
Validation of a discrete choice model of walking behavior

Th. Robin, M. Bierlaire, G. Antonini and J. Cruz

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

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

Modeling elements

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

- Strategic: 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

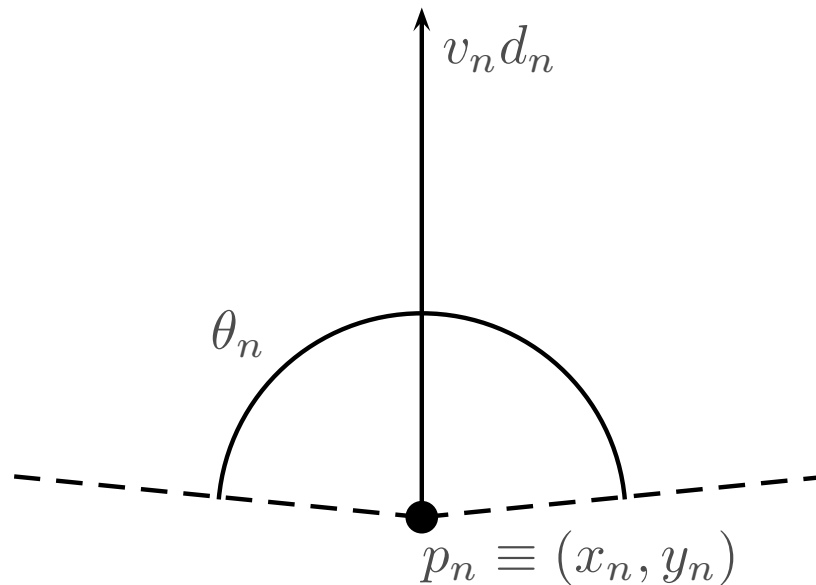
In our context

- 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

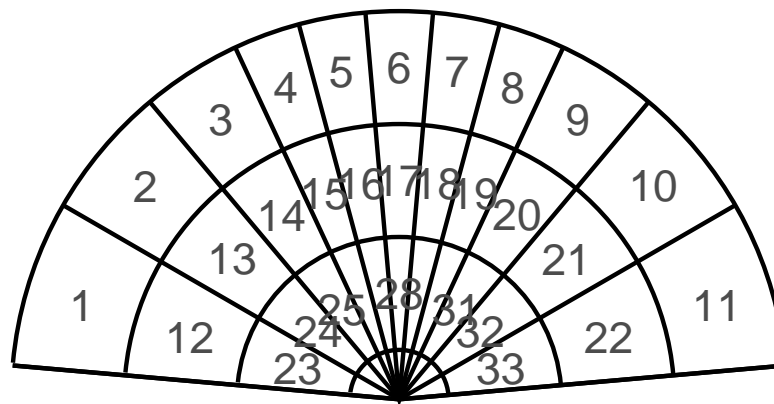
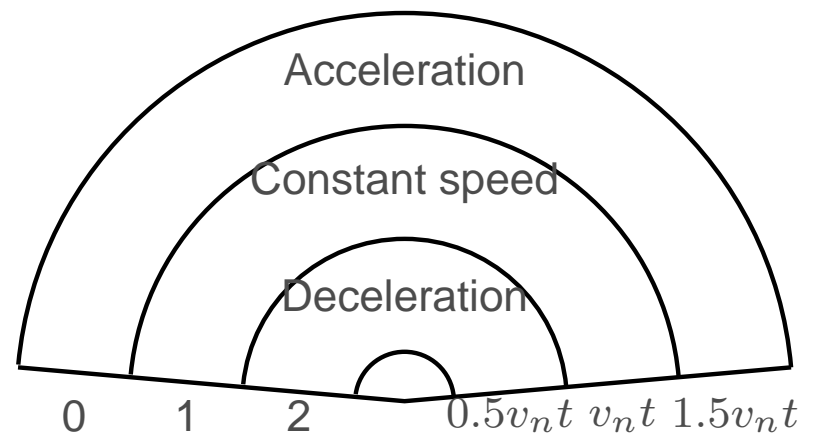
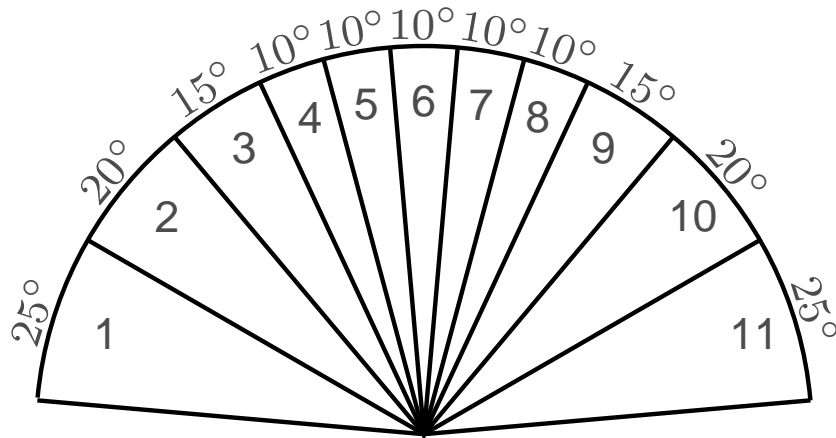
Given

- 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 $\theta_n = 170^\circ$



Modeling elements

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



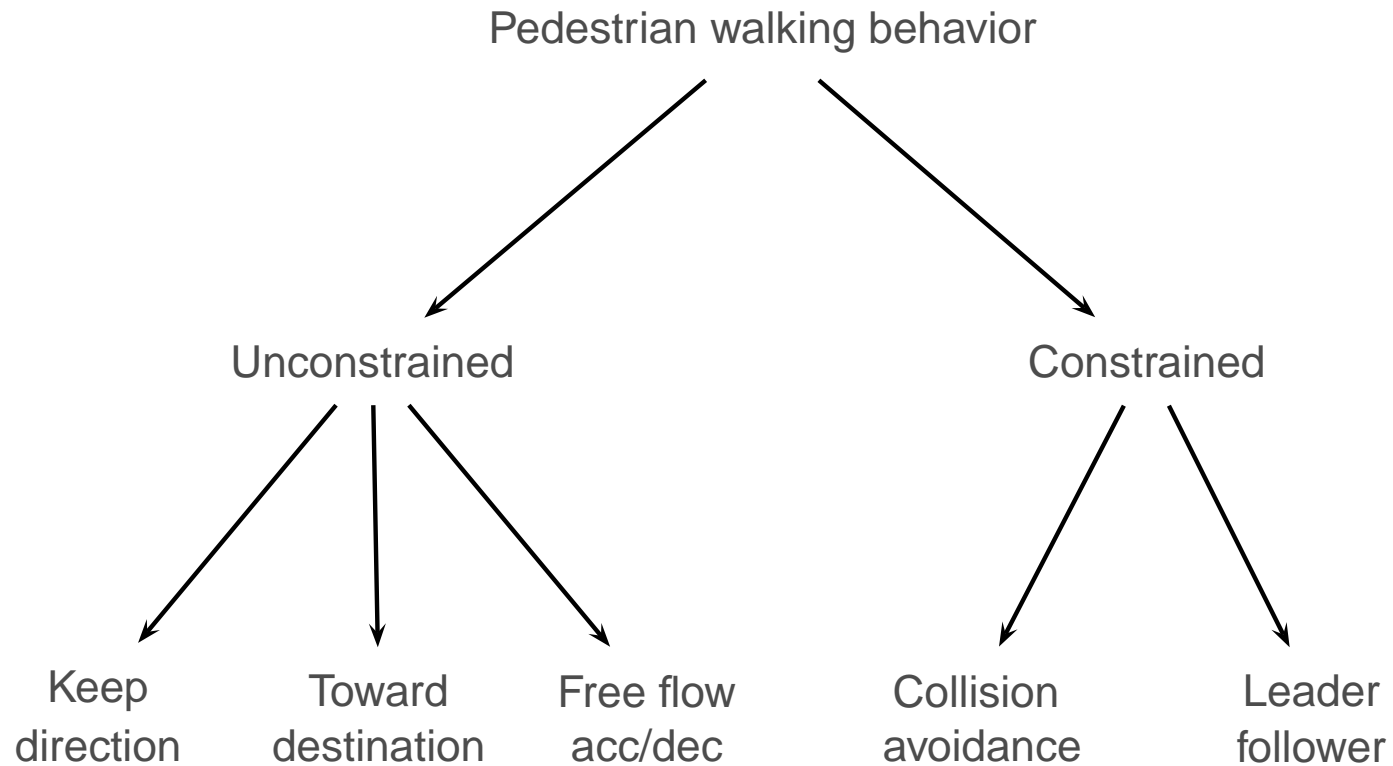
Modeling elements

- 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

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

Modeling elements

Behavioral elements



Outline

- ✓ Modeling elements
 - Model specification
 - Estimation data
 - Estimation results
 - Validation
 - Conclusion

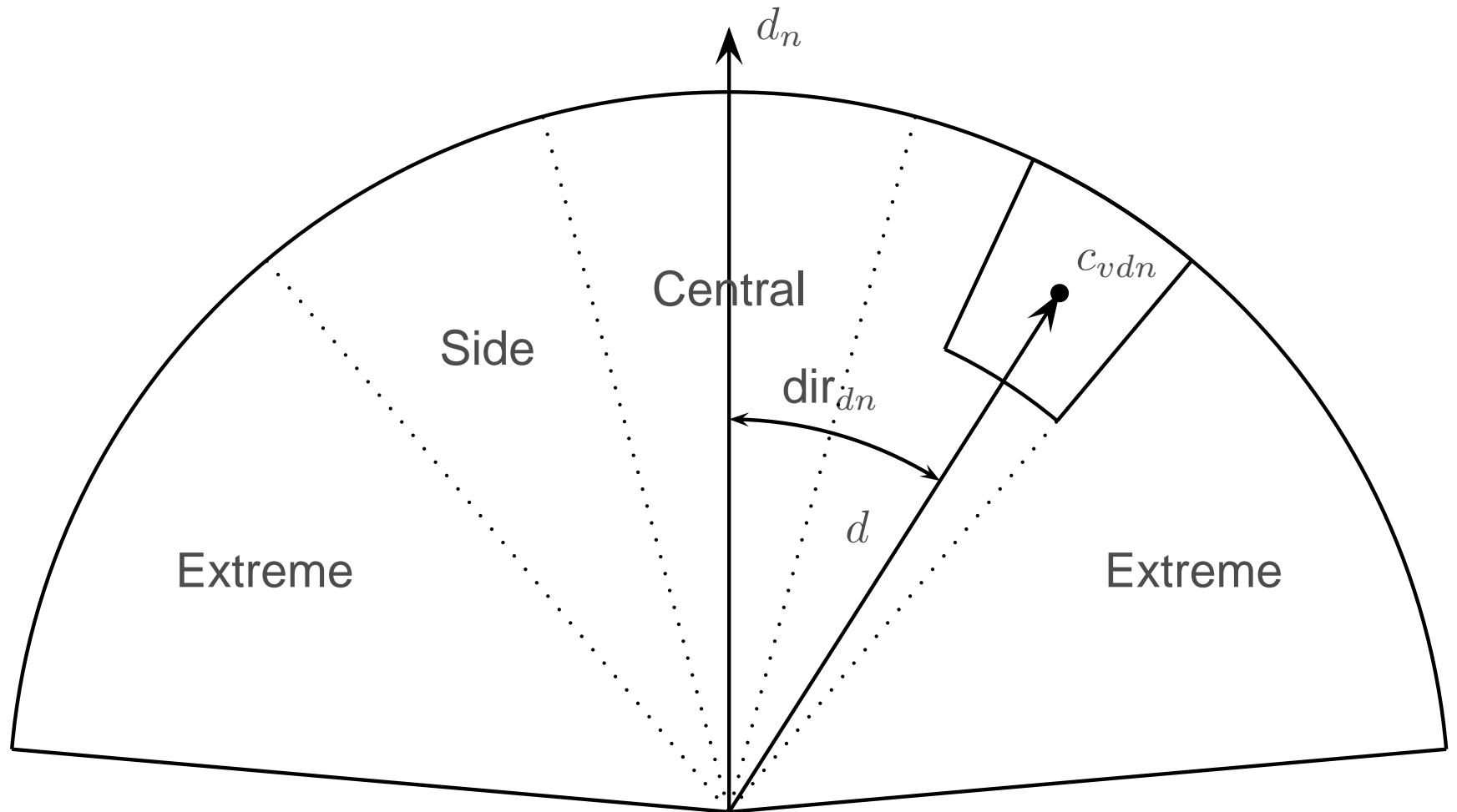
Model specification

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

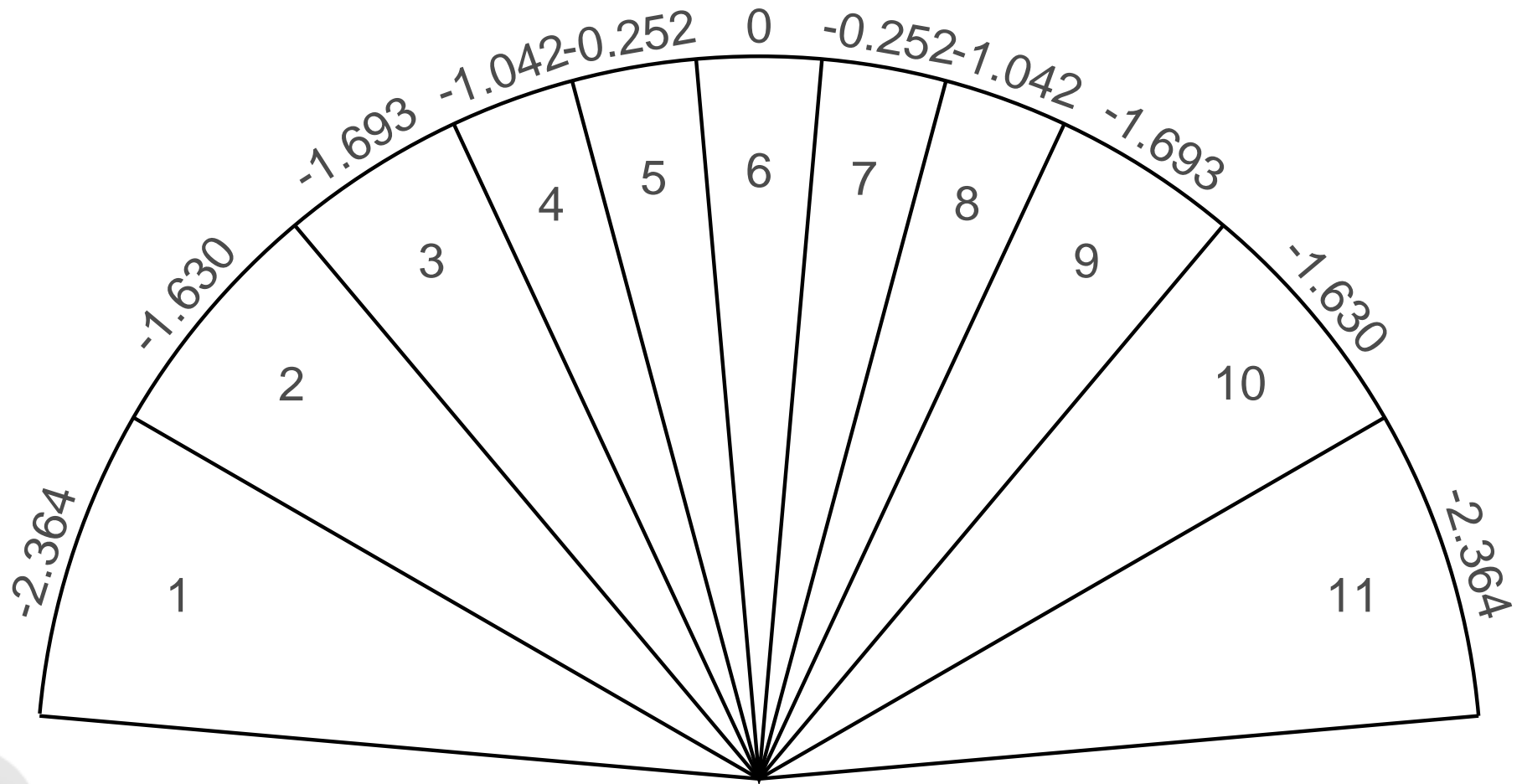
- The greater the angle, the lower the utility
- Not necessarily in a pure proportional way
- We include the following terms in the utility function

$$\beta_{\text{dir_central}} \text{dir}_{dn} I_{\text{central}} + \beta_{\text{dir_side}} \text{dir}_{dn} I_{\text{side}} + \beta_{\text{dir_extreme}} \text{dir}_{dn} I_{\text{extreme}}$$

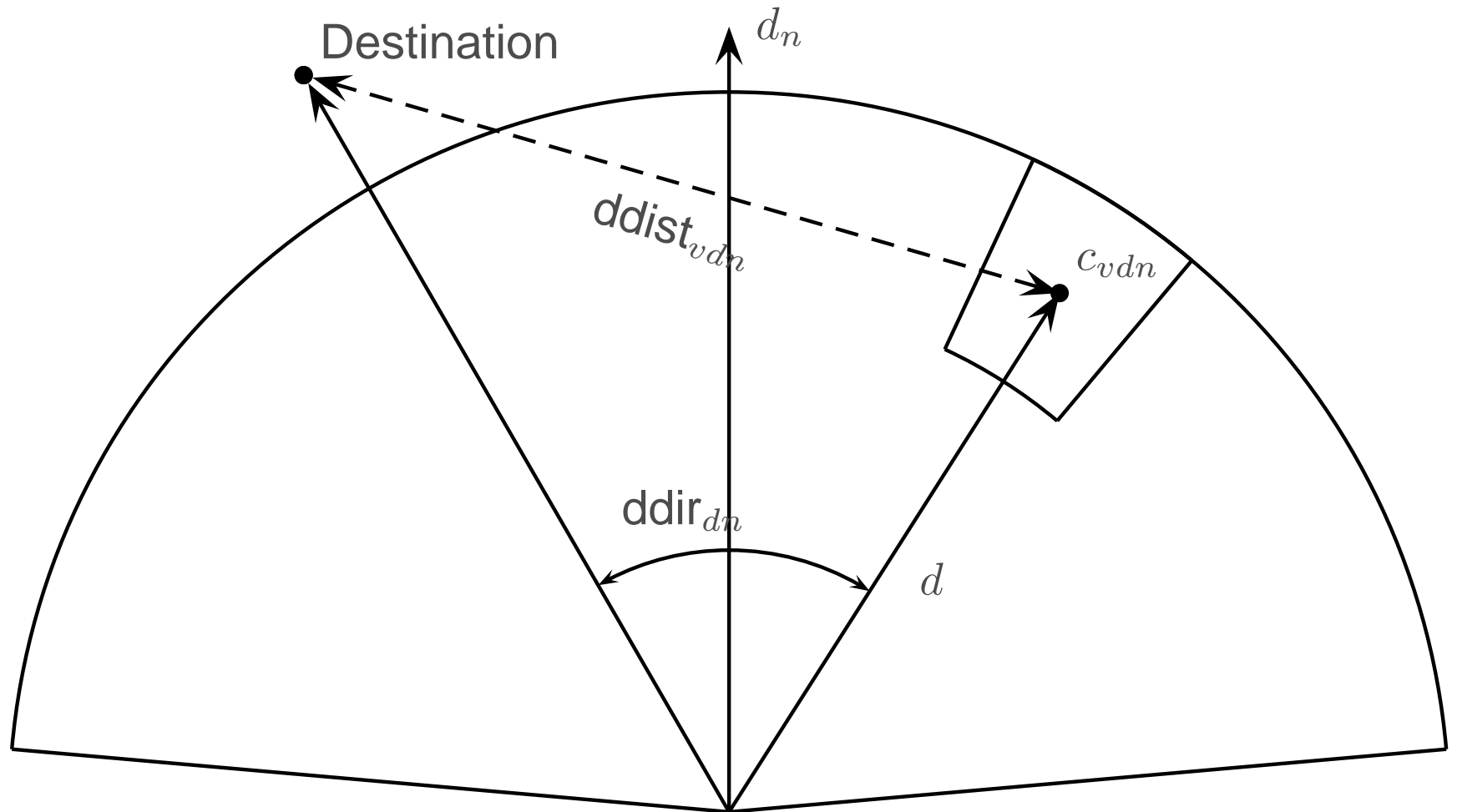
- Only one of the terms is non zero
- I_k is 1 if the alternative belongs to zone k
- β . are unknown parameters to be estimated from the data
- We expect them to be negative

Model specification: keep direction

Estimated contributions to the utility



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

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

- β . are unknown parameters to be estimated from the data
- We expect them to be negative
- Results:

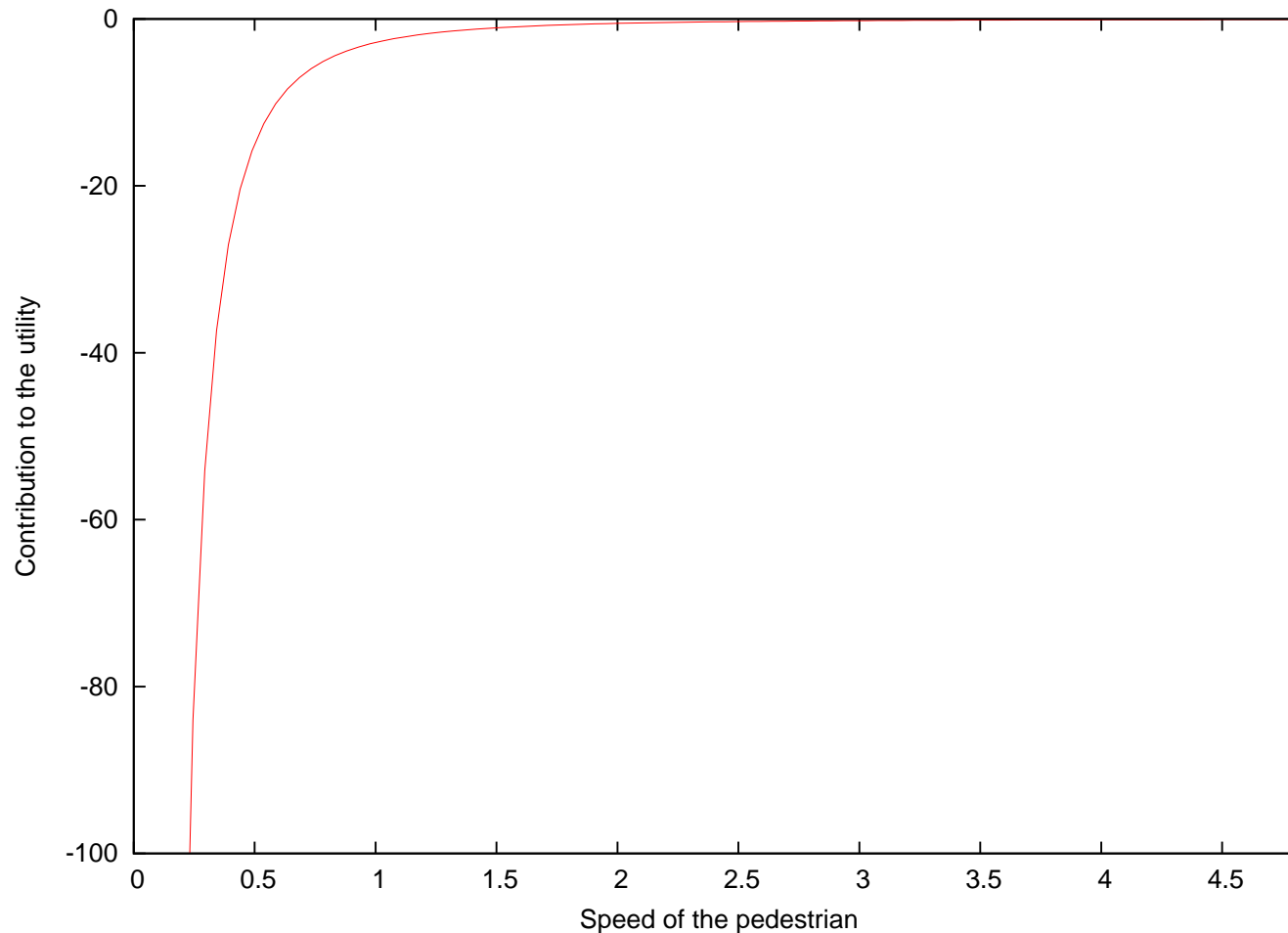
$$-1.55 \text{ ddist}_{vdn} - 0.079 \text{ ddir}_{dn}$$

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

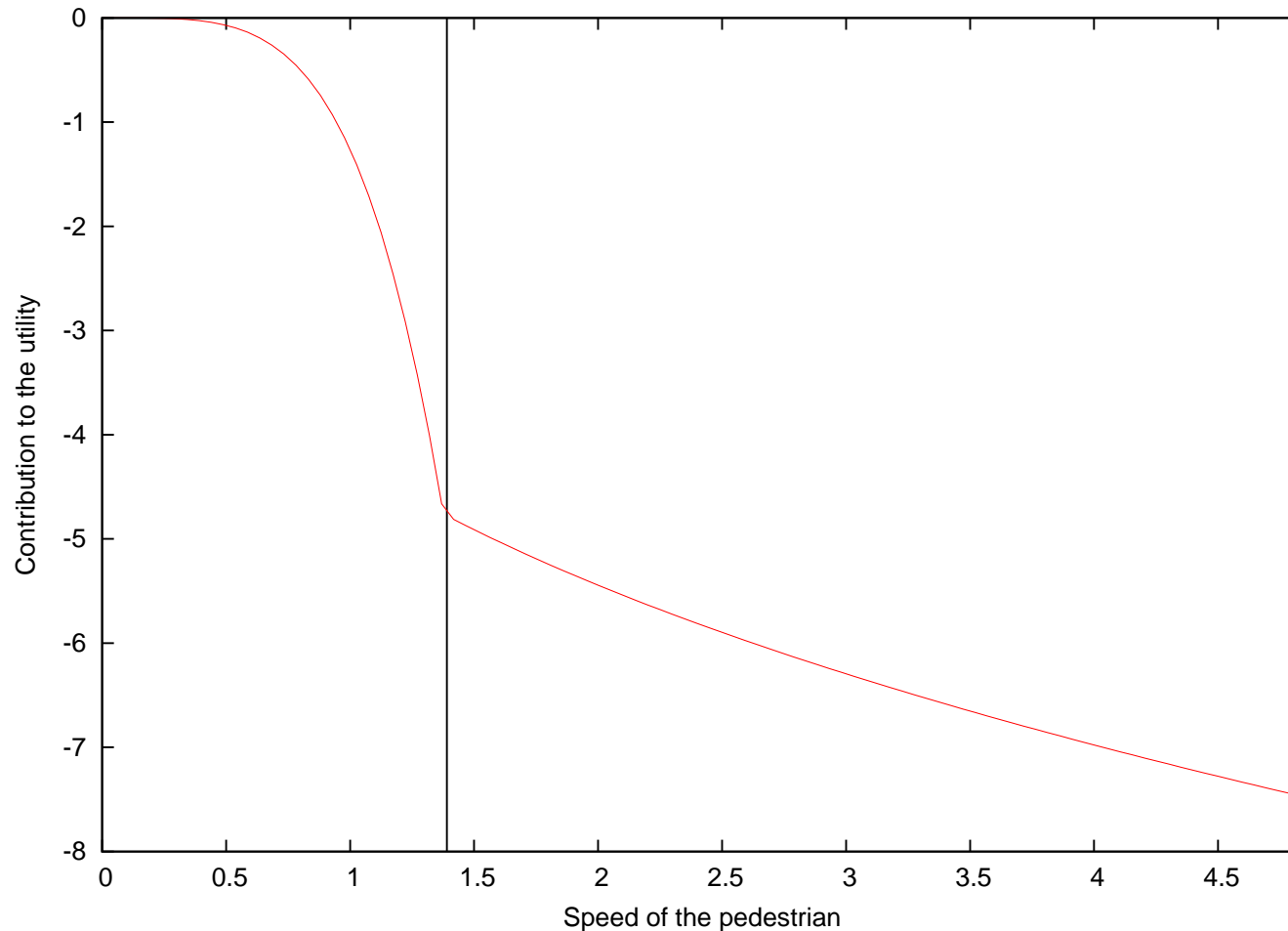
Model specification: free flow acceleration

Penalty for alternatives corresponding to deceleration



Model specification: free flow acceleration

Penalty for alternatives corresponding to acceleration



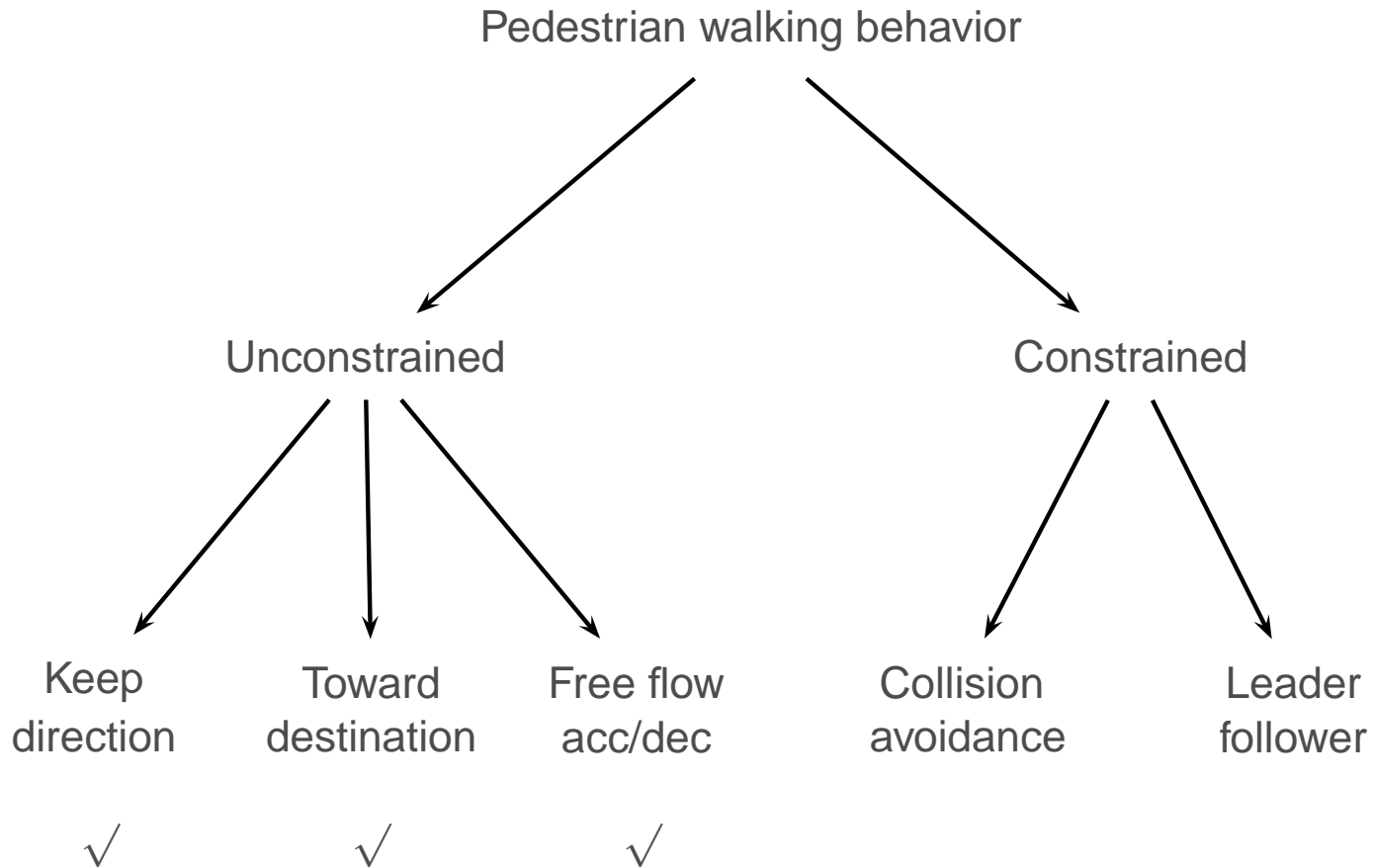
Model specification: free flow acceleration

- We include the following terms in the utility function

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

- Maximum one term is not zero for each alternative
- $I_{v,\text{dec}}$ and $I_{v,\text{acc}}$ indicates if the alternative corresponds to acceleration or deceleration
- I_{LS} and I_{HS} indicates low speed (≤ 1.39) and high speed
- β . and λ . 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)

Model specification



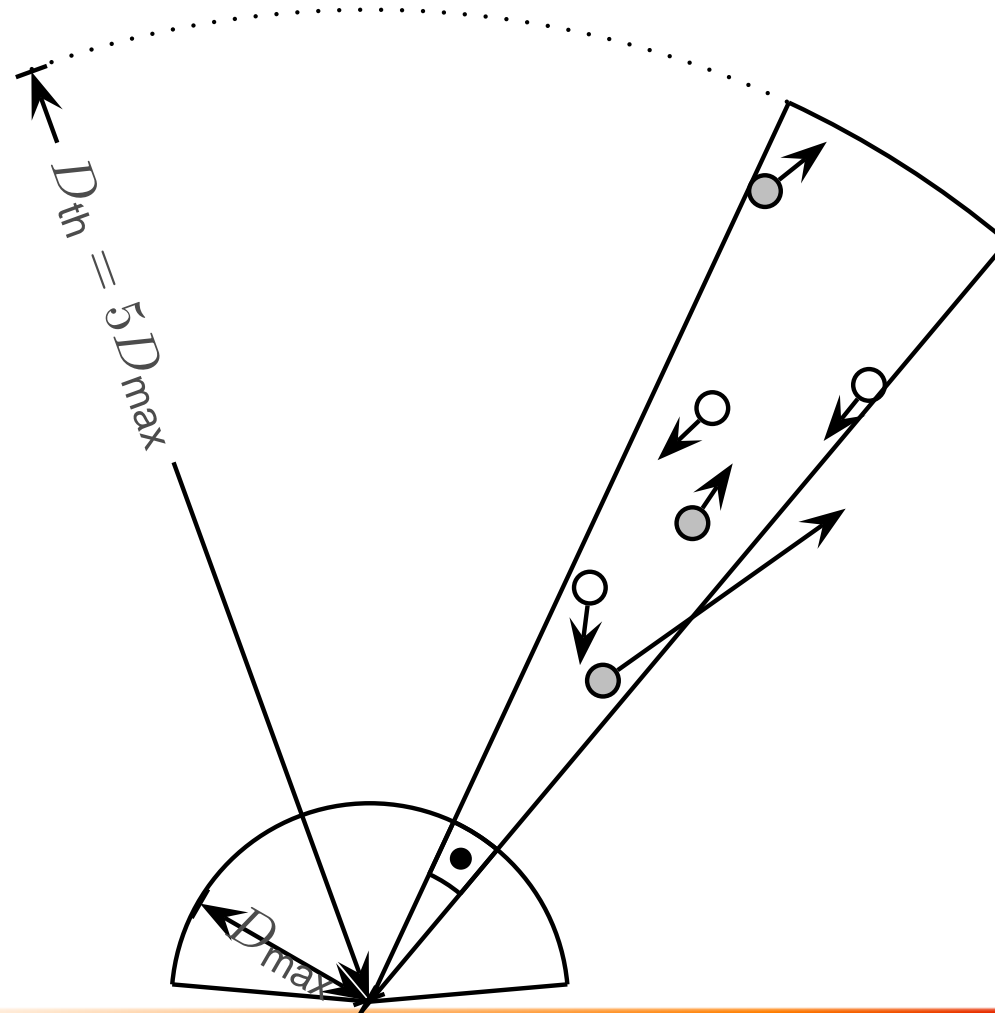
Model specification: leader-follower

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

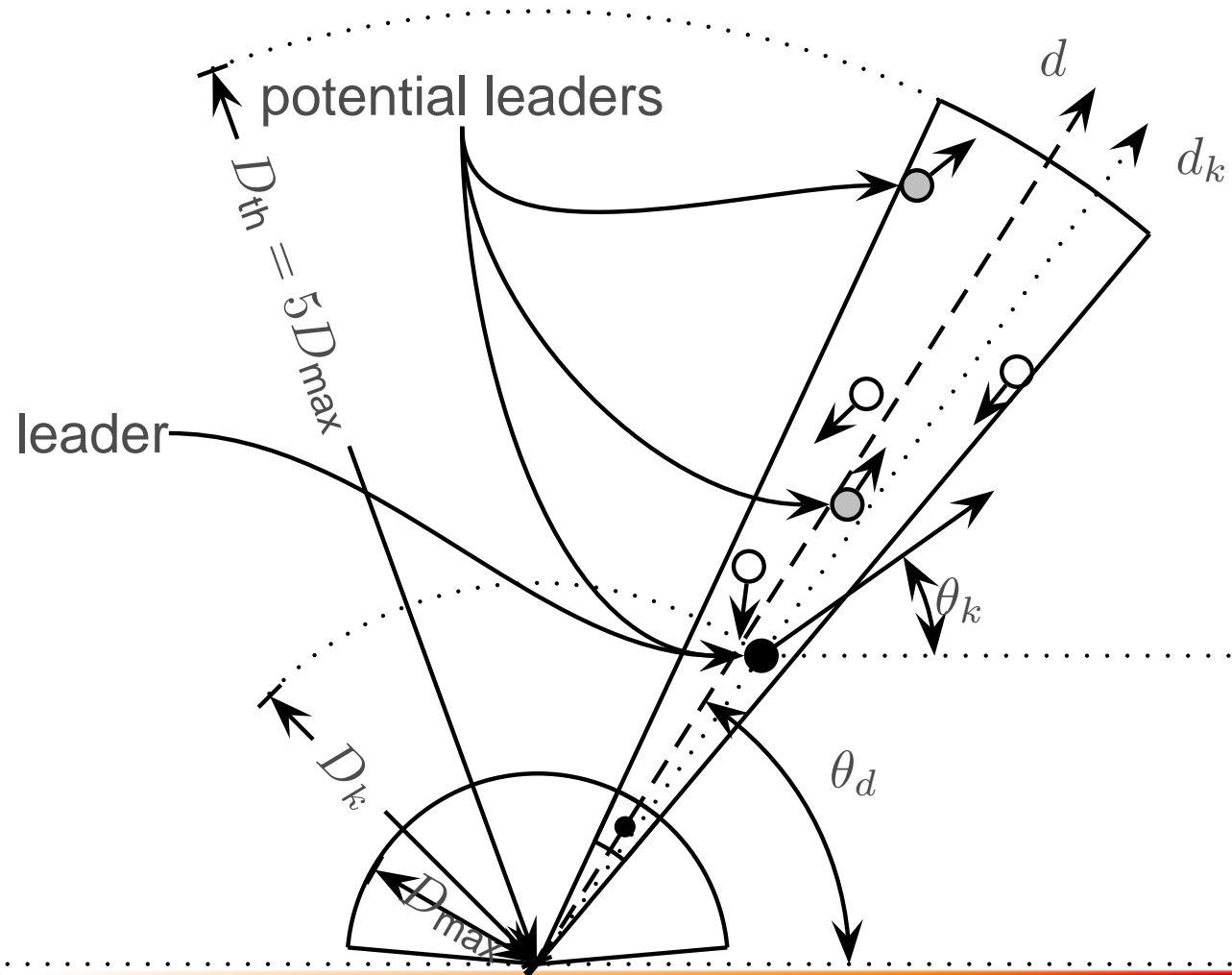
$$\left\{ \begin{array}{l} \text{if } d_l \leq d_k \leq d_r \text{ (is in the cone),} \\ \text{and } 0 < D_k \leq D_{th} \text{ (not too far),} \\ \text{and } 0 < |\Delta\theta_k| \leq \Delta\theta_{th} \text{ (walking in almost the same direction),} \end{array} \right.$$

- 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

$$\begin{array}{l}
 I_{v,acc} I_{acc}^L \alpha_{acc}^L D_L^{\rho_{acc}^L} \Delta v_L^{\gamma_{acc}^L} \Delta \theta_L^{\delta_{acc}^L} \\
 I_{v,dec} I_{dec}^L \alpha_{dec}^L D_L^{\rho_{dec}^L} \Delta v_L^{\gamma_{dec}^L} \Delta \theta_L^{\delta_{dec}^L}
 \end{array} +$$

- Indicators
- Sensitivity
- Stimulus
- α^L , ρ^L , γ^L and δ^L are unknown parameters to be estimated from the data

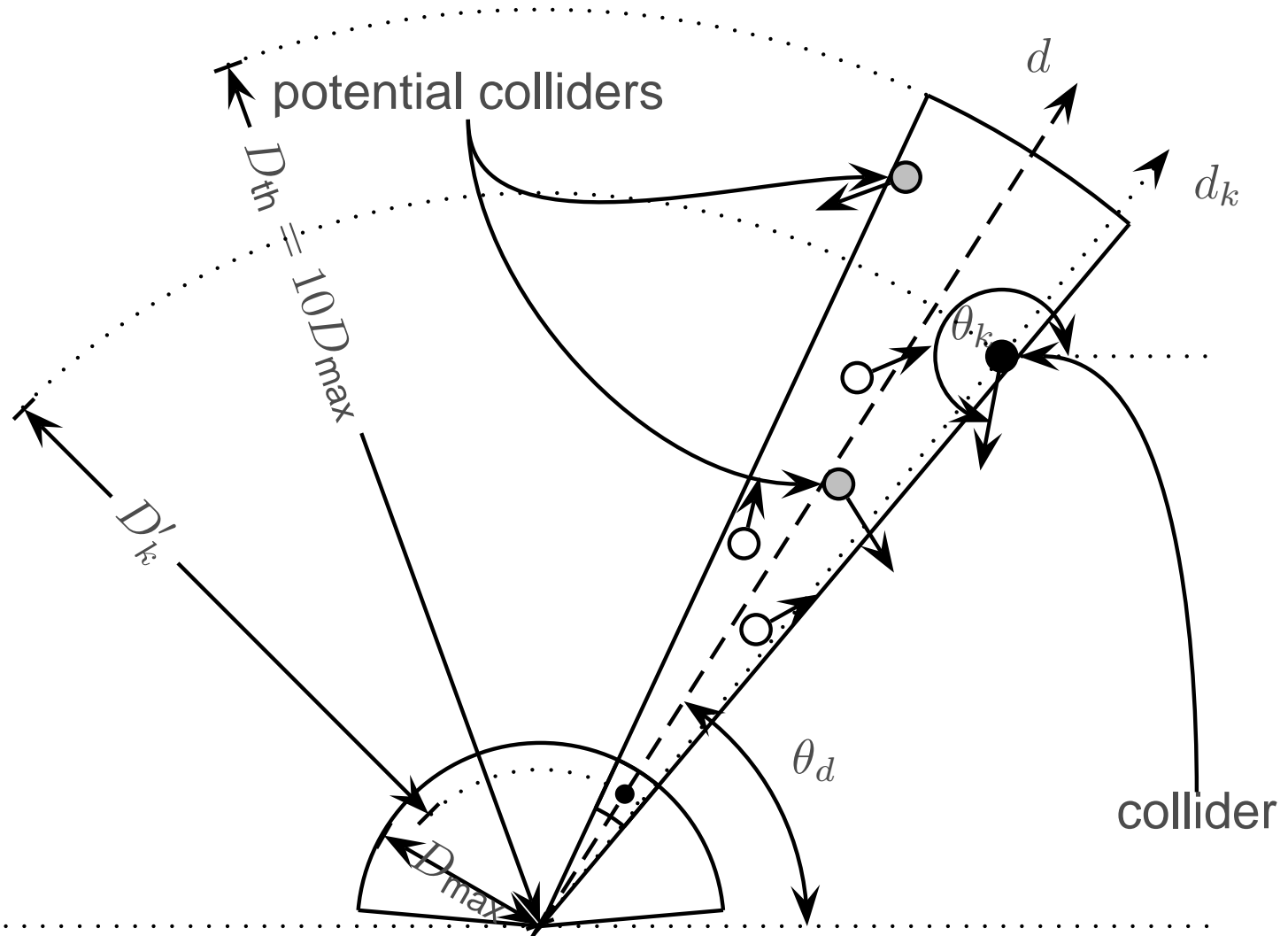
Model specification: collision avoidance

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

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

- 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

$$I_d, d_n I_C \alpha_C e^{\rho_C D_C} \Delta v_C^{\gamma_C} \Delta \theta_C^{\delta_C}.$$

- 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

$$\begin{aligned}
 V_{v_{dn}} = & \beta_{\text{dir_central}} \text{dir}_{dn} I_{\text{central}} & + & \\
 & \beta_{\text{dir_side}} \text{dir}_{dn} I_{\text{side}} & + & \\
 & \beta_{\text{dir_extreme}} \text{dir}_{dn} I_{\text{extreme}} & + & \\
 & & & \left. \vphantom{\beta_{\text{dir_central}}} \right\} \text{keep direction} \\
 & \beta_{\text{ddist}} \text{ddist}_{v_{dn}} & + & \\
 & \beta_{\text{ddir}} \text{ddir}_{dn} & + & \\
 & & & \left. \vphantom{\beta_{\text{ddist}}} \right\} \text{toward destination} \\
 & \beta_{\text{dec}} I_{v,\text{dec}} (v_n / v_{\text{max}})^{\lambda_{\text{dec}}} & + & \\
 & \beta_{\text{accLS}} I_{\text{LS}} I_{v,\text{acc}} (v_n / v_{\text{maxLS}})^{\lambda_{\text{accLS}}} & + & \\
 & \beta_{\text{accHS}} I_{\text{HS}} I_{v,\text{acc}} (v_n / v_{\text{max}})^{\lambda_{\text{accHS}}} & + & \\
 & & & \left. \vphantom{\beta_{\text{dec}}} \right\} \text{free flow acceleration} \\
 & I_{v,\text{acc}} I_{\text{acc}}^L \alpha_{\text{acc}}^L D_L^{\rho_{\text{acc}}^L} \Delta v_L^{\gamma_{\text{acc}}^L} \Delta \theta_L^{\delta_{\text{acc}}^L} & + & \\
 & I_{v,\text{dec}} I_{\text{dec}}^L \alpha_{\text{dec}}^L D_L^{\rho_{\text{dec}}^L} \Delta v_L^{\gamma_{\text{dec}}^L} \Delta \theta_L^{\delta_{\text{dec}}^L} & + & \\
 & & & \left. \vphantom{I_{v,\text{acc}}} \right\} \text{leader-follower} \\
 & I_{d,d_n} I_C \alpha_C e^{-\rho_C D_C} \Delta v_C^{\gamma_C} \Delta \theta_C^{\delta_C} & & \left. \vphantom{I_{d,d_n}} \right\} \text{collision avoidance}
 \end{aligned}$$

Model specification

Utility:

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

- ✓ Specification of V_{vdn} to capture the behavioral elements
 - Specification of ε_{vdn} to capture the spatial correlation

The cross-nested logit model

- Nests for directions
- Nests for speed regimes
- Each alternative belongs to two nests

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

- ✓ Modeling elements
- ✓ Model specification
 - Estimation data
 - Estimation results
 - Validation
 - Conclusion

Estimation data

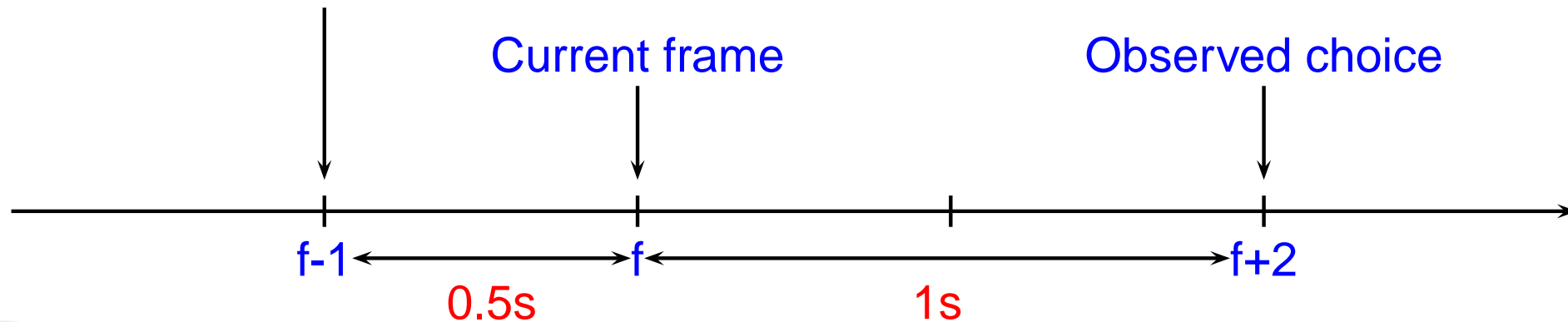


Sendai, Japan, August 2000 (K. Teknomo)

Estimation data

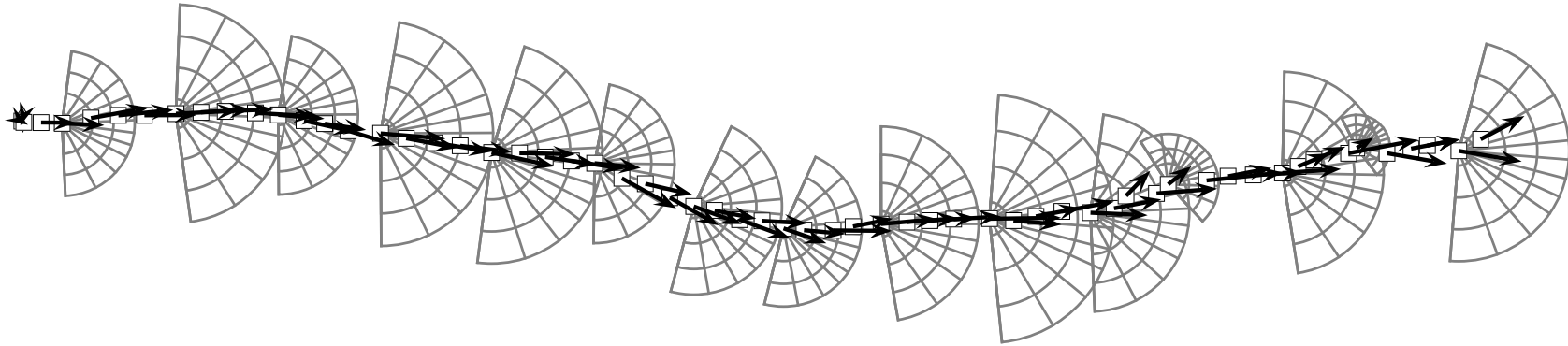
- 190 pedestrian trajectories
- manually tracked, frame by frame
- 10200 positions
- Two frames per second
- Data from Arsenal Research

Frame used to compute
speed and direction



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

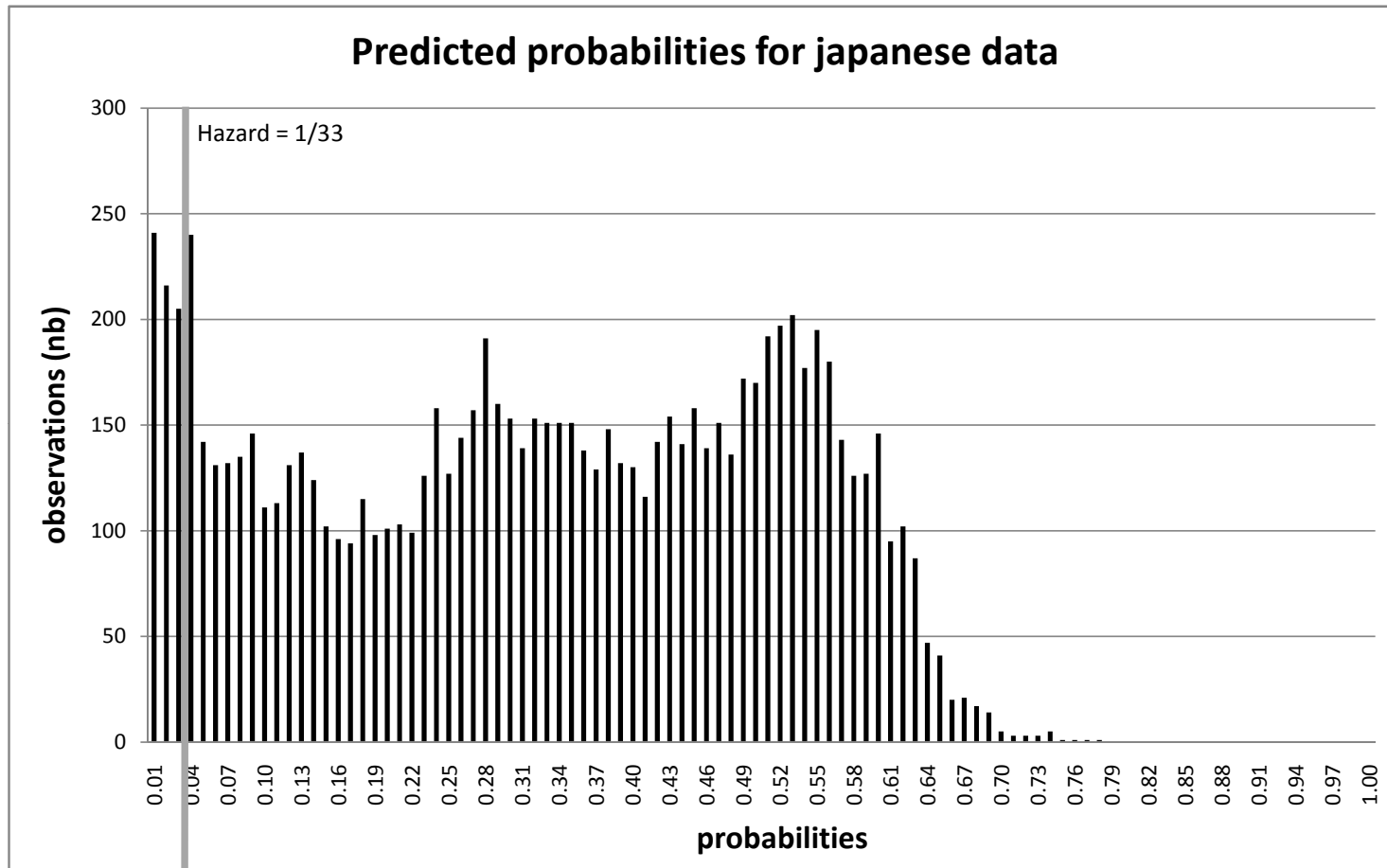
Outline

- ✓ Modeling elements
- ✓ Model specification
- ✓ Estimation data
- ✓ Estimation results
 - Validation
 - Conclusion

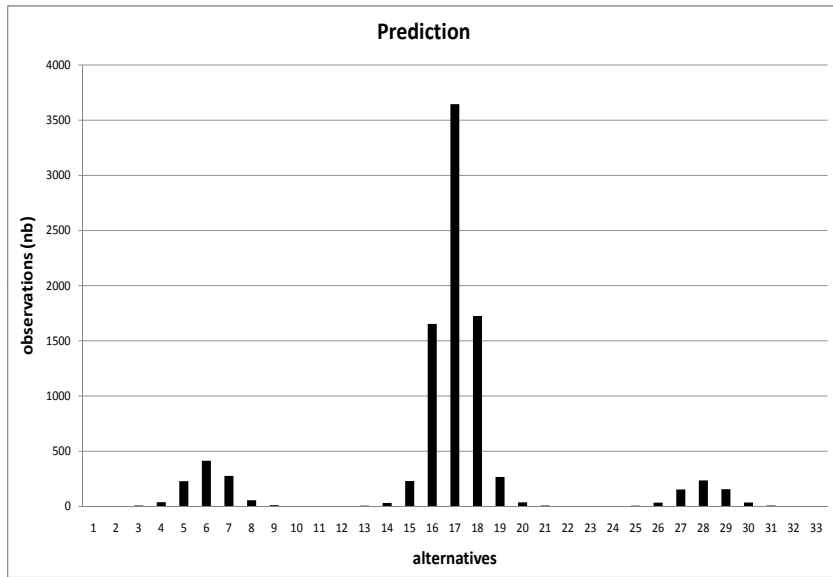
Validation

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

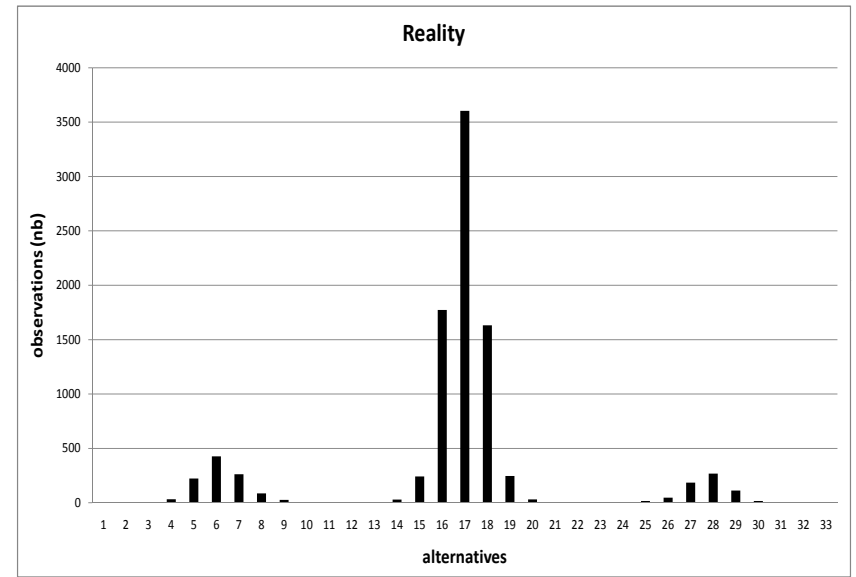
Japanese data: predicted probabilities



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%

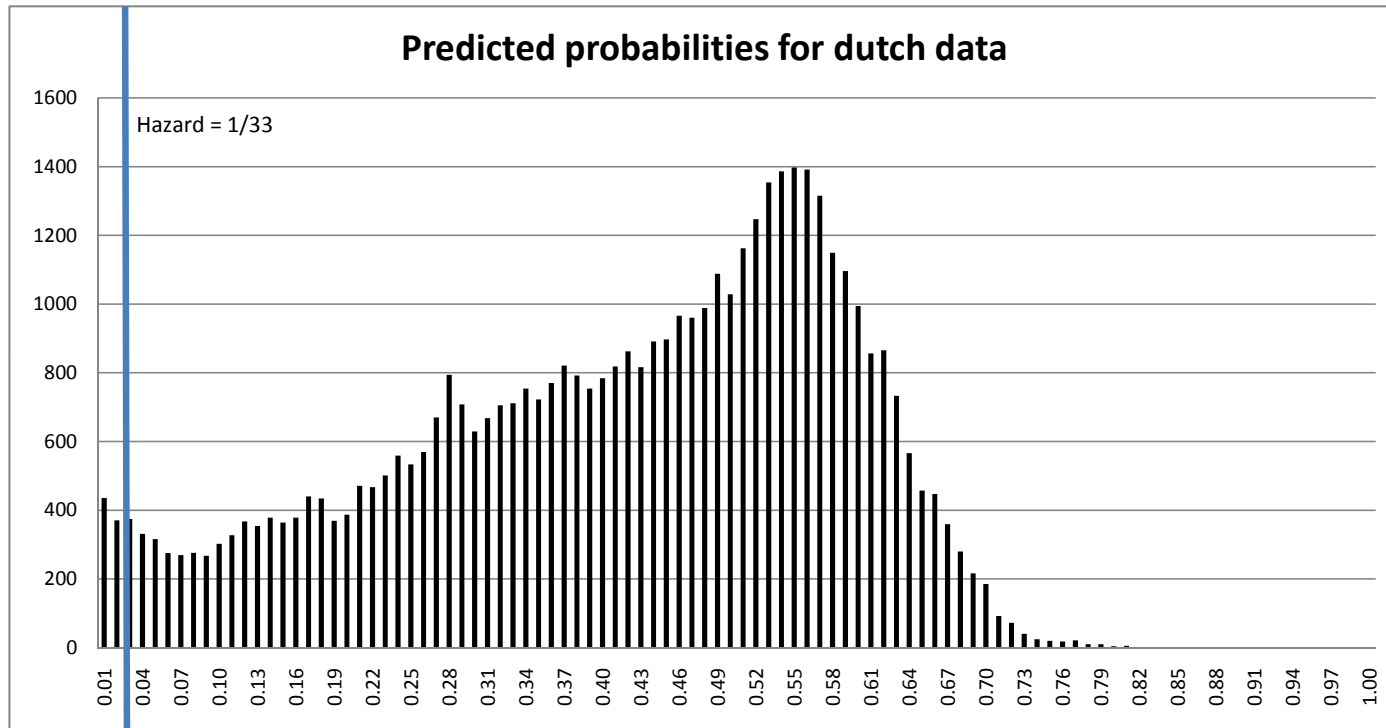
Model	Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5
Proposed spec.	8.78%	6.36%	7.60%	7.87%	5.87%
Constant only	20.79%	20.70%	17.13%	19.88%	18.64%

Dutch data

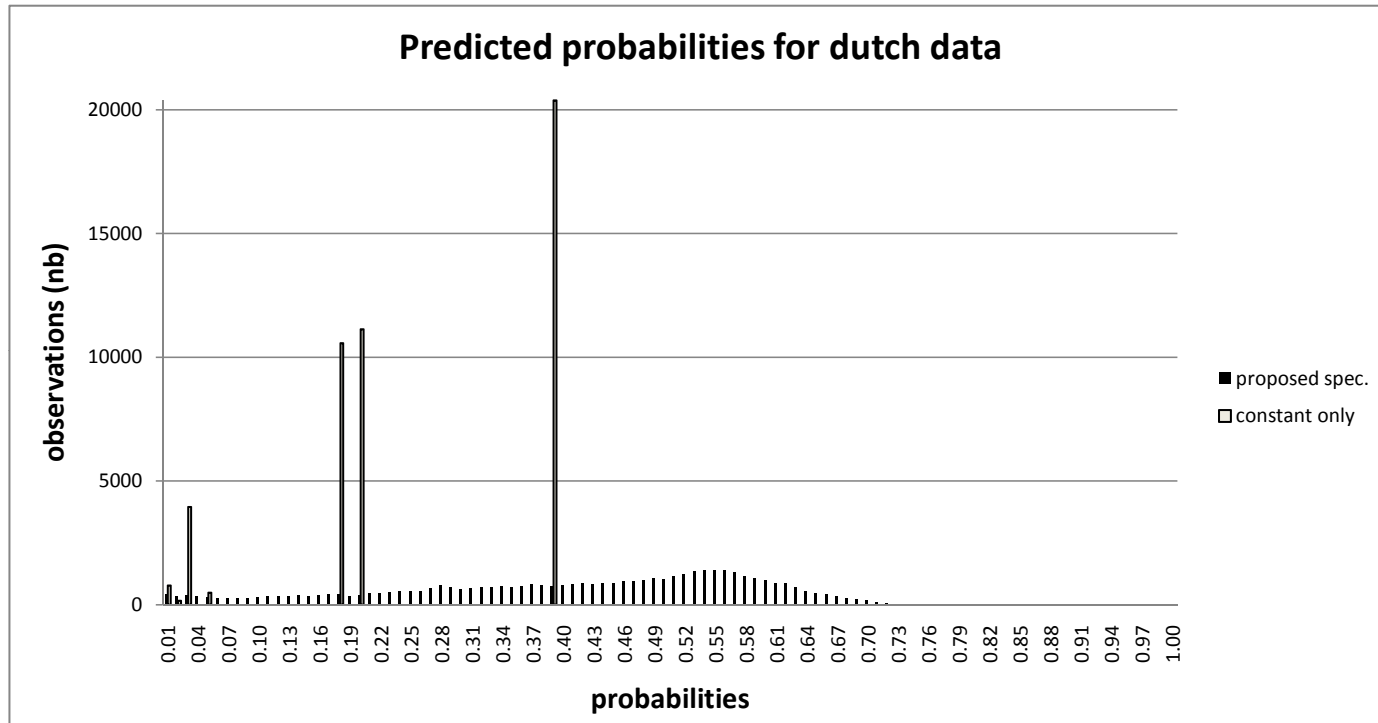
- Collected at TU Delft, 2000-2001 (Hoogendoorn & Daamen)
- Controlled experiment with volunteer pedestrians



Dutch data

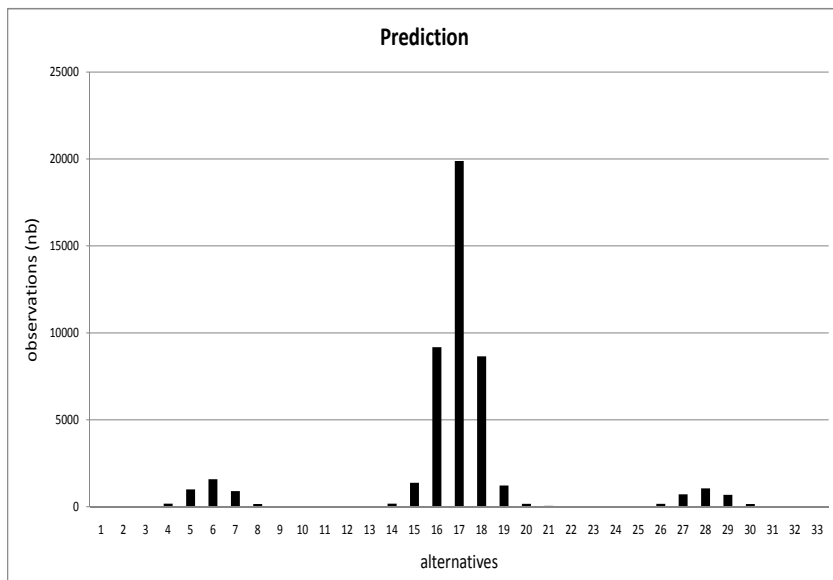


Dutch data

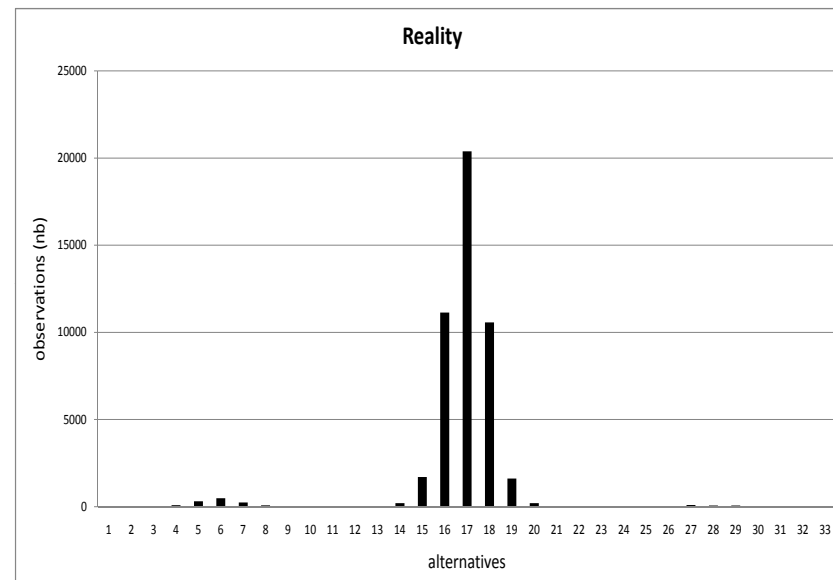


Comparison with a constant-only model

Dutch data



Predicted



Observed

Dutch data

Cone	Γ	M_Γ	R_Γ	$(M_\Gamma - R_\Gamma)/R_\Gamma$
Front	5 – 7, 16 – 18, 27 – 29	43619.98	43374	0.0057
Left	3, 4, 14, 15, 25, 26	1968.79	2089	-0.0575
Right	8, 9, 19, 20, 30, 31	1764.39	1972	-0.1053
Extreme left	1, 2, 12, 13, 23, 24	45.86	27	0.6985
Extreme right	10, 11, 21, 22, 32, 33	81.97	19	3.3144

Predicted (M_Γ) and observed (R_Γ) shares for alternatives grouped by directions with the Dutch data set.

Dutch data

Area	Γ	M_Γ	R_Γ	$(M_\Gamma - R_\Gamma)/R_\Gamma$
acceleration	1 – 11	3892.35	1273	2.0576
constant speed	12 – 22	40733.53	45869	-0.112
deceleration	23 – 33	2855.12	339	7.4222

Predicted (M_Γ) and observed (R_Γ) shares for alternatives grouped by speed regime with the Dutch data set.

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)