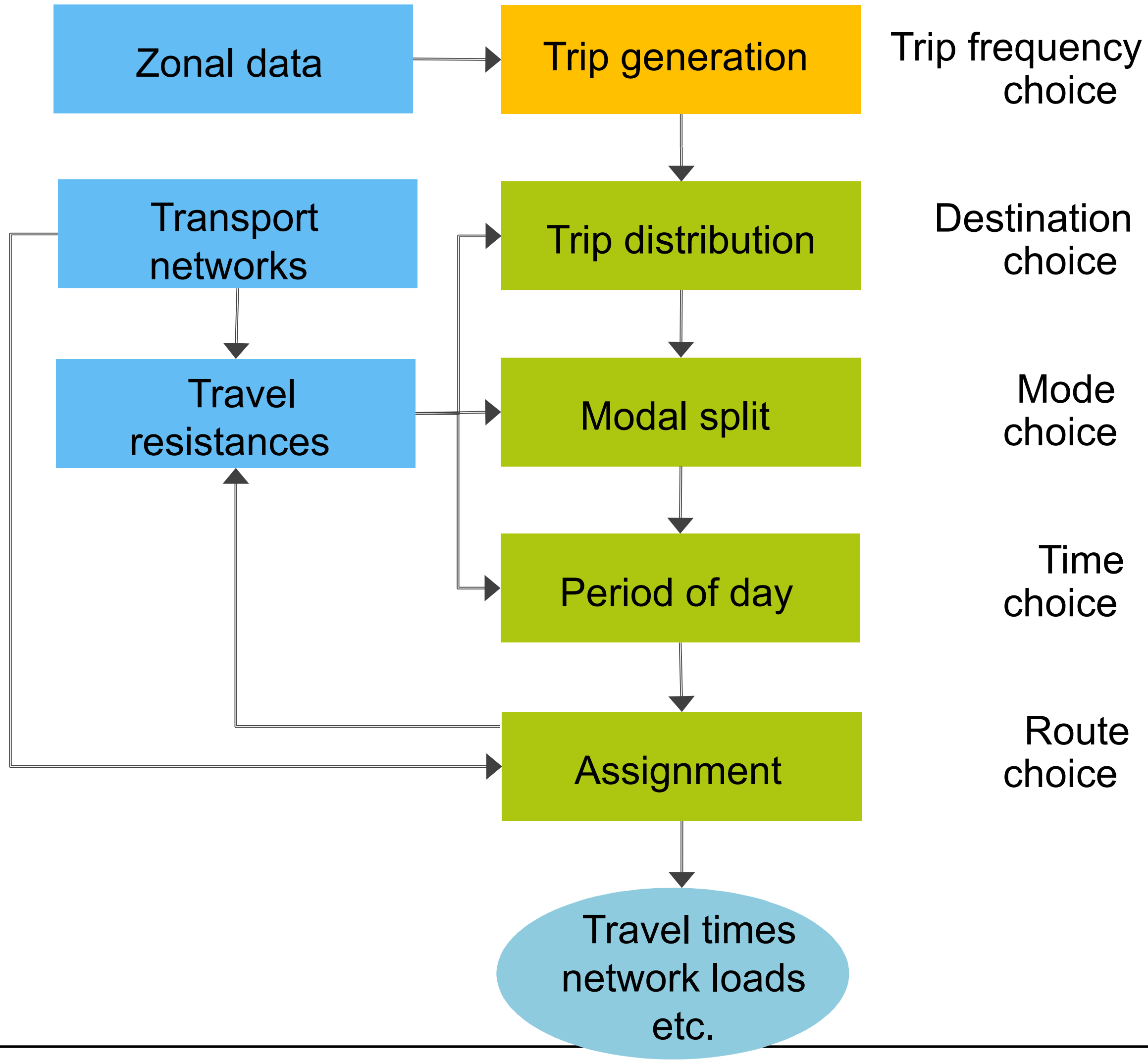
Rob van Nes, Transport & Planning

31-08-18

# 2.

*Trip generation*

# Trip generation

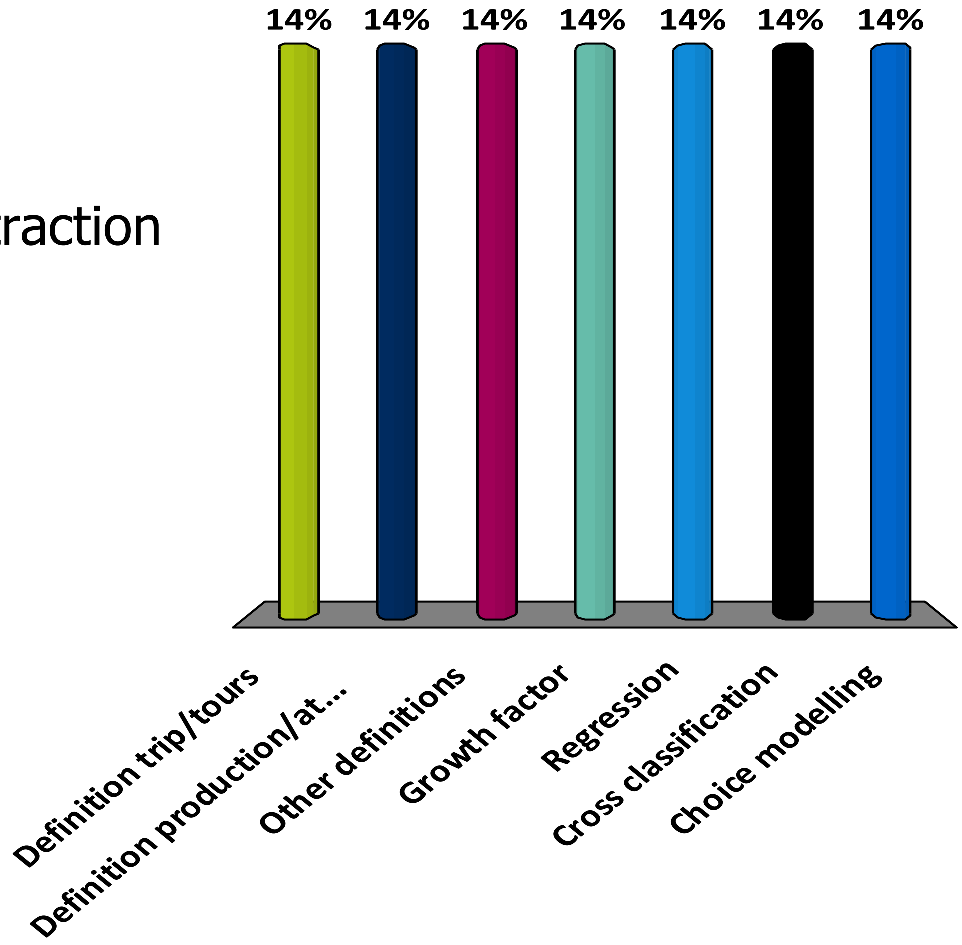


# Topics to study: Trip generation

- What does this modelling component do? What's its output and what's its input? How does it fit in the framework?
- Do you understand the definitions?
    - Trips, tours
    - Generation, production, attraction, departures, arrivals
    - Home-based, non-home-based, intra-zonal, inter-zonal
- Do you understand the modelling methods?
    - Growth factor
    - Regression analysis
        - Zone level (total or means) and household level
        - Dummies or segmentation
    - Cross classification
    - Choice modelling
- Are these models appropriate?

# Which topics would need discussion in class?

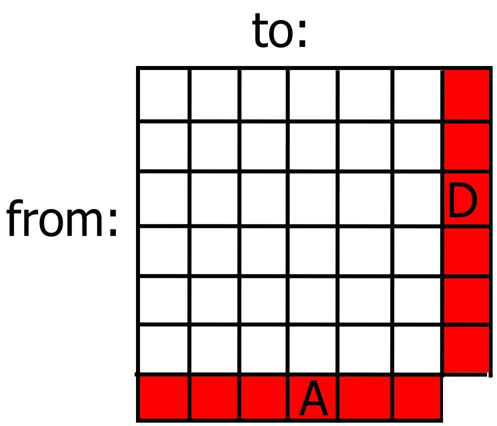
A. Definition trip/tours
B. Definition production/attraction
C. Other definitions
D. Growth factor
E. Regression
F. Cross classification
G. Choice modelling

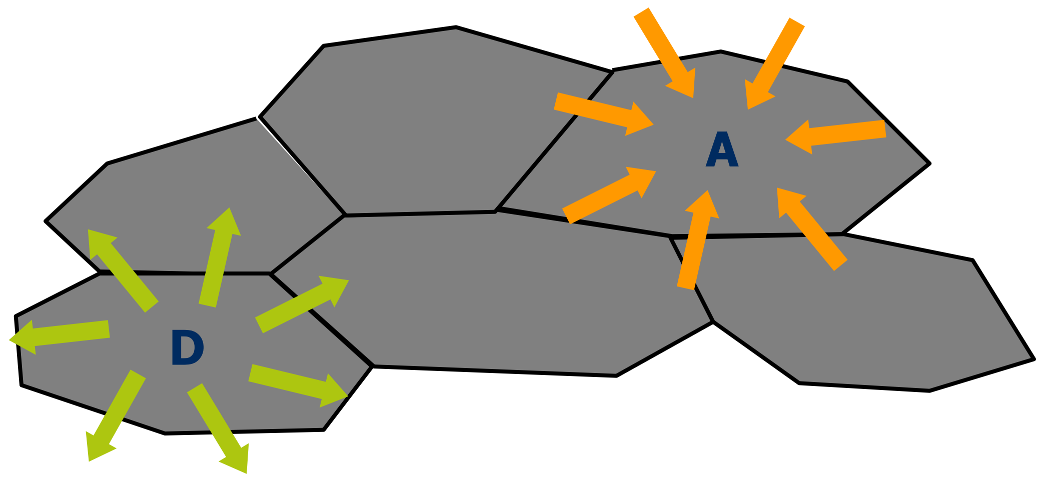


# 2.1

*Trip generation*
*Definitions*

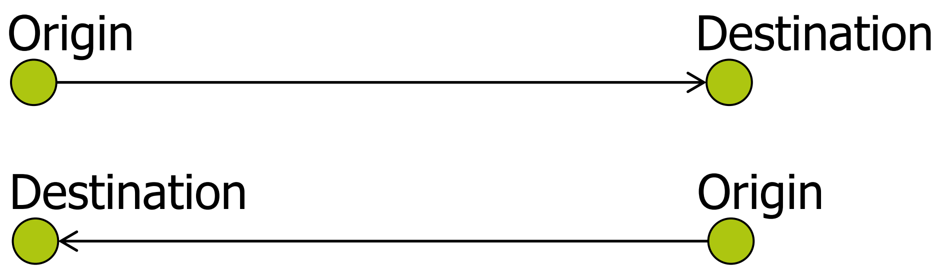
# Main idea of trip generation



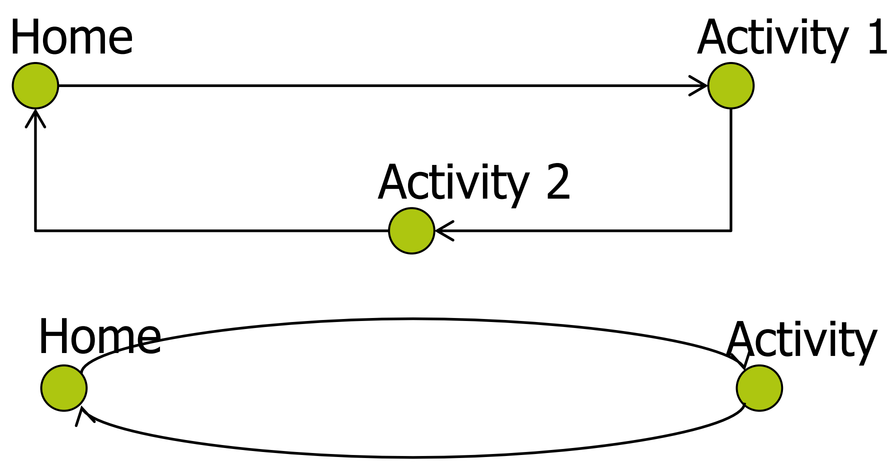



# What is generated: trips or tours?

- Trip




- Tour



# Some statistics on tours (Netherlands)

- Percentages for number of trips per tour:
  - 1 trip 10%
  - 2 trips 70%
  - 3 trips or more 20%
- Percentage home based tours: 94%
- Percentage non home based trips: 20%
  - Especially combinations of trip purposes Shopping and Social, and of Work and Business

# Ambiguous definitions in literature

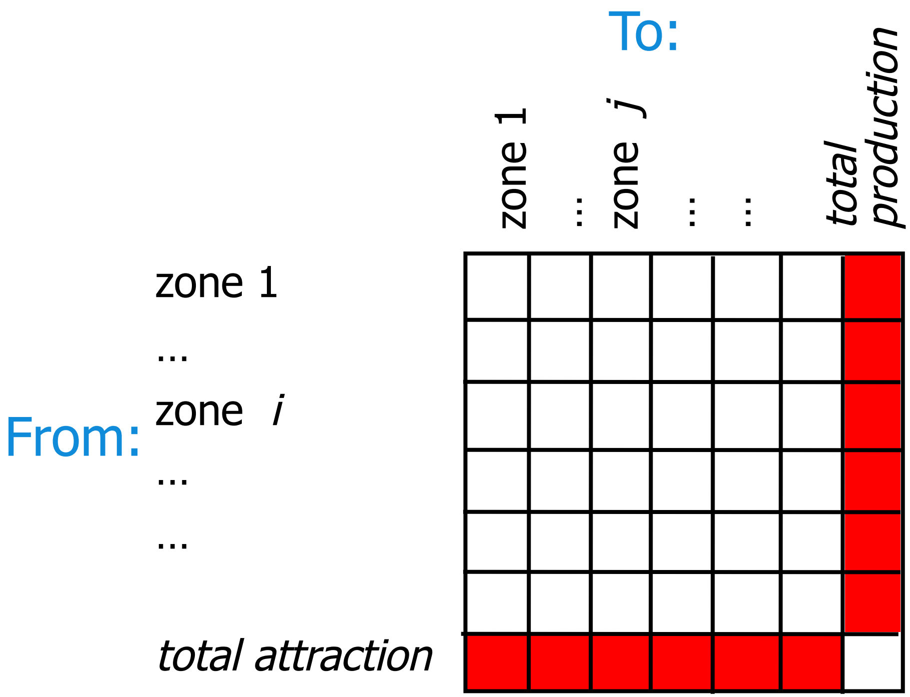
- Definition 1
  - Production: person related (home end or origin)
  - Attraction: activity related (activity end or destination)
- Problem with definition 1
  - Non-home based trips (about 20% of all trips):
    How to allocate these trips to zones?
- Definition 2
  - Production: departures
  - Attraction: arrivals

# Dominant definition in this course

Given: A map with zones and zonal data

Determine:

- The number of trips departing at each zone (production)
- The number of trip arriving at each zone (attraction)



# Possible classifications

## Trip purpose

- compulsory / mandatory trips
  - working trips
  - education trips
- discretionary / optional trips
  - social trips
  - recreational trips

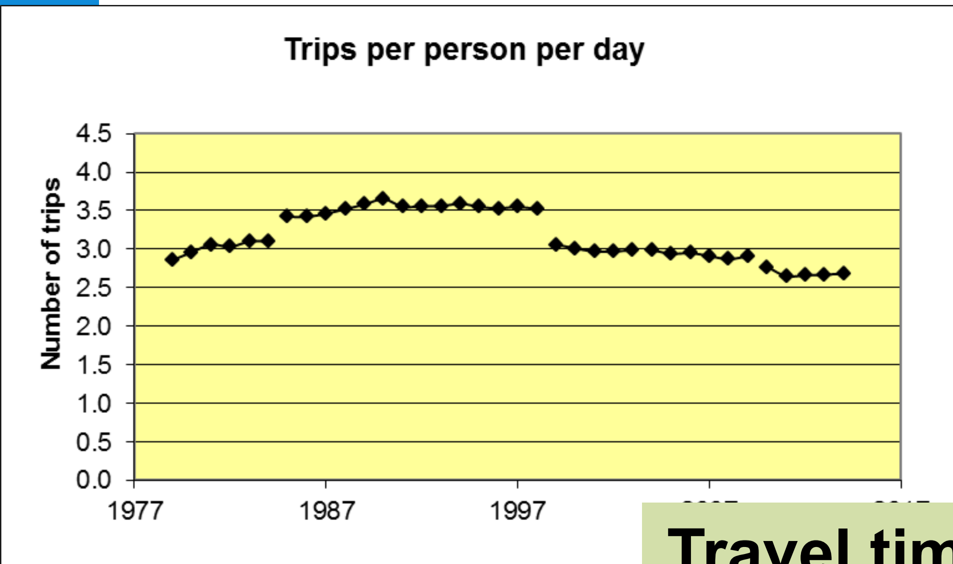
## Time of day

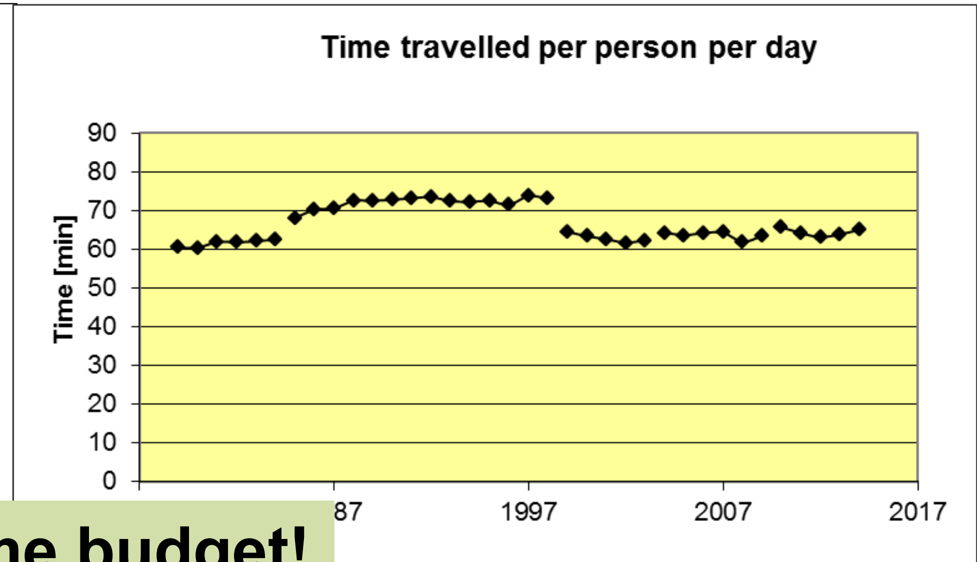
- peak period
- off-peak period

## Person type

- income level
- car ownership
- household size

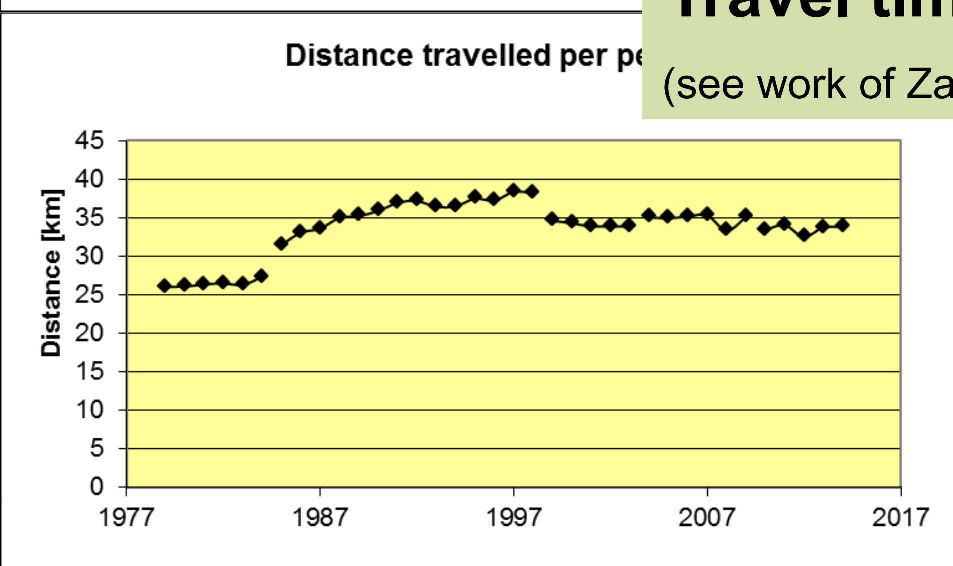
# Travel characteristics Netherlands (MON)

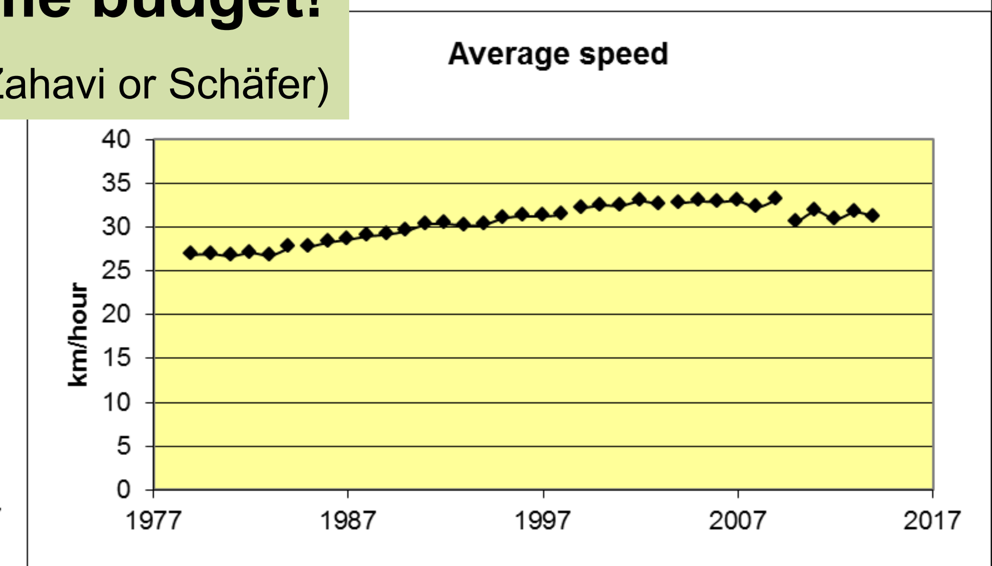


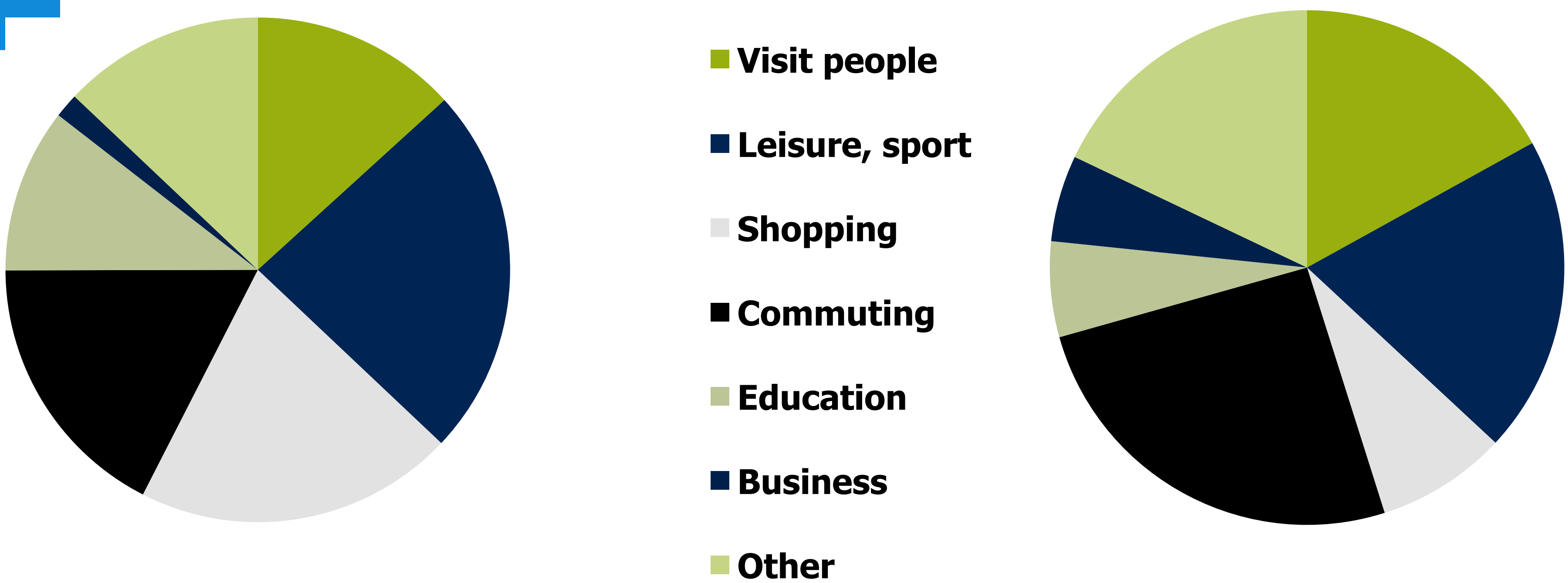

**Travel time budget!**

(see work of Zahavi or Schäfer)



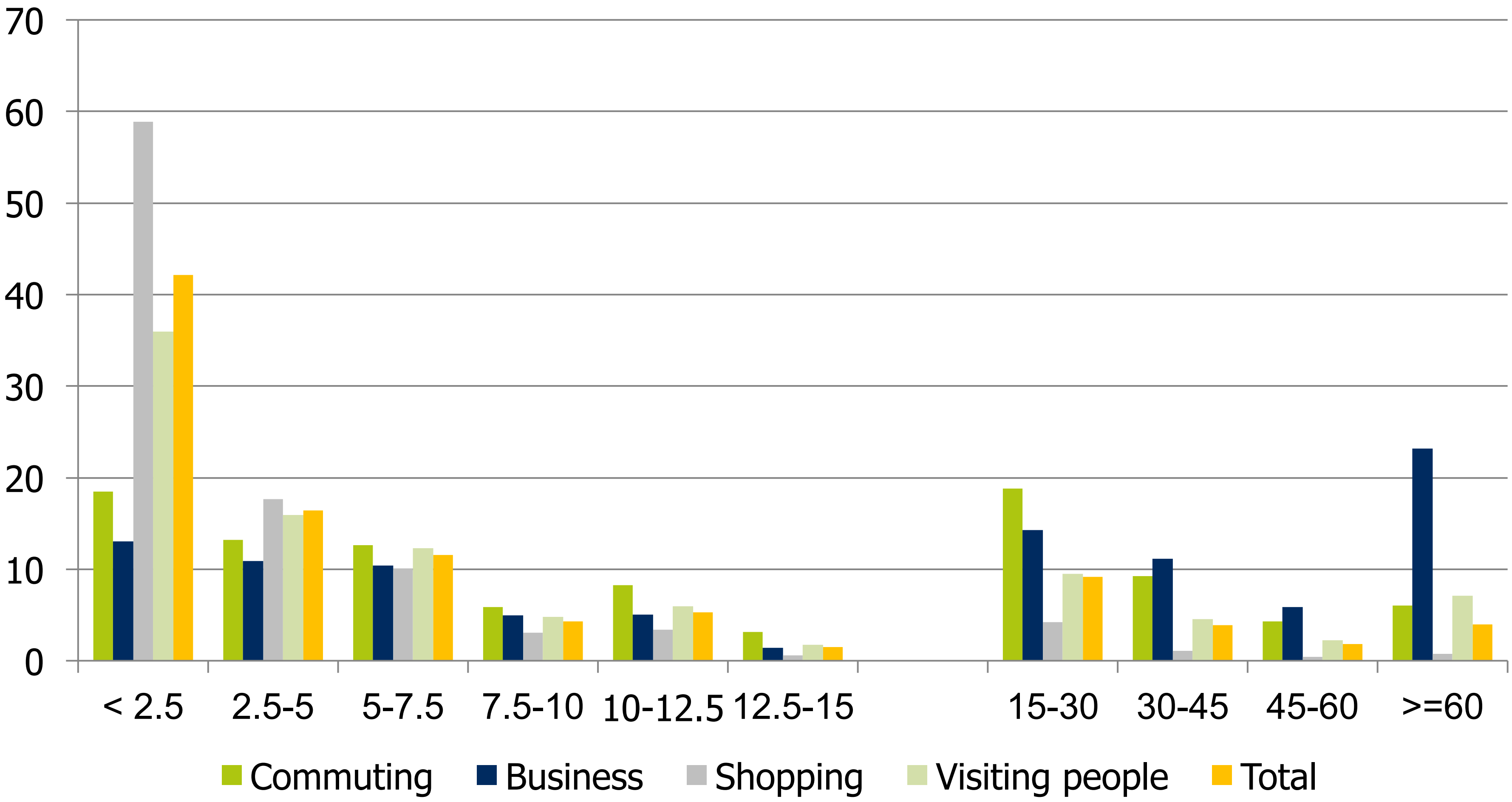



# Trip purpose (Netherlands)
## Trips and trip kilometres for an average day




Trip purpose is defined by the activity at the destination, except when the destination is home, then the activity at the origin is decisive
(note the similarity with definition 1 for attraction)

# Trip length distributions (Netherlands)



# 2.2

*Explanatory factors*

# Possible factors traveller production

Factors affecting the production:

- income
- car ownership
- household structure
- family size
- value of land
- residential density

*personal*

*household*

*zone*

Note the type of definition that is used

# Possible factors traveller attraction

Factors affecting the attraction:

- office space
- retail space (shops)
- employment (jobs)
- households
- “student capacity”

# 2.3.1

## *Trip generation models*
## *Method 1: Growth factors*

# Growth factor

- Simple approach:
  - Multiplier based on a relative change of one or more zone attributes
- Usually applied for external zones (cordon model)

# 2.3.2

*Trip generation models*
*Method 2: Regression analysis*

# Regression models

$$Y = \sum_k \theta_k X_k$$ (linear regression)

$Y$ Endogenous variable

e.g. number of trips produced by a zone or household

$X_k$ Exogenous (explanatory) variables

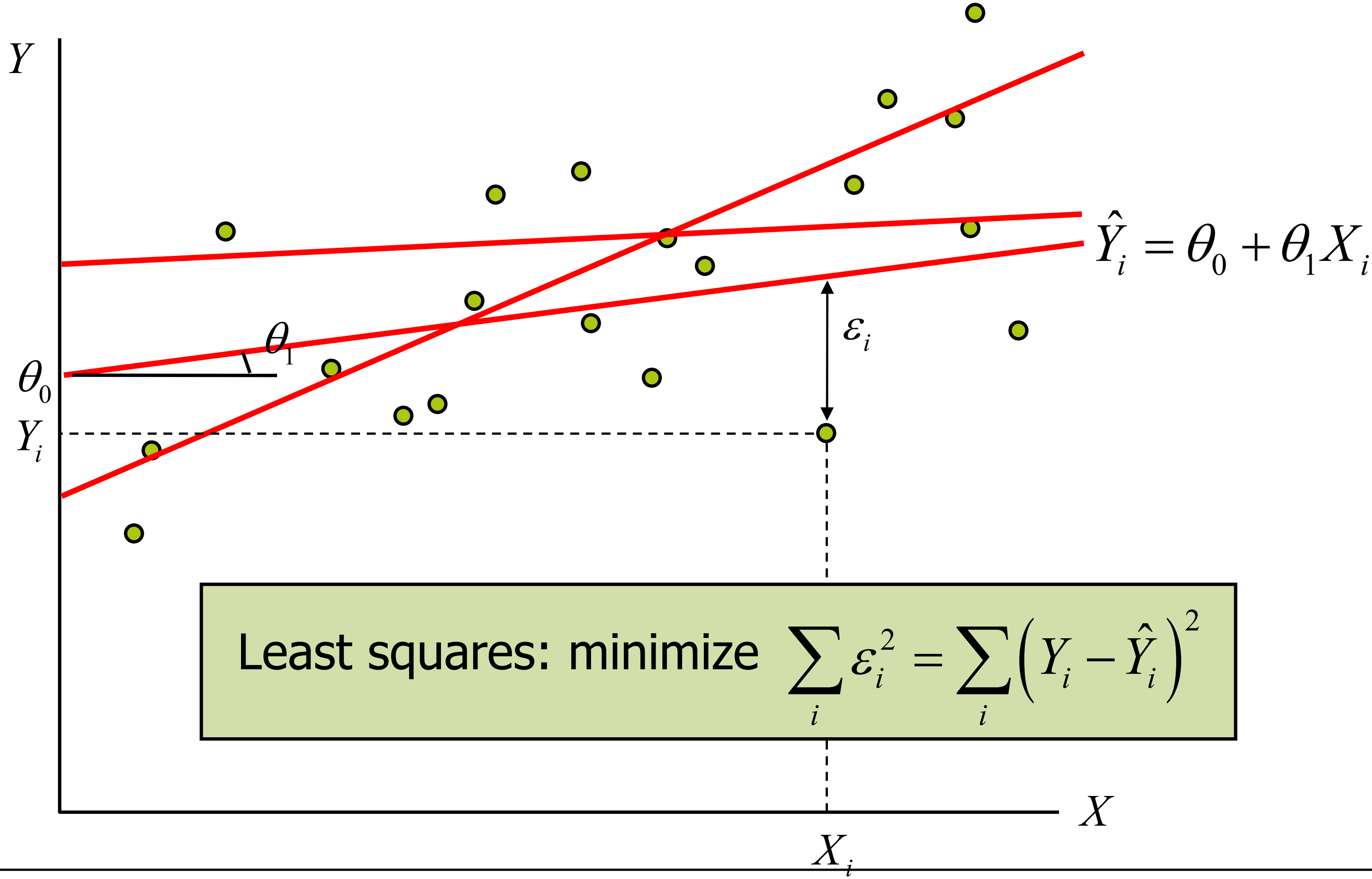
e.g. number of inhabitants, household size, education

$\theta_k$ Parameters

$$Y = 0.91 + 1.44 X_1 + 1.07 X_2$$

number of trips number of workers number of cars

# Regression models




Least squares: minimize $\sum_i \varepsilon_i^2 = \sum_i \left(Y_i - \hat{Y}_i\right)^2$

# Regression models

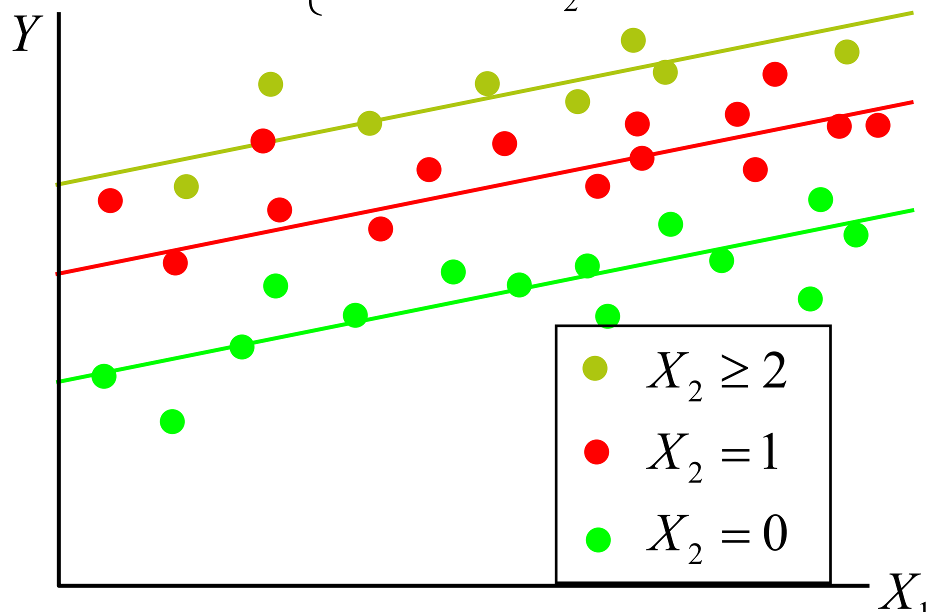
$Y$ = # trips per household
$X_1$ = # workers per household
$X_2$ = # cars per household

**Linear regression:**

$$Y = 0.91 + 1.41X_1 + 1.07X_2$$

Nonlinearity problem:
The parameter for $X_2$
is not constant.

$$Y = 1.41X_1 + \begin{cases} 0.84 & \text{if } X_2 = 0 \\ 1.59 & \text{if } X_2 = 1 \\ 3.98 & \text{if } X_2 \geq 2 \end{cases}$$




**Regression with dummy variables:**

$$Y = 0.84 + 1.41X_1 + \\ + 0.75D_1 + 3.14D_2$$

$D_1 = 1$ if 1 car, 0 otherwise
$D_2 = 1$ if 2 or more cars, 0 otherwise

# Regression at zonal or household level

- Zone level doesn't capture the variety in travel behaviour
- Differences in zone size may affect the regression result
  - So you could use zone rates, e.g. zone total per household
- Regression at household level allows for more detail in modelling variety in travel behaviour
- Zone totals can then be obtained by multiplying the trip rate with the number of households
- However, what about attractions (i.e. activity ends) …….

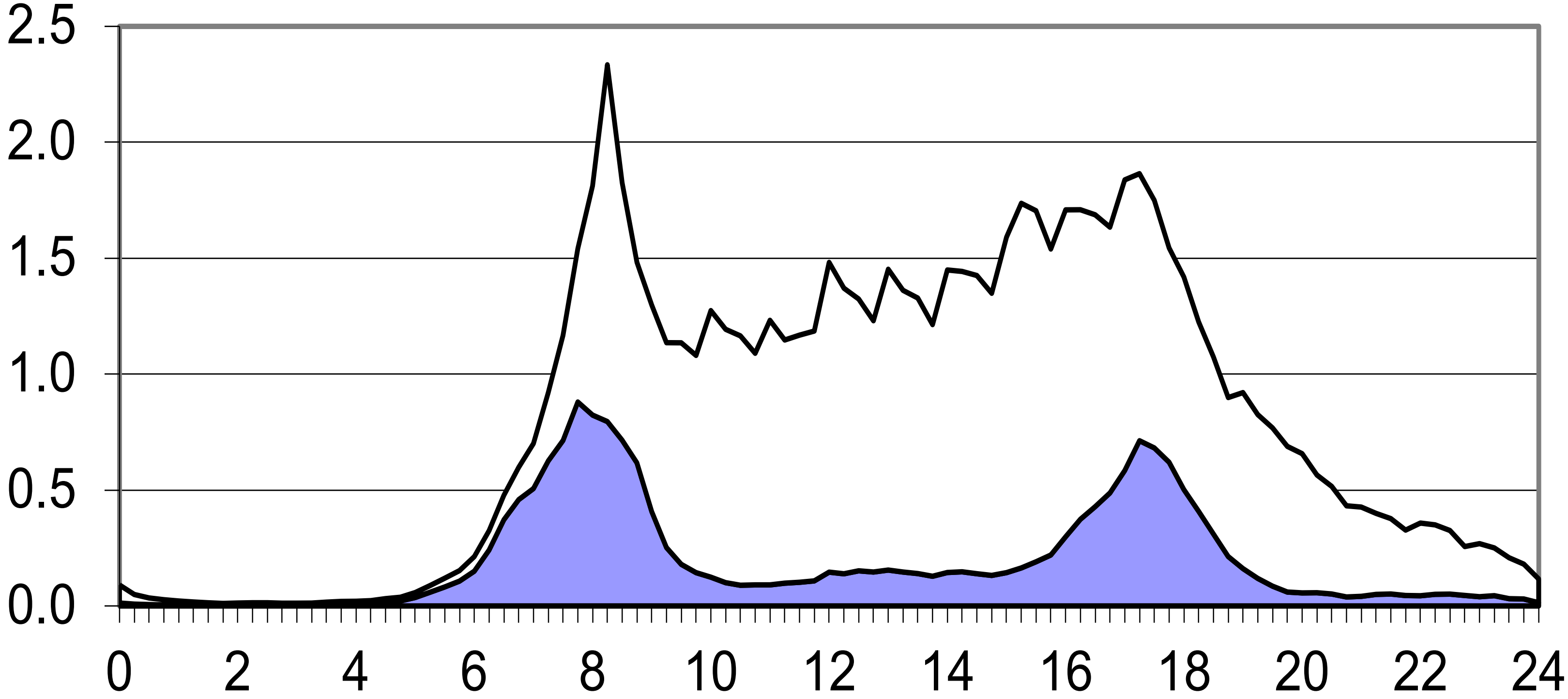
# Example for 24-hour model

| Trip purpose | Regression formula |
|---|---|
| Work | 0.9 * working population + 0.9 * jobs |
| Business | 0.5 * jobs |
| Education | 0.2 * households + 1.9 * student capacity |
| Shopping | 1.0 * households + 15.6 * retail jobs |
| Other | 3.5 * households |
| Total | 6.5 * households + 2.9 * jobs |

Note that in this case it is assumed that production equals attraction

How would the production and attraction formula for work look like for a morning peak period?

# Number of trips over the day
## Total and commuting (Netherlands)



# Two issues to consider

- When is a trip 'generated' in the peak?
  - Departure time
  - Arrival time
  - Both
  - Time spend travelling in peak period
  - Percentage of travel time in peak period

- A 24-hour model contains both outbound and homebound trips, each of these having a distribution over the day
  - Assuming simple tours, outbound equals homebound
  - Departures outbound depend on workforce, arrivals on jobs
  - Departures homebound depend on jobs, arrivals on workforce
  - Morning peak is dominated by outbound trips, evening peak by home bound trips

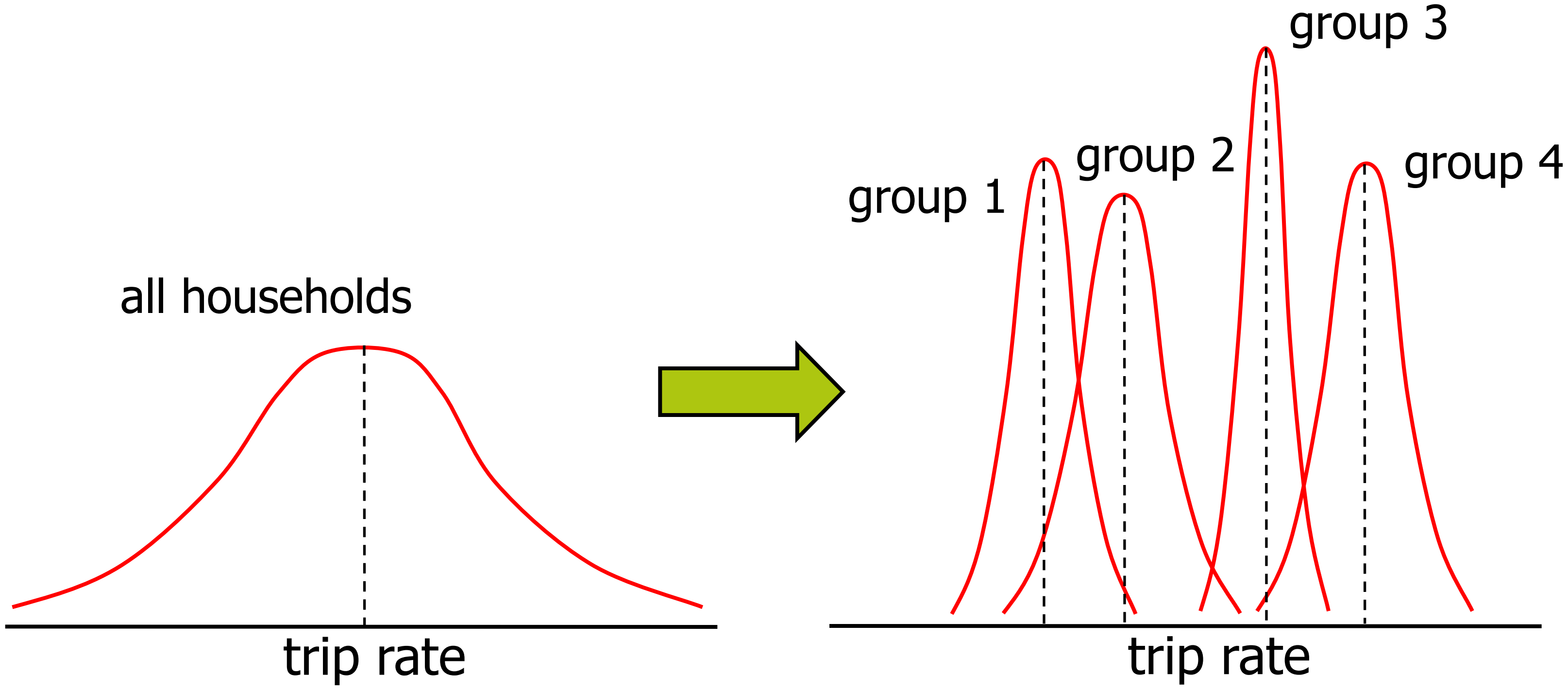
# 2.3.3

## *Trip generation models*
## *Method 3: Cross classification*

# Cross-classification models

Classify households in homogenous groups

e.g. number of people in household, number of cars, and combinations



# Cross-classification models

Advantages:

- Groupings are independent of zone system
- Relationships do not need to be linear
- Each group can have a different form of relationship

Disadvantages:

- No extrapolation beyond the calibrated groups
- Large samples are required (at least 50 obs. per group)
- What is the best grouping?

# Cross-classification models

| household size | 0 cars | 1 car | $\geq$2 cars | |
|---|---|---|---|---|
| 1 person | $n = 20,\ \mu = 0.2$ | $n = 10,\ \mu = 0.5$ | $n = 0,\ \mu = ?$ | $n = 30,\ \mu = 0.3$ |
| 2 persons | $n = 85,\ \mu = 0.5$ | $n = 150,\ \mu = 0.9$ | $n = 20,\ \mu = 1.5$ | $n = 255,\ \mu = 0.8$ |
| $\geq$ 3 persons | $n = 25,\ \mu = 0.7$ | $n = 40,\ \mu = 1.2$ | $n = 30,\ \mu = 2.0$ | $n = 95,\ \mu = 1.3$ |
| | $n = 130,\ \mu = 0.5$ | $n = 200,\ \mu = 0.9$ | $n = 50,\ \mu = 1.8$ | |

Problem is cells having limited or no observations

# Completing a cross classification table Multiple class analysis (MCA)

| household size | 0 cars | | 1 car | | ≥2 cars | | |
|---|---|---|---|---|---|---|---|
| 1 person | 0.0 | 0.2 | 0.3 | 0.5 | 1.2 | ? | –0.6 |
| 2 persons | 0.4 | 0.5 | 0.8 | 0.9 | 1.7 | 1.5 | –0.1 |
| ≥ 3 persons | 0.9 | 0.7 | 1.3 | 1.2 | 2.2 | 2.0 | +0.4 |
| | –0.4 | | 0.0 | | +0.9 | | $\mu = 0.9$ |

Cell values are derived from statically more reliable row and column averages
Estimated cell value = grand mean + Δ row + Δ column

# Application

Zone *i* having 253 households yields

| household size | 0 cars | 1 car | ≥2 cars | |
|---|---|---|---|---|
| 1 person | 50 0.0 | 30 0.3 | 1 1.2 | |
| 2 persons | 25 0.4 | 70 0.8 | 10 1.7 | |
| ≥ 3 persons | 10 0.9 | 44 1.3 | 13 2.2 | |
| | | | | |

Trip production of zone *i*: 188

# Two questions

- What about attractions (i.e. activity ends)?
- Are the differences in trip rates large enough?
    - Car trips
    - All modes

# 2.3.4

*Trip generation models*
*Method 4: Choice modelling*

# Discrete choice models

## Binary logit

Alternative 0: do not make a tour $V_0 = 0$

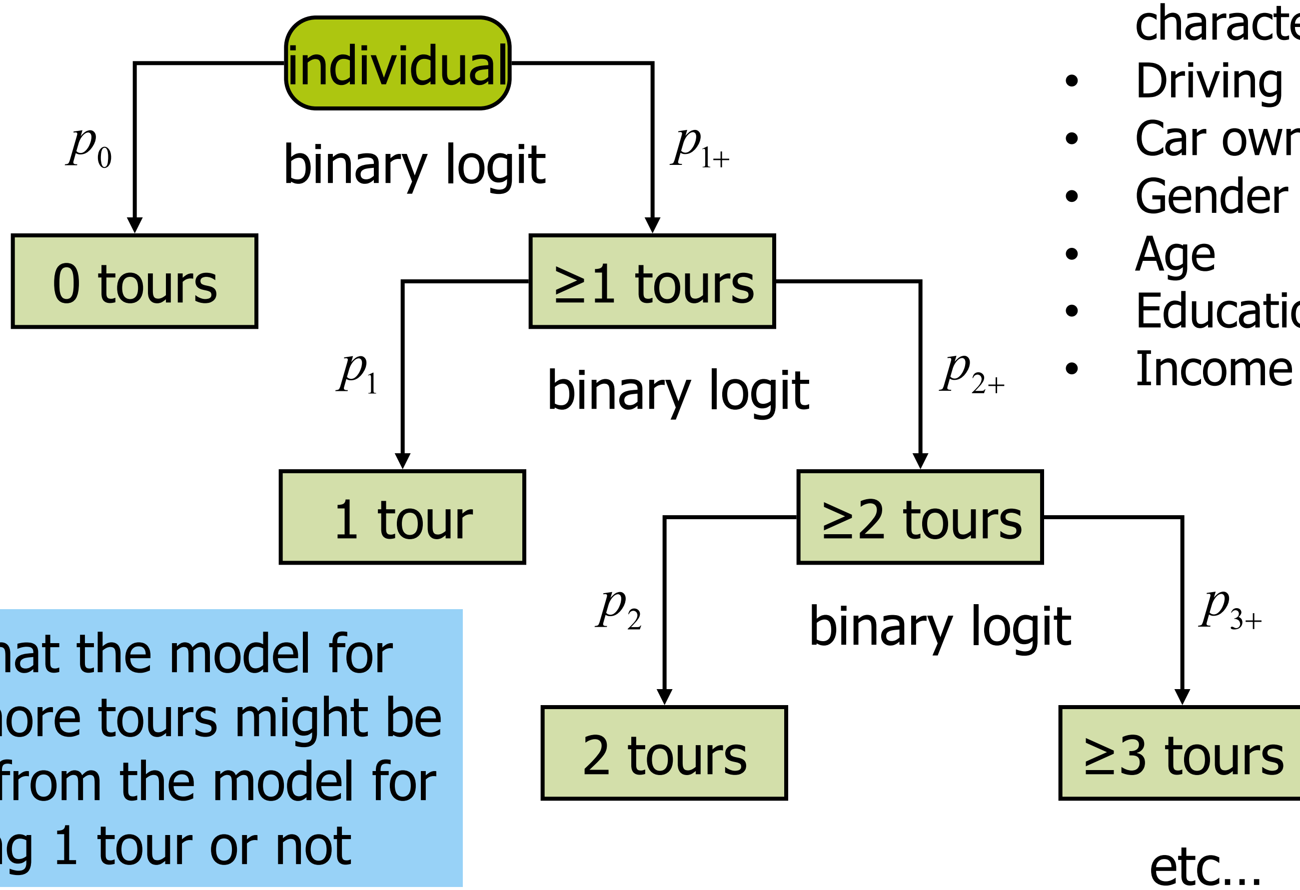
Alternative 1: make one or more tours $V_1 = \ldots$

Probability of making at least one tour:

$$p_{1+} = \frac{\exp(\beta V_1)}{\exp(\beta V_0) + \exp(\beta V_1)} = \frac{1}{\exp(-\beta V_1) + 1}$$

# Discrete choice model

Stop/repeat-model




Possible attributes:

- Household characteristics
- Driving license
- Car ownership
- Gender
- Age
- Education
- Income

Note that the model for making more tours might be different from the model for making 1 tour or not

# Discrete choice model for trips

- How would a stop/repeat model look like if you would use trips instead of tours?

# 3.

## *Trip generation*
## *Practical issues*

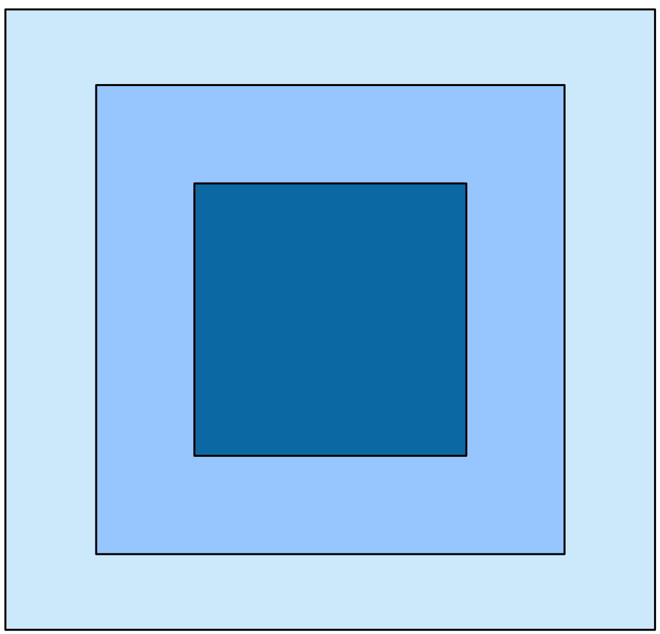
# Practical issues

- What about external zones?
- Role of intra-zonal trips?
- Segmentation?
- Role of accessibility?
- Does total of departures equal total of attractions?

# Difference between model types

Cordon model:

- Departures and arrivals are based on counts or external models



Model having influence area and external zones:

- Departures and arrivals are based on zone data

# Role of intra-zonal trips

- Usually all trips are modelled

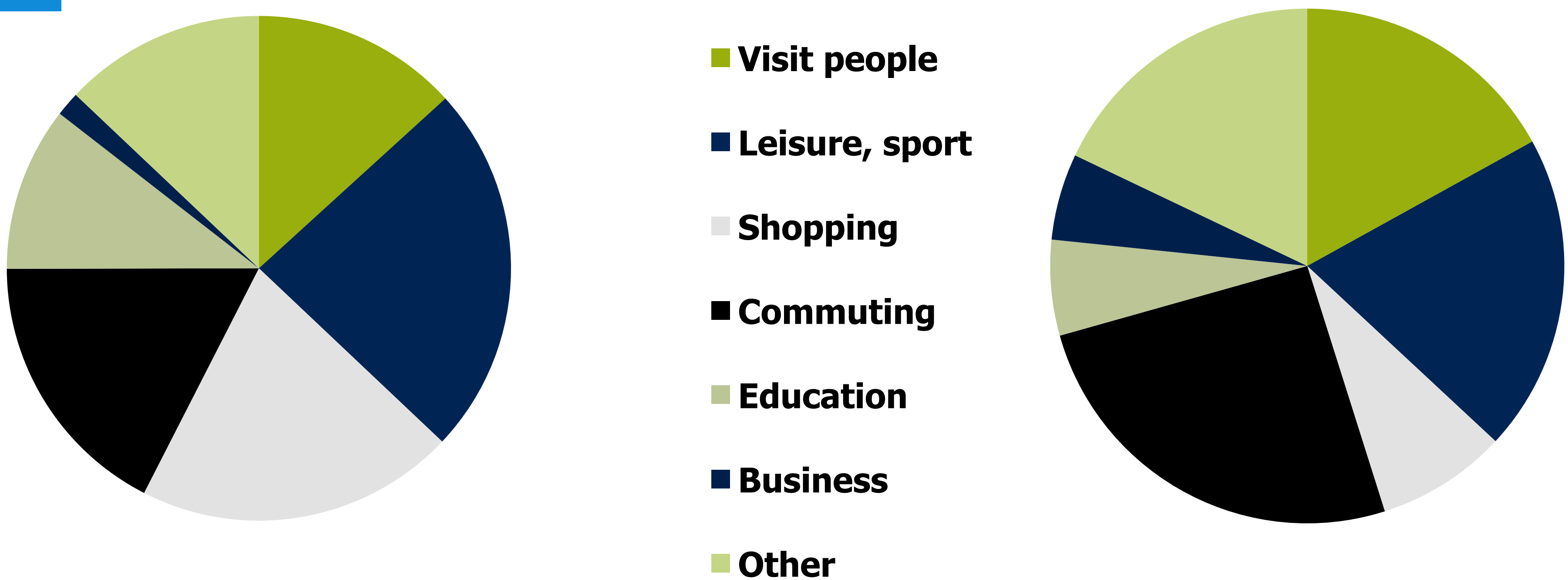
  Intra-zonal trips are determined in distribution model

- Some models however exclude intra-zonal trips in the distribution model

  In that case intra-zonal trips should be excluded when modelling trip generation

# Segmentation: Trip purpose (NL)

## Trips and trip kilometres for an average day




Modelling usually focuses on commuting, business, education, shopping and other

# Role of accessibility?

- MON: ‘fixed’ trip rate per person per day
- Recent model estimations: small effect for some trip types
- In case of unimodal models?

Substantial impact, especially for public transport

# Does trip production equal trip attraction?

- We have determined the trip production (=departures)
- We have determined the trip attraction (=arrivals)

| | | | |
|---|---|---|---|
| | | | 120 |
| | | | 130 |
| | | | 100 |
| 93 | 198 | 59 | 350 |

350

350

Sum of row totals doesn't equal sum of column totals, while the concept of OD-matrices requires equal sums

What to do?

Trip balancing:
your choice to decide which of the two is the constraint

# Categorisation of methods

<table>
<tr><th>Home-end</th><th>Activity-end</th><th>Comment</th></tr>
<tr><td>Regression at zone level</td><td>Regression at zone level</td><td>Different trip purposes</td></tr>
<tr><td>Regression at household level</td><td></td><td>Dummy variables for user categories, e.g. number of cars</td></tr>
<tr><td>Cross classification</td><td></td><td rowspan="2">Different categories of households or individuals</td></tr>
<tr><td>Choice models (Stop & repeat)</td><td></td></tr>
</table>

# Topics to study: Trip distribution

- What does this modelling component do? What's its output and what's its input? How does it fit in the framework?
- Do you understand the definitions?
  - OD-matrix, production, attraction, generalised costs, deterrence function
- Do you understand the modelling methods?
  - Growth factor: singly and doubly constrained
  - Gravity model
    - Logic for "borrowing" from Newton or Entropy maximisation?
- Do you understand the iterative algorithm?
- Do you understand the calibration of the model?
  - Deterrence function (Hymann's method)
- Do you understand tri-proportional fitting ("bins")?
- Are these models appropriate?