Modeling the walking behavior of pedestrians

A discrete choice approach

Michel Bierlaire

TRANSP-OR Transport and Mobility Laboratory

Ecole Polytechnique Fédérale de Lausanne

transp-or.epfl.ch

Context

- Pedestrian walking behavior
- in normal conditions
- as a function of other pedestrians
- *Click here for an example*

Context

Objectives:

- Specify a mathematical model to forecast the walking behavior
- Estimate the model parameters on real data
- Validate the model with real data, not involved in the estimation

Applications:

- Pedestrian simulation [*Click here*][*With background*]
- Pedestrian tracking [*Click here*]
  - Tracking without model [*Click here*]
  - Tracking with model [*Click here*]

Outline

- Mathematical framework
- Modeling elements
- Model specification
- Estimation data
- Estimation results
- Validation
- Conclusion

Mathematical framework

Econometric model

$$y = f(x, \beta, \varepsilon)$$

where

- y is the dependent variable (e.g. position of the next step)
- x is a vector of independent or explanatory variables (e.g. position and speed of other pedestrians, etc.)
- β is a vector of unknown parameters to be estimated from data
- ε is a (vector of) random variable(s) capturing the errors related to modeling simplifications, data collection, etc.
- f is derived from an underlying "economic" theory

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Mathematical framework

- Objective: predict where a pedestrian chooses to put her next step
- Theoretical tool: discrete choice theory
- Operational tool: random utility model
- Derivation:
  1. Identification of the choice set
  2. Characteristics of the decision-maker
  3. Attributes of the alternatives
  4. Behavioral assumptions
- Illustrative example: choice of transportation mode to go to work

Introduction to choice modeling

Identification of the choice set $\mathcal{C}_n$

- Finite and discrete
- Example: { car as driver, car as passenger, train, bus }

Characteristics S_n of the decision-maker n

- Age, sex, income, etc.
- Trip purpose, value-of-time, etc.

Attributes z_{in} of the alternative i for individual n

- Cost, travel time, walking time, frequency, etc.
- Level of comfort, reliability, etc.

It is common to merge everything in one vector of variables

$$x_{in} = (z_{in}, S_n)$$

Introduction to choice modeling

Behavioral assumptions: utility theory

- Individual n associates a utility with each available alternative i
- The utility is a function of S_n and z_{in}.
- To capture the uncertainty, an error term is involved.
- Typical formulation

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

  where
  - V_{in} is deterministic
  - ε_{in} is the error term
- Most common specification:

$$V_{in} = \sum_k \beta_k (x_{in})k$$

Introduction to choice modeling

Behavioral assumptions: utility theory

- Individual chooses the alternative with the largest utility
- Choice model:

$$P(i|\mathcal{C}_n) = \Pr(U_{in} \geq U_{jn} \; \forall j \in \mathcal{C}_n)$$

- If ε_{in} are assumed i.i.d. extreme value distributed, we obtain the multinomial logit model

$$P(i|\mathcal{C}_n) = \frac{e^{V_{in}}}{\sum_{j \in \mathcal{C}_n} e^{V_{jn}}}$$

- In the context of pedestrian modeling, a more complex model will be used.

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Outline

- $\surd$ Mathematical framework
- Modeling elements
- Model specification
- Estimation data
- Estimation results
- Validation
- Conclusion

Modeling elements

Pedestrian movement based on a hierarchical framework (Daamen, 2004)

- Strategical: list of activities
- Tactical: activity schedule
  - Time and location of activities
  - Choice of itinerary
- Operational: short-term walking behavior
  - The "next step" decision
  - Direction, speed
  - Collision avoidance
  - Leader-follower

Modeling elements

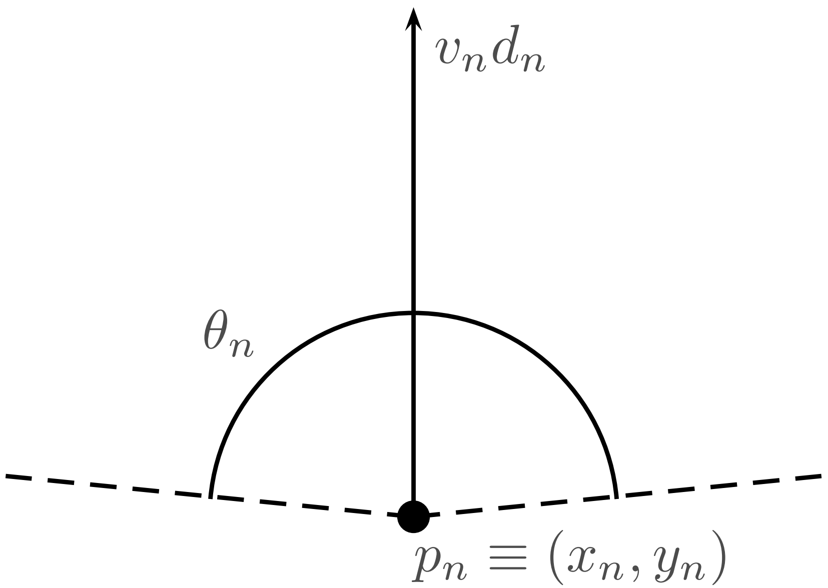
In our context

- strategical and tactical decisions are exogenous
- current intermediary destination is known (“next door”)
- we focus of a “myopic” behavior
- reactions to the immediate environment, mainly other pedestrians

Modeling elements

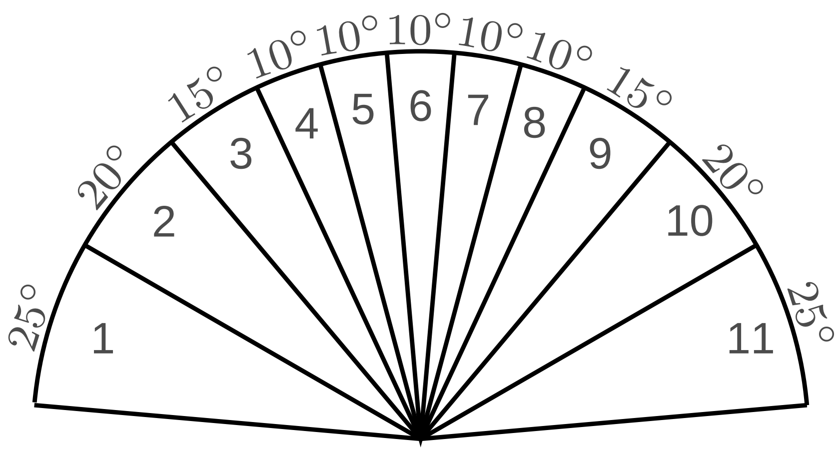
Given

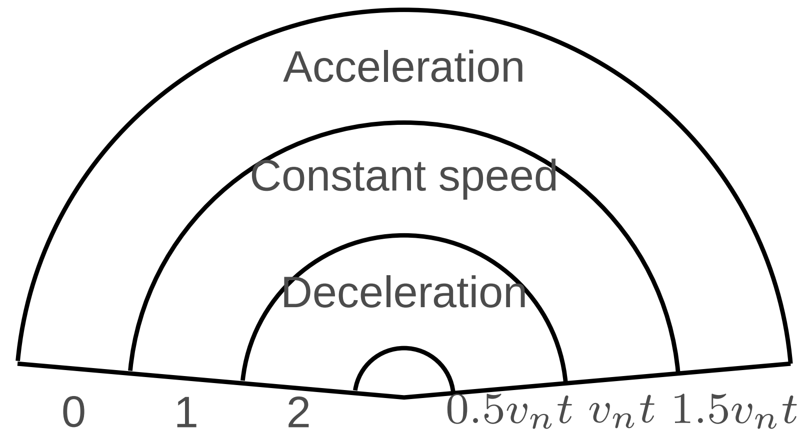
- the current position $p_n = (x_n, y_n)$
- the current speed v_n (m/sec)
- the current direction d_n, $d_n \in \mathbb{R}^2$, $\|d_n\| = 1$
- a visual angle $\theta_n = 170°$

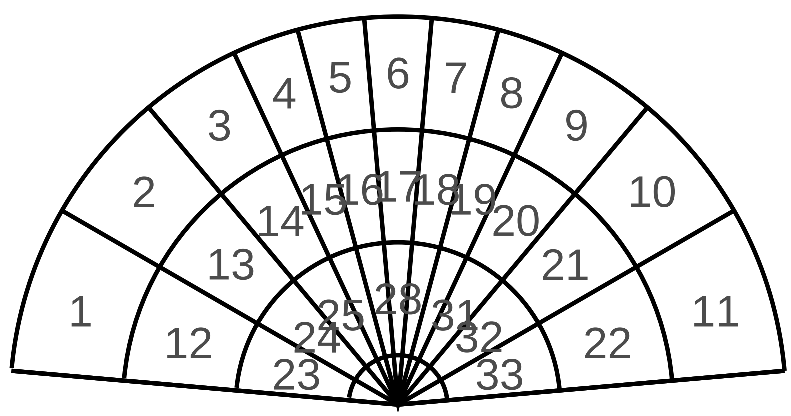


Modeling elements

Choice set $\mathcal{C}_n$: individual-specific discretization







Modeling elements

- 11 directions, relative to d_n
- 3 speed regimes: $0.5v_n$, v_n, $1.5v_n$
- 33 alternatives
- Each alternative is a combination of a direction d and a speed regime v
- Each alternative corresponds to the physical position of the next step

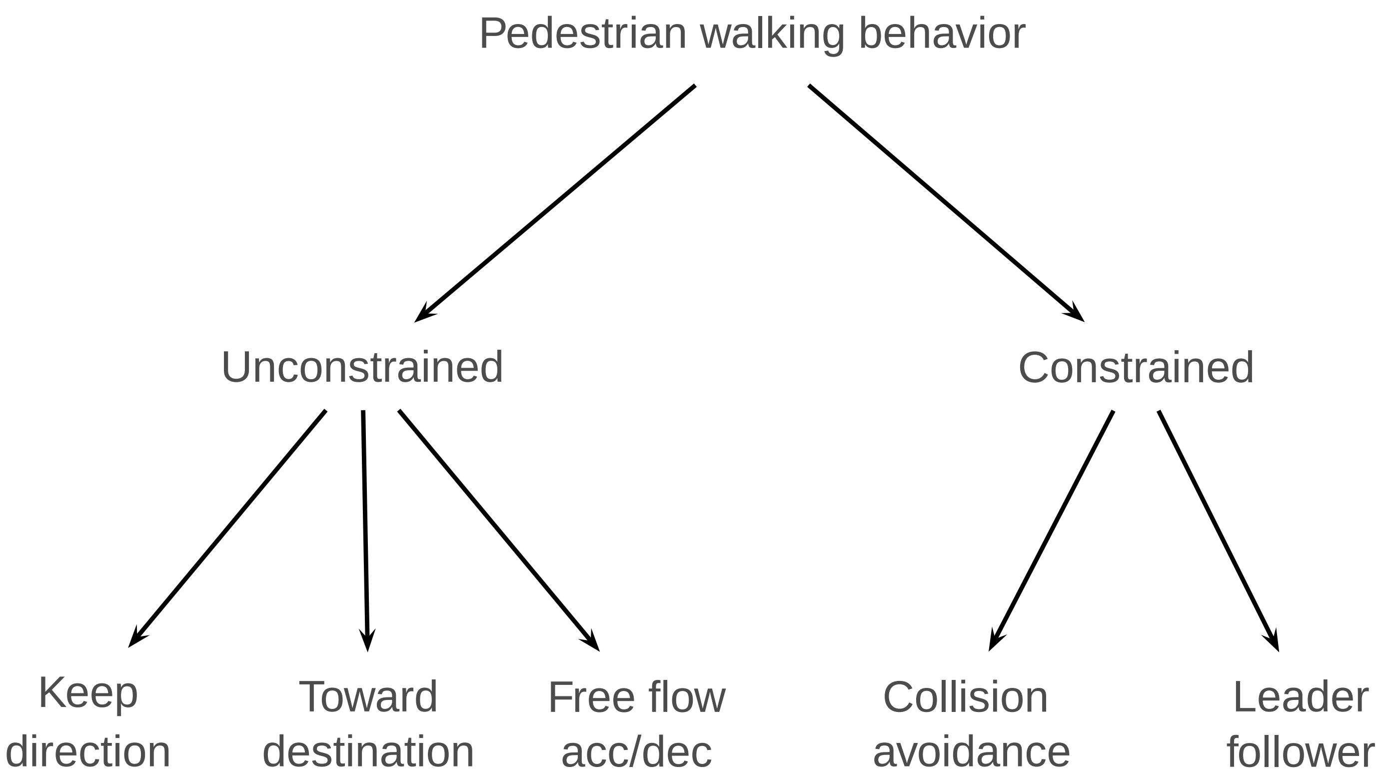
$$c_{vd} = p_n + vtd,$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Modeling elements

Behavioral elements



Outline

- $\surd$ Mathematical framework
- $\surd$ Modeling elements
- Model specification
- Estimation data
- Estimation results
- Validation
- Conclusion

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

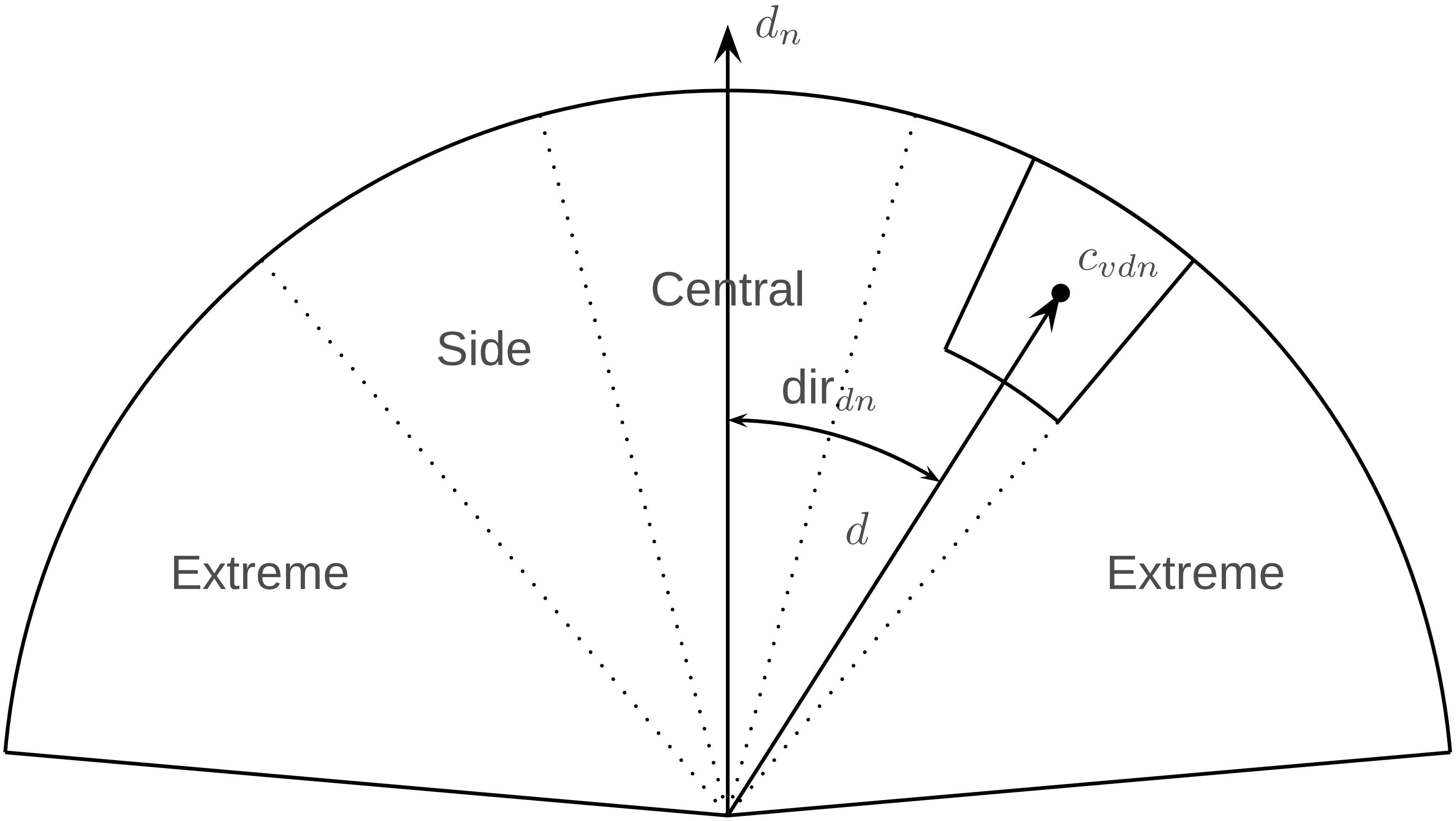
Model specification

Utility:

$$\begin{aligned} U_{in} &= V_{in} + \varepsilon_{in} \\ U_{vdn} &= V_{vdn} + \varepsilon_{vdn} \end{aligned}$$

1. Specification of V_{vdn} to capture the behavioral elements
2. Specification of ε_{vdn} to capture the spatial correlation

Model specification: keep direction



Model specification: keep direction

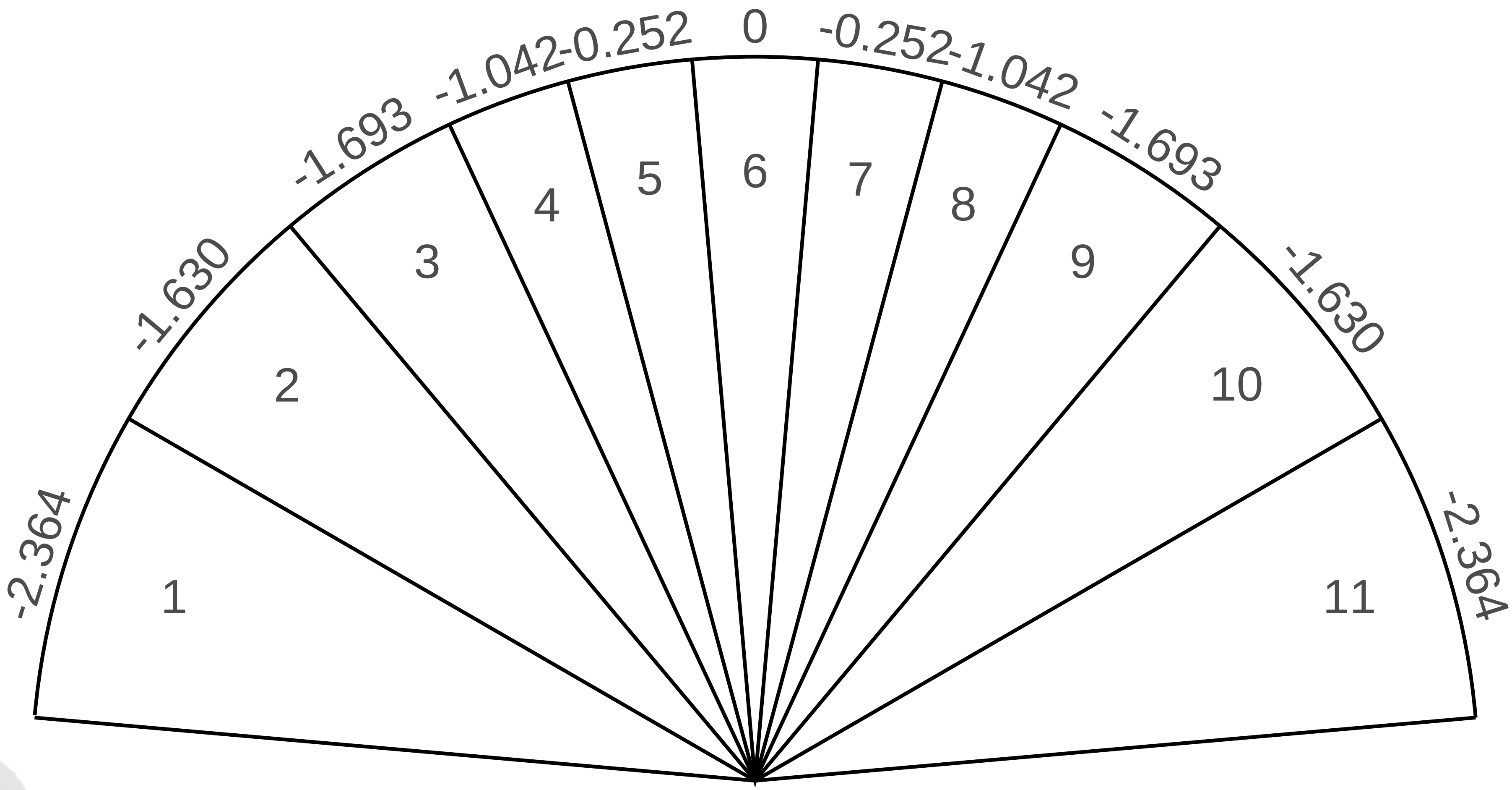
- The greater the angle, the lower the utility
- Not necessarily in a pure proportional way
- We include the following terms in the utility function

$$\beta_{\text{dir\_central}} \text{dir}_{dn} I_{\text{central}} + \beta_{\text{dir\_side}} \text{dir}_{dn} I_{\text{side}} + \beta_{\text{dir\_extreme}} \text{dir}_{dn} I_{\text{extreme}}$$

- Only one of the terms is non zero
- I_k is 1 if the alternative belongs to zone k
- $\beta.$ are unknown parameters to be estimated from the data
- We expect them to be negative

Model specification: keep direction

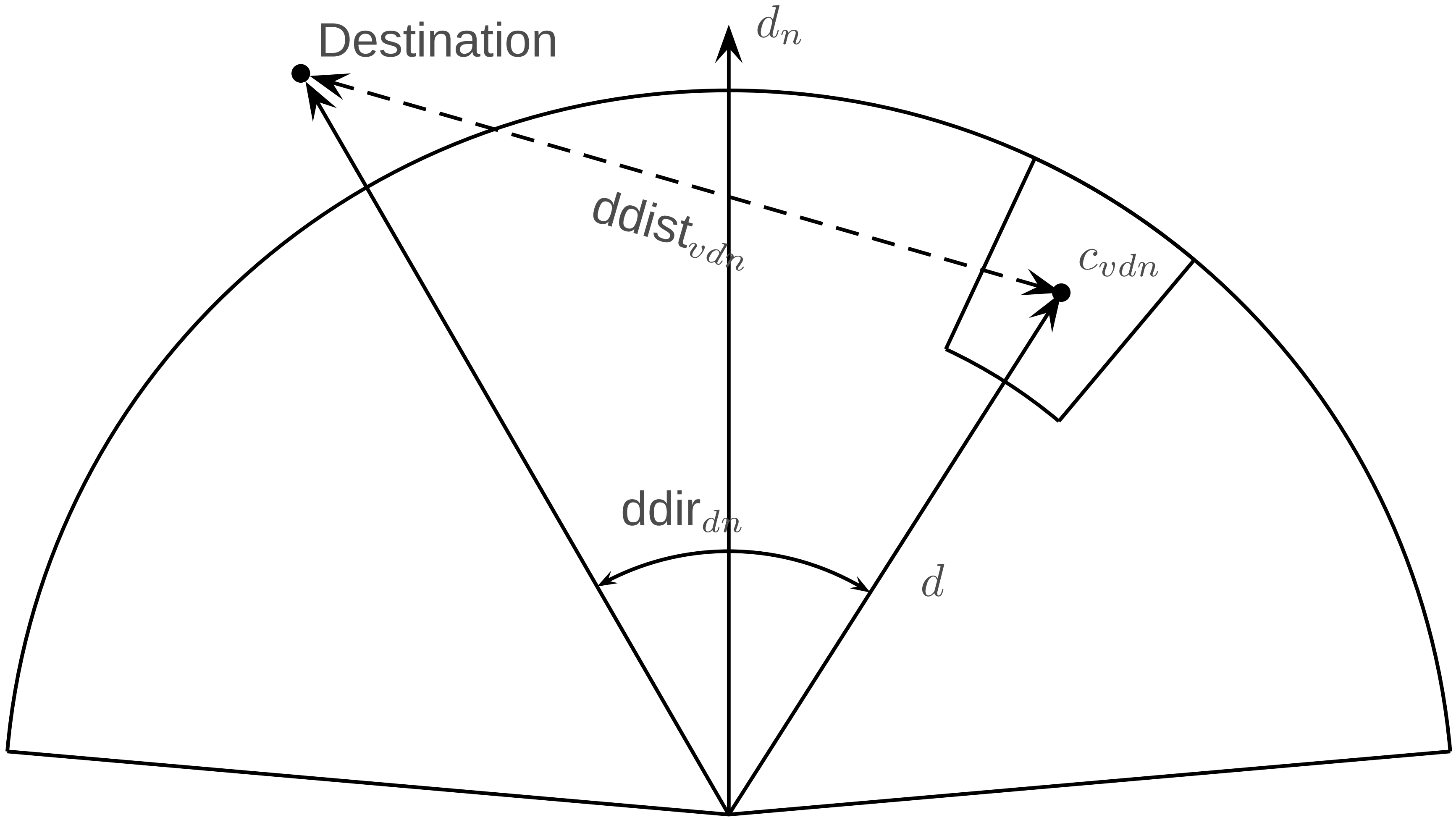
Estimated contributions to the utility



TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model specification: toward destination



Model specification: toward destination

- The destination is exogenously given
- Two effects: the distance and the angle
- We include the following terms in the utility function

$$\beta_{\text{ddist}} \text{ddist}_{vdn} + \beta_{\text{ddir}} \text{ddir}_{dn}$$

- $\beta_{.}$ are unknown parameters to be estimated from the data
- We expect them to be negative
- Results:

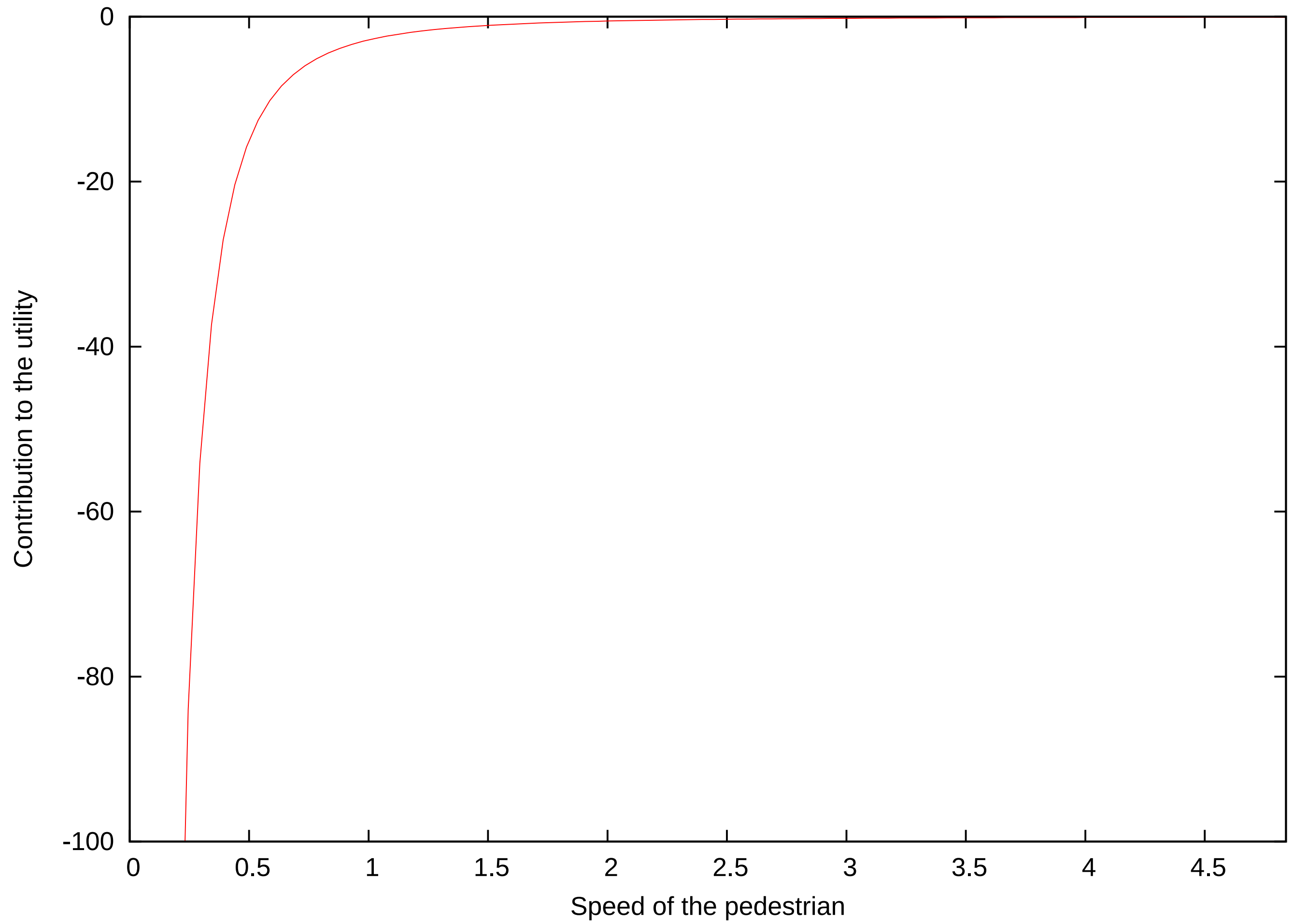
$$-1.55 \text{ ddist}_{vdn} - 0.079 \text{ ddir}_{dn}$$

Model specification: free flow acceleration

- Constant speed is assumed to be the most comfortable
- However, pedestrians accelerate and decelerate to achieve a desired speed
- The desired speed is unknown to the analyst
- Alternatives corresponding to acceleration and deceleration are penalized
- The penalty varies with the current speed
- If the speed is already low, deceleration is less likely
- If the speed is already high, acceleration is less likely

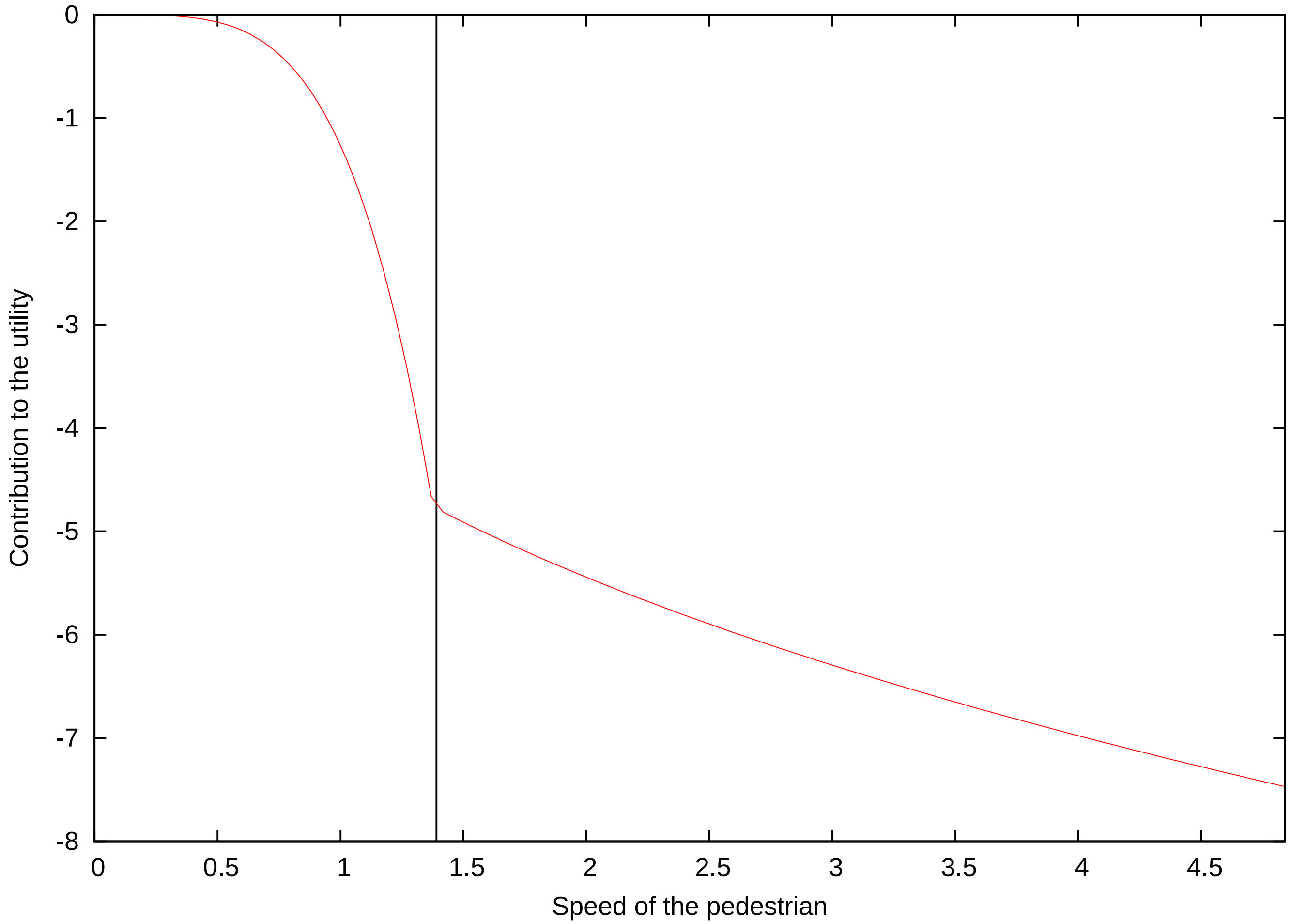
Model specification: free flow acceleration

Penalty for alternatives corresponding to deceleration



Model specification: free flow acceleration

Penalty for alternatives corresponding to acceleration



Model specification: free flow acceleration

- We include the following terms in the utility function

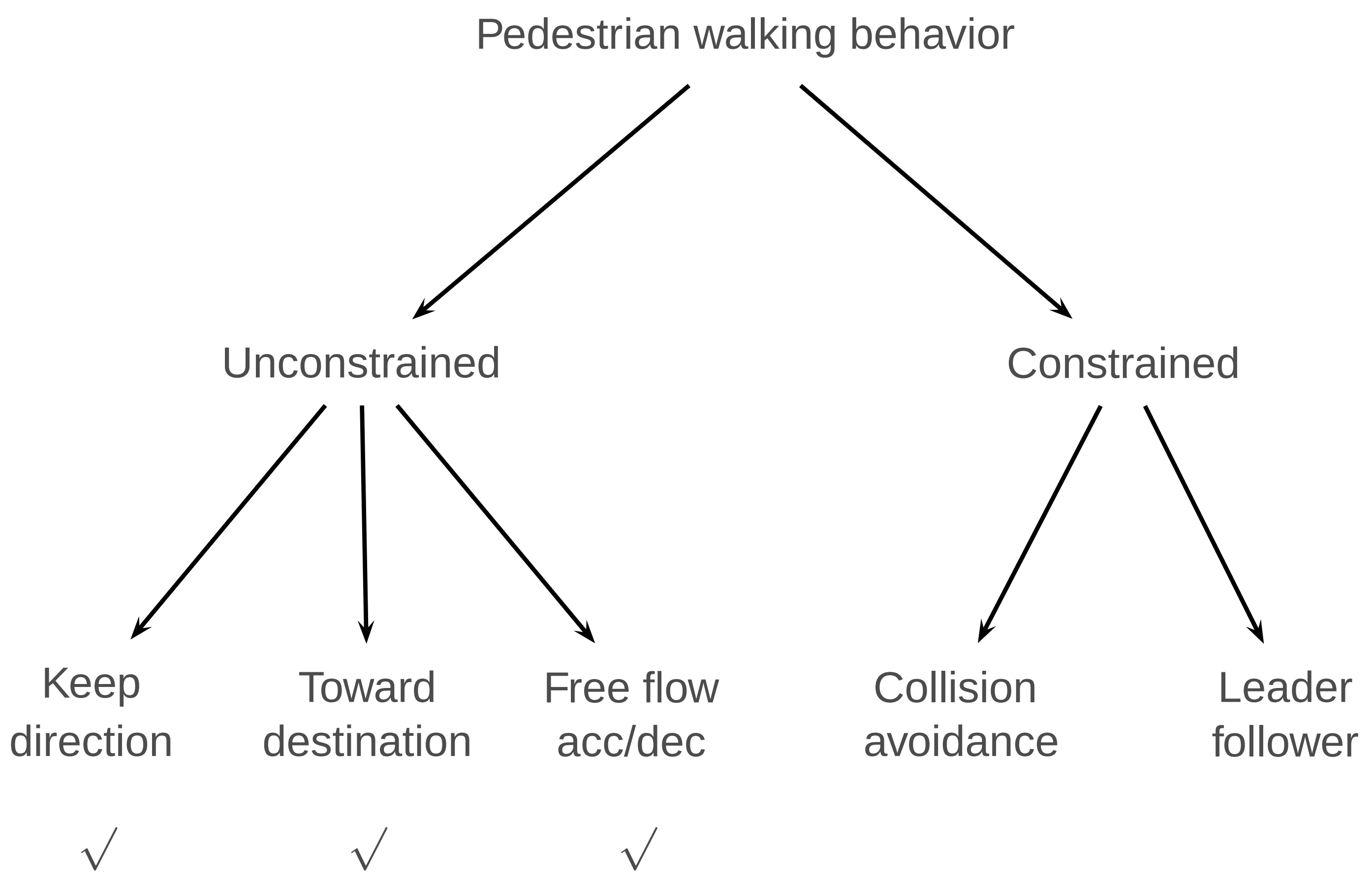
$$
\begin{aligned}
\beta_{\text{dec}} I_{\text{v,dec}} (v_n / v_{\text{max}})^{\lambda_{\text{dec}}} & \quad + \\
\beta_{\text{accLS}} I_{\text{LS}} I_{\text{v,acc}} (v_n / v_{\text{maxLS}})^{\lambda_{\text{accLS}}} & \quad + \\
\beta_{\text{accHS}} I_{\text{HS}} I_{\text{v,acc}} (v_n / v_{\text{max}})^{\lambda_{\text{accHS}}} &
\end{aligned}
$$

- Maximum one term is not zero for each alternative
- $I_{\text{v,dec}}$ and $I_{\text{v,dec}}$ indicates if the alternative corresponds to acceleration or deceleration
- I_{LS} and I_{HS} indicates low speed (≤ 1.39) and high speed
- $\beta.$ and $\lambda.$ are unknown parameters to be estimated from the data
- Normalization: $v_{\text{max}} = 4.84$, $v_{\text{maxLS}} = 1.39$ (1.39 m/s = 5km/h)

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model specification



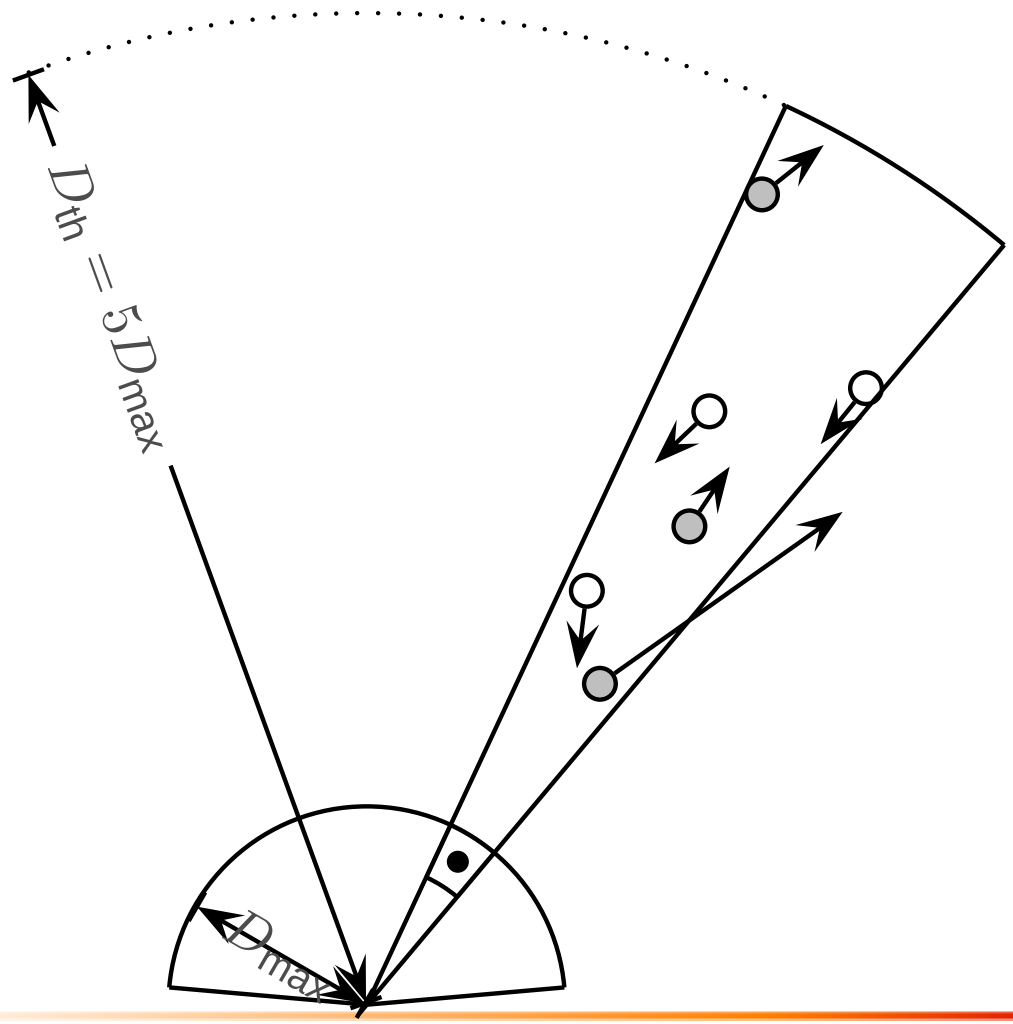
Model specification: leader-follower

- Tendency to follow individuals going in about the same direction
- In each cone, we identify potential leaders
- Individual k is a potential leader

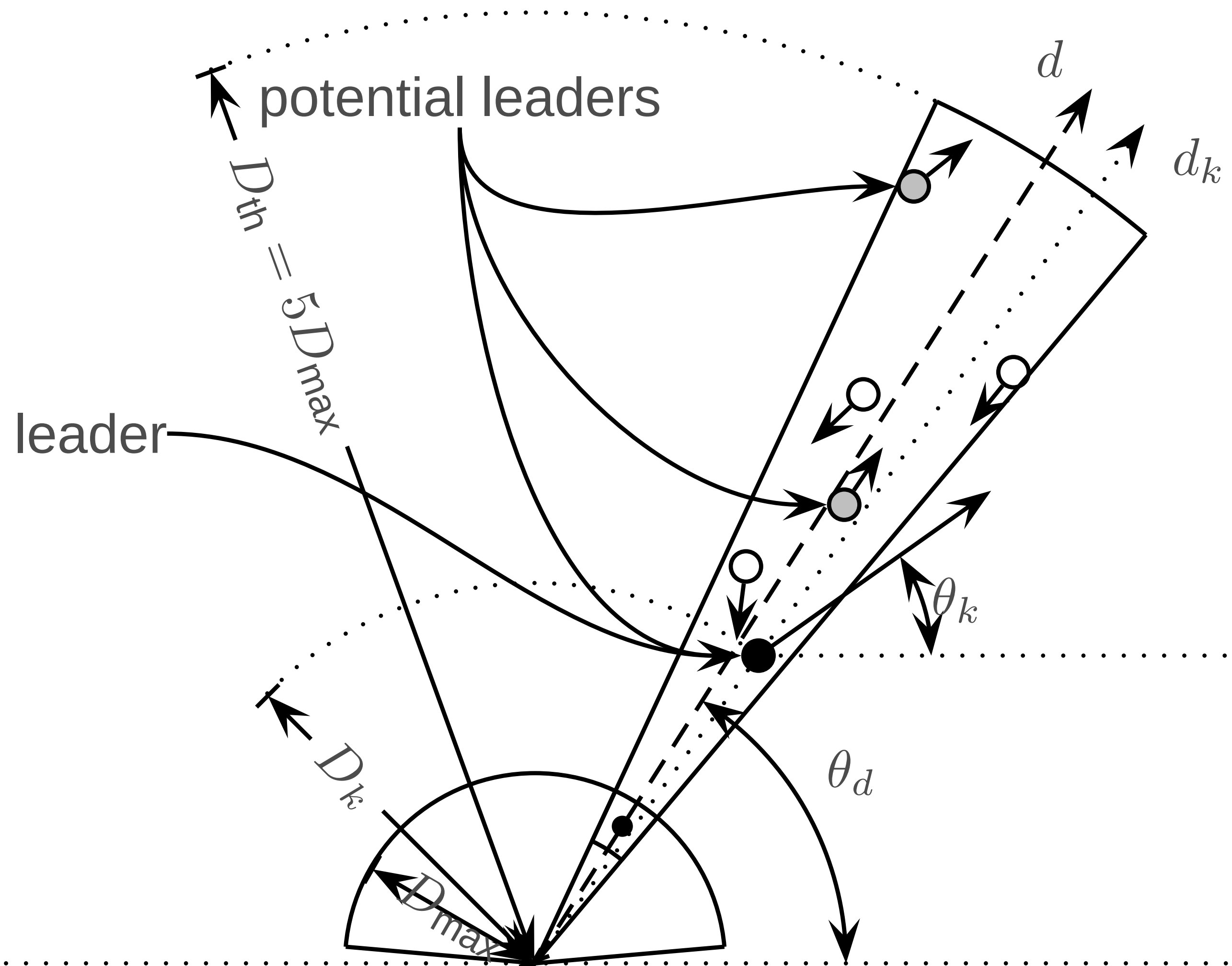
$$\begin{cases} \text{if } d_l \leq d_k \leq d_r \ \text{(is in the cone),} \\ \text{and } 0 < D_k \leq D_{th} \ \text{(not too far),} \\ \text{and } 0 < |\Delta\theta_k| \leq \Delta\theta_{th} \ \text{(walking in almost the same direction),} \end{cases}$$

- Among them, the individual k who is the closest is selected as the leader
- Her speed and direction are recorded
- Her presence may trigger a change of speed
- ...with different effects for acceleration and deceleration

Model specification: leader-follower



Model specification: leader-follower



Model specification: leader-follower

- We include the following terms in the utility function

$$
\begin{aligned}
& I_{\text{v,acc}} I^L_{\text{acc}} \alpha^L_{\text{acc}} D_L^{\rho^L_{\text{acc}}} \Delta v_L^{\gamma^L_{\text{acc}}} \Delta\theta_L^{\delta^L_{\text{acc}}} \qquad + \\
& \underbrace{I_{\text{v,dec}} I^L_{\text{dec}}}\, \underbrace{\alpha^L_{\text{dec}}}\, \underbrace{D_L^{\rho^L_{\text{dec}}} \Delta v_L^{\gamma^L_{\text{dec}}} \Delta\theta_L^{\delta^L_{\text{dec}}}}.
\end{aligned}
$$

- Indicators
- Sensitivity
- Stimulus
- $\alpha^L_{\cdot}$, $\rho^L_{\cdot}$, $\gamma^L_{\cdot}$ and $\delta^L_{\cdot}$ are unknown parameters to be estimated from the data

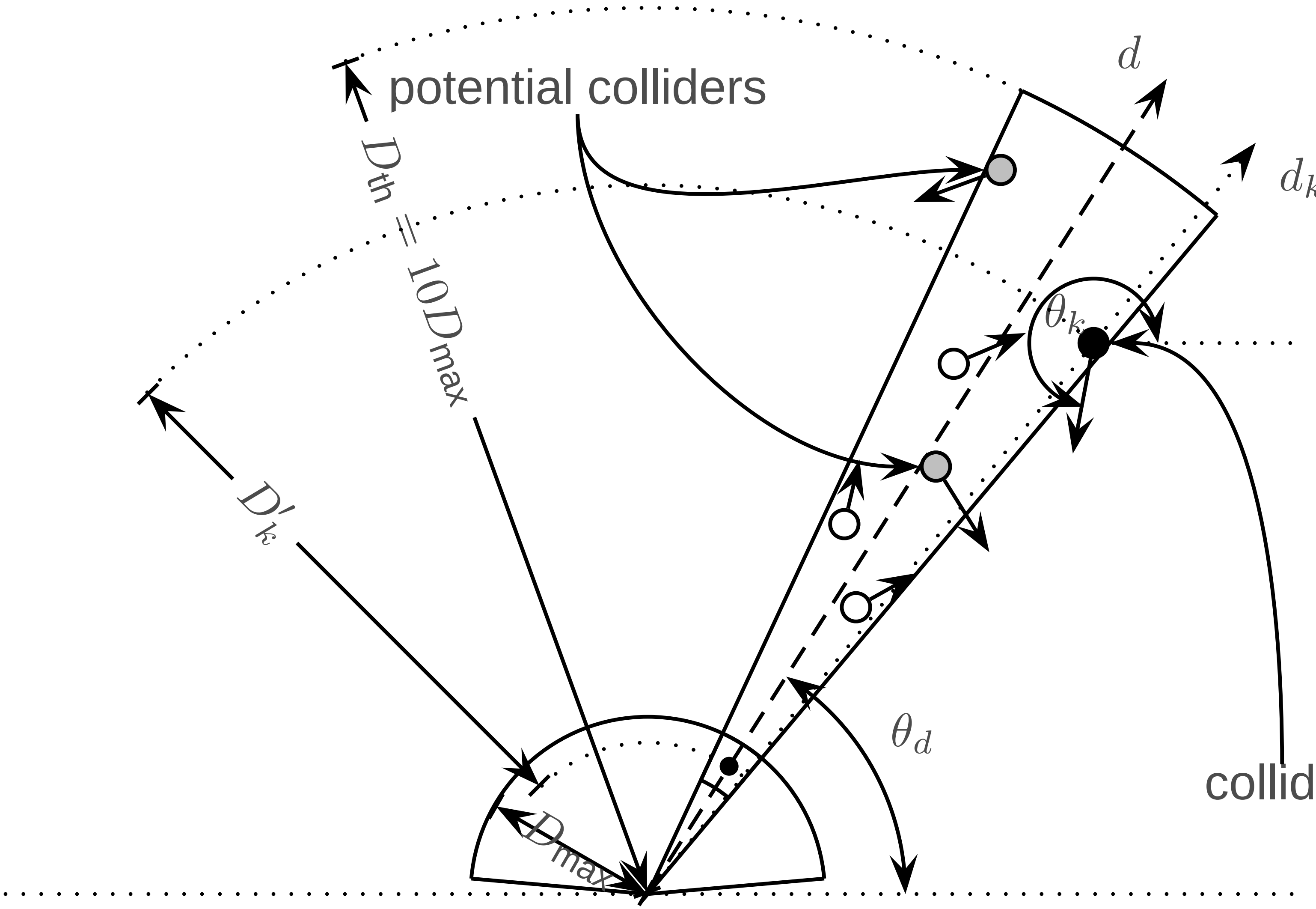
Model specification: collision avoidance

- Tendency to avoid individuals coming in the opposite direction
- In each cone, we identify potential "colliders"
- Individual k is a potential collider

$$\begin{cases} \text{if } d_l \leq d_k \leq d_r & \text{(is in the cone),} \\ \text{and } 0 < D_k \leq D'_{th} & \text{(not too far),} \\ \text{and } \frac{\pi}{2} \leq |\Delta\theta_k| \leq \pi & \text{(walking in the other direction).} \end{cases}$$

- Among them, the collider is identified as the individual k whose walking direction is the closest to the opposite direction, that is the one with $|\Delta\theta_k|$ closest to π.

Model specification: collision avoidance



Model specification: collision avoidance

We include the following terms in the utility function

$$\underbrace{I_{\mathsf{d},d_n} I_C}_{\text{Indicators}} \underbrace{\alpha_C e^{\rho_C D_C}}_{\text{Sensitivity}} \underbrace{\Delta v_C^{\gamma_C} \Delta\theta_C^{\delta_C}}_{\text{Stimulus}}.$$

- Indicators
- Sensitivity
- Stimulus
- I_{d,d_n} = 1 if $d \neq d_n$, otherwise, that is the term is zero for alternatives corresponding to walking straight ahead
- I_C = 1 if there is a collider in the cone, 0 otherwise.

Model specification

$$
\left.
\begin{aligned}
V_{vdn} = {} & \beta_{\text{dir\_central}} \text{dir}_{dn} I_{\text{central}} + \\
& \beta_{\text{dir\_side}} \text{dir}_{dn} I_{\text{side}} + \\
& \beta_{\text{dir\_extreme}} \text{dir}_{dn} I_{\text{extreme}} +
\end{aligned}
\right\} \quad \textit{keep direction}
$$

$$
\left.
\begin{aligned}
& \beta_{\text{ddist}} \text{ddist}_{vdn} + \\
& \beta_{\text{ddir}} \text{ddir}_{dn} +
\end{aligned}
\right\} \quad \textit{toward destination}
$$

$$
\left.
\begin{aligned}
& \beta_{\text{dec}} I_{\text{v,dec}} (v_n / v_{\text{max}})^{\lambda_{\text{dec}}} + \\
& \beta_{\text{accLS}} I_{\text{LS}} I_{\text{v,acc}} (v_n / v_{\text{maxLS}})^{\lambda_{\text{accLS}}} + \\
& \beta_{\text{accHS}} I_{\text{HS}} I_{\text{v,acc}} (v_n / v_{\text{max}})^{\lambda_{\text{accHS}}} +
\end{aligned}
\right\} \quad \textit{free flow acceleration}
$$

$$
\left.
\begin{aligned}
& I_{\text{v,acc}} I^L_{\text{acc}} \alpha^L_{\text{acc}} D_L^{\rho^L_{\text{acc}}} \Delta v_L^{\gamma^L_{\text{acc}}} \Delta \theta_L^{\delta^L_{\text{acc}}} + \\
& I_{\text{v,dec}} I^L_{\text{dec}} \alpha^L_{\text{dec}} D_L^{\rho^L_{\text{dec}}} \Delta v_L^{\gamma^L_{\text{dec}}} \Delta \theta_L^{\delta^L_{\text{dec}}} +
\end{aligned}
\right\} \quad \textit{leader-follower}
$$

$$
\left.
I_{\text{d},d_n} I_C \alpha_C e^{-\rho_C D_C} \Delta v_C^{\gamma_C} \Delta \theta_C^{\delta_C}
\right\} \quad \textit{collision avoidance}
$$

Model specification

Utility:

$$\begin{aligned} U_{in} &= V_{in} &+ \varepsilon_{in} \\ U_{vdn} &= V_{vdn} &+ \varepsilon_{vdn} \end{aligned}$$

- $\surd$ Specification of V_{vdn} to capture the behavioral elements
- Specification of ε_{vdn} to capture the spatial correlation

The cross-nested logit model

- Bierlaire, M. (2006). A theoretical analysis of the cross-nested logit model, Annals of Operations Research 144(1):287-300. doi:10.1007/s10479-006-0015-x [Click here]

Model specification

- Choice model:

$$P(i|C) = \sum_{m=1}^{M} \frac{\left(\sum_{j\in C} \alpha_{jm}^{\mu_m} e^{\mu_m V_j}\right)^{\frac{1}{\mu_m}}}{\sum_{n=1}^{M} \left(\sum_{j\in C} \alpha_{jn}^{\mu_n} e^{\mu_n V_j}\right)^{\frac{1}{\mu_n}}} \frac{\alpha_{im}^{\mu_m} e^{\mu_m V_i}}{\sum_{j\in C} \alpha_{jm}^{\mu_m} e^{\mu_m V_j}}$$

- $\mu_{.}$ are unknown parameters to be estimated from data
- $\alpha_{.}$ are all fixed to 0.5 in this context.

Model specification

In summary:

- The context is described by various variables
- The variables are used to associate a utility with each cell
- The utilities are used to associate a probability with each cell

Outline

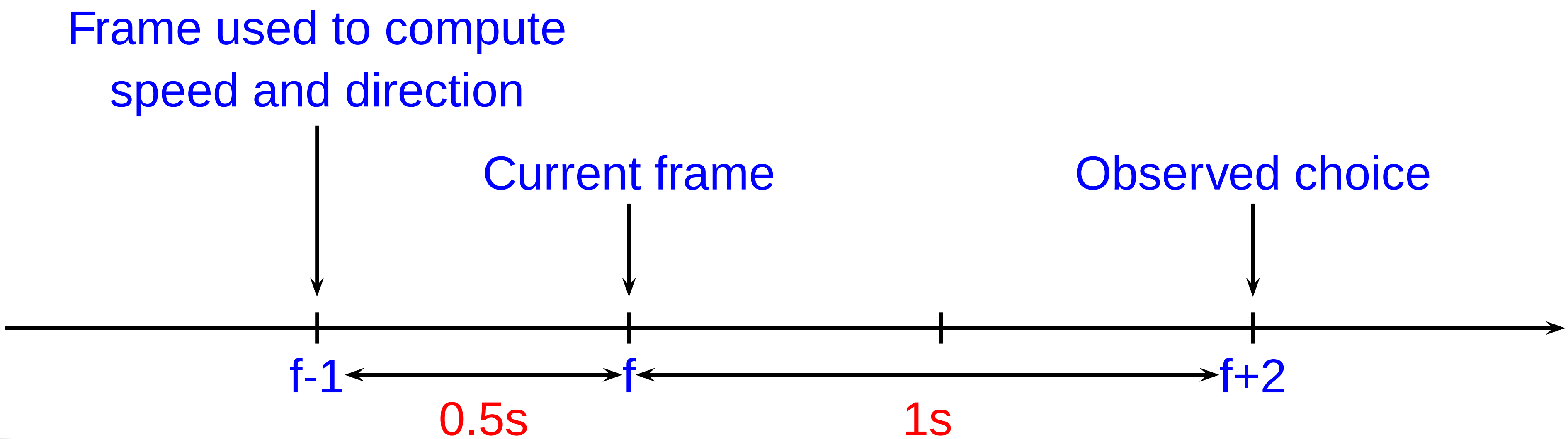
- √ Mathematical framework
- √ Modeling elements
- √ Model specification
- Estimation data
- Estimation results
- Validation
- Conclusion

Estimation data

Sendaï, Japan, August 2000 (K. Teknomo)

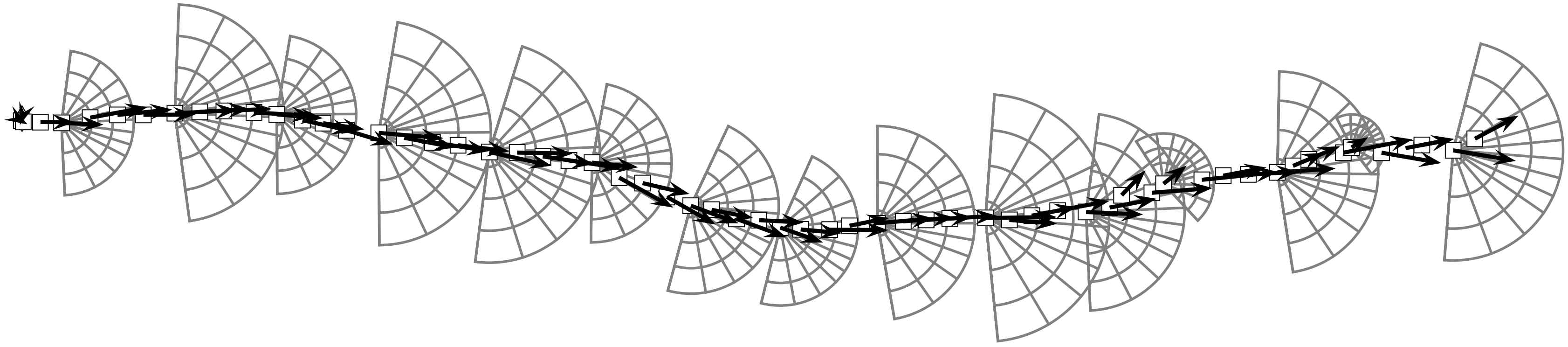
Estimation data

- 190 pedestrian trajectories
- manually tracked, frame by frame
- 10200 positions
- Two frames per second
- Data from Arsenal Research



Estimation data

Example of a trajectory with some choice sets



Estimation results

- Model estimated with biogeme
- Number of estimated parameters: 24
- Signs of the parameters consistent with our expectation

Outline

- $\surd$ Mathematical framework
- $\surd$ Modeling elements
- $\surd$ Model specification
- $\surd$ Estimation data
- $\surd$ Estimation results
- Validation
- Conclusion

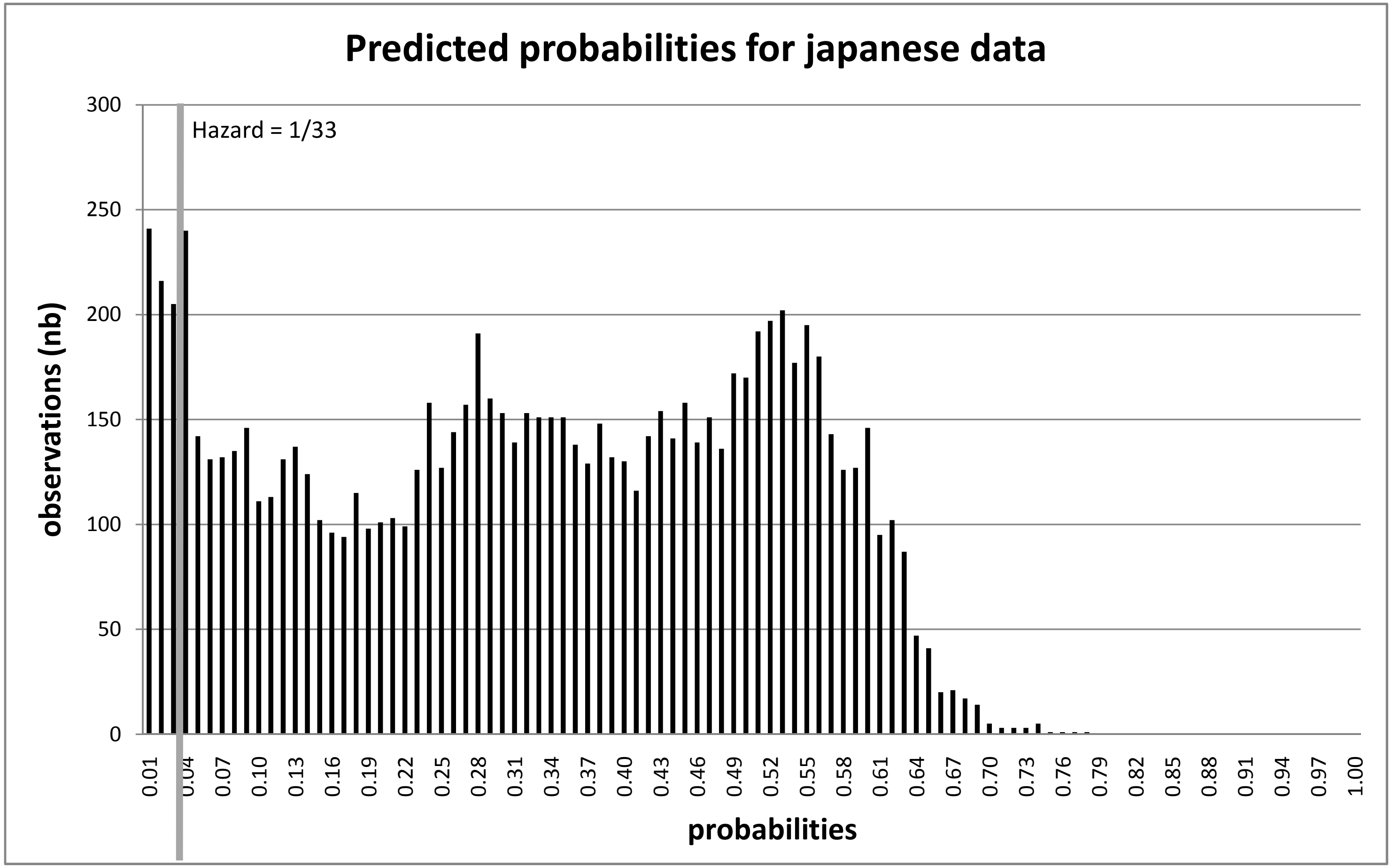
Validation

- Two data sets
  - Japanese: used for model estimation
  - Dutch: not used for model estimation
- Cross-calibration
- Compare predicted and observed choices

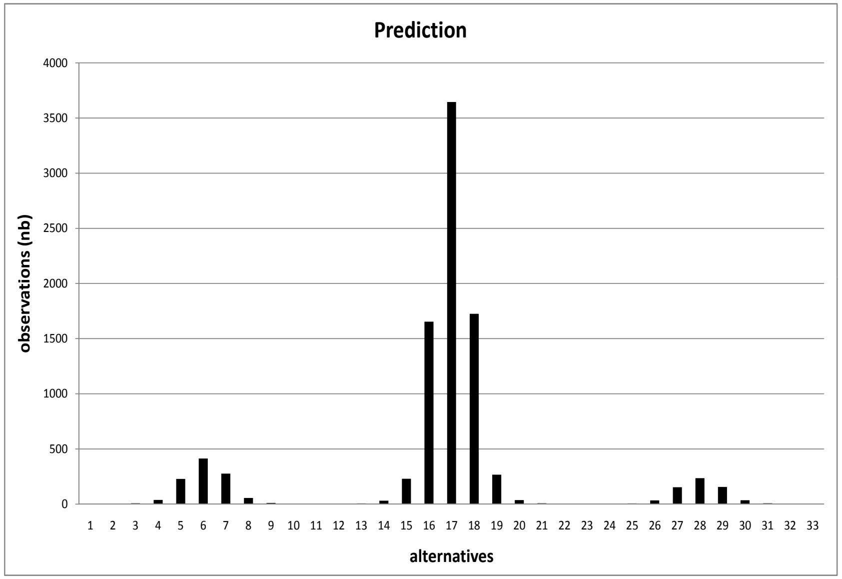
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

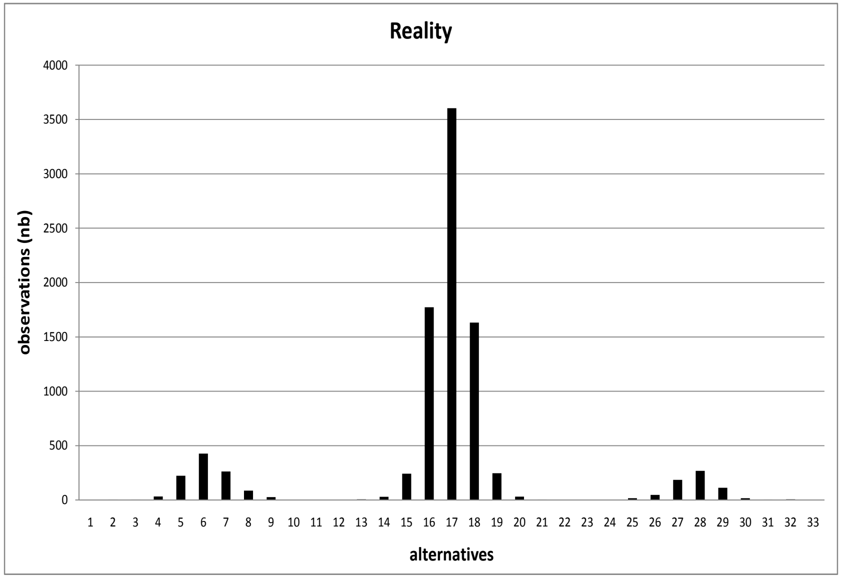
Japanese data: predicted probabilities



Japanese data: predicted probabilities



Predicted



Observed

Japanese data: cross-validation

- Verify the robustness of the specification
- Re-estimate the model on 80% of the data
- Apply it on the remaining 20%
- Do the same with a simple model which exactly replicates the shares in the data
- Outliers with full model: 7.13%
- Outliers with constant-only model: 19.90%

Model	Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5
Proposed spec.	8.78%	6.36%	7.60%	7.87%	5.87%
Constant only	20.79%	20.70%	17.13%	19.88%	18.64%

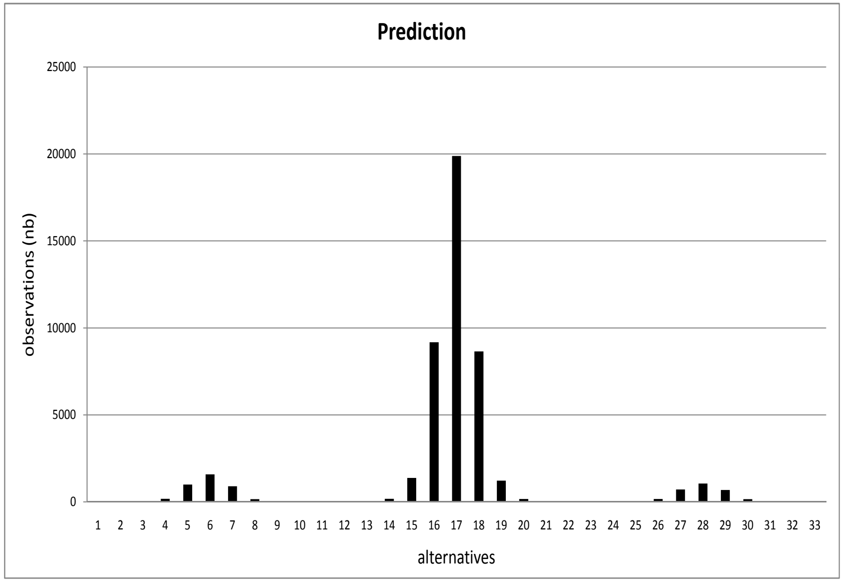
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

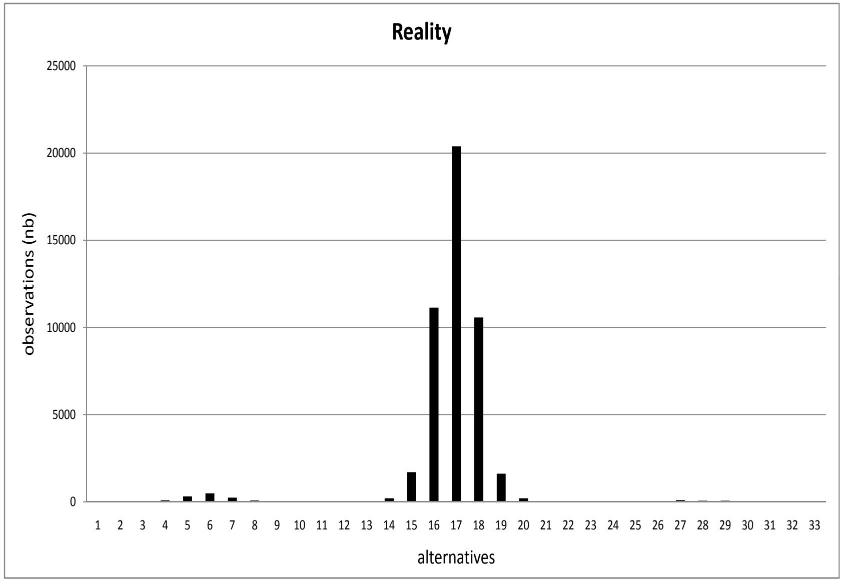
Dutch data

- Collected at TU Delft, 2000-2001
- Controlled experiment with volunteer pedestrians

Dutch data



Predicted



Observed

Conclusion

- Model for pedestrian walking behavior
- New methodological framework
- Discrete choice model – random utility model
- Specification of the utility to capture key behavioral aspects
- Parameters estimated on real data
- Model has been successfully validated on experimental data collected in TU Delft (The Netherlands)