

# **New Challenges in Disaggregate Behavioral Modeling: Emotions, Investments and Mobility**

THÈSE N° 5040 (2011)

PRÉSENTÉE LE 27 MAI 2011  
À LA FACULTÉ SCIENCES DE BASE  
LABORATOIRE TRANSPORT ET MOBILITÉ  
PROGRAMME DOCTORAL EN MATHÉMATIQUES

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

**Thomas ROBIN**

acceptée sur proposition du jury:

Prof. T. Mountford, président du jury  
Prof. M. Bierlaire, directeur de thèse  
Dr G. Antonini, rapporteur  
Dr S. Hess, rapporteur  
Prof. J.-Ph. Thiran, rapporteur



ÉCOLE POLYTECHNIQUE  
FÉDÉRALE DE LAUSANNE

Suisse  
2011



# Acknowledgments

First, I would like to thank my supervisor Prof. Michel Bierlaire who helps and supports me all along this work. I would like to underline his perfectionism and human qualities. His intuition crucially contributes to the research developed in this thesis. I would like also to thank him for having offered me a position in the *Transport and Mobility Laboratory* in 2006, and the possibility to start a PhD.

I wish also to thank Dr. Gianluca Antonini and Prof. Jean-Philippe Thiran for the great collaboration on the work related to the Chapter 2 of the thesis. I thank the other members of the thesis committee, Dr. Stephane Hess and Prof. Thomas Mountford who gave me some precious advices and help, for improving the quality of this document.

I give thanks to Dr. Matteo Sorci for the collaboration we did on the early stage of the Chapter 3, and for the codes he provides us for collecting data and analyzing videos. I thank also Dr. Frank Crittin from Lombard Odier, for his collaboration on Chapter 2.

This thesis was supported by the Swiss National Science Foundation grant number 200021-117823.

I thank all my colleagues from the laboratory for the great work atmosphere. I specially thank my colleague from the seat behind, and one of my best friend, Javier Cruz-Mota, alias “El Profesor”. We intensively collaborate on Chapters 2 and 3, and this work would not have been possible without him. I am also very grateful to Aurélie Glerum and Dr. Michaël Thémans, my officemates who kindly support me in the last months of the work. I thank also Ricardo and Gunnar for the virtual band “the irrelevant alternatives”, which may become true.

I thank, of course, my parents Papa and Manman, as well as my sister MC\* and my brother Pierrot for their constant support and love in all my student life. They help me a lot in managing the work pressure.

I give thanks to my dear Sarah, as well as her parents Jean-Luc and Rose-Marie, her brothers Raphaël, Jérémie, her sister Adrienne, as well as Vèreène, and Jenny for their welcome, support and the great time in Lausanne and in the Valais.

Finally, I would like to thank my colleagues and friends from the nautic center team: Mitch, Ben, Franck and Anoy for the great leisure time, and the every day pasta on the side of the Geneva lake. I enjoyed a lot their company, as well as the different sport and teaching activities we performed together.



# Abstract

This thesis tackles new challenges associated with the disaggregate modeling of the human behavior. Decision-aid tools help in making decisions, by providing quantitative insights on the decisions and associated consequences. They are useful in complex situations where human actors are involved. Inside decision-aid tools, there is a need for explicitly capturing and predicting the human behavior. The prediction of human actions is done through models. Models are simplified representations of the reality, which provide a better understanding of it and allow to predict its future state. They are often too simplistic, with bad prediction capabilities. This is an issue as they generate the outcome of the decision-aid tools, which influence decisions. Good models are required in order to adequately capture the complexity of human actions. Behavioral models appear to be relevant. They allow to translate behavioral assumptions into equations, which make their strength but also their complexity. They have been mainly used in transportation and marketing.

Many advances have been recently achieved. On one hand, emerging technologies allow to collect various and detailed data about the human behavior. On the other hand, new modeling techniques have been proposed to handle complex behaviors. Estimation softwares are now available for their estimation. The combination of these advances open open opportunities in the field of the behavioral modeling.

The motivations of the proposed work are the investigation of the challenges associated with non-traditional applications of the behavioral modeling, the emphasis of multi-disciplinarity, the handling of the behavioral complexity and the development of operational models. Different applications are considered where these challenges appear. The applications are the investors' behavior, the walking behavior and the dynamic facial expression recognition. Challenges are addressed in the different tasks of the modeling framework, which are the data collection, the data processing, the model specification, estimation and validation.

The modeling of the investors' behavior consists in characterizing how individuals are taking financial decisions. It is relevant for predicting monetary gains and regulating the market. We propose an hybrid discrete choice framework for modeling decisions of investors performed on stock markets. We focus on the choice of action (buy or sell) and the duration until the next action. The choice of action is handled with a binary logit model with latent classes, while a Weibull regression model is used for the duration until the next action. Both models account for the risk perception

and the dynamics of the phenomenon. They are simultaneously estimated by maximum likelihood using real data. The predictive performance of the models are tested by cross-validation. The forecasting accuracy of the action model is studied more in details. Parameters of both models are interpretable and emphasize interesting behavioral mechanisms related to investors' decisions. The good prediction capabilities of the action model in a real context makes it operational.

The modeling of the walk apprehends how a person is choosing her next step. It is useful to simulate the behavior of crowds, which is relevant for the urban planning and the design of infrastructures. We specify, estimate and validate a model for pedestrian walking behavior, based on discrete choice modeling. Two main types of behavior are identified: *unconstrained* and *constrained*. By unconstrained, we refer to behavior patterns which are independent from other individuals. The constrained patterns are captured by a *leader-follower* model and by a *collision avoidance* model. The spatial correlation between the alternatives is captured by a cross nested logit model. The model is estimated by maximum likelihood on a real data set of pedestrian trajectories, manually tracked from video sequences. The model is successfully validated using another data set of bi-directional pedestrian flows.

The dynamic facial expression recognition consists in characterizing the facial expression of a subject in a video. This is relevant in human machine interfaces. We model it using a discrete choice framework. The originality is based on the explicit modeling of causal effects between the facial features and the recognition of the expression. Five models are proposed, based on different assumptions. The first assumes that only the last frame of the video triggers the choice of the expression. In the second model, one frame is supposed to trigger the choice. The third model is an extension of the second model. It assumes that the choice of the expression results from the average of expression perceptions within a group of frames. The fourth and fifth models integrate the panel effect inherent to the estimation data and are respectively based on the first and second models. The models are estimated by maximum likelihood using facial videos. Parameters are interpretable. Labeling data on the videos has been obtained using an internet survey. The prediction capabilities of the models are studied and compared, by cross-validation using the estimation data. The results are satisfactory, emphasizing the relevance of the models in a real context.

The thesis contributes to fields. Challenges of the behavioral modeling have been investigated in complex contexts. Original and multi-disciplinary modeling approaches have been successfully proposed for each application. Model specifications have been developed to handle the behavioral complexity, allowing to quantify behavioral mechanisms. Operational models are proposed. Complex behavioral models are used in a predictive context and a detailed validation methodology has been set.

**Keywords:** human behavior, discrete choice model, latent class , application, finance, pedestrian, facial expression recognition , data collection, specification, validation, computer vision , challenge, decision-aid tool, multi-disciplinarity, dynamics

# Résumé

Cette thèse porte sur les nouveaux défis de la modélisation du comportement humain. Les outils d'aide à la décision assistent les individus dans leurs choix, en fournissant des informations quantitatives sur les alternatives possibles et leurs conséquences. Ils sont utiles dans les situations complexes impliquant des acteurs humains. La compréhension et la prédiction du comportement humain servent à l'élaboration d'outils d'aide à la décision. La prédiction se fait par le biais de modèles qui sont des représentations simplifiées de la réalité. En général, ces modèles sont simplistes et caractérisés par un faible pouvoir prédictif. De bons modèles sont nécessaires pour prendre en compte la complexité des actions humaines. Les modèles comportementaux sont adaptés à ce contexte. Leur force provient du fait que les comportements sont directement traduits en équations. Mais cela les rend également très complexes. Ils ont montré leur potentiel dans le domaine du transport et du marketing.

De nombreuses avancées ont été récemment réalisées dans différents domaines. D'une part, les nouvelles technologies permettent de collecter des informations détaillées sur les actions humaines; d'autre part, de nouvelles techniques de modélisation ont été proposées, ainsi que des logiciels permettant leur estimation. La combinaison des deux génèrent de nouveaux défis dans le domaine de la modélisation du comportement humain.

Ce travail est motivé par plusieurs objectifs: l'investigation des défis de la modélisation du comportement dans les applications non traditionnelles, la mise en avant de la multi-disciplinarité, la modélisation de phénomènes complexes ainsi que le développement de modèles opérationnels. Nous considérons différentes applications où les nouveaux défis de la modélisation du comportement se manifestent. Ils sont explorés dans chacune des étapes du processus de modélisation. Les applications considérées sont la finance, le mouvement des piétons et la reconnaissance dynamique d'expressions faciales.

En finance, nous modélisons le comportement des investisseurs sur les marchés d'actions. Ceci est utile pour la prédiction de gains ou la régulation des marchés. Nous nous intéressons à la décision d'acheter ou de vendre, ainsi qu' à la durée entre deux décisions consécutives. Nous proposons une approche dans laquelle le processus de décision est considéré dans son ensemble. Un modèle de choix à classes latentes est utilisé pour le choix entre achat et vente, alors qu'une régression de Weibull s'avère appropriée pour la durée. Les deux modèles prennent en compte la perception du

risque, ainsi que la dynamique du phénomène. Ils sont estimés simultanément par maximum de vraisemblance sur des données réelles. Le pouvoir de prédiction des modèles est testé par validation croisée et le modèle de choix est étudié plus en détails. Les paramètres des deux modèles sont interprétables et mettent en avant des mécanismes comportementaux intéressants. Par sa qualité, le modèle de choix est directement opérationnel.

La modélisation des mouvements de piétons consiste à analyser comment des personnes choisissent leurs prochains pas. Ceci est utile dans la planification urbaine et la conception d'infrastructures. Nous proposons, estimons et validons un modèle de mouvement de piétons basé sur un modèle de choix discret. Deux types de comportement sont pris en compte dans le modèle: contraint et non contraint. Le comportement contraint est relatif aux interactions avec les autres piétons, au contraire du non contraint. Un modèle cross nested logit est proposé pour prendre en compte la corrélation spatiale entre les alternatives de pas. Le modèle est estimé par maximum de vraisemblance en utilisant des trajectoires de piétons extraites de vidéos. Il est ensuite validé avec succès en utilisant un second jeu de données.

La reconnaissance dynamique d'expressions faciales consiste à caractériser l'expression faciale d'un sujet dans une vidéo. Ceci s'avère utile dans les interfaces homme-machine. Nous la modélisons en utilisant des modèles de choix discret. L'originalité de l'approche est basée sur la modélisation explicite des relations de causes à effets entre les mesures faciales et l'expression identifiée, ainsi que la prise en compte de l'ambiguïté des expressions. Cinq modèles sont proposés, basés sur différentes hypothèses concernant la dynamique du processus de décision. Les modèles sont estimés par maximum de vraisemblance en utilisant des vidéos de visages. La labélisation de ces vidéos avec des expressions s'est faite par le biais d'une enquête internet. Les paramètres des modèles sont interprétables et reflètent le jugement humain. Les pouvoirs prédictifs des modèles sont étudiés et comparés par validation croisée. Les résultats étant satisfaisants, les modèles sont utilisables dans des applications réelles.

Plusieurs contributions sont apportées dans cette thèse. De nouveaux défis de la modélisation du comportement humain ont été investigués. Pour chaque application, des approches originales et multi-disciplinaires ont été conduites. Des modèles détaillés sont proposés, et les mécanismes comportementaux quantifiés. Les modèles développés sont opérationnels. Des modèles complexes ont été utilisés dans un cadre prédictif. Une méthode de validation détaillée est proposée et appliquée.

**Mots-clés:** modélisation du comportement humain, modèle de choix discret, applications, défis, modèle à classe latente, finance, piétons, reconnaissance dynamique d'expressions faciales, collecte de données, spécification, vision assistée par ordinateur, validation, dynamique, outil d'aide à la décision, multi-disciplinarité.



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Context . . . . .	1
1.2	Motivation . . . . .	2
1.3	General framework . . . . .	3
1.4	Applications . . . . .	4
1.5	Contributions . . . . .	6
1.6	Thesis outline . . . . .	8
<b>2</b>	<b>The financial behavior: modeling of investors' decisions</b>	<b>9</b>
2.1	Introduction . . . . .	9
2.2	Motivation . . . . .	9
2.3	Raw data . . . . .	12
2.4	Notations . . . . .	16
2.4.1	Variables . . . . .	18
2.4.2	Parameters . . . . .	19
2.5	Time discretization . . . . .	20
2.6	Explanatory variables . . . . .	22
2.7	Model specification . . . . .	27
2.7.1	The risk perception . . . . .	29
2.7.2	The action model . . . . .	31
2.7.3	The duration model . . . . .	32
2.7.4	The likelihood function . . . . .	34
2.7.5	Alternative specifications . . . . .	35
2.8	Model estimation . . . . .	35
2.8.1	The action model . . . . .	35
2.8.2	The duration model . . . . .	37
2.9	Model prediction . . . . .	38
2.9.1	Prediction on the estimation data . . . . .	38
2.9.2	Cross-validation . . . . .	39
2.9.3	Simulation . . . . .	41
2.10	Contributions . . . . .	43
2.11	Conclusion . . . . .	43

<b>3</b>	<b>The walking behavior: modeling of pedestrian movements</b>	<b>51</b>
3.1	Introduction . . . . .	51
3.2	Motivation . . . . .	51
3.3	Modeling framework . . . . .	53
3.4	The model . . . . .	55
3.4.1	Keep direction . . . . .	57
3.4.2	Toward destination . . . . .	58
3.4.3	Free flow acceleration . . . . .	59
3.4.4	Leader-follower . . . . .	60
3.4.5	Collision avoidance . . . . .	63
3.4.6	The error term . . . . .	65
3.4.7	Intermediary specification steps . . . . .	66
3.5	Data . . . . .	66
3.6	Estimation results . . . . .	70
3.7	Model validation . . . . .	71
3.7.1	Japanese data set: validation of the model . . . . .	71
3.7.2	Japanese data set: validation of the specification . . . . .	72
3.7.3	Dutch data set: validation of the model . . . . .	74
3.8	Contributions . . . . .	82
3.9	Conclusions . . . . .	82
<b>4</b>	<b>The judgmental behavior: dynamic facial expression recognition</b>	<b>85</b>
4.1	Introduction . . . . .	85
4.2	MOTIVATION . . . . .	85
4.3	Data . . . . .	87
4.4	Models specification . . . . .	96
4.4.1	The reduced model . . . . .	96
4.4.2	The latent model . . . . .	98
4.4.3	The smoothed model . . . . .	102
4.4.4	Models with panel effect . . . . .	105
4.5	Intermediary specification steps . . . . .	108
4.6	Estimation of the models . . . . .	109
4.6.1	The reduced model . . . . .	110
4.6.2	The latent model . . . . .	110
4.6.3	The smoothed model . . . . .	111
4.6.4	Models with panel effect . . . . .	111
4.6.5	Comparison of the five models . . . . .	112
4.7	Prediction capability . . . . .	113
4.7.1	Aggregate prediction . . . . .	113
4.7.2	Cross-validation . . . . .	115
4.7.3	Disaggregate prediction . . . . .	117
4.8	Contributions . . . . .	123
4.9	Conclusions and Perspectives . . . . .	124

<b>5 Conclusion</b>	<b>127</b>
5.1 Review of the main results . . . . .	127
5.2 Future research . . . . .	130
<b>Appendices</b>	
<b>A The modeling of investors' decisions</b>	<b>133</b>
<b>B The dynamic facial expression recognition</b>	<b>137</b>
B.1 Notations . . . . .	137
B.2 Estimation results . . . . .	138
<b>Bibliography</b>	<b>147</b>



# List of Tables

1.1	Tasks performed for each application in the thesis . . . . .	6
2.1	Number of transactions, number of companies and percentage of transactions per fund in the raw data . . . . .	13
2.2	Number of actions, number of companies and percentages of actions per fund in the processed data . . . . .	22
2.3	Generic and significant correlations between the action variable and the explanatory variables . . . . .	26
2.4	Generic and significant correlations between the action variable and the explanatory variables, for a high level of VIX . . . . .	27
2.5	General estimation results of the action and duration models . . . . .	36
2.6	Starting dates, ending dates and number of actions per subset of data for the cross-validation . . . . .	39
2.7	Results of the cross-validation performed on the estimation data . . . . .	40
2.8	Percentages of badly predicted observations per experience of cross-validation . . . . .	40
2.9	Results of the simulations performed with the action model on the period going from 2009.07.24 to 2010.09.13 . . . . .	41
2.10	Percentages of error (see Equation (2.39)) based on simulations of Table 2.9 . . . . .	42
3.1	Speed statistics(m/sec) . . . . .	68
3.2	Number of chosen steps in each nest for the real data set . . . . .	69
3.3	<b>CNL</b> estimation results for the Japanese data set . . . . .	71
3.4	Predicted ( $M_{\Gamma}$ ) and observed ( $R_{\Gamma}$ ) shares for alternatives grouped by directions with the Japanese data set . . . . .	73
3.5	Predicted and observed shares for alternatives grouped by speed regime with the Japanese data set. . . . .	74
3.6	Summary of the cross-validation performed on the Japanese data set . . . . .	74
3.7	Number of chosen steps in each nest for Dutch data . . . . .	75
3.8	Comparison between Japanese and Dutch data sets for the observations proportions in the direction's cones . . . . .	76

3.9	Predicted ( $M_\Gamma$ ) and observed ( $R_\Gamma$ ) shares for alternatives grouped by directions with the Dutch data set. . . . .	77
3.10	Predicted ( $M_\Gamma$ ) and observed ( $R_\Gamma$ ) shares for alternatives grouped by speed regime with the Dutch data set. . . . .	78
3.11	Average pedestrian speed in the data sets . . . . .	80
3.12	Predicted ( $M_\Gamma$ ) using the constant-only model and observed ( $R_\Gamma$ ) shares for alternatives grouped by direction with the Dutch data set. . . . .	81
3.13	Predicted ( $M_\Gamma$ ) using the constant-only model and observed ( $R_\Gamma$ ) shares for alternatives grouped by speed regime with the Dutch data set. . . . .	81
3.14	Loglikelihood of each model applied to the two data sets . . . . .	81
4.1	Correspondences between measures on masks 4.5(b) and 4.5(a) . . . . .	94
4.2	Expressions Descriptive Units . . . . .	95
4.3	General estimation results . . . . .	113
4.4	Percentages of badly predicted observations on the estimation data . . . . .	115
4.5	Percentages of badly predicted observations calculated over the validation subsets, obtained when cross-validating the models . . . . .	115
A.1	Estimated parameters of the action model ( $\beta$ ) . . . . .	134
A.2	Estimated parameters of the risk model associated to the action model ( $\omega_A$ ) . . . . .	134
A.3	Estimated parameters of the duration model ( $\theta$ ) . . . . .	135
A.4	Estimated parameters of the risk model associated to the duration model ( $\omega_D$ ) . . . . .	135
B.1	Summary of the mentioned acronyms . . . . .	138
B.2	Estimation results of the constants for <b>reduced model</b> . . . . .	138
B.3	Estimation results and description of the specification of <b>reduced model</b> . . . . .	139
B.4	Estimation results of the constants for the <b>latent model</b> , associated the expression perception model . . . . .	139
B.5	Estimation results and description of the specification of the <b>latent model</b> , associated to the expression perception model . . . . .	140
B.6	Estimation results of the <b>latent model</b> , associated to the memory effects parameters . . . . .	140
B.7	Estimation results and description of the specification of the <b>latent model</b> , associated to the model which detects the most meaningful frame . . . . .	141
B.8	Estimation results of the constants for the <b>smoothed model</b> , associated to the expression perception model . . . . .	141
B.9	Estimation results and description of the specification of the <b>smoothed model</b> , associated to the expression perception model . . . . .	142

B.10	Estimation results and description of the specification of the <b>smoothed model</b> , associated to the model related to the detection of the first frame of the relevant group of frames . . . . .	142
B.11	Estimation results of the constants for <b>reduced model with panel effect</b> . . . . .	142
B.12	Estimation results and description of the specification of <b>reduced model with panel effect</b> . . . . .	143
B.13	Estimation results of the constants for the <b>latent model with panel effect</b> , associated the expression perception model . . . . .	144
B.14	Estimation results and description of the specification of the <b>latent model with panel effect</b> , associated to the expression perception model . . . . .	144
B.15	Estimation results of the <b>latent model with panel effect</b> , associated to the memory effects parameters . . . . .	145
B.16	Estimation results and description of the specification of the <b>latent model with panel effect</b> , associated to the model which detects the most meaningful frame . . . . .	145





# List of Figures

1.1	The behavioral modeling framework adopted in this thesis . . . . .	4
2.1	Evolution of the number of transactions <i>buy</i> and <i>sell</i> contained in the raw data . . . . .	14
2.2	Examples of the indicator variations for one company . . . . .	15
2.3	Evolution of the VIX during the considered period of time . . . . .	16
2.4	Aggregation of isolated transactions . . . . .	21
2.5	Aggregation of neighboring observations . . . . .	21
2.6	Evolution of the number of actions <i>buy</i> and <i>sell</i> contained in the processed data . . . . .	23
2.7	Calculation of the dynamic attributes . . . . .	24
2.8	Transformation of the raw attributes . . . . .	25
2.9	The process of investors' decisions . . . . .	28
2.10	The general modeling framework . . . . .	46
2.11	Distribution of the observed durations, expressed in weeks (5 days) . . . . .	47
2.12	Distribution of the predicted action choice probabilities calculated on the estimation data . . . . .	48
2.13	Comparison between the distribution of the residuals $\{z_{c,t}\}$ , related to the duration model (histogram) and the theoretical distribution (curve) . . . . .	49
3.1	Conceptual framework for pedestrian walking behavior . . . . .	54
3.2	The basic geometrical elements of the space structure . . . . .	55
3.3	The spatial discretization. . . . .	56
3.4	Choice set representation, with numbering of alternatives . . . . .	57
3.5	The elements capturing the <i>keep direction</i> and <i>toward destination</i> behaviors . . . . .	58
3.6	Impact of free flow acceleration terms on utility (x axis: the speed, y axis: the utility contribution ) . . . . .	61
3.7	Leader and potential leaders in a given cone . . . . .	62
3.8	Collider and potential colliders in a given cone . . . . .	64
3.9	A frame from the Japanese video . . . . .	67
3.10	Speed histogram . . . . .	68
3.11	Example of one manually tracked trajectory with choice sets . . . . .	69

3.12	Revealed choices histograms . . . . .	70
3.13	Predicted probabilities of the Japanese data . . . . .	72
3.14	Predicted and observed shares for the Japanese data set . . . . .	73
3.15	A representative frame from the video sequences used for data collection	75
3.16	Comparison between the Japanese and Dutch normalized observation distributions across the alternatives . . . . .	76
3.17	Prediction with the proposed model . . . . .	77
3.18	Prediction with the constant-only and proposed model . . . . .	78
3.19	Choice histogram predicted by the model against revealed choices in the Dutch data set . . . . .	79
3.20	Distribution of speed in the two data sets . . . . .	80
4.1	Snapshot of a FEED database video: neutral face (subject N°2) . . . .	88
4.2	Snapshot of a FEED database video: expression produced by the TV program (subject N°2) . . . . .	88
4.3	Numbers of considered videos per subject . . . . .	89
4.4	Mask tracked by AAM along a video sequence . . . . .	89
4.5	a) Facial landmarks (55 points); b) the geometrical relationship of facial feature points; c) some facial descriptors; . . . . .	90
4.6	Sample of AU . . . . .	90
4.7	Examples of synthesized faces obtained varying the first $C$ parameter from the mean face ( $\pm 3\text{std}$ ). . . . .	91
4.8	Snapshot of the internet survey screen (subject N°15) . . . . .	92
4.9	Distribution of the collected labels among expressions . . . . .	93
4.10	The dynamic process of the <b>latent model</b> . . . . .	99
4.11	Examples of the variations of $y_{2,t,o}$ , associated to the height of the mouth (“ <i>mouth_h</i> ”), for a video . . . . .	114
4.12	Cumulative distributions of the choice probabilities predicted by the five proposed models, on the estimation data . . . . .	116
4.13	Example of a detailed prediction of the <b>reduced model</b> . . . . .	118
4.14	Example of detailed prediction of the <b>latent model</b> . . . . .	119
4.15	Example of detailed prediction of the <b>smoothed model</b> . . . . .	120
4.16	Example of detailed prediction of the <b>reduced model with panel effect</b> . . . . .	121
4.17	Example of detailed prediction of the <b>Latent model with panel effect</b>	122

# 1. Introduction

## 1.1 Context

Decision making is difficult in complex situations, as many trade-offs are usually considered. Consequences of decisions may be difficult to predict. Decision-aid tools are relevant in these circumstances. They give insight into the decisions and their impact, by providing quantitative information. They allow to forecast scenarios. The forecasting often requires the prediction of human actions. Indeed, the majority of complex decision contexts involve human actors. This is the case in markets, where humans exchange goods and services. This is also the case in systems, where humans interact with facilities, such as in transportation.

Models enable the prediction of human actions. They provide a simplified representation of reality, a better understanding of it and allow to predict its future state. Models embedded in decision-aid tools are often simplistic. This is due to three reasons: simple models are the most tractable, they do not require advanced knowledge and their development is not time-consuming. For example, aggregate methods are commonly used, based on linear regressions or time series analysis. An issue is raised for simplistic models, as their prediction capabilities are limited. This is problematic as they directly influence the output of decision-aid tools.

Models with high prediction accuracy are necessary. Disaggregate behavioral models are relevant in this context. They have been developed since the late 50's and are designed to capture the behavior of single individuals. In these models, the behavioral mechanisms are directly translated into equations, which is what makes their strength but also their complexity.

On one hand, emerging technologies (e.g. smart phones, computers or video cameras) enable the collection of various and detailed human behavioral data. The cost of data storage has significantly decreased in the last decade. Huge databases are easily stored and accessible. These data are highly relevant for characterizing human behavior. On the other hand, new modeling techniques have been proposed to deal with complex behaviors ( J.L.Walker (2001), Ben-Akiva et al. (2002) and Ben-Akiva (2010)). They capture detailed causal effects, but their estimation involves the optimization of complex mathematical functions. For many years, the estimation effort has limited their development. This is no longer the case, and software is available for estimating such models.

The combination of new modeling techniques, software and data open the field of behavioral modeling to new applications, but also challenges. Challenges have to be identified in the behavioral modeling process, and solutions need to be proposed.

Traditionally, behavioral models have been developed in transportation for modeling travel demand (Ben-Akiva and Lerman, 1985). They are key to the classical transportation planning approach. Marketing is the other historical field (Louviere et al., 2000). For obvious reasons, the understanding and prediction of the behavior of consumers is crucial in economy.

This thesis is focused on the challenges of behavioral modeling for non-traditional applications.

## 1.2 Motivation

The work is motivated by several objectives summarized in what follows.

- **Identify challenges of non-traditional applications:** Detailed knowledge about human behavior has been acquired when conducting analysis in traditional applications. Moreover, accurate modeling methodologies have been proposed in these circumstances. The transferability of this flourishing of literature has to be investigated in the context of non-traditional applications. Adaptations need to be done and new methods should be proposed. Challenges have to be identified and addressed in every task of the modeling process.
- **Emphasize multi-disciplinarity:** Three types of knowledge should be combined when developing an accurate model. First, application-specific knowledge is needed to understand the behavioral context. Second, mathematical skills are essential for the development of a rigorous model. Third, good judgment is required of the analyst in order to translate the behavioral understanding into equations. The analyst's judgment is also crucial for interpreting estimation results of the model, as well as evaluating its prediction capabilities.
- **Focus on behavioral complexity:** Non traditional applications are characterized by complex behaviors. For example, dynamic behaviors are common, meaning that current decisions are influenced by previous decisions and information. In addition, individuals often consider complex interactions between information, when making decisions in these contexts. The translation of these phenomena into equations is challenging and remains difficult. Innovative mathematical formulations have to be proposed. Their relevance has to be demonstrated in real applications.
- **Investigate non-traditional applications of behavioral modeling:** Behavioral models have shown their relevance mainly in transportation and marketing. New application fields have been identified, where the understanding

and prediction of human behavior are important. It is now possible to consider quantitative models thanks to new data collection technologies. The development of behavioral models will allow to capture the complexity of the decision process with regard to these applications.

- **Develop operational models:** New behavioral modeling techniques have been proposed in the literature. So far, they have been mostly used to quantify the behavioral understanding. Few attempts have been carried out to use them for prediction. This is important, as forecasting is the finality of the modeling process.

### 1.3 General framework

We now present the general modeling framework that we have adopted throughout the thesis. It is displayed in Figure 1.1. Tasks are represented by squares. Ellipses represent inputs or outputs to tasks. The analyst is involved in every task.

The analysis starts with the choice of an application. Experts of the application field are contacted to help the analyst in understanding the application context. Data are required to characterize the studied behavior. Experts help in setting a data collection method and collecting data. Data collection consists of observing and reporting actions performed by subjects in the application context. Raw data are processed. For each observation, the decision variable is computed, as well as the explanatory variables. Data are used to understand, specify, estimate and validate the model. The specification of the model is the translation of the behavioral mechanisms into equations. The estimation consists of fitting the parameters of the theoretical model with real data. Tests are carried out to validate the estimation results: parameter interpretations are checked and statistical tests are performed to ensure the significance of the parameters. If the results are good, the process goes to validation, otherwise a new model specification is provided. Validation assesses the prediction capabilities of the model and is carried out with a different dataset than that of the estimation. Measures are calculated and analyzed to judge the prediction capabilities of the model. If results are satisfactory, the model is finalized, otherwise, it is respecified.

The focus on non-traditional applications raises challenges in the different tasks of the modeling framework. In this thesis, we will address several of them. Data collection involves the use of emerging technologies, leading to the development of specific experimental designs. The collected data require specific and heterogeneous data processing techniques. Regarding the specification, the behavioral complexity should be accounted for in the model using precise mathematical formulations. For this purpose, modeling insights from traditional applications have to be adapted, or new formulations need to be proposed. The estimation should ensure the quality of

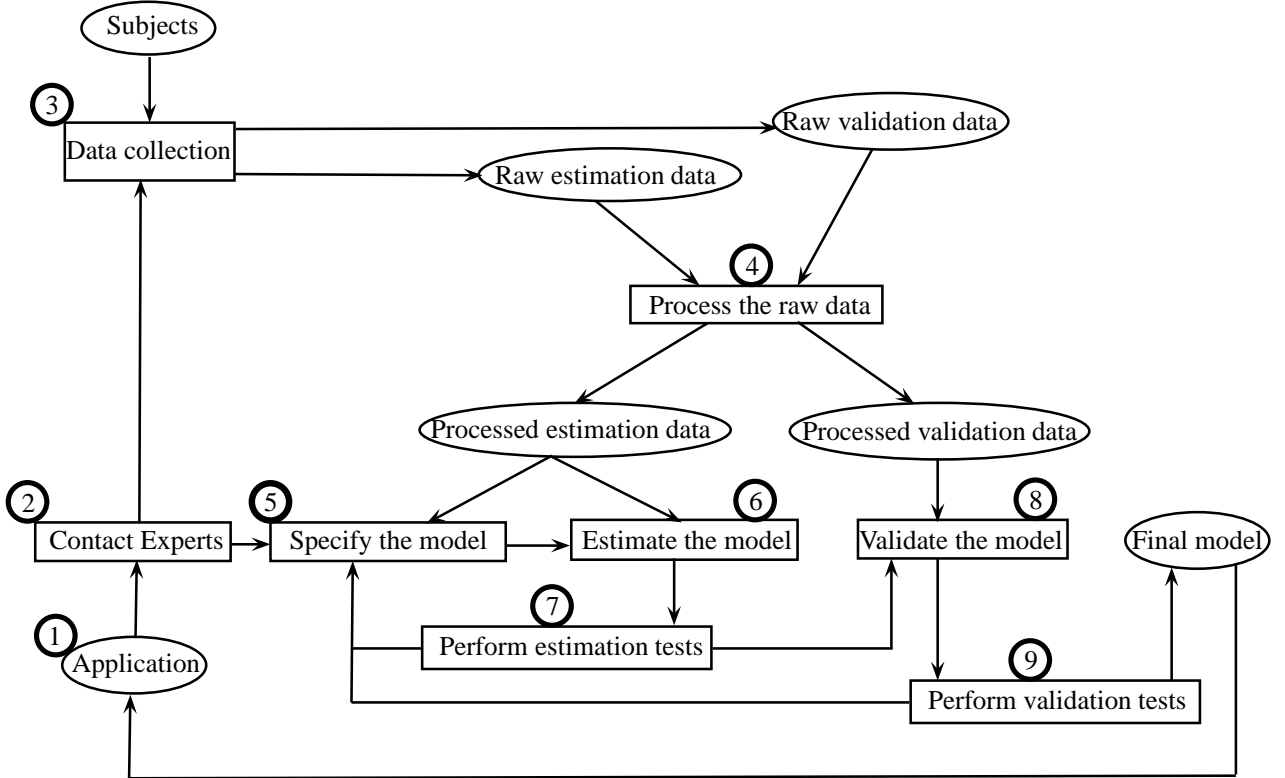


Figure 1.1: The behavioral modeling framework adopted in this thesis

the proposed specification by checking the consistency of parameters. Concerning the validation, proper methodologies have to be set up and applied. This task is particularly important in proving that the developed model is operational.

In this thesis, we explore the challenges associated with behavioral modeling in the context of three specific applications.

## 1.4 Applications

We consider three contexts where the human dimension plays an important role: pedestrian and crowd management, automatic facial expression recognition, and financial decisions. They have been selected among a long list of applications because of the availability of real and relevant data. The three applications and their corresponding challenges are presented in what follows.

The modeling of financial decisions consists in characterizing how individuals be-

have in the financial market. From the point of view of investors, this model is relevant for predicting monetary gains. From the point of view of authorities, the model predictions can help to regulate the market. The evolution of the stock market depends on decisions taken by several financial actors, including asset managers, firms, long- and short-term investors, or non-professional individuals. The financial actors base their decisions on eclectic information from heterogeneous sources. Information is related to stocks and underlying companies, or to the market. News and governmental announcements are considered. Information about the previous days is also relevant. The stock market is a complex system, hard to understand and model.

We are interested in modeling financial decisions made by investors in stock markets. Data provided by a Swiss private bank are used to characterize their behavior. We focus on the choice of action (buy or sell), and the duration between two actions. The translation of behavioral complexity into a model is challenging. Numerous explanatory variables are available and causalities have to be identified. We must account for the market risk which has a large influence on the decisions. Moreover, the process is dynamic, as investors account for previous actions when making decisions. Estimation is challenging as well, due to the complexity of the model. Another challenge concerns the prediction capabilities. In the literature, the proposed models are often limited. There is room for improvement.

Pedestrian behavioral models describe how a person walks. It is an important topic in various contexts. For the design of buildings, architects want to understand how individuals move in order to design for optimal space. In transportation, engineers conceive facilities with particular emphasis on the safety of pedestrians. Recent dramatic events increased the interest for video surveillance systems able to monitor pedestrians in public spaces, and detect suspicious behavior. Models also help in testing evacuation scenarios. The walking behavior is composed of mainly three items: choice of destination, choice of route and choice of step. The destination choice determines where the pedestrian goes. The route choice concerns the path to reach the destination. The step choice is related to the walk along the path. Pedestrian behavior depends on several factors, such as density, interactions with other pedestrians, or characteristics of the pedestrian.

We are interested in modeling the step choice. We use pedestrian trajectory data extracted from videos to characterize pedestrian behavior. We start from the model of Antonini, Bierlaire and Weber (2006), which is a discrete step choice model. The model integrates various behavioral patterns, such as the destination's attractiveness and some pedestrian interactions (*leader-follower* and *collision avoidance*). The validation of such a model is challenging. In the literature, few pedestrian models have been quantitatively validated. Most of the time, there is no validation, or only a qualitative validation. Two main reasons are involved. Data reflecting the dynamics of pedestrians are difficult to collect, and no proper validation methodology has been proposed.

Automatic facial expression recognition consists in characterizing the expression of a human face. Facial expressions are essential in conveying emotions and represent a powerful way for human beings to relate to one another. In human machine interfaces, computers must account for human emotions. For example, it is useful in the context of “aware vehicles”, for the development of systems managing interior car features. Moreover, emotions influence numerous choice processes, making it an important decision variable. This is the case in marketing or transportation, and facial expression is one of the main indicators of emotion.

We model the dynamic facial expression recognition (DFER), which consists of capturing how a person labels the facial expression of a subject in a video. We start from the model of Sorci, Antonini, Cruz, Robin, Bierlaire and Thiran (2010) who propose a discrete choice model for the recognition of facial expressions in images. There are several challenges. We need to design an experiment for collecting data which reflects the heterogeneity of the perception of expressions. Indeed, perception of facial expressions is subjective, differing from one person to another. Computer vision techniques must be used to extract information about the faces. Regarding the model specification, accurate explanatory variables should be identified, requiring the use of psychological concepts. Complex causal effects must be translated into equations. Explicit accounting for dynamics of the process is also challenging. Hypotheses need to be set up and translated into equations. Estimation is intricate due to both model and data complexity. Validation is not trivial, because both a quantitative and a qualitative analysis are required. Indeed, even if parameter interpretations are verified after estimation, predictions related to the dynamic parts of the models should be carefully studied.

We summarize the tasks performed for each application in Table 1.1.

Modeling task	Investors' behavior	Walking behavior	DFER
Contact experts	×		×
Collect data			×
Process data	×	×	×
Specify a model	×		×
Estimate a model	×	×	×
Validate a model	×	×	×

Table 1.1: Tasks performed for each application in the thesis

## 1.5 Contributions

This thesis contributes to various fields. These contributions are summarized as follows.



- **Multi-disciplinarity:** We propose a behavioral modeling framework where knowledge related to different research fields is integrated. Relevance of the concept is proved by conducting detailed analysis for three different applications. In every analysis, methodologies associated with the discrete choice modeling enter the picture. Regarding the behavior of investors, finance is considered. For walking behavior, computer vision and pedestrian science are used. Concerning the DFER, computer vision and psychology appear to be meaningful. In each application, the knowledge combination added to the analyst's judgment, reinforce the proposed analysis.
- **Quantification of behavioral mechanisms:** Advanced models are proposed to capture the complexity of behaviors under investigation. Specific mathematical formulations are proposed and estimated in each analysis. The parameters are meaningful and emphasize complex behavioral mechanisms. In finance, an integrated approach is developed to handle simultaneously two decision variables. Complex causalities are quantified, as well as process dynamics and risk perception. Regarding walking behavior, the behavioral patterns considered are relevant. For the DFER, the translation of psychological concepts into equations is meaningful. Moreover, the proposed models help in characterizing the dynamics of the recognition process.
- **Modeling of dynamic behaviors:** Dynamics of the decision process are handled with specific techniques. For the investors' behavior, a specification is proposed to account for the previous decisions, both in the action choice and duration models. Concerning the DFER, discrete choice models with latent classes are developed to capture it. The dynamics are also accounted for in the utilities. The different formulations proposed in the thesis are relevant for future behavioral analysis.
- **Handling of ambiguities:** For the DFER, the developed approach allows to handle the ambiguity of facial expressions. Indeed, the perceptions of several individuals are considered when estimating the models.
- **Use of complex behavioral models in a predictive context:** For the investors' behavior and the DFER, discrete choice models with latent classes are proposed. The prediction capabilities of the models are studied in detail. In the literature, few articles report the use of such models in a predictive context.
- **Validation methodology:** For the walking behavior, a complete validation methodology of the model is proposed and performed. It can be reused in future analysis.
- **Applications:** New approaches are developed in different application fields. In each chapter, a detailed analysis is performed. The different tasks of the

modeling framework are addressed with originality. The work helps in the understanding of behavior. For researchers, it gives insights for future analysis. For each application, the models are validated. Prediction capabilities are satisfactory and emphasize the robustness of the models. As a consequence, they can be implemented in decision-aid tools for real-life applications. The work is useful for practitioners.

## 1.6 Thesis outline

This thesis is structured around three papers corresponding to the three chapters. Each chapter corresponds to an application. The outline of the thesis is presented in the following. For each chapter we make reference to the corresponding publication.

- Chapter 2 addresses the modeling of the investors' behavior. We aim at modeling financial decisions of investors on stock markets. This work has been submitted to the *International choice modelling conference 2011*. This chapter has been published as:

Robin, T., and Bierlaire, M. (2011). Modeling the behavior of investors. Technical report. Transport and Mobility Laboratory, ENAC, EPFL.

- Chapter 3 presents the modeling of pedestrian behavior. We model how a person chooses her next step when walking. A detailed procedure is performed to validate the model. This chapter has been published as:

Robin, T., Antonini, G., Bierlaire, M. and Cruz, J.(2009). Specification, estimation and validation of a pedestrian walking behavior model, *Transportation Research Part B: Methodological* **43**(1): 36–56.

*Ranked first in the TOP 25 hottest articles of Transportation Research Part B for October-December 2008 as well as for January-March 2009.*

- Chapter 4 focuses on the modeling of dynamic facial expression recognition. We model a person who has to recognize the facial expression of a subject in a video. This chapter is a follow-up on Sorci, Antonini, Cruz, Robin, Bierlaire and Thiran (2010) and has been published as:

Robin, T., Bierlaire, M. and Cruz, J., (2010). Dynamic facial expression recognition with a discrete choice model, Technical Report TRANSP-OR 100423, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne.

- Chapter 5 provides conclusions and future research perspectives.

## 2. The financial behavior: modeling of investors' decisions

### 2.1 Introduction

We propose an hybrid discrete choice framework for modeling decisions made by investors on stock markets. We focus on the choice of action (buy or sell) and the duration until the next action. The choice of action is handled with a binary logit model with latent classes characterizing the perception of the risk, while a Weibull regression model is used for the duration until the next action. The duration model also accounts for the risk perception. Both models consider the dynamic nature of the underlying phenomenon. They are merged in a single model called combined model. It is estimated using data from a Swiss private bank consisting of 25989 observations of transactions performed between January 2005 and September 2010, in 6 different funds. The predictive performance of the models are tested. A cross-validation analysis is performed. The forecasting accuracy of the action model is studied more in details. Parameters of both models are interpretable and emphasize interesting behavioral mechanisms related to investors' decisions. The good prediction capabilities of the action model in a real context makes it operational.

This chapter contains mainly the developments proposed by Robin and Bierlaire (2011). This work has been submitted to the *international choice modelling conference 2011*.

### 2.2 Motivation

The prediction of the evolution of the stock market is crucial for investors in order to forecast their monetary gains. For authorities, this topic is important for regulating the market. The evolution of the stock market depends on the decisions taken by numerous financial actors, including asset managers, firms, long-term and short-term investors, or unprofessional individuals. In addition, automatic trading based on algorithms is also used for taking advantage of the instantaneous price variations of stocks. A recent study done by the Aite Group (Aite, 2010) shows that in 2009, slightly more than half of European equity volume is executed electronically. They

## 10.2. *THE FINANCIAL BEHAVIOR: MODELING OF INVESTORS' DECISIONS*

forecast an increase of this share in a near future. The actors' decisions are based on information coming from very eclectic sources and vary across actors. Stocks price play a key role as well as information associated with the underlying company and the whole stock market. News and governmental announces are considered in these decisions. As the process is dynamic, the past information are fully relevant. The variations of the price of stocks reflect the interaction between the different actors' decisions, according to the supply and demand rule, making the stock market an extremely complex system difficult to model and predict.

Many financial problems have been analyzed in the literature. A major field of research is the portfolio optimization. Perold (1984) develop an algorithm in case of large-scale portfolio. Chang et al. (2000) present heuristics for optimizing it. El-Ghaoui et al. (2003) work on the robust optimization using conic programming. Dentcheva and Ruszczyński (2006) propose other methods for this optimization. Equilibrium problem appear to be relevant in finance. Werner (1985) analyze equilibrium in economies with particular financial markets. Basak and Cuoco (1998) solve the price equilibrium problem in a pure-exchange, continuous-time economy in which the financial actors face costs or other types of information preventing them from investing in the stock market. Veronesi (1999) model the overreaction of financial actors toward different stimuli using a rational expectations equilibrium model of prices. The prediction of return is another interesting problem. Lamoureux and Lastrapes (1990) provide empirical support to autoregressive conditional heteroskedastic (ARCH) models in predicting return of stocks. Campbell and Shiller (2001) study the prediction capabilities of price earnings ratio on stock markets for long time periods. The analysis of the risk is well studied. Blume (1971) examine the statistical properties of one measure of risk based on the returns. Bollerslev and Mikkelsen (1996) use time series model to predict the volatility of the stock market, which characterize the level of risk. McNeil and Frey (2000) propose a method to estimate the value at risk (VaR) and related risk measures. Rockafellar and Uryasev (2000) optimize the VaR. Artzner et al. (1999) review and study different measures of risk in financial and non-financial fields. Frittelli and Gianin (2002), Acerbi (2002) develop several risk measures. Another topic is the prediction of firm bankruptcies which helps financial actor in taking decisions. Johnsen and Melicher (1994) uses logit models to predict firm financial distress and bankruptcy. Hensher and Jones (2004) develop a mixed logit model to predict firm financial distress. Shin et al. (2005), Min et al. (2006) applicate machine learning methods (support vector machine (SVN) and genetic algorithm) for predicting bankruptcy.

Since the early 90's the behavioral finance has a growing interest. The evolution of financial market is caused by the behavior of the underlying actors, which appears to be highly heterogeneous. Shiller (2003) explain the transition between theories about market efficiencies and the behavioral finance. Shiller (1999) discuss the different human behaviors observed in financial systems. Cont and Bouchaud (1997) propose a model to account for the herd behavior in financial markets. Lo (2004) present a

framework for conciling market efficiency and behavioral alternatives using the evolution theory. De Grauwe and Grimaldi (2004) develop a simple behavioral model in which agents optimize their portfolio based on rules, accounting for the reaction toward bubbles and crashes. Baker and Wurgler (2007) investigate the influence of the investor sentiments about risk on the stock market. Some attempts have been made to model the financial scenes. Arthur et al. (1996) propose a theory and simulate the behavior of heterogeneous agents in an artificial stock market. Chen and Yeh (2001) develop a genetic algorithm to model traders in an artificial stock markets. Barberis and Thaler (2003) summarize applications of behavioral finance (including individual trading behavior). Hommes (2006) review disaggregate and dynamic agent models in economics and finance.

Different methodologies have been applied and developed. Specific time series techniques are common. Bollerslev et al. (1992) review theory and applications of ARCH model in finance. Mikosch and Stărică (2004) develop models in order to account for non stationarity in financial time series. Rule-based models are largely used. Zopounidis (1999) study the contribution of multicriteria analysis in solving financial decision problems in a realistic context. Wagner et al. (2002) propose a guidance for specifying rules of expert systems in the financial field. Operation research techniques appear to be useful for optimization. El-Yaniv (1998) survey results concerning on-line algorithms for solving problems related to the management of money. Machine learning methods are well adapted to the financial context, due to the huge amount of aggregate data available on financial markets. Giles et al. (2001) develop neural networks (NN) for predicting noisy financial time series. Kim (2003) use SVN to forecast financial time series. Huang et al. (2004) compare artificial intelligence techniques for the credit rating analysis (SVN against NN). Other methods have been used. For example, Fermanian and Scaillet (2004) discuss the use of copulas for modeling cross-dependences in financial applications, or Embrechts and Schmidli (1994) review methodologies used to model stochasticity in finance.

In this chapter, we investigate the development of behavioral models designed to capture the behavior of specific actors in the financial sector. In that purpose, we propose to use an hybrid discrete choice framework to model the behavior of investors. Discrete choice models (DCM) are well adapted in the context of disaggregate modeling (Ben-Akiva and Lerman, 1985). They have the advantage to explicitly capture causal effects. Interestingly, few articles report the use of DCM in the financial context. Johnsen and Melicher (1994) and Hensher and Jones (2004) developed DCM for predicting the firm financial distress. de Palma et al. (2008) consider the inclusion of the risk perception in random utility models, although they do not focus on finance per se.

We are interested in modeling the behavior of professional investors. Each fund is managed by a person called the fund manager, with the help of a dedicated team

of analysts. The objective of the fund manager is to adjust his portfolio in order to maximize the returns. For each stock present in his investment universe, he decides to buy, sell or wait. In case of buying and selling, he has to decide the involved quantity of money. Then, the order is passed to a trader in charge to implement this decision. The trader translates the decision into transactions. Several transactions can reflect the same decision as the trader chooses the right moments in the day for taking advantage of the market, or even split the order over a few days to reduce the impact of the decision on the market.

The decisions made by a fund manager are based on many heterogeneous information like volatility and price of each stock, fundamental data reported by companies, and the global state of the market often provided by global indices. Other information such as news about politics, fusion of companies or governmental announcements are also relevant. Another event triggers the decision made by a fund manager. When a customer invests in a fund, an equivalent amount of money is transferred to the fund and the fund manager needs to deploy this inflow buying some stocks. On the contrary, if there is an outflow, meaning that a customer withdraw a certain amount of money, the fund manager needs to sell some stocks to provide the necessary cash.

We focus on two behavioral aspects which are the choice of action (buy or sell) and the duration between two actions. Regarding the choice of action, the developed model is inspired from the work of J.L.Walker (2001) and Greene and Hensher (2003) about DCM with latent classes. In our case, the latent classes account for the risk perception. For the duration, we considered the models presented by den Berg (2001) and Bauwens and Veredas (2004). A Weibull model appears to be appropriate. It accounts for a different risk perception than the model of action choice.

This chapter is organized into various sections as follows. In Section 2.3, we present the raw data, in Section 2.4 the notations, in Section 2.5 the time discretization, in Section 2.6 the explanatory variables. Section 2.7 details the model. In Section 2.8, the estimation results are shown. Section 2.9 validates the model.

## 2.3 Raw data

We have access to transactions related to six long-only funds of the Swiss private bank **Lombard Odier**. For confidentiality reasons, the data are anonymous and we will call these funds 1,2,3,4,5 and 6.

The raw data consists in 25989 observations of transactions passed by traders. The time period goes from 2005.01.03 to 2010.09.13. Each transaction is characterized by a date, a company, a fund and an amount of money. The amount is positive if stocks have been bought and negative if they have been sold. Stocks of 1236 companies are considered. The number of observations and companies per fund are shown in Table

2.1. Some companies appear in several funds. The shares between the transactions *buy* and *sell* are also shown. The number of transactions are equally spread between *buy* and *sell* for all the funds, except fund 3.

Fund	Nb of transactions	Nb of companies	% buy	% sell
1	4354	160	44.63	55.37
2	1189	64	55.82	47.18
3	6427	363	78.70	21.30
4	2018	560	54.21	45.79
5	6935	55	45.84	54.16
6	5066	185	57.26	42.74

Table 2.1: Number of transactions, number of companies and percentage of transactions per fund in the raw data

The data are dynamic per nature. The evolution of the number of transactions per fund is presented in Figure 2.1. The number of *buy* transactions is stable in average, but with strong fluctuations. Regarding *sell* transactions, two phases appear. The first goes from 2005 to 2008, which is characterized by stability and weak fluctuations. The second goes from 2008 to 2010, it is stable in average but with stronger fluctuations compared to phase 1. The average in phase 2 is higher than in phase 1. These two observations are linked to the fact that the starting of phase 2 coincides with the financial crisis of 2008.

Five indicators are considered as explanatory variables. These indicators have been designed and computed by a quantitative team of **Lombard Odier**. They are called *quality*, *sentiment*, *technic*, *value* and *price*. *Quality* measures the fundamental quality of a company by examining specific economic and financial data published by the company. *Sentiment* is a number based on a combination of estimates of the analysts covering the company, as for example the next year earnings estimates. *Technic* is a combination of indicators that analyze the company's activity on the market by identifying chart pattern of prices, as for example the momentum reversal. The momentum is defined by the difference between two prices of a same stock for a chosen time horizon. *Value* is an objective value of a company based on classic valuation metrics, like for example price to earning ratio. *Price* characterizes the price of the stock associated to the company. Note that portfolio information is not available in the data, neither data about fund managers.

For a certain company, we display the variations of the five indicators in Figure 2.2. These values are scores and have no unit. For the sake of reading, the values have been normalized between 0 and 1 by adding the observed minimum per fund

14 2. THE FINANCIAL BEHAVIOR: MODELING OF INVESTORS' DECISIONS

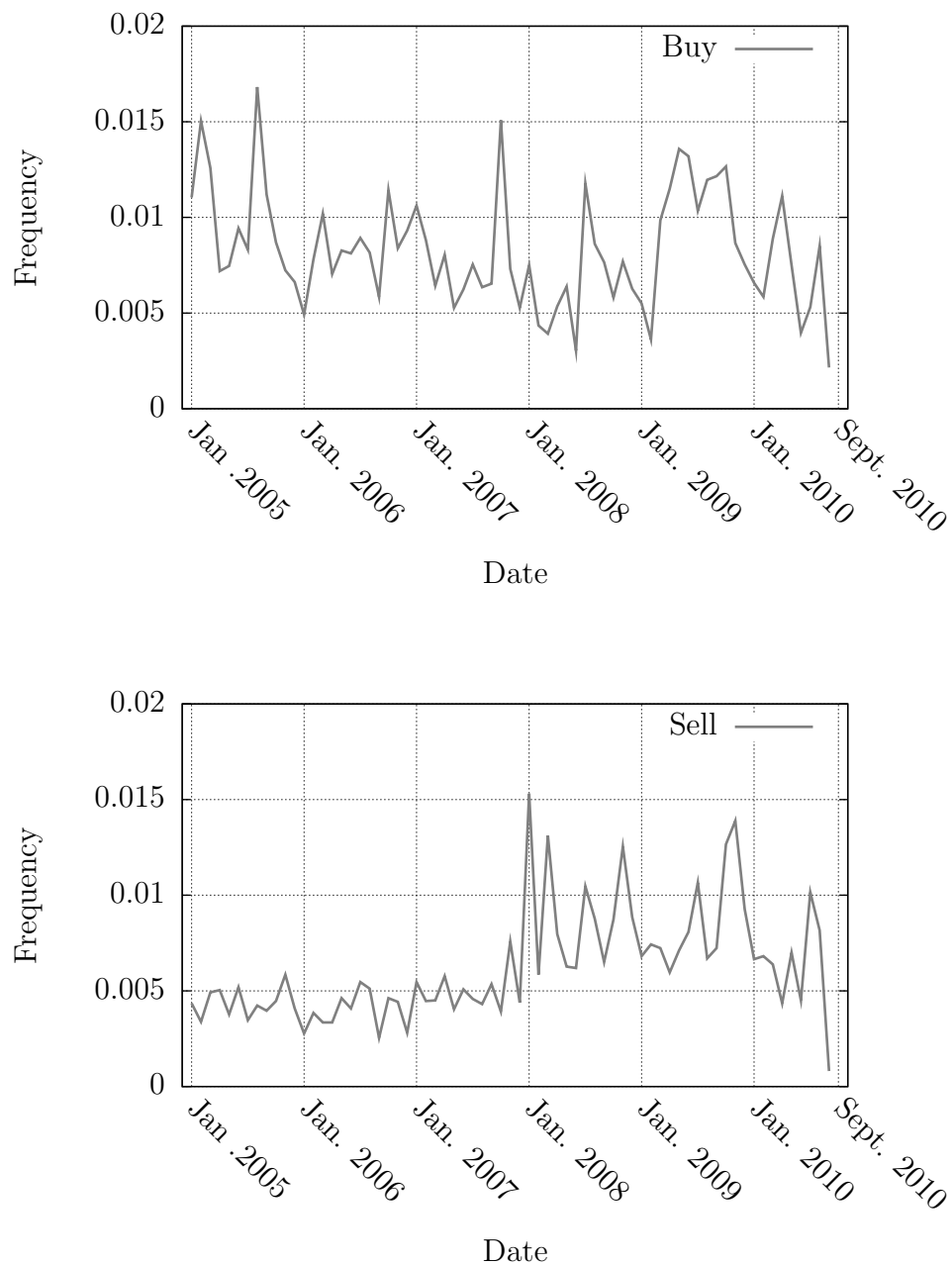


Figure 2.1: Evolution of the number of transactions *buy* and *sell* contained in the raw data

and dividing by the observed maximum per fund. A correlation analysis performed between the different indicators did not evidence strong links between them. This



is logical because, they have been built to reflect complementary information about the company. The quality and sentiment are constant over small time period by definition, generating levels on their associated curves (see Figures 2.2(a) and 2.2(e)). This is not the case for the other indicators which present continuous variations.

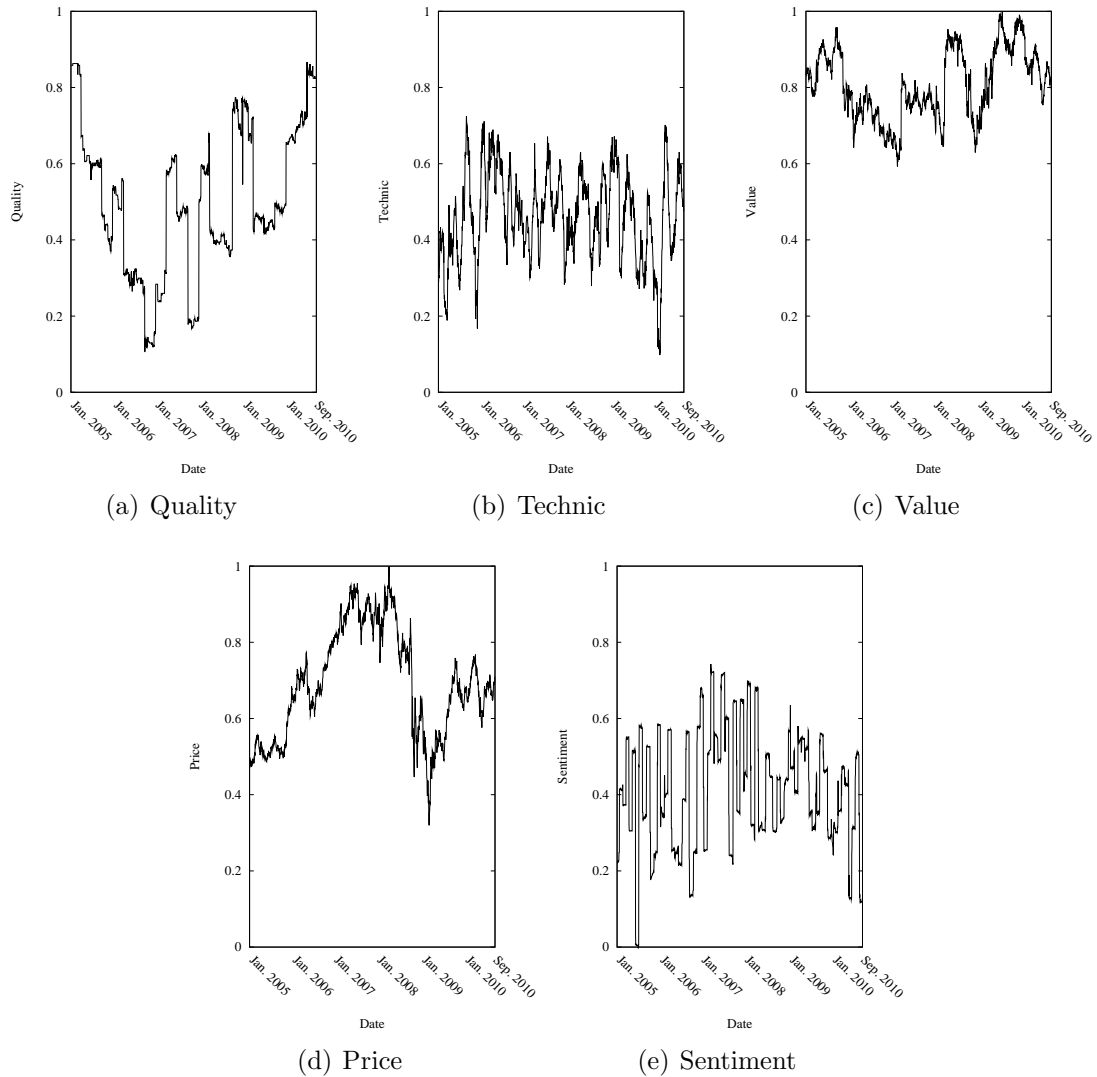


Figure 2.2: Examples of the indicator variations for one company

The decision context depends on the state of the stock market. The VIX (symbol for the Chicago board options exchange market volatility index) has been retained. This is a popular measure of the implied volatility of the S&P 500 index. It represents one measure of the market's expectation of the stock volatility and characterizes quite well the market risk. It has been plotted for the considered period of time in Figure 2.3. The financial crisis of the end 2008 appears clearly. It is characterized by the highest peak on the curve. The horizontal line is set for the VIX equal to 25.

This value, arbitrarily chosen following a discussion with a fund manager, defines a limit between a high volatility market and a low volatility market, and consequently generates two complementary decisional behaviors.

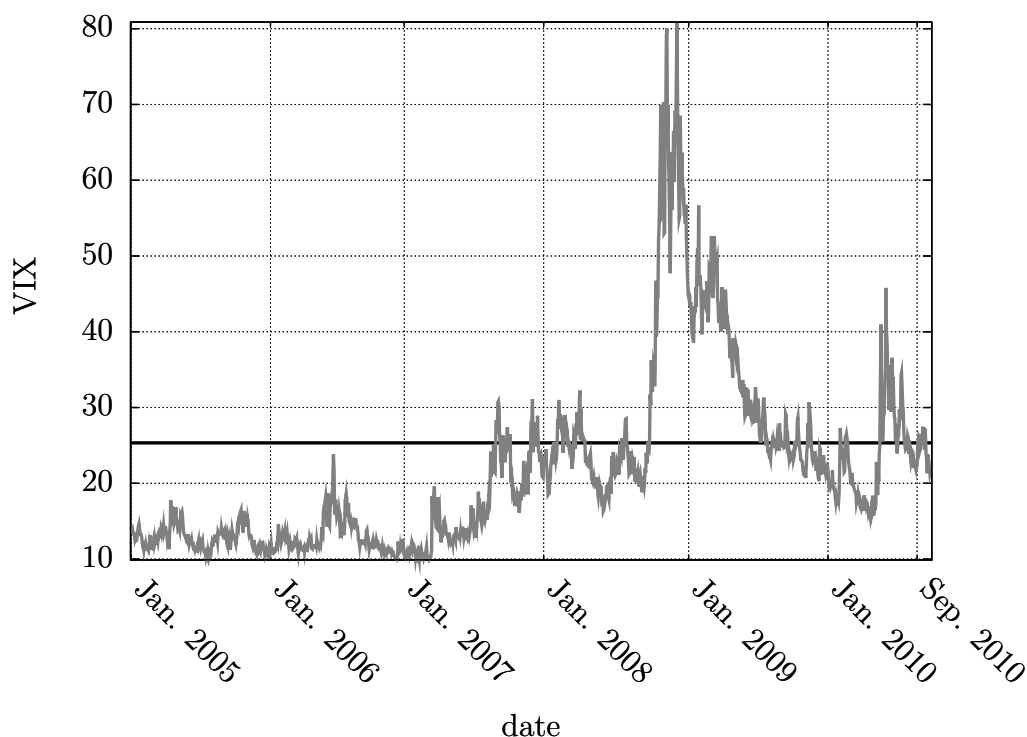


Figure 2.3: Evolution of the VIX during the considered period of time

## 2.4 Notations

We introduce the notations that are used along the chapter, specially in Sections 2.5, 2.6 and 2.7. They are used and not redefined afterward. The modeling concepts are defined specifically in Section 2.7.

- DCM: discrete choice models;
- i.i.d.: independent and identically distributed;
- $EV(0, 1)$ : standardized extreme value distribution;

- $N(0, 1)$ : standardized normal distribution;
- $c$ : company;
- $C$ : the number of companies;
- $t$ : day;
- $T$ : the length of the entire time period;
- $t_H$ : time horizon in days;
- $H$ : vector of considered  $t_H$ ,  $H = \{1 \dots 5, 10, 15, 20, 25, 30, 60, 90, 180, 270, 360\}$ ;
- $f$ : fund,  $f \in F = \{1, 2, 3, 4, 5, 6\}$ ;
- $F$ : vector of funds;
- $g$ : group of funds,  $g = 1$  groups funds 1, 2, 3,  $g = 2$  groups funds 4, 5, 6;
- $t' = t + D(c, t) + 5$ : the day of the action performed after  $A(c, t)$ ,  $A(c, t)$  and  $D(c, t)$  are defined in Section 2.4.1;
- $B$ : action *buy*;
- $S$ : action *sell*;
- $R$ : *risky* situation;
- $N$ : *normal* situation, as opposed to  $R$ ;
- $e$ : experience of the cross-validation;
- $t_{0,e}$ : starting date of the simulation set of experience  $e$ ;
- $T_e$ : ending date of the simulation set of experience  $e$ ;
- $R_e^2$ :  $R^2$  predicted by the duration model on the simulation set of experience  $e$ ;
- $z_{c,t}$ : standardized residual of the duration model.

Remaining notations are organized per thematics.

### 2.4.1 Variables

Notations for dependent and explanatory variables are summarized below.

- $O(c, t)$ : transaction observed on  $t$  for  $c$  (money), if it is positive, stocks have been bought; if it is negative, stocks have been sold;
- $A(c, t)$ : action decided on  $t$  for  $c$ ,  $A(c, t) \in \{B, S\}$ ;
- $D(c, t)$ : time duration between  $A(c, t)$  and  $A(c, t')$  in weeks (5 days);
- $\hat{D}(c, t)$ : predicted duration;
- $\bar{D}(c, t)$ : duration mean calculated over  $t$  and  $c$ .
- $r_A$ : risk in the action model,  $r_A \in \{\mathbf{N}, \mathbf{R}\}$ ;
- $r_D$ : risk in the duration model,  $r_D \in \{\mathbf{N}, \mathbf{R}\}$ ;
- $qual_{c,t}$ : quality associated to  $c$  on  $t$ ;
- $tech_{c,t}$ : technic associated to  $c$  on  $t$ ;
- $sent_{c,t}$ : sentiment associated to  $c$  on  $t$ ;
- $pric_{c,t}$ : price associated to  $c$  on  $t$ ;
- $valu_{c,t}$ : value associated to  $c$  on  $t$ ;
- $x_{c,t} = \{qual_{c,t}, tech_{c,t}, sent_{c,t}, pric_{c,t}, valu_{c,t}\}$ : vector containing the 5 fundamental indicators for  $c$  on  $t$ ;
- $K_x$ : the length of  $x_{c,t}$
- $VIX_t$ : VIX on  $t$ ;
- $Perf(x_{c,t}(k), t_H)$ : performance of  $x_{c,t}(k)$ , calculated on  $t_H$  (Equation (2.1));
- $Long(x_{c,t}(k), t_H)$ : long-term value of  $x_{c,t}(k)$ , calculated on  $t_H$  (Equation (2.2));
- $Short(x_{c,t}(k), t_H)$ : short-term value of  $x_{c,t}(k)$ , calculated on  $t_H$  (Equation (2.3));
- $Sigm(x_{c,t}(k), t_H)$ : standard-error of  $x_{c,t}(k)$ , calculated on  $t_H$  (Equation (2.4));
- $X_{c,t}$ : vector of raw and transformed values of  $\{x_{c,t}(k)\}_{k=1 \dots K_x}$  (Equation (2.5));
- $Y_t$ : vector of raw and transformed values of the  $VIX_t$  (Equation (2.6)).

### 2.4.2 Parameters

Two models are introduced in Section 2.7. The parameters of the models are denoted as follows.

#### The action model

- $\beta$ : vector of parameters (Equation (2.24));
- $\mu_f$ : scale of the random variables  $\varepsilon_{B,r_A,c,f,t'}$  and  $\varepsilon_{S,r_A,c,f,t'}$ ;
- $\omega_A$ : vector of parameters associated to the risk perception (Equation (2.8));
- $\beta_B$ : vector of parameters associated to the explanatory variables;
- $K_B$ : size of  $\beta_B$ ;
- $ASC_{B,r_A}$ : constant parameter;
- $\alpha_{B,r_A}$ : parameter associated to the deterministic utility of B in the previous action;
- $\lambda_{B,r_A}$ : parameter weighting the influence of the deterministic utility of B in the previous action.

#### The duration model

- $\theta$ : vector of parameters (Equation (2.28));
- $\eta_D$ : shape parameter of the Weibull distribution;
- $\theta_D$ : vector of parameters associated to the explanatory variables;
- $K_D$ : size of  $\theta_D$ ;
- $\omega_D$ : vector of parameters associated to the risk perception (Equation (2.13));
- $ASC_{D,r_D,g}$ : constant parameter;
- $\alpha_{D,r_D,g}$ : parameter capturing the effect of the previous duration;
- $\theta_{B,r_D}$ : parameter capturing the influence of the deterministic utility of B.

## 2.5 Time discretization

In the raw data, we do not observe the investors' decisions but direct consequences of them. For a given stock, once the investor has decided an action and an associated amount of money, a trader implements the decision. Traders have a tendency to split the investors' decisions in several successive transactions in order to decrease the influence of the decisions on the underlying stocks, in terms of price, due to the supply and demand rule. These transactions constitute the data. As a consequence, different transactions in the raw data could reflect the same decision. The date of the decision and the date of the first transaction coincides. Transactions have to be aggregated for each stock in order to represent the investors' decisions.

For given stock, we group transactions in sets. A transaction belongs to a set if at least one transaction inside the set is separated from the considered transaction from less than five days. Then, the time period of each set is split in consecutive time windows of five days. The transactions within the time windows of five days constitute subsets. Within subsets, the transactions are aggregated by summing their associated amounts of money. If the sum is positive, it is a *buy* action; if it is negative, a *sell* action appears. The information of the first day of the subset are the explanatory variables. Actions involving a small quantity of money have been discarded.

A time window of 5 days is used, because it corresponds to a working week. The aggregation procedure starts the first day with an observation, which is the first working day of 2005 (2005.01.03). The 5 days period can cross week-ends in case of legal holidays or market closures. Anyway, this period should be interpreted as a buffer, where investors do not revise their decisions. These are major assumptions which have been validated by the involved investors. In addition, other aggregation techniques have been tested (on different time periods and with different rules) and the results presented in Section 2.8 appeared to be stable, showing the robustness of the approach. In practice, we observed that the summing of the actions always imply transactions of the same sign, validating the choice of the 5 days period. Two situations appear when aggregating the transactions, they are illustrating in the following:

1. **A transaction with no near neighbor.** This means that the set contains only one transaction. It is considered as an investor's decision. An example of this situation is presented in Figure 2.4. In that case, there is no aggregation.
2. **A time period with neighboring transactions.** This means that a set contains at least two transactions. The time period is covered by non-overlapping time windows of five days. Transactions are aggregated within each subset and information about the first date of the subset are considered. An example is shown in Figure 2.5.

Regarding this example, the starting of the aggregation on day  $t$  could be conditioned by the first day of the collected data. This is the case between 2005

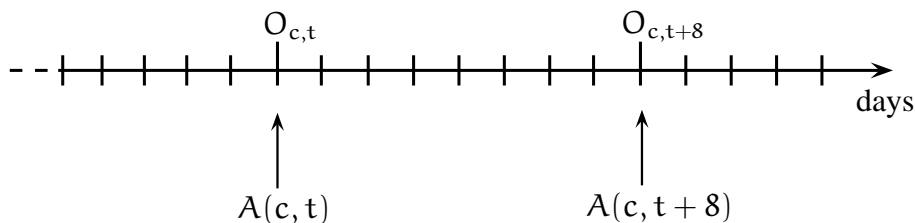


Figure 2.4: Aggregation of isolated transactions

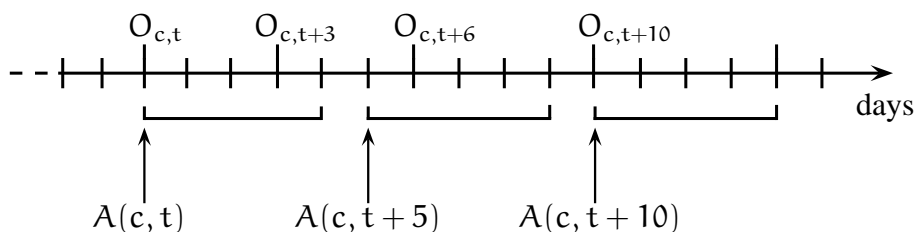


Figure 2.5: Aggregation of neighboring observations

and 2007 (Figure 2.6 compared to Figure 2.1). As mentioned previously, this has been validated with the involved investors and by performing aggregation tests.

Note that a minimum of five days separate two actions. Actions with small quantities of money have been removed because they correspond to adjustments representing noise in this modeling context. Following a discussion with a fund manager, it has been decided to arbitrarily remove 25% of the actions (with the smallest associated amounts of money). After processing the transactions, the data contains 9178 observations of actions performed on stocks of 1121 companies. The details are presented in Table 2.2. The shares between actions are equally distributed and rather the same than for the transactions presented in Table 2.1, except for the fund 3. For this fund, the distribution is much more balanced between the actions, compared to the transactions.

The repartition of the actions across time is presented in Figure 2.6. Compared to Figure 2.1, the aggregation has been quite strong for the *buy* transactions from 2005 to 2007. Lots of situations 2 (periods with neighboring observations) appeared (see Figure 2.5). Regarding other periods, the graphs are similar, showing the higher propensity of situations 1 (transactions with no near neighbor, see Figure 2.4).

Fund	Nb of decisions	Nb of companies	% buy	% sell
1	1461	145	54.96	45.04
2	913	58	53.34	46.66
3	508	50	59.45	40.55
4	3738	505	51.66	48.34
5	1659	316	43.40	56.60
6	899	175	45.05	54.95

Table 2.2: Number of actions, number of companies and percentages of actions per fund in the processed data

## 2.6 Explanatory variables

Fund managers often base their decision on the dynamic of the variables (Figures 2.2 and 2.3) when taking decisions. New variables have been computed based on the five fundamental indicators and the VIX for reflecting this dynamic. The framework of the dynamic data calculation is presented in Figure 2.7. Two consecutive actions are represented  $A(c, t)$  and  $A(c, t + 5)$ . As explained in Section 2.5, two consecutive actions are separated by a minimum of five days.

The dynamic variables are calculated as follows. We call *performance* the relative variation

$$\text{Perf}(x_{c,t}(k), t_H) = \frac{x_{c,t}(k) - x_{c,t-t_H}(k)}{x_{c,t-t_H}(k)}. \quad (2.1)$$

The mean calculated over  $t_H$  is called the *long-term value*

$$\text{Long}(x_{c,t}(k), t_H) = \frac{1}{t_H} \sum_{l=t-t_H}^t x_{c,l}(k). \quad (2.2)$$

The difference between the current value  $x_{c,t}(k)$  and the long-term value  $\text{Long}(x_{c,t}(k), t_H)$  is called the *short-term value*

$$\text{Short}(x_{c,t}(k), t_H) = x_{c,t}(k) - \text{Long}(x_{c,t}(k), t_H). \quad (2.3)$$

Finally, we define the *standard-error* as

$$\text{Sigm}(x_{c,t}(k), t_H) = \sqrt{\frac{1}{t_H} \sum_{t=t-t_H}^t (x_{c,t}(k) - \text{Long}(x_{c,t}(k), t_H))^2}. \quad (2.4)$$

It characterizes the variations of  $x_{c,t}(k)$  within  $t_H$ . We explicit



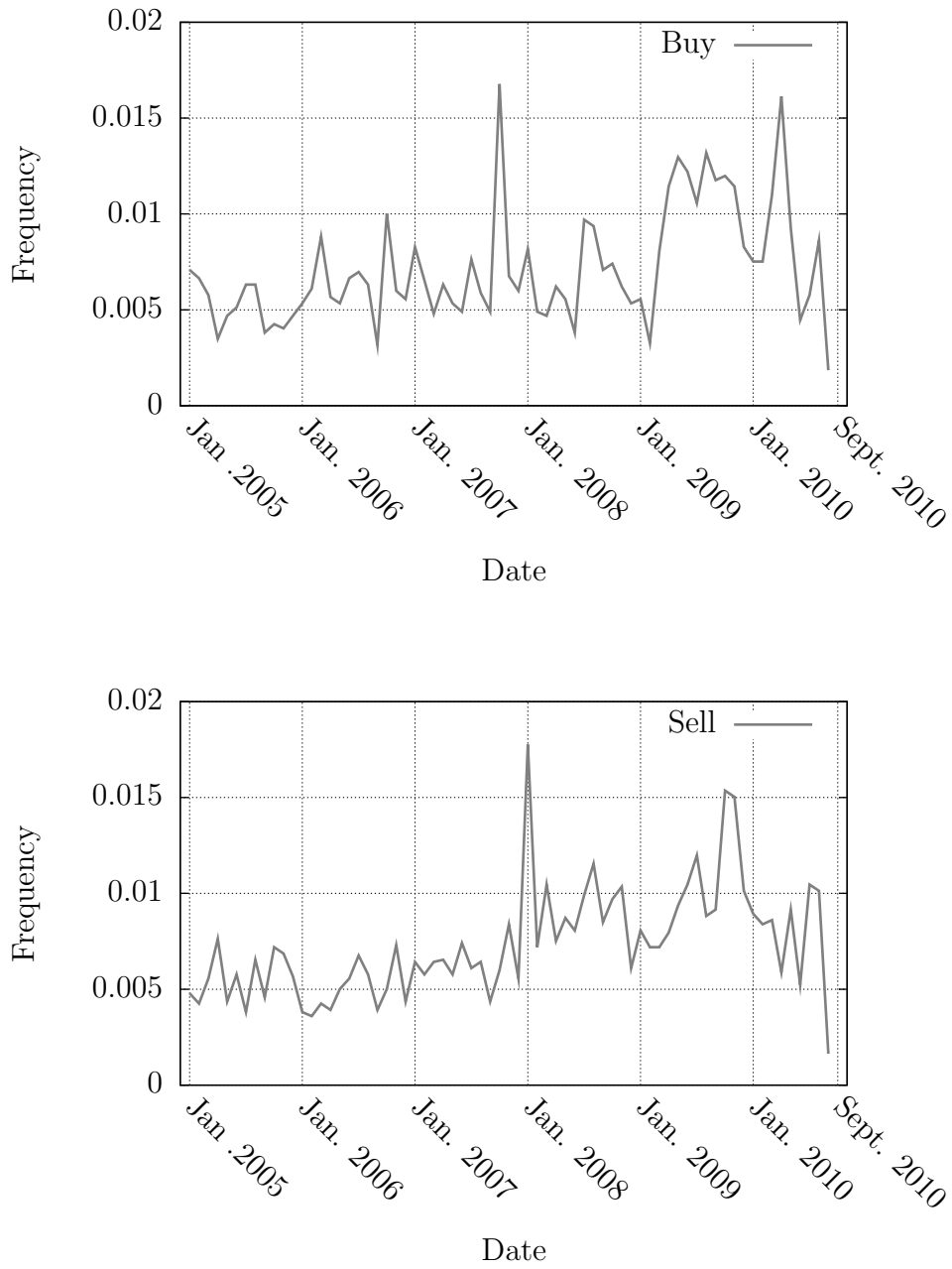


Figure 2.6: Evolution of the number of actions *buy* and *sell* contained in the processed data

$$X_{c,t} = \{x_{c,t}(k), \text{Perf}(x_{c,t}(k), t_H), \text{Long}(x_{c,t}(k), t_H), \text{Short}(x_{c,t}(k), t_H), \text{Sigm}(x_{c,t}(k), t_H)\}_{t_H \in H, k=1 \dots K_x}, \quad (2.5)$$

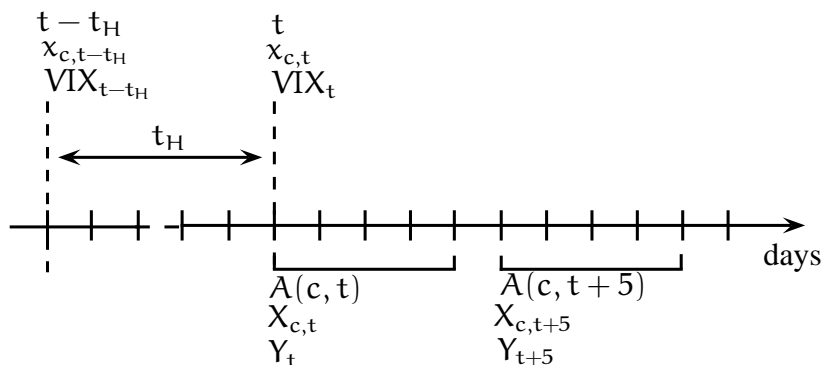


Figure 2.7: Calculation of the dynamic attributes

and

$$Y_t = \{VIX_t, Perf(VIX_t, t_H), Long(VIX_t, t_H), Short(VIX_t, t_H), Sigm(VIX_t, t_H)\}_{t_H \in H}, \quad (2.6)$$

Heterogeneous  $t_H$  are considered,  $t_H \in H = \{1 \dots 5, 10, 15, 20, 25, 30, 60, 90, 180, 270, 360\}$ . This is motivated by the fact that investors consider short-term to long-term dynamic of variables when making decisions. 366 variables are considered in total (5 indicators and the VIX, 4 transformations, 15 time horizons,  $366 = 6 \times (1 + 4 \times 15)$ ).

The variables have been normalized per fund (except the VIX) by sequentially subtracting the minimum and dividing by the maximum. Then the variables are in  $[0, 1]$ . The normalization has been done per fund because fund managers are considering the entire fund when taking decisions. The VIX has not been normalized because it is not fund specific, its interval of variation is manageable, and for interpretation of the proposed models (see Section 2.7).

For the considered time period, examples of the evolution of the dynamic variables are presented in Figure 2.8. The raw indicator is the price shown in Figure 2.8(a), which is the same than in Figure 2.2(d).  $t_H = 1$  for the calculation of the *performance*, *long-term value* and *short-term value*. The *performance* and *short-term value* capture the immediate variations of the variable (Figures 2.8(b) and 2.8(d)), whereas the *long-term value* is the smoothed version of the the raw variable (Figure 2.8(c)). Note that the smoothing degree is  $t_H$ . Due to the small value of  $t_H$ , it is qualitatively similar to the variation of the raw variable (Figure 2.8(a) and Figure 2.8(c)). The evolution of the *standard-error* is the most regular, because  $t_H = 60$  for its calculation.  $t_H = 1$  is used to illustrate the variables used in the action model (Section 2.7.2),  $t_H = 60$  appeared to be relevant for the duration model (Section

2.7.3).

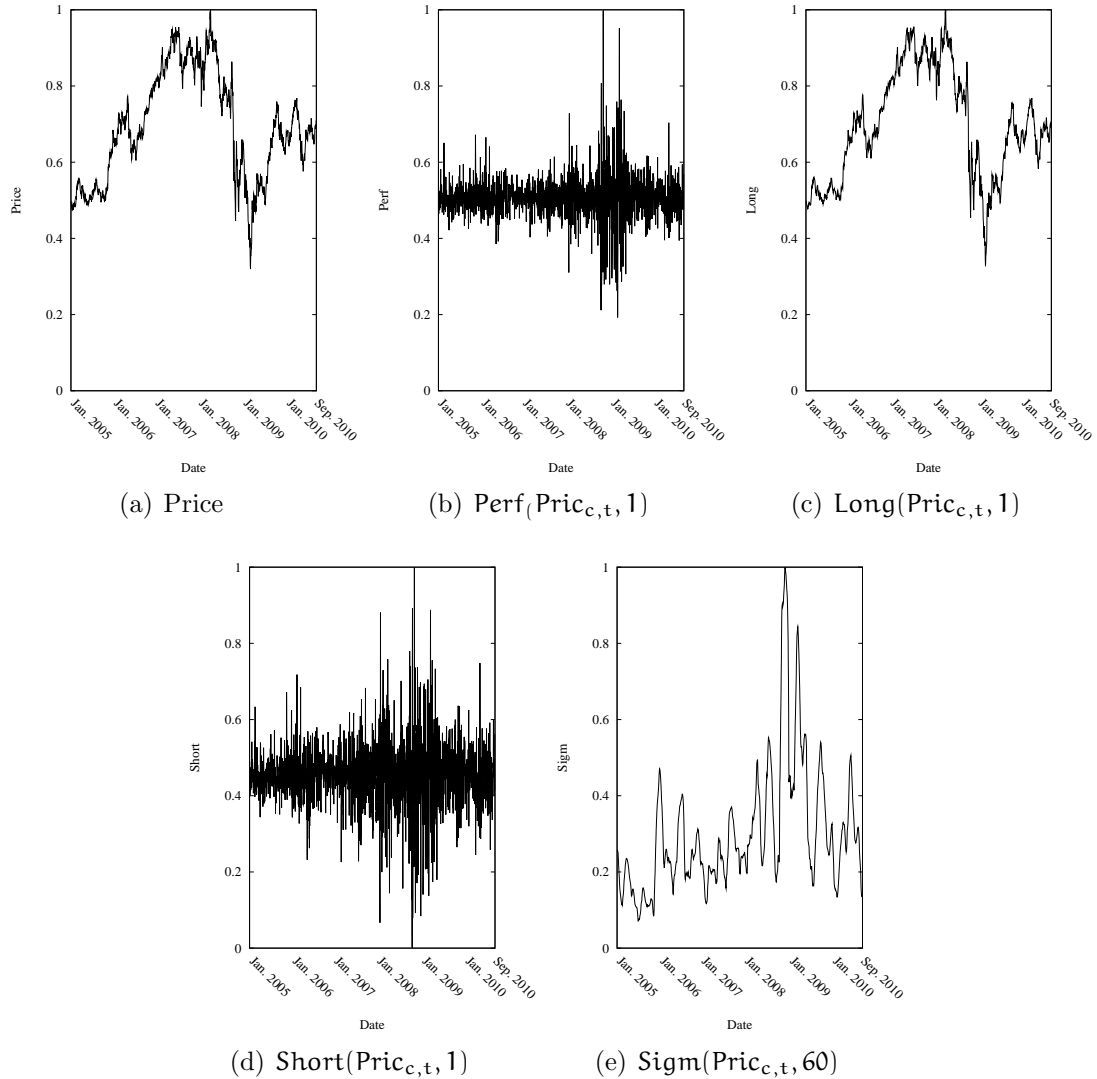


Figure 2.8: Transformation of the raw attributes

A correlation analysis has been performed between the action variable  $A(c, t)$  and the explanatory variables. Results are summarized in Table 2.3.  $B$  is coded 0, and  $S$  is coded 1. If the correlation is positive *sell* is favored, otherwise it is *buy*. Note that a Pearson test has been performed for each correlation. Only significant effects have been kept. The correlations are not very high, the highest value is 0.449 and is observed for the *standard-error* of the VIX for  $t_H = 360$ , in fund 3. The number of significant and generic correlations across funds is low, compared to the number of variables (366). This points out the difference of financial management between funds, and emphasizes the specificity of the fund managers' behavior within

## 26 2. THE FINANCIAL BEHAVIOR: MODELING OF INVESTORS' DECISIONS

each fund. Nevertheless, generic and significant correlations provide information. Regarding variables associated with the companies and stocks, time horizons are low, showing the propensity of investors to account for immediate information in their decisions. A difference of behavior appears between funds 1, 2, 3 and 4, 5, 6. Except for the variables associated to the VIX, correlations have the same signs within the two fund groups, and are opposed between the two groups. This difference is partly explained by the fact that in the two fund groups, a team manage two funds (two teams per groups of three funds). In each fund group, one team manages one fund, and the other team manages the two remaining funds.

Transform	Variable	$t_H$	Fund 1	Fund 2	Fund 3	Fund 4	Fund 5	Fund 6
Perf()	Price	1	0.065	0.139	0.127	-0.279	-0.238	-0.329
Short()	Value	3	-0.090	-0.084	-0.099	0.227	0.170	0.277
Sigm()	VIX	360	0.126	0.303	0.449	-0.062	-0.114	0.213
Short()	Technic	1	-0.065	-0.110	-0.113	0.213	0.181	0.257

Table 2.3: Generic and significant correlations between the action variable and the explanatory variables

A correlation analysis has been also performed by splitting the processed data into two parts according to the level of VIX. This has been done in order to test if there is a significant difference of behavior in volatile and non-volatile markets. The considered threshold is 25 (see Figure 2.3). In case of low VIX, the significant and generic correlations are the same than in Table 2.3. Interpretations remain the same. In case of high VIX, only the variable  $\text{Short}(\text{valu}_{c,t}, 60)$  stands out. This underlines the specificity of the investors' behavior within each fund in *risky* situations. Two reasons are evoked. First, the nature of the funds is different, which conducts to specific managements in *risky* situations. Second, the financial styles of the fund managers are emphasized and predominate over established rules. This is logical as in panic situations, emotions tend to overcome conventions. Correlations are displayed in Table 2.4. The difference between the two groups of fund appear and is consistent with Table 2.3. The time horizon is higher (60 days compared to 3 days), which is logical. In *risky* situations, investors have more tendency to consider long-term information.

These correlation analysis help us to get intuition about the data, but are limited due to their univariate nature.

Transform	Variable	$t_H$	Fund 1	Fund 2	Fund 3	Fund 4	Fund 5	Fund 6
Short()	Value	60	-0.210	-0.125	-0.200	0.206	0.095	0.232

Table 2.4: Generic and significant correlations between the action variable and the explanatory variables, for a high level of VIX

## 2.7 Model specification

We aim at understanding and modeling the financial decisions of an investor in a given time horizon (expressed in days), regarding a set of stocks. Given that an action is performed on day  $t$  for stocks of company  $c$ , we assume that the investor decides the type of action (buy or sell), and the duration until the next action performed on the same stock. The decision about the duration is not supposed to be revised, once it has been taken. An overview of the decision process is shown in Figure 2.9.  $D(c, t)$  is the duration between  $A(c, t)$  and the next action  $A(c, t')$ , with  $t' = t + D(c, t) + 5$ . If  $D(c, t) = 0$ , the duration between the two consecutive actions is five days, which is the minimum duration according to the aggregation presented in Section 2.5. We model the decisions taken in  $t'$ ,  $A(c, t')$  and  $D(c, t')$ , conditionally on  $A(c, t)$  and  $D(c, t)$ .

The hypothesis of non-revision of the duration is strong. Two reasons explain it. First, this is necessary as important explanatory variables are not present in the data for explaining this decision. They are money inflows and outflows. Second, the hypothesis has been adopted to operationalize the model. The model is already complicated and a revision hypothesis significantly complicates the formulation of the likelihood function (see Section 2.7.4)

The model associated to  $A(c, t')$  is called action model. The model associated to  $D(c, t')$  is called duration model. The combined model groups the action and duration models. The general modeling framework is presented in Figure 2.10. Square shapes represent observed variables, whereas round shapes are latent variables. Shapes with dotted lines are for random variables, whereas plain lines are associated to deterministic variables. Arrows stand for causal links between variables, each arrow is associated to an equation. Plain arrows stand between variables of  $t'$ , dotted arrows link variables in  $t$  to variables in  $t'$ . Note that some variables can be both latent and deterministic, as the deterministic parts of utilities in a discrete choice model. We define the modeling concepts of the scheme per models. For the action model, we specify:

- $V_A(R, c, t' | \omega_D)$ : the measure for the risk  $R$  in the action model (Equation (2.7));
- $W_A(r, c, t' | \omega_A)$ : the randomized measure of the market risk  $r_A$  (Equation (2.9));

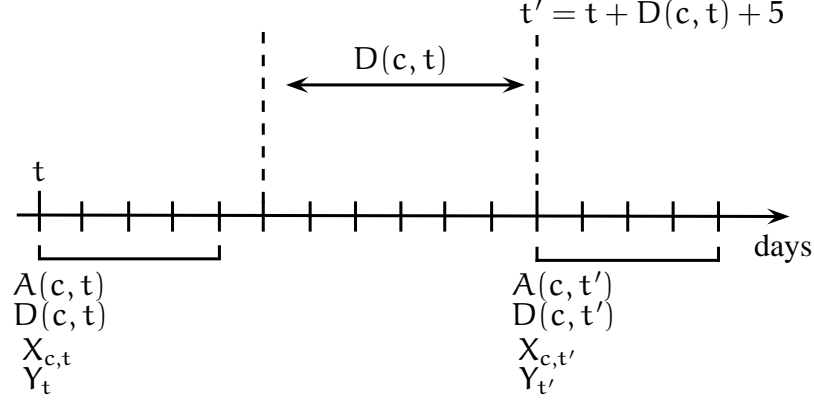


Figure 2.9: The process of investors' decisions

- $V_B(c, t'|r_A, \beta)$ : the deterministic utility associated with the alternative B (Equation (2.18));
- $M_B(c, t'|r_A, \beta)$ : the term capturing the effect of the previous action on the current choice of action (Equation (2.19));
- $U_B(c, t'|r_A, \beta)$ : the random utility of the alternative B (Equation (2.17));
- $U_S(c, t'|r_A, \beta)$ : the utility of the alternative S (Equation (2.17));
- $\varepsilon_{A,r_A,c,t'}$ : a random variable associated to the risk  $r_A$  (Equation (2.10));
- $\varepsilon_{B,r_A,c,f,t'}$ : a random variable associated to B, under  $r_A$  (Equation (2.20));
- $\varepsilon_{S,r_A,c,f,t'}$ : a random variable associated to S under  $r_A$  (Equation (2.20)).

For the duration model we specify:

- $V_D(R, t'|\omega_D)$ : the measure for the risk R (Equation (2.12));
- $W_D(r_D, t'|\omega_D)$ : the randomized measure of the market risk  $r_D$  (Equation (2.14));
- $m_D(c, t'|r_D, \theta, \beta)$ : a utility (Equation (2.27));
- $\lambda_D(c, t'|\theta, \omega_D, \beta)$ : scale parameter of the Weibull distribution (Equation (2.29));
- $\varepsilon_{D,r_D,t'}$ : a random variable associated to the risk  $r_D$  in the duration model (Equation (2.15));
- $\epsilon_{D,t'}$ : a random variable,  $D(c, t)$  is assumed to be its mean (Equation (2.25)).

Descriptive statistics shown in Table 2.3 allow to underline a significant difference of behavior between investors managing funds 1, 2, 3 and 4, 5, 6. We account for this difference in the specification of the models. In Section 2.7.1 the risk perception in both models are detailed. In Section 2.7.2 the action model is presented, and in Section 2.7.3 it is the duration model.

### 2.7.1 The risk perception

Day and Huang (1990) define three market types: *bear*, *bull* and *sheep* markets. A bear market corresponds to a decreasing confidence of the investors in the market, generating an increase of the risk in terms of returns. This is the contrary for the bull market. The sheep market represents an intermediary position between the bull and bear markets. In that case, the majority of the investors follow the market tendencies. According to the investors implicated in the observed decisions, two types of behavior occur depending on the market risk. The first behavior is called *normal*, corresponding to the bull and sheep markets. The second is called *risky* corresponding to the bear market. The risk perception is not directly observed, and has a strong influence on the investors' decisions. Two models for the risk perception have been developed, one associated to the action model and one to the duration model. No characteristics of the investors were available in the data, so the risk perception only depends on attributes of the decision context.

#### The risk perception in the action model

The risk  $r_A$  is a discrete variable. A model for risk classification is developed. A logit function is used. The deterministic measure of the risk  $R$  is

$$\begin{aligned} V_A(R, c, t' | \omega_A) &= ASC_{W_A} + \omega_{A,1} VIX_{t'} I_{c,g=1} + \omega_{A,2} VIX_{t'} I_{c,g=2} \\ &+ \omega_{A,3} \text{Sigm}(\text{Sent}_{c,t'}, 5), \end{aligned} \quad (2.7)$$

where  $I_{c,g}$  is an indicator equal to 1 if  $c$  belongs to  $g$ , 0 otherwise. We explicit

$$\omega_A = \{ASC_{W_A}, \omega_{A,1}, \omega_{A,2}, \omega_{A,3}\}, \quad (2.8)$$

The randomized measure of risks  $N$  and  $R$  are

$$\begin{aligned} W_A(N, c, t' | \omega_A) &= \varepsilon_{A,N,c,t'}, \\ W_A(R, c, t' | \omega_A) &= V_A(R, c, t' | \omega_A) + \varepsilon_{A,R,c,t'}, \end{aligned} \quad (2.9)$$

and assuming

$$\varepsilon_{A,r_A,c,t'} \stackrel{\text{i.i.d}}{\sim} \text{EV}(0, 1), \text{ for } r_A \in \{N, R\}, \quad (2.10)$$

a binary logit model is derived where the alternatives are the risks N and R. The associated probabilities are

$$\begin{aligned} P_A(N, c, t' | \omega_A) &= \frac{1}{1 + e^{V_A(R, c, t' | \omega_A)}}, \\ P_A(R, c, t' | \omega_A) &= \frac{1}{1 + e^{-V_A(R, c, t' | \omega_A)}}. \end{aligned} \quad (2.11)$$

We expect  $V_A(R, c, t' | \omega_A)$  (Equation (2.7)) to increase when the risk increases, so  $\omega_{A,1}$ ,  $\omega_{A,2}$  and  $\omega_{A,3}$  should be positively estimated. The risk perception depends both on the market and on  $c$ , due to presence of the VIX and  $\text{Sigm}(\text{Sent}_{c,t'}, 5)$  (attribute described in equation 2.4). This specification is motivated by the fact that the VIX can be seen as a risk measure as such.  $\text{Sent}$  reflects analyst opinions, and the standard-deviation characterizes its variation level. It seems logical to expect an increase of the risk perception, when analysts change frequently their mind. The perception of the VIX is supposed to differ between the two fund groups (see the correlation analysis in Section 2.6). Some tests have been performed to split other parameters between the two fund groups, but the a priori perceptions of risk, as well as the perception of  $\text{Sigm}(\text{Sent}_{c,t'}, 5)$  appeared to be generic (Section 2.8.1).

### The risk perception in the duration model

The risk  $r_D$  is a discrete variable. A model for risk classification is developed. A logit function is used. The deterministic measure of the risk R is

$$\begin{aligned} V_D(R, t' | \omega_D) &= \text{ASC}_{W_D,1} I_{c,g=1} + \text{ASC}_{W_D,2} I_{c,g=2} \\ &+ \omega_{D,1} \text{VIX}_{t'} I_{c,g=1} + \omega_{D,2} \text{VIX}_{t'} I_{c,g=2}, \end{aligned} \quad (2.12)$$

where  $I_{c,g}$  is an indicator equal to 1 if  $c$  belongs to  $g$ , 0 otherwise. We explicit

$$\omega_D = \{\text{ASC}_{W_D}, \omega_{D,1}, \omega_{D,2}\} \quad (2.13)$$

The randomized measure of the risks N and R are

$$\begin{aligned} W_D(N, t' | \omega_D) &= \varepsilon_{D,N,t'}, \\ W_D(R, t' | \omega_D) &= V_D(R, t' | \omega_D) + \varepsilon_{D,R,t'}, \end{aligned} \quad (2.14)$$

and assuming

$$\varepsilon_{D,r_D,t'} \stackrel{\text{i.i.d}}{\sim} \text{EV}(0, 1), \text{ for } r_D \in \{N, R\}, \quad (2.15)$$



a binary logit model is derived where the alternatives are the risks  $\mathbf{N}$  and  $\mathbf{R}$ . The associated probabilities are

$$\begin{aligned} P_{\mathbf{D}}(\mathbf{N}, t' | \omega_{\mathbf{D}}) &= \frac{1}{1 + e^{V_{\mathbf{D}}(\mathbf{N}, t' | \omega_{\mathbf{D}})}}, \\ P_{\mathbf{D}}(\mathbf{R}, t' | \omega_{\mathbf{D}}) &= \frac{1}{1 + e^{-V_{\mathbf{D}}(\mathbf{R}, t' | \omega_{\mathbf{D}})}}. \end{aligned} \quad (2.16)$$

We expect  $V_{\mathbf{D}}(\mathbf{R}, t' | \omega_{\mathbf{D}})$  (Equation (2.12)) to increase when the risk increases, so  $\omega_{\mathbf{D},1}$ ,  $\omega_{\mathbf{D},2}$  should be positively estimated. Contrary to  $r_{\mathbf{A}}$ , the risk perception depends only on the market. Some tests have been performed with  $\text{Sent}$ , but the results were not satisfactory (see Section 2.8.2). However, the difference between the two fund groups was relevant for the a priori perceptions and for the  $\text{VIX}$  (see Section 2.8.2).

### 2.7.2 The action model

The choice of action is a discrete choice situation. We develop a binary logit model with two latent classes corresponding to the two risk situations. The random utilities of the two alternatives  $\mathbf{B}$  and  $\mathbf{S}$  are

$$\begin{aligned} U_{\mathbf{B}}(\mathbf{c}, t' | r_{\mathbf{A}}, \beta) &= V_{\mathbf{B}}(\mathbf{c}, t' | r_{\mathbf{A}}, \beta) \\ &+ M_{\mathbf{B}}(\mathbf{c}, t' | r_{\mathbf{A}}, \beta) \\ &+ \varepsilon_{\mathbf{B}, r_{\mathbf{A}}, \mathbf{c}, f, t'}, \\ U_{\mathbf{S}}(\mathbf{c}, t' | r_{\mathbf{A}}, \beta) &= \varepsilon_{\mathbf{S}, r_{\mathbf{A}}, \mathbf{c}, f, t'}, \end{aligned} \quad (2.17)$$

with

$$V_{\mathbf{B}}(\mathbf{c}, t' | r_{\mathbf{A}}, \beta) = \text{ASC}_{\mathbf{B}, r_{\mathbf{A}}} + \sum_{g \in \{1, 2\}} I_{g, \mathbf{c}} \sum_{k=1}^{K_{\mathbf{B}}} \beta_{\mathbf{B}, k} \sum_{l=1}^{K_{\mathbf{X}}} I_{\mathbf{B}, g, k, l, r_{\mathbf{A}}} X_{\mathbf{c}, t'}(l) \quad (2.18)$$

where  $I_{\mathbf{B}, g, k, l, r_{\mathbf{A}}}$  is an indicator equal to 1, if the parameter  $\beta_{\mathbf{B}, k}$  is associated to the attribute  $X_{\mathbf{c}, t'}(l)$ , to the group of fund  $g$  and appear under the risk  $r_{\mathbf{A}}$ .  $I_{g, \mathbf{c}}$  is an indicator equal to 1 if  $\mathbf{c}$  belongs to  $g$ .

$$M_{\mathbf{B}}(\mathbf{c}, t' | r_{\mathbf{A}}, \beta) = \alpha_{\mathbf{B}, r_{\mathbf{A}}} V_{\mathbf{B}}(\mathbf{c}, t | r_{\mathbf{A}}, \beta) e^{\lambda_{\mathbf{B}, r_{\mathbf{A}}} D(\mathbf{c}, t)} \quad (2.19)$$

is a term accounting for the effect of the previous action, represented by the deterministic utility of  $\mathbf{B}$  in the previous action, weighted by  $D(\mathbf{c}, t)$ , the duration between the last and the current action performed on stocks of  $\mathbf{c}$ . The assumptions about the random terms are

$$\varepsilon_{\mathbf{B}, r_{\mathbf{A}}, \mathbf{c}, f, t'}, \varepsilon_{\mathbf{S}, r_{\mathbf{A}}, \mathbf{c}, f, t'} \stackrel{\text{i.i.d.}}{\sim} \text{EV}(0, \mu_f), \text{ for } r_{\mathbf{A}} \in \{\mathbf{N}, \mathbf{R}\}, \quad (2.20)$$

The probabilities of the actions B and S under the risk  $r_A$  are

$$\begin{aligned} P_B(c, t'|r_A, \beta) &= \frac{1}{1 + e^{-\mu_f V'_B}}, \\ P_S(c, t'|r_A, \beta) &= 1 - P_B(c, t'|r_A, \beta), \end{aligned} \quad (2.21)$$

where

$$V'_B = V_B(c, t'|r_A, \beta) + M_B(c, t'|r_A, \beta). \quad (2.22)$$

After having summed on the risks N and R, the probabilities of the actions come

$$\begin{aligned} P_B(c, t'|\beta, \omega_A) &= P_B(c, t'|N, \beta)P_A(N, c, t'|\omega_A) \\ &\quad + P_B(c, t'|R, \beta)P_A(R, c, t'|\omega_A), \\ P_S(c, t'|\beta, \omega_A) &= 1 - P_B(c, t'|\beta, \omega_A), \end{aligned} \quad (2.23)$$

where  $P_A(N, c, t'|\omega_A)$ ,  $P_A(R, c, t'|\omega_A)$  are the probabilities to be in the risk R defined in equation (2.11). The vector of parameters  $\beta$  is then

$$\beta = \{\{\beta_{B,k}\}_{k=1\dots K_B}, \{ASC_{B,r}, \alpha_{B,r}, \lambda_{B,r}\}_{r=N,R}, \{\mu_f\}_{f=2\dots 6}\}. \quad (2.24)$$

Only attributes for  $t_H = 1$  are used in the deterministic utility shown in Equation (2.18). This results from the statistical analysis presented in Table 2.3 and explained in Section 2.6. Indeed, dynamic variables calculated with small  $t_H$  are significantly correlated with the action choice. This means that investors have tendency to make decisions using short-range information, showing their reactivity. In a multivariate context  $t_H = 1$  appears to be the most appropriate (see Section 2.8).  $M_B(c, t'|r_A, \beta)$  is the memory effect. In this term, the deterministic utility of the previous *buy* alternative has been chosen to represent the previous action because it is tractable for prediction. We expect  $\lambda_{B,r_A}$  to be negative, as we suppose the impact of the previous action to decrease when the duration between the previous and the current action increases. A scale parameter is associated to each fund in order to account for the behavioral specificity of the investors within funds. Note that  $\mu_1$  has been fixed to 1 because all the  $\mu_f$  are identifiable, except one.

### 2.7.3 The duration model

After the action, we model the duration until the next action. This duration is supposed to depend on this latter decision. We expect the investor not to change his mind when waiting for the next action. This assumption is made for easing the mathematical formulation of the likelihood function, and consequently the estimation of the combined model. In this section, we model the survival of the action  $A(c, t')$ . The distribution of the observed duration between two actions is presented in Figure

2.11, proving the survival nature of the underlying phenomenon. A lifetime model has been chosen to handle the duration. The Weibull regression model has been retained, because it mimics the exponential model, with a more flexible formulation.

$D(c, t')$  is  $\bar{\epsilon}_{D,t'}$ , the mean of the random variable  $\epsilon_{D,t'}$ .  $\epsilon_{D,t'}$  is assumed to follow a Weibull distribution.  $\lambda_D(c, t'|\theta, \omega_D, \beta)$  is the scale parameter of the Weibull distribution, and  $\eta_D$  the shape parameter.

$$D(c, t') = \bar{\epsilon}_{D,t'} \text{ with } \epsilon_{D,t'} \sim W(\lambda_D(c, t'|\theta, \omega_D, \beta), \eta_D). \quad (2.25)$$

We define a utility

$$\begin{aligned} m_D(c, t'|r_D, \theta, \beta) &= \sum_{g \in \{1,2\}} I_{g,c} ASC_{D,r_D,g} + \sum_{g \in \{1,2\}} I_{g,c} \sum_{k=1}^{K_D-1} \theta_{D,k} \sum_{l=1}^{K_X} I_{D,g,k,l,r_D} X_{c,t'}(l) \\ &+ \theta_{D,K_D} \text{Sigm}_{VIX_t', 360} I_{r_D=N} \\ &+ \alpha_{D,N,1} I_{c,g=1} I_{r_D=N} D(c, t) \\ &+ \alpha_{D,N,2} I_{c,g=2} I_{r_D=N} D(c, t) \\ &+ \alpha_{D,R,2} I_{c,g=2} I_{r_D=R} D(c, t) \\ &+ \theta_{B,r_D} V_B(c, t'|r_D, \beta) \end{aligned} \quad (2.26)$$

where  $I_{D,g,k,l,r_D}$  is an indicator equal to 1 if the parameter  $\theta_{D,k}$  is associated to the attribute  $X_{c,t'}(l)$ , to the group of funds  $g$ , and associated to the risk  $r_D$ .  $I_{g,c}$  is an indicator equal to 1 if the company  $c$  belongs to the group of funds  $g$ .  $I_{r_D=N}$  is an indicator equal to 1 if  $r_D = N$ , 0 otherwise.  $I_{r_D=R} = 1 - I_{r_D=N}$ . In order to get rid of the risk  $r_D$  in  $m_D(c, t'|r_D, \theta)$ , we need to sum on levels of  $r_D$

$$\begin{aligned} m_D(c, t'|\theta, \omega_D, \beta) &= m_D(c, t'|N, \theta, \beta) P_D(N, t'|\omega_D) \\ &+ m_D(c, t'|R, \theta, \beta) P_D(R, t'|\omega_D), \end{aligned} \quad (2.27)$$

where  $P_D(N, t'|\omega_D)$  and  $P_D(R, t'|\omega_D)$  are shown in equation (2.16). The vector of parameters  $\theta$  is

$$\theta = \{ASC_{D,N}, ASC_{D,R}, \{\theta_{D,k}\}_{k=1 \dots K_D}, \alpha_{D,N,1}, \alpha_{D,N,2}, \alpha_{D,R,2}, \{\theta_{B,r_D}\}_{r_D=N,R}\}. \quad (2.28)$$

We have

$$\lambda_D(c, t'|\theta, \omega_D, \beta) = \frac{1}{e^{m_D(c, t'|\theta, \omega_D, \beta)}}, \quad (2.29)$$

consequently, if  $\Gamma(\cdot)$  denotes the gamma function

$$\begin{aligned}
D(c, t') &= \bar{\epsilon}_{D,t'} \\
&= \frac{1}{\lambda_D(c, t'|\theta, \omega_D)} \Gamma\left(1 + \frac{1}{\eta_D}\right) \\
&= e^{m_D(c, t'|\theta, \omega_D, \beta)} \Gamma\left(1 + \frac{1}{\eta_D}\right). \tag{2.30}
\end{aligned}$$

The density of the Weibull distribution calculated for  $D(c, t')$  is

$$\begin{aligned}
&f(D(c, t')|\lambda_D(c, t'|\theta, \omega_D, \beta), \eta_D) \\
&= \eta_D \lambda_D(c, t'|\theta, \omega_D, \beta)^{\eta_D} D(c, t')^{\eta_D - 1} e^{-(\lambda_D(c, t'|\theta, \omega_D, \beta) D(c, t'))^{\eta_D}}. \tag{2.31}
\end{aligned}$$

In Equation (2.26), only attributes for  $t_H = 60$  are used. They give the best model fit and provide interpretable parameters. Note that we did not present any descriptive statistics regarding  $D(c, t)$  in Section 2.6, because no generic correlation across funds appeared during the univariate analysis. In Equation (2.26), the influence of the previous duration  $D(c, t)$  for  $g = 1$  and  $r_D = R$  has been discarded because the associated parameter did not appear to be significant (see Section 2.8).

#### 2.7.4 The likelihood function

The likelihood of the action model is

$$\begin{aligned}
l_A(\beta, \omega_A) &= \prod_{c=1}^C \prod_{t'=2}^T (P_B(c, t'|\beta, \omega_A)^{z_{B,c,t'} I_{c,t'}} \\
&\quad \times P_S(c, t'|\beta, \omega_A)^{(1-z_{B,c,t'}) I_{c,t'}}), \tag{2.32}
\end{aligned}$$

$C$  is the number of companies.  $z_{B,c,t'}$  is an indicator equal to 1 if stocks of  $c$  have been bought on  $t'$ , 0 otherwise.  $I_{c,t'}$  is an indicator equal to 1 if an action has been observed on  $t'$  for  $c$ , 0 otherwise. Concerning the duration model, the likelihood function is

$$l_D(\theta, \omega_D, \beta) = \prod_{c=1}^C \prod_{t'=1}^{T-1} f(D(c, t')|\lambda_D(c, t'|\theta, \omega_D, \beta), \eta_D)^{I_{c,t'}}, \tag{2.33}$$

where  $I_{c,t'}$  is an indicator equal to 1 if an action has been performed for  $c$  on  $t'$ . Then, the joint likelihood function is

$$l(\beta, \omega_A, \theta, \omega_D) = l_A(\beta, \omega_A) l_D(\theta, \omega_D, \beta) \tag{2.34}$$

and the log-likelihood function

$$L(\beta, \omega_A, \theta, \omega_D) = \log(l(\beta, \omega_A, \theta, \omega_D)) \tag{2.35}$$

### 2.7.5 Alternative specifications

In this chapter, we focus on the modeling of the action choice and the duration between two actions. Other tests have been done to ensure the quality of the proposed approach. In the early stage of the analysis, a discrete choice model with three alternatives has been developed. The three alternatives were *buy*, *sell*, and *wait*. The assumption was that every day an investor decides to buy or sell stocks, or wait for investing. This previous approach allowed to get rid of the duration model. This model has not been kept because it had some problems, due to the prevalence of *wait* actions in the data. The estimation results showed a dominance of the alternative specific constants and some very weak elasticities associated to the explanatory variables.

The invested quantity of money is important to model in order to have a complete picture of the investors behavior. We did not include it in the final analysis. Some attempts have been conducted, but very limited causalities could be captured, with extremely low prediction capabilities. The reason of these limitations is the lack of precious information in the data. Portfolio information are crucial to explain the invested quantity of money, but they are not available in the data for obvious sensitivity reasons.

## 2.8 Model estimation

The combined model is estimated by maximum likelihood using the biogeme software (Bierlaire (2003a) and Bierlaire and Fetiarison (2009)). The log-likelihood for the entire model is presented in equation (2.35). The processed data were used for estimation (see Sections 2.5 and 2.6). General estimation results are displayed in Table 2.5. The 61 parameters of the combined model are split between the action and duration models (respectively 29 and 32 parameters). The number of observations for the duration model is equal to the number of actions (9178), minus the number of companies (1121). For the last observed decision associated to a company, we do not know the duration until the next action. The duration model has a big impact on the log-likelihood of the combined model. The  $R^2$  of the duration model is low. Interpretations of the parameters are explained in the following for the action and duration models.

### 2.8.1 The action model

Parameters values and associated t-tests are presented in Tables A.1 and A.2.

- **the risk perception in the action model:**  $ASC_{W_A}$  was not significantly different from minus the VIX threshold defined in Figure 2.3. This value was the starting value for estimation.  $\omega_{A,1}$  and  $\omega_{A,2}$  are the two parameters associated to the VIX, respectively for the groups of funds 1 and 2. Logically,

	Action choice model	Duration model	Combined model
Nb parameters	29	32	61
Nb observations	9178	8057	9178
Null Log-likelihood	-6361.705		
Final Log-likelihood	-5635.580	-19711.528	-25347.109
$\bar{\rho}^2/R^2$	0.109	0.048	

Table 2.5: General estimation results of the action and duration models

both parameters are positive meaning that the risk increases with the VIX, as described in Sections 2.7.1 and 2.7.1.  $\omega_{A,1} > \omega_{A,2}$ , so investors managing funds in group 1 are more sensitive to risk than investors managing group 2.  $\omega_{A,3}$  is associated to the standard error of sentiment calculated for  $t_H = 5$  (days). It is positive as expected, showing that the increase of the fluctuations in the analyst opinions increases the perception of risk, which is logical.

- **The alternative specific constants:**  $ASC_{B,N}$  and  $ASC_{R,N}$  are negative, meaning that in the two risk situations, *buy* is penalized.
- **The parameters associated with the explanatory variables:** A statistical analysis has been conducted (see Table 2.3) and several model specifications have been tested. The variables which appeared to be the best appropriated to this analysis, were calculated with  $t_H = 1$  (day), meaning that the investors base their choices on short-term information. No  $\beta_{B,k}$  ( $k = 1 \dots 15$ ) is generic across risk situations and groups of funds. The generic parameters across groups of funds are  $\beta_{B,4}$ ,  $\beta_{B,5}$ ,  $\beta_{B,6}$  and  $\beta_{B,15}$ .  $\beta_{B,6}$  and  $\beta_{B,15}$  are positive, Under the risk N, the increase of the *short-term values* of *quality* and *value* favor the *buy* alternative.  $\beta_{B,5}$  is negative, an increase of the *long-term value* of *quality* favor the *sell* alternative.  $\beta_{B,4}$  is positive, so under the risk R, the increase of the *short-term value* of *price* increases the probability to buy, underlying the tendency of investors for taking advantage of immediate fluctuations of variables.
- **The memory effect:**  $\alpha_{B,N}$  and  $\alpha_{B,R}$  are both negative showing that for a given stock, investors have not the tendency to perform consecutively two *buy* actions. This is stronger in under the risk R than in under the risk N. Investors are more likely to bet on short-term returns in under R, which is logical. As expected, this effect is attenuated with the increase of the duration between two consecutive actions, as shown by the negative value of  $\lambda_{B,N}$  and  $\lambda_{B,R}$ . The attenuation is higher in a *normal* situation than in a *risky* situation.
- **The scale parameters:**  $\mu_1$  is fixed to 1 for identification reasons, and  $\{\mu_f\}_{f \in \{2 \dots 6\}}$

are significantly different from  $\mu_1$ , showing the specificities of the investors' behavior within each fund. The variance of the error terms  $\varepsilon_{B,RA,c,f,t'}$  and  $\varepsilon_{S,RA,c,f,t'}$  (introduced in equation (2.17)) is

$$\text{Var} = \frac{\pi^2}{6\mu_f^2}, \quad (2.36)$$

as  $\mu_f > \mu_1$  for  $f \in \{2, 3, 4, 5, 6\}$ , the variance associated to the choice of action is higher for the fund 1 compared to the others, meaning that the decisions taken in funds 2, 3, 4, 5, 6 are apparently more rational than the decisions taken in fund 1.

### 2.8.2 The duration model

Parameters values and associated t-tests are shown and in Tables A.3 and A.4.

- **the risk perception in the duration model:**  $ASC_{WD,1}$  and  $ASC_{WD,2}$  are negative, as expected. Investors of fund groups 1 have a stronger a priori toward the risk N, compared to investors of group 2 ( $ASC_{WD,1} < ASC_{WD,2}$ ).  $\omega_{D,1}$  and  $\omega_{D,2}$  are positive as expected. Investors of fund group 1 are more sensitive to risk than investors of group 2 ( $\omega_{D,1} > \omega_{D,2}$ ).
- **The shape parameter:**  $\eta_D$  is the shape parameter of the Weibull regression model, which has been estimated under 1, showing the proximity to a an exponential regression model, but with a distribution characterized by a less heavy right tail.
- **The constants:**  $ASC_{D,N,1}$ ,  $ASC_{D,N,2}$ ,  $ASC_{D,R,1}$ ,  $ASC_{D,R,2}$  are all positive. They characterize the average duration in pure risk situations and are specific to the fund groups, all other attributes being equal to 0. The average of the predicted duration is presented in Equation (2.30). The average of the duration in the situation N for the fund group 1 is 68.9 weeks, and for group 2, 57.6 weeks. The average of the duration under risk R for the fund group 1 is 2.5 weeks, and for group 2, 2.9 weeks. This seems logical, because investors have to react much more faster in *risky* situations than in *normal* situations.
- **The parameters associated with the explanatory variables:** No parameter is generic across fund groups and risk situations, which underline the strong specificities of the investors' behaviors.  $\theta_{D,16}$  is associated to the *standard error* of the VIX calculated over 360 days, for the fund group 2 under the risk N. It is negative, which is logical because when the variation of the principle risk indicator of the market increases, investors have tendency to perform actions more often.  $\theta_{B,N}$  and  $\theta_{B,R}$  are negative, so when the utility of the *buy* alternative increases, the duration decreases under both risk N and R. In any case,

when money is invested in a company, the vigilance of the investor toward the company increases, and he is more likely to adjust his decision in a near future.

- **The influence of the previous duration:**  $\alpha_{D,N,1}$ ,  $\alpha_{D,R,1}$  and  $\alpha_{D,N,2}$  are positive. This shows the stability of the phenomenon, in the sense that the increase of the previous duration generates an increase of the current duration.

To conclude this section, the parameters of the combined model are significant and interpretable. The interpretations have been discussed with the involved investors. The variables presented in the Section 2.6 are adapted to explain the investors' behavior. Behavioral specificities within each fund appear clearly.

## 2.9 Model prediction

In this section we study the prediction accuracy of the combined model. We start by examining the model prediction on the estimation data and perform a cross-validation. For the action model, we conduct a simulation analysis.

### 2.9.1 Prediction on the estimation data

The frequencies of the predicted probabilities of the observed actions are shown in Figure 2.12. If the model was perfect, all predicted probabilities should be equal to 1. This corresponds to a log-likelihood equal to 0. This is of course never the case, but a significant shift of the distribution on the right is observed. In addition 66.25% of the actions are predicted with a probability higher than 0.5, represented by the grey bin. The model is compared to a simple binary logit model. It contains only two parameters which are the constants in the deterministic utilities of *buy* and *sell*, and without risk perception. It has the property to reproduce the aggregated shares of actions of the estimation data, when used for prediction. In that case, there are approximately as much actions of type *Buy* (4530) than action of type *Sell* (4648). This simple model predicts a quasi equal probability for the two actions ( $\sim 0.5$ , for the two actions). Consequently for 66.25% of the decisions, the proposed model makes a better prediction than this simple model.

The standardized residual of the duration model  $z_{c,t}$  is defined as

$$z_{c,t} = \eta_D(\log(D(c, t)) - m_D(c, t|\theta, \omega_D, \beta)) \sim EV(0, 1), \quad (2.37)$$

which makes the parallel with the lognormal regression where the standardized residuals calculated with the logarithm of the dependent variable are supposed normally distributed  $N(0, 1)$ . The distribution of  $\{z_{c,t}\}$  is plotted in Figure 2.13 (histogram), as well as the theoretical distribution  $EV(0, 1)$  (curve). The observed distribution is near from the theoretical curve, but the model tends to over-predict the duration



for some observations. A bi-modality appears. The over-predictions concerns very small durations. The duration model works well for situations presented in Figure 2.4 where actions are sparse, but not for situations displayed in Figure 2.5, where actions are concentrated. Several regression models have been tested, such as the lognormal, Poisson, negative binomial, exponential and Rayleigh (these two latter are particular cases of the Weibull). In terms of fit and residual analysis, the Weibull appeared to be the best. Regarding the specification, many trials have been done. The improvement of the residual distribution is possible by the using of distribution mixtures and the integration of supplementary data, such as money flows.

This prediction analysis is performed on the estimation data. It reinforces the estimation results (see Section 2.8), but the models have not been yet tested for forecasting.

### 2.9.2 Cross-validation

We need to check the prediction capability of the combined model. The cross-validation consists in estimating the combined model on a part of the data, and simulate on the remaining part. The total time horizon is divided in five periods of equal duration. The starting date and ending date of each period as well as the number of actions per period are shown in Table 2.6. The estimation of the combined model is done on 4 subsets and the simulation on the remaining subset. The experience is repeated five times in order to cover all the possibilities.

Validation set	1	2	3	4	5
Starting date	2005.01.03	2006.03.02	2007.04.25	2008.06.05	2009.07.24
Ending date	2006.03.01	2007.04.24	2008.06.04	2009.07.23	2010.09.13
Nb actions	1363	1511	1766	2149	2399

Table 2.6: Starting dates, ending dates and number of actions per subset of data for the cross-validation

For each experience, we calculate statistics revealing the prediction accuracy of the models on the simulation subset. Concerning the action model, the predicted log-likelihood is calculated (Predicted  $\mathcal{L}$ ). For the duration model,  $R_e^2$  is the predicted  $R^2$  for experience  $e$ . These values are respectively compared to the log-likelihood (Estimated  $\mathcal{L}$ ) and the  $R^2$  (Estimated  $R^2$ ) obtained when applying the combined model (estimated on the entire data) on the simulation subset.

The results are presented in Table 2.7. Logically the estimated  $R^2$  and  $\mathcal{L}$  are always higher than  $R_e^2$  and the predicted  $\mathcal{L}$ . Regarding the action model, the log-likelihood

increases chronologically because the volume of decisions is also increasing (see Table 2.6). The higher difference between the estimated and predicted log-likelihood is observed for the experience 5 (235.141). However for every experiences, both log-likelihoods have the same order. This shows the stability of the action model. This is not the case for the duration model. We explicit  $R_e^2$ .  $t_{0,e}$  is the starting date of the simulation subset for experience  $e$ , and  $T_e$  is the ending date.  $D(c, t)$  is the observed duration for day  $t$  and company  $c$ ,  $\hat{D}(c, t)$  is the predicted duration, and  $\bar{D}(c, t)$  is the duration mean calculated over  $t$  and  $c$ .

$$R_e^2 = 1 - \frac{\sum_{c=1}^C \sum_{t=t_{0,e}}^{T_e} (\hat{D}(c, t) - D(c, t))^2}{\sum_{c=1}^C \sum_{t=t_{0,e}}^{T_e} (D(c, t) - \bar{D}(c, t))^2} \quad (2.38)$$

$R_e^2$  is negative for  $e \in \{3, 4, 5\}$ . The predictions are worse than those of the simple model predicting the duration mean on the considered time period. The experience 5 is even not well predicted by the duration model estimated on the entire data. Two explanations are invoked: this is an heterogeneous period in terms of risk and behavior, compared to the other periods. The market tends to return to a normal risk situation, but with lots of aftershocks related to the 2008 crisis. This is shown by significant picks on the VIX curve in Figure 2.3. At the beginning of the period, fund managers reinvest a lot, as shown in Figure 2.6. Then, the average of action volumes decrease, but adjustments are still performed.

Experience	1	2	3	4	5
$R_e^2$	0.034	0.047	-0.003	-0.003	-1.790
Estimated $R^2$	0.053	0.065	0.010	0.025	-0.464
Predicted $\mathcal{L}$	-394.436	-431.504	-1095.921	-1269.583	-1588.811
Estimated $\mathcal{L}$	-383.671	-420.055	-1052.993	-1220.002	-1353.67

Table 2.7: Results of the cross-validation performed on the estimation data

Specifically to the action model and for each experience, we have calculated the percentage of observations predicted with a probability less than 0.5, which are considered badly predicted. The results are displayed in Table 2.8. For the action model estimated on the entire data, there are 33.75% of bad predictions. For every experience, the percentage is similar to this value. This underlines the stability of this model.

Experience	1	2	3	4	5
Action model	36.90	38.65	31.65	30.15	41.31

Table 2.8: Percentages of badly predicted observations per experience of cross-validation

The financial crisis of 2008 appears in the subset 4 (see Table 2.6). In this period, the prediction of the duration model are bad ( $R_4^2 < 0$ ). This is not the case for the action model (30.15% of badly predicted observation), which shows its robustness. The cross-validation is a first step toward forecasting, the results are worth for the action model and limited for the duration model.

### 2.9.3 Simulation

In this section, we present a concrete forecasting application of the action model. The experience 5 of the cross-validation is considered (see Table 2.6). The action model is estimated on the calibration subset and applied on the simulation subset. We hypothesize that the action days are fixed. Five simulations are performed on the simulation subset. In each simulation and for each action day, an action is drawn from the predicted probability distribution. The number of *buy* and *sell* actions are aggregated per month. Results are presented in Table 2.9. For each month, the number of *buy* actions,  $nb\_buy\_sim$ , and *sell* actions  $nb\_sell\_sim$ , are shown as “ $nb\_buy\_sim/nb\_sell\_sim$ ”. The observed shares are also displayed in the column “Reality” (“ $nb\_buy\_obs/nb\_sell\_obs$ ”).

Month	Reality	Simul. 1	Simul. 2	Simul. 3	Simul. 4	Simul. 5
Jul. 09	21/31	20/32	25/27	26/26	22/30	18/34
Aug. 09	121/81	109/93	96/106	96/106	94/108	108/94
Sep. 09	108/84	91/101	106/86	116/76	101/91	108/84
Oct. 09	110/141	121/130	130/121	116/135	116/135	112/139
Nov. 09	105/138	109/134	112/131	117/126	113/130	104/139
Dec. 09	76/93	75/94	86/83	77/92	75/94	58/111
Jan. 10	69/82	71/80	70/81	67/84	69/82	70/81
Feb. 10	69/77	68/78	61/85	56/90	68/78	67/79
Mar. 10	101/79	87/93	90/90	92/88	94/86	91/89
Apr. 10	148/54	96/106	112/90	104/98	101/101	108/94
May 10	85/84	76/93	85/84	75/94	83/86	86/83
Jun. 10	41/48	34/55	38/51	41/48	49/40	40/49
Jul. 10	53/96	65/84	77/72	78/71	72/77	81/68
Aug. 10	79/93	80/92	86/86	64/108	86/86	74/98
Sep. 10	17/15	23/9	14/18	16/16	20/12	15/17

Table 2.9: Results of the simulations performed with the action model on the period going from 2009.07.24 to 2010.09.13

For each month, the simulated and observed shares are compared. We define the

percentage of error

$$\text{Err} = \frac{\text{n\_buy\_obs} - \text{n\_buy\_sim}}{\text{n\_buy\_obs} + \text{n\_sell\_obs}} \times 100, \quad (2.39)$$

it corresponds to the number of false simulated actions divided by the total number of actions within the month. As the action days are fixed, we have  $\text{n\_buy\_obs} + \text{n\_sell\_obs} = \text{n\_buy\_sim} + \text{n\_sell\_sim}$ . The percentages of error are shown in Table 2.10. The highest values are observed for April 2010, otherwise no percentage of error is above 20%. Moreover, 81.33% of the percentages of error are under 10%, showing the forecasting accuracy of the action model. The aggregation of the results has been done per month, which is relatively detailed specially for long-term investments. This simulation emphasizes the good quality and usefulness of the model in real-life applications.

Month	Simul. 1	Simul. 2	Simul. 3	Simul. 4	Simul. 5
Jul. 09	1.92	-7.69	-9.62	-1.92	5.77
Aug. 09	5.94	12.38	12.38	13.37	6.44
Sep. 09	8.85	1.04	-4.17	3.65	0.00
Oct. 09	-4.38	-7.97	-2.39	-2.39	-0.80
Nov. 09	-1.65	-2.88	-4.94	-3.29	0.41
Dec. 09	0.59	-5.92	-0.59	0.59	10.65
Jan. 10	-1.32	-0.66	1.32	0.00	-0.66
Feb. 10	0.68	5.48	8.90	0.68	1.37
Mar. 10	7.78	6.11	5.00	3.89	5.56
Apr. 10	25.74	17.82	21.78	23.27	19.80
May 10	5.33	0.00	5.92	1.18	-0.59
Jun. 10	7.87	3.37	0.00	-8.99	1.12
Jul. 10	-8.05	-16.11	-16.78	-12.75	-18.79
Aug. 10	-0.58	-4.07	8.72	-4.07	2.91
Sep. 10	-18.75	9.38	3.13	-9.38	6.25

Table 2.10: Percentages of error (see Equation (2.39)) based on simulations of Table 2.9

Note that a simulation using the combined model has been tested. The obtained results were not convincing. The reason is the limited predictive power of the duration model (see Section 2.9.2). Consequently we focused on the action model.

## 2.10 Contributions

The analysis of the investors' behavior has raised several contributions.

In general, the lack of disaggregate data valuable for studying the investors' behavior, obliges analysts to work with synthetic data (De Grauwe and Grimaldi, 2004). In these works, most of the models are rule-based. This approaches are interesting and allow to go deep in the simulation. The main drawback is the distance to the reality, in terms of understanding and applications. In this work, we have the chance to work with real disaggregate data. Behavioral hypothesis have been built by observing investors and exploring data with descriptive statistics. These hypothesis have been confirmed by estimating the proposed model on the data. In addition, the behavioral finance mainly focuses on the qualification of the behaviors, but not on their quantifications (Baker and Wurgler, 2007). Our approach allows to measure the causalities, which is an important contribution. We mainly showed that investors have tendency to use short-range information to make decisions, they consider previous actions when making decisions, and they are highly influenced by their risk perception.

Finance is characterized by an abundance of aggregate data, which are very noisy. Machine learning methods have been naturally applied in this field, due to the huge amount of available data, and the difficulty for identifying causalities (Kim, 2003). Their main drawback is the over-fitting. The proposed action model presents good prediction capabilities, which emphasizes the added-value of the behavioral hypothesis setting. Hypothesis allow to compensate the intrinsic noise of the data.

Few discrete choice models have been developed in finance. We have translated and prove the relevance of mathematical formulations which have been mainly proposed in transportation and marketing.

## 2.11 Conclusion

We developed models capturing the investors' behavior. The data provided by the Swiss private bank **Lombard Odier** have been processed to infer the decisions performed by investors. Variables are computed based on the five indicators and the VIX, for reflecting the dynamics of the observed behavior. A correlation analysis allows to understand the links between the decisions and the decisional context. An integrated approach has been proposed to model simultaneously the choice of action (*buy* or *sell*) and the duration between two actions. A binary logit model with latent classes is developed for the choice of action, where the latent classes correspond to the risk situations. Two situations are considered, *normal* and *risky*. The duration is handled using a Weibull regression, which also accounts for the two risk situations. But, the perception is different compared to the action model. Both models account explicitly for the dynamics of the behavioral phenomenon. The two models are linked, because the perception of the *buy* alternative enters in the duration model.

The combination of the action and duration models is estimated simultaneously, using the processed data. Parameters are interpretable. Results have been discussed with investors. They explicit causalities and reveal behavioral mechanisms. The accounting of the dynamics has sense, investors consider their previous decisions when taking current decisions. The specificity of the investors' behavior within each fund appears. The hypothesis about the risk perception is valid in the action and duration models. In *risky* situations, the duration between two consecutive actions is shorter than in *normal* situations, which is logical, as investors tend to adjust more often their portfolio in *risky* situations. Regarding the action choice model, the specificity of the behavior per fund is more important in *risky* than in *normal* situations. This is logical, as individualities and emotions are emphasized in panic situations.

Predictions of the models have been checked and the combined model has been cross-validated on the estimation data. The predictive accuracy of the action model is good, and limited for the duration model. This is not surprising for the duration model because the associated modeling assumptions are moved from the reality. This is due to the fact that we have decided to develop the most realistic models remaining operational. In addition, crucial data are missing about money inflows and outflows to characterize this duration. The relevance of the action model has been shown by the simulation which is the practical way to use it. As it is, the action model can be embedded in a simulator for forecasting aggregated action shares based on market scenarios. For fund managers and investors, this could be relevant decision-aid tool as well as a potential starting point for a quantitative investment strategy.

There are several perspectives to this work. In terms of modeling, the risk perception can be refined for both models. More than two latent classes could be considered. The integration of relevant supplementary data can also be considered, such as portfolio information, money inflows and outflows, and investors' characteristics. Dedicated data collection could be conducted in order to point out precisely the information used by investors when taking decisions and refine the explanatory variables. Then, the proposed models can be improved. Regarding the duration model, mixtures of distribution can be considered. In addition, the assumption about the non revision of the duration once the decision has been taken, can be relaxed. The forecasting accuracy of the action model has been tested by simulation. The same thing with the combined model could be considered with a better duration model.

In this analysis, we focused on the investors' decisions because they are the sources of the observed trades. It would be interesting to model the behavior of the traders who are implementing the investors' decisions, in order to obtain a full picture of the decision process inside the bank. Then, It would be also useful to compare the investors' behavior with the traders' behavior.

The prediction of the stock prices evolution is a further perspective. The developed

model is not sufficient, other financial actors should be modeled, in order to capture the entire financial scene. This requires to have specific and detailed behavioral data about all the actors and data about the scene, which is an utopian task. Then, accurate behavioral models should be developed. Finally, the different models could be embedded into a single simulator in order to predict the stock prices evolution.

A first step would be to focus on stocks, for which the actions performed by professional investors dominate those of other financial actors. This will ease the prediction of the price evolution, as the proposed model can make the major part of it.

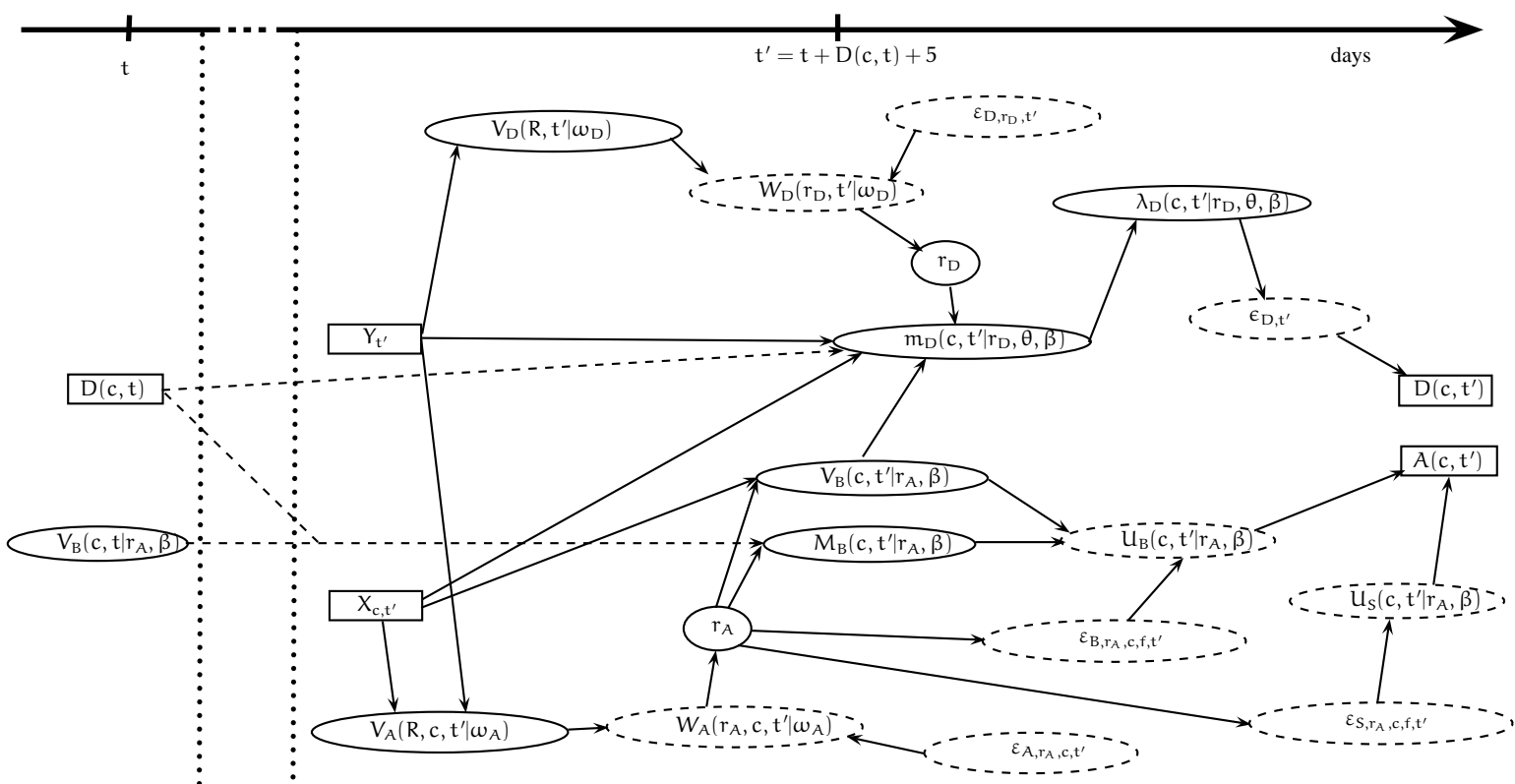


Figure 2.10: The general modeling framework



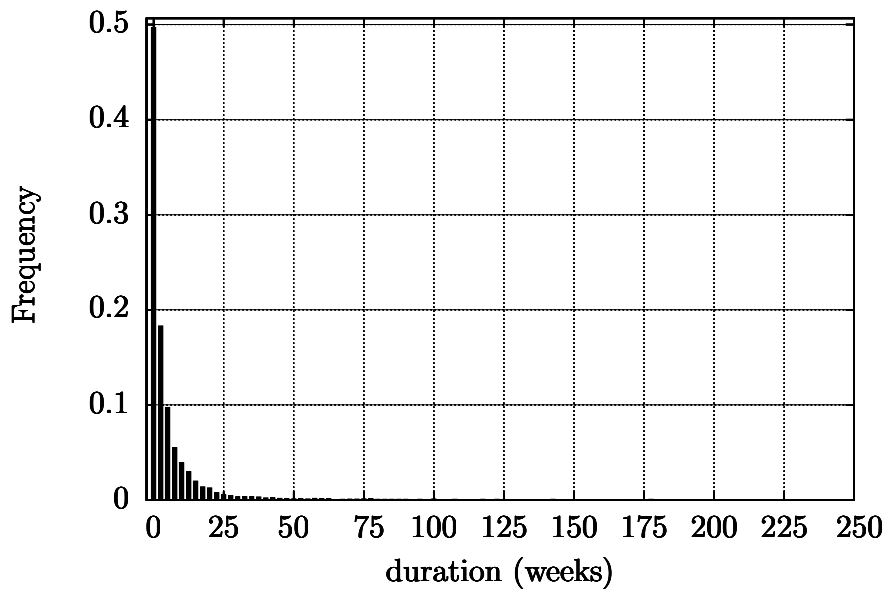


Figure 2.11: Distribution of the observed durations, expressed in weeks (5 days)

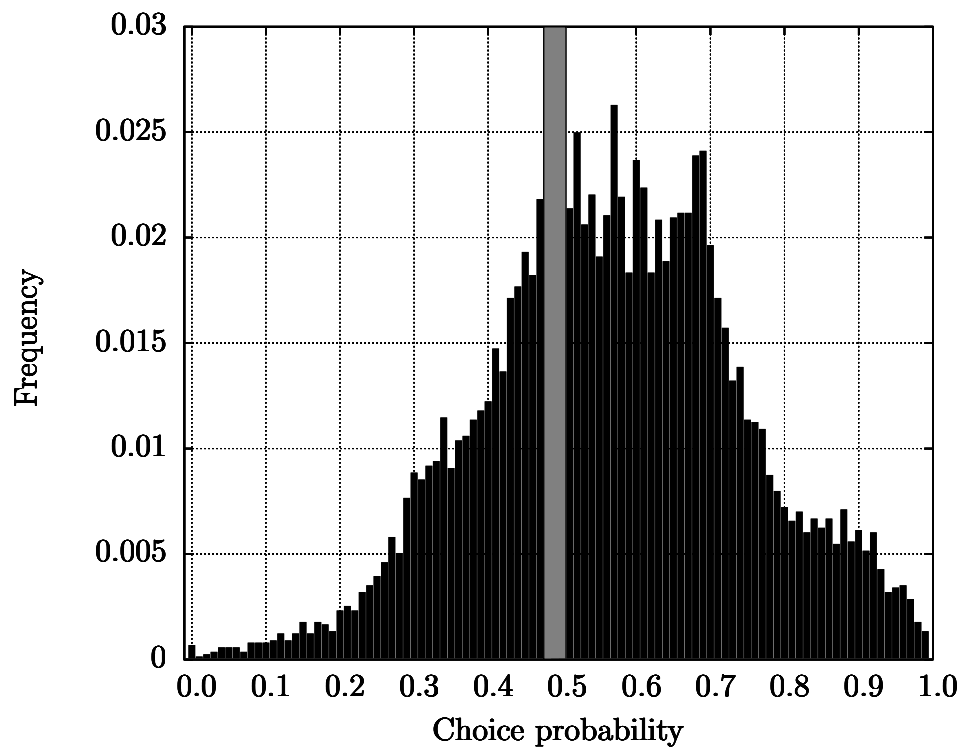


Figure 2.12: Distribution of the predicted action choice probabilities calculated on the estimation data

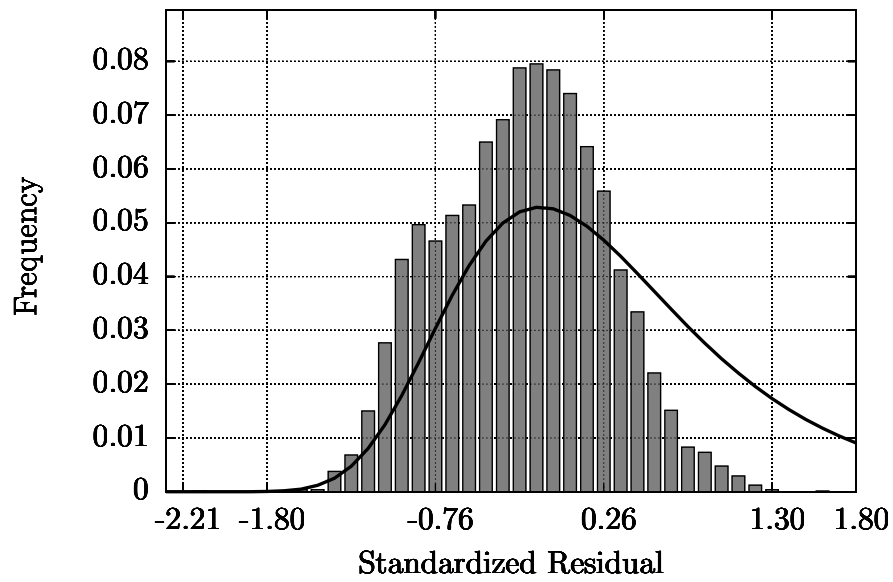


Figure 2.13: Comparison between the distribution of the residuals  $\{z_{c,t}\}$ , related to the duration model (histogram) and the theoretical distribution (curve)



# 3. The walking behavior: modeling of pedestrian movements

## 3.1 Introduction

We propose and validate a model for pedestrian walking behavior, based on discrete choice modeling. Two main types of behavior are identified: *unconstrained* and *constrained*. By unconstrained, we refer to behavior patterns which are independent from other individuals. The constrained patterns are captured by a *leader-follower* model and by a *collision avoidance* model. The spatial correlation between the alternatives is captured by a cross nested logit model. The model is estimated by maximum likelihood on a real data set of pedestrian trajectories, manually tracked from video sequences. The model is successfully validated using a bi-directional flow data set, collected in controlled experimental conditions at Delft university.

This chapter contains mainly the developments proposed by Robin et al. (2009).

## 3.2 Motivation

Pedestrian behavior modeling is an important topic in different contexts. Architects are interested in understanding how individuals move into buildings to create optimal space designs. Transport engineers face the problem of integration of transportation facilities, with particular emphasis on safety issues for pedestrians. Recent tragic events have increased the interest for automatic video surveillance systems, able to monitor pedestrian flows in public spaces, throwing alarms when abnormal behavior occurs. Special emphasis has been given to more specific evacuation scenarios, for obvious reasons. In this spirit, it is important to define mathematical models based on behavioral assumptions, tested by means of proper statistical methods. Data collection for pedestrian dynamics is particularly difficult and only few models presented in the literature have been calibrated and validated on real data sets.

Previous methods for pedestrian behavior modeling can be classified into two main categories: *microscopic* and *macroscopic* models. In the last years much more attention has focused on microscopic modeling, where each pedestrian is modeled as an agent. Examples of microscopic models are the *social forces* model in Helbing and

Molnar (1995) and Helbing et al. (2002) where the authors use Newtonian mechanics with a continuous space representation to model long-range interactions, and the multi-layer utility maximization model by Hoogendoorn et al. (2002) and Daamen (2004). Blue and Adler (2001) and Schadschneider (2002) use cellular automata models, characterized by a static discretization of the space where each cell in the grid is represented by a state variable. Another microscopic approach is based on space syntax theory where people move through spaces following criteria of space visibility and accessibility (see Penn and Turner, 2002) and minimizing angular paths (see Turner, 2001). Finally, Borgers and Timmermans (1986), Whyne et al. (1996) and Dellaert et al. (1998) focus on destination and route choice problems on network topologies. For a general literature review on pedestrian behavior modeling we refer the interested reader to Bierlaire et al. (2003). For applications of pedestrian models in image analysis, we refer the reader to our previous work (Antonini et al., 2004, Venegas et al., 2005, Antonini, 2005 and Antonini, Venegas, Bierlaire and Thiran, 2006)

Leader-follower and collision avoidance behavior play a major role in explaining pedestrian movements. Existing literature has shown the occurrence of self-organizing processes in crowded environments. At certain levels of density, interactions between people give rise to lane formation (Helbing et al., 2005, Hoogendoorn and Daamen, 2005). Collision avoidance (e.g. Collett and Marsh, 1974) and leader-follower (e.g. Li et al., 2001) have been widely studied. In order to include these aspects in our model, we took inspiration from previous car following models in transport engineering (including Newell, 1961, Herman and Rothery, 1965, Lee, 1966, Ahmed, 1999). The main idea in these models is that two vehicles are involved in a car following situation when a subject vehicle follows a leader, normally represented by the vehicle in front, reacting to its actions. In general, a sensitivity-stimulus framework is adopted. According to this framework a driver reacts to stimuli from the environment, where the stimulus is usually the leader's relative speed. Different models differ in the specification of the sensitivity term. This modeling idea is extended here and adapted to the more complex case of pedestrian behavior. We want to stress the fact that in driver behavior modeling a distinction between acceleration and direction (or lane) is almost natural (see Toledo, 2003 and Toledo et al., 2003), being suggested by the transport facility itself, organized into lanes. The pedestrian case is more complex, since movements are two-dimensional on the walking plane, where acceleration and direction changes are not easily separable. Constrained behavior in general, and collision avoidance in particular are also inspired by studies in human sciences and psychology, leading to the concept of *personal space* (see Horowitz et al., 1964, Dosey and Meisels, 1969 and Sommer, 1969). Personal space is a protective mechanism founded on the ability of the individual to perceive signals from the physical and social environment. Its function is to create spacing patterns that regulate distances between individuals and on which individual behaviors are based (Webb and Weber, 2003). Helbing and Molnar (1995) in their social forces model use the term "territorial effect". Several studies in psychology and sociology show how individual

characteristics influence the perception of space and interpersonal distance. Brady and Walker (1978) found for example that 'anxiety states' are positively correlated with interpersonal distance. Similarly, Dosey and Meisels (1969) found that individuals establish greater distances in high-stress conditions. Hartnett et al. (1974) found that male and female individuals approached short individuals more closely than tall individuals. Other studies (Phillips, 1979 and Sanders, 1976) indicate that an other person's body size influences space.

The validation of pedestrian walking models is a difficult task, and has not been extensively reported in the literature. Berrou et al. (2007) and Kretz et al. (2008) validate their model by comparing real and simulated flows and densities at bottlenecks. Brogan and Johnson (2003) compare real walking paths with simulated paths using three different metrics: the distance error, that is the mean distance between the real and the simulated path for all simulation time steps, the area error, that is the area between the two paths, and the speed error, that is the mean difference in speed between the two paths for all simulation time steps.

### 3.3 Modeling framework

In this work we refer to the general framework for pedestrian behavior described by Daamen (2004). Individuals make different decisions, following a hierarchical scheme: *strategical*, *tactical* and *operational*. Destinations and activities are chosen at a strategical level; the order of the activity execution, the activity area choice and route choice are performed at the tactical level, while instantaneous decisions such as walking and stops are taken at the operational level. In this chapter, we focus on pedestrian walking behavior, naturally identified by the operational level of the hierarchy just described. We consider that strategic and tactical decisions have been exogenously made, and are interested in modeling the short range behavior in *normal* conditions, as a reaction to the surrounding environment and to the presence of other individuals. By "normal" we mean non-evacuation and non-panic situations.

The motivations and the soundness of discrete choice methods have been addressed in our introductory work (Bierlaire et al., 2003, Antonini, Bierlaire and Weber, 2006, Antonini and Bierlaire, 2007). The objective of this chapter is twofold. First, we aim to provide an extended disaggregate, fully estimable behavioral model, calibrated on real pedestrian trajectories manually tracked from video sequences. Second, we want to test the coherence, interpretability and generalization power of the proposed specification through a detailed validation on external data. Compared with Antonini, Bierlaire and Weber (2006), we present three important contributions: (i) we estimate the model using significantly more data representing revealed walking behavior, (ii) the model specification explicitly captures leader-follower and collision-avoidance patterns and (iii) the model is successfully validated both using cross-validation on the estimation data set, and forecasting validation on another experimental data set, not involved in the estimation process.

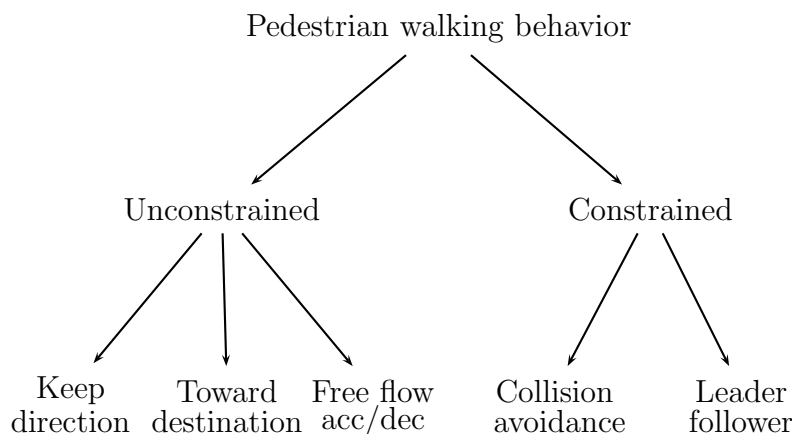


Figure 3.1: Conceptual framework for pedestrian walking behavior

We illustrate in Figure 3.1 the behavioral framework. Unconstrained decisions are independent of the presence of other pedestrians and are generated by subjective and/or unobserved factors. The first of these factors is represented by the individual's destination. It is assumed to be exogenous to the model. The second factor is represented by the tendency of people to keep their current direction, minimizing their angular displacement. Finally, unconstrained acceleration and deceleration are dictated by the individual's desired speed. The implementation of these ideas is made through the three unconstrained patterns indicated in Figure 3.1.

We assume that behavioral constraints are induced by interactions with other individuals nearby. The *collision avoidance* pattern is designed to capture the effects of possible collisions on the current trajectory of the decision maker. The *leader-follower* pattern is designed to capture the tendency of people to follow another individual in a crowd, in order to benefit from the space she creates.

The discrete choice model introduced by Antonini, Bierlaire and Weber (2006) is extended here. The basic elements are the same and summarized below. Pedestrian movements and interactions take place on the horizontal walking plane. The spatial resolution depends on the current speed vector of the individuals. The geometrical elements of the space model are illustrated in Figure 3.2.

In a given coordinate system, the current position of the decision maker  $\mathbf{n}$  is  $\mathbf{p}_n \equiv (\mathbf{x}_n, \mathbf{y}_n)$ , her current speed  $\mathbf{v}_n \in \mathbb{R}$ , her current direction is  $\mathbf{d}_n \in \mathbb{R}^2$  (normalized such that  $\|\mathbf{d}_n\| = 1$ ) and her visual angle is  $\theta_n$  (typically,  $\theta_n = 170^\circ$ ). The region of interest is situated in front of the pedestrian, ideally overlapping with her visual field. An individual-specific and adaptive discretization of the space is obtained to generate a set of possible places for the next step. Three speed regimes are considered. The individual can accelerate to 1.5 times her speed, decelerate to half time her speed, or maintain her current speed. Therefore, the next position will lie in one of the



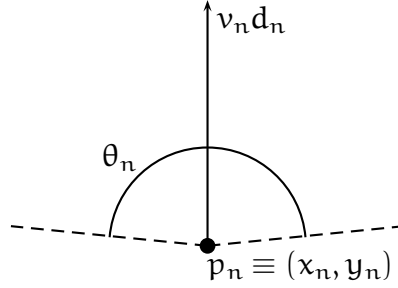


Figure 3.2: The basic geometrical elements of the space structure

zones, as depicted in Figure 3.3(b). For a given time step  $t$  (typically, 1 second), the *deceleration* zones range from  $0.25v_n t$  to  $0.75v_n t$ , with the center being at  $0.5v_n t$ , the *constant speed* zones range from  $0.75v_n t$  to  $1.25v_n t$ , with the center being at  $v_n t$ , and the *acceleration* zones range from  $1.25v_n t$  to  $1.75v_n t$ , with the center being at  $1.5v_n t$ . With respect to the direction, a discretization into 11 radial directions is used, as illustrated in Figure 3.3(a), where the angular amplitudes of the radial cones are reported in degrees.

A choice set of 33 alternatives is generated where each alternative corresponds to a combination of a speed regime  $v$  and a radial direction  $d$ , as illustrated in Figure 3.4. Each alternative is identified by the physical center of the corresponding cell in the spatial discretization  $c_{vd}$ , that is

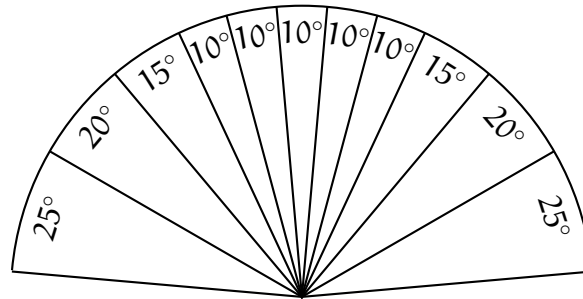
$$c_{vd} = p_n + vt d, \quad (3.1)$$

where  $t$  is the time step. The choice set varies with direction and speed and so does the distance between an alternative's center and other pedestrians. As a consequence, differences in individual speeds are naturally mapped into differences in their relative interactions. Note that the presence of physical obstacles can be modeled by declaring the corresponding cells as not available.

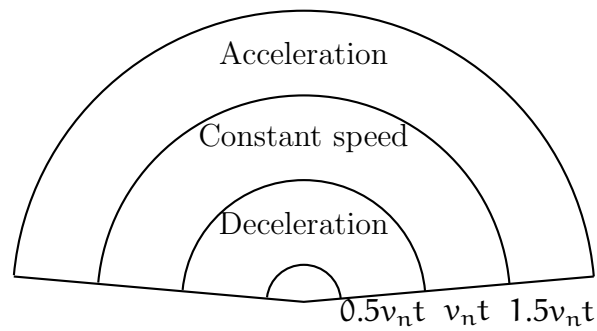
### 3.4 The model

Individuals walk on a 2D plane and we model two kinds of behavior: changes in direction and changes in speed, i.e. accelerations. Five behavioral patterns are defined. In a discrete choice context, they have to be considered as terms entering the utility functions of each alternative, as reported in Equation 3.2. The utilities describe the space around the decision maker and under the assumption of rational behavior, the individual chooses the location (alternative) with the maximum utility. In the following, we discuss the different patterns and the associated assumptions in more details.

Following the framework proposed in Figure 3.1 we report here the systematic



(a) Discretization of directions



(b) Discretization of speed regimes

Figure 3.3: The spatial discretization.

utility as perceived by individual  $\mathbf{n}$  for the alternative identified by the speed regime

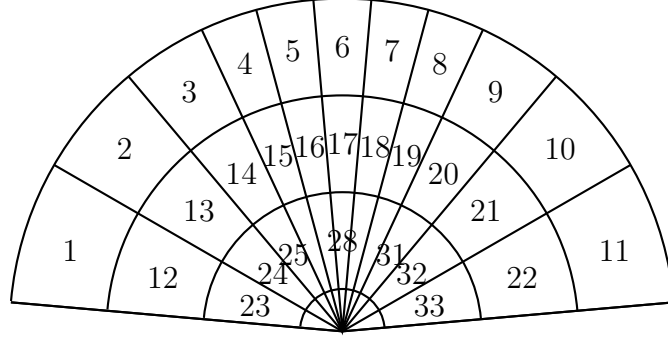


Figure 3.4: Choice set representation, with numbering of alternatives

$v$  and direction  $d$ .

$$\begin{aligned}
 V_{v,dn} = & \left. \begin{array}{l} \beta_{dir\_central} dir_{dn} I_{d,central} \\ \beta_{dir\_side} dir_{dn} I_{d,side} \\ \beta_{dir\_extreme} dir_{dn} I_{d,extreme} \end{array} \right\} \textit{keep direction} \\
 & \left. \begin{array}{l} \beta_{ddist} ddist_{v,dn} \\ \beta_{ddir} ddir_{dn} \end{array} \right\} \textit{toward destination} \\
 & \left. \begin{array}{l} \beta_{dec} I_{v,dec} (v_n/v_{max})^{\lambda_{dec}} \\ \beta_{accLS} I_{n,LS} I_{v,acc} (v_n/v_{maxLS})^{\lambda_{accLS}} \\ \beta_{accHS} I_{n,HS} I_{v,acc} (v_n/v_{max})^{\lambda_{accHS}} \end{array} \right\} \textit{free flow acceleration} \\
 & \left. \begin{array}{l} I_{v,acc} I_{d,acc}^L \alpha_{acc}^L D_L^{\rho_{acc}^L} \Delta v_L^{\gamma_{acc}^L} \Delta \theta_L^{\delta_{acc}^L} \\ I_{v,dec} I_{d,dec}^L \alpha_{dec}^L D_L^{\rho_{dec}^L} \Delta v_L^{\gamma_{dec}^L} \Delta \theta_L^{\delta_{dec}^L} \end{array} \right\} \textit{leader-follower} \\
 & \left. \begin{array}{l} I_{d,c} \alpha_c e^{\rho_c D_c} \Delta v_c^{\gamma_c} \Delta \theta_c^{\delta_c} \end{array} \right\} \textit{collision avoidance}
 \end{aligned} \tag{3.2}$$

where all the  $\beta$  parameters as well as  $\lambda_{acc}$ ,  $\lambda_{dec}$ ,  $\alpha_{acc}^L$ ,  $\rho_{acc}^L$ ,  $\gamma_{acc}^L$ ,  $\delta_{acc}^L$ ,  $\alpha_{dec}^L$ ,  $\rho_{dec}^L$ ,  $\gamma_{dec}^L$ ,  $\delta_{dec}^L$ ,  $\alpha_c$ ,  $\rho_c$ ,  $\gamma_c$ ,  $\delta_c$  are unknown and have to be estimated. We explain in the following the different terms of the utilities.

### 3.4.1 Keep direction

This part of the model captures the tendency of people to avoid frequent variation of direction. People choose their next position in order to minimize the angular displacement from their current direction of movement. In addition to the behavioral motivation of this factor, it also plays a smoothing role in the model, avoiding drastic changes of direction from one time period to the next. In order to capture the non-linearity of this pattern, we include a different term for each group of directions. The

“central” group, identified by the indicator  $I_{d,\text{central}}$ , contains the cones 5, 6 and 7 (see Figure 3.3), the “side” group, identified by the indicator  $I_{d,\text{side}}$ , contains the cones 3, 4, 8 and 9, and the “extreme” group, identified by the indicator  $I_{d,\text{extreme}}$ , contains the cones 1, 2, 10 and 11.

The associated terms in the utility function are

$$\beta_{\text{dir\_central}} \text{dir}_{\text{dn}} I_{d,\text{central}} + \beta_{\text{dir\_side}} \text{dir}_{\text{dn}} I_{d,\text{side}} + \beta_{\text{dir\_extreme}} \text{dir}_{\text{dn}} I_{d,\text{extreme}} \quad (3.3)$$

where the variable  $\text{dir}_{\text{dn}}$  is defined as the angle in degrees between the direction  $\mathbf{d}$  and the direction  $\mathbf{d}_n$ , corresponding to the current direction, as shown in Figure 3.5. Note that the indicators guarantee that only one of these three terms is nonzero for any given alternative. We expect the  $\beta$  parameters to be negative.

We preferred this specification to a continuous specification, in terms of the direction cones, in order to capture the specificity of each of them.

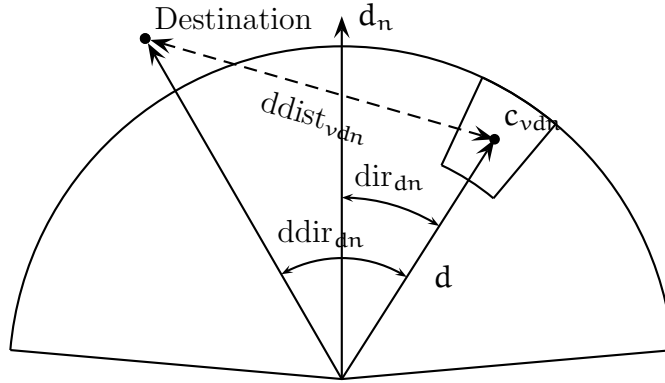


Figure 3.5: The elements capturing the *keep direction* and *toward destination* behaviors

### 3.4.2 Toward destination

The destination is defined as the final location that the pedestrian wants to reach. To be coherent with the general framework introduced in Section 4.2, we assume that the destination choice is performed at the strategical (or possibly tactical) level in the hierarchical decision process, and is therefore exogenous in this model. Such a higher level choice is naturally reflected on short term behavior as the tendency of individuals to choose, for the next step, a spatial location that minimizes both angular displacement and the distance to the destination.

This behavior is captured by the term

$$\beta_{\text{ddist}} \text{d\_dist}_{\text{vdn}} + \beta_{\text{ddir}} \text{d\_dir}_{\text{dn}} \quad (3.4)$$

where the variable  $\text{ddist}_{\text{v}_{\text{dn}}}$  is defined as the distance (in meters) between the destination and the center of the alternative  $\text{C}_{\text{v}_{\text{dn}}}$ , while  $\text{ddir}_{\text{dn}}$  is defined as the angle in degrees between the destination and the alternative's direction  $\mathbf{d}$ , as shown in Figure 3.5. We expect a negative sign for both the  $\beta_{\text{ddir}}$  and  $\beta_{\text{ddist}}$  parameters.

### 3.4.3 Free flow acceleration

In free flow conditions the behavior of the individual is driven by her desired speed. The acceleration is then a function of the difference between current speed and desired speed. However, this variable is unobserved and it cannot be introduced explicitly in the model. As a consequence, we assume that the utility for acceleration is dependent on the current speed. Increasing speed corresponds to decreasing utility for further accelerations. In order to reflect that a parameter varies with speed  $\mathbf{v}_{\text{n}}$ , we use the specification

$$\beta = \bar{\beta} \left( \frac{\mathbf{v}_{\text{n}}}{\mathbf{v}_{\text{ref}}} \right)^\lambda. \quad (3.5)$$

Note that

$$\lambda = \frac{\partial \beta}{\partial \mathbf{v}_{\text{n}}} \frac{\mathbf{v}_{\text{n}}}{\beta}$$

can be interpreted as the elasticity of the parameter  $\beta$  with respect to the speed  $\mathbf{v}_{\text{n}}$ . The value of  $\mathbf{v}_{\text{ref}}$  is arbitrary, and determines the reference speed corresponding to  $\bar{\beta}$ .

In our context, we define such a term for the parameters associated with deceleration

$$\beta_{\text{dec}} \text{I}_{\text{v,dec}}(\mathbf{v}_{\text{n}}/\mathbf{v}_{\text{max}})^{\lambda_{\text{dec}}} \quad (3.6)$$

where  $\text{I}_{\text{v,dec}}$  is one if  $\mathbf{v}$  corresponds to a deceleration, and zero otherwise, and the reference speed is selected to be the maximum speed observed  $\mathbf{v}_{\text{max}} = 4.84$  (m/s). The impact of this term on the utility is illustrated in Figure 3.6(a) (the estimated values of the parameters have been used to generate Figure 3.6). It shows that the utilities of the alternatives associated with deceleration are very low when the pedestrian is already walking slowly. For higher speeds, this term has basically no impact on utility.

For the acceleration, we have introduced two terms, one for lower speeds (less than or equal to  $5\text{km/h} = 1.39$  m/s), and one for higher speeds.

$$\beta_{\text{accLS}} \text{I}_{\text{n,LS}} \text{I}_{\text{v,acc}}(\mathbf{v}_{\text{n}}/\mathbf{v}_{\text{maxLS}})^{\lambda_{\text{accLS}}} + \beta_{\text{accHS}} \text{I}_{\text{n,HS}} \text{I}_{\text{v,acc}}(\mathbf{v}_{\text{n}}/\mathbf{v}_{\text{max}})^{\lambda_{\text{accHS}}} \quad (3.7)$$

where  $\text{I}_{\text{n,LS}}$  is one if the individual's current speed is less than or equal to 1.39 and zero otherwise,  $\text{I}_{\text{n,HS}} = 1 - \text{I}_{\text{n,LS}}$ , and the reference speed for low speeds  $\mathbf{v}_{\text{maxLS}} = 1.39$ . The indicator  $\text{I}_{\text{v,acc}}$  is 1 if the alternative corresponds to an acceleration and 0 otherwise. We expect negative signs for  $\beta_{\text{accHS}}$ ,  $\beta_{\text{accLS}}$ ,  $\beta_{\text{dec}}$  and  $\lambda_{\text{dec}}$  parameters, while a positive sign is expected for  $\lambda_{\text{accLS}}$  and  $\lambda_{\text{accHS}}$ . Indeed, we assume a depreciation of utilities associated to accelerated alternatives when the speed increases. Similarly we

assume a depreciation of utilities associated to decelerated alternatives when the speed decreases. The impact of this term on utility is illustrated on Figure 3.6(b), where the two parts of the curve (low and high speed) are represented. It appears clearly that the role of the second part is to avoid a too dramatic penalty of acceleration for high speeds.

### 3.4.4 Leader-follower

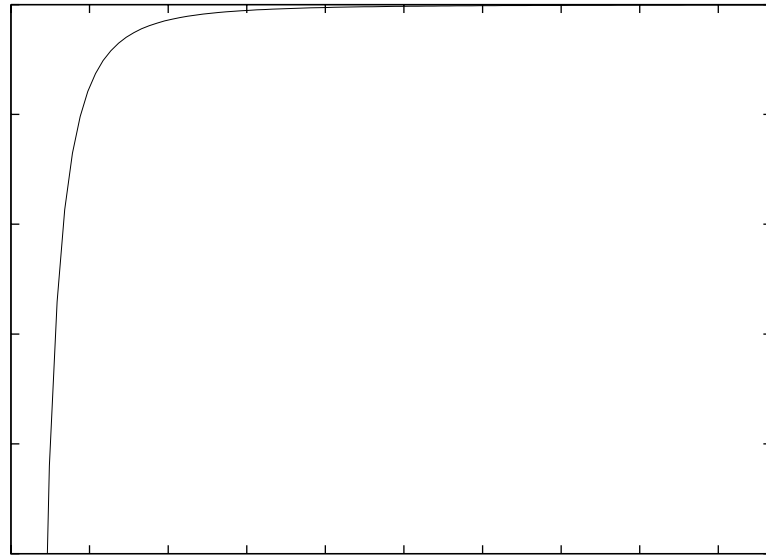
We assume that the decision maker is influenced by leaders. In our spatial representation 11 radial cones partition the space (see Figure 3.3). In each of these directions a possible leader can be identified among a set of *potential leaders*. A potential leader is an individual who is inside a certain region of interest, *not so far* from the decision maker and with a moving direction *close enough* to the direction of the radial cone where she is. Among the set of potential leaders for each radial direction, one of them is selected as leader for that direction (the closest to the decision maker). Once identified, the leader induces an attractive interaction on the decision maker. Similarly to car following models, a leader acceleration corresponds to decision maker acceleration. The leader-follower model is given by the following terms

$$I_{v,acc} I_{d,acc}^L \alpha_{acc}^L D_L^{\rho_{acc}^L} \Delta v_L^{\gamma_{acc}^L} \Delta \theta_L^{\delta_{acc}^L} + I_{v,dec} I_{d,dec}^L \alpha_{dec}^L D_L^{\rho_{dec}^L} \Delta v_L^{\gamma_{dec}^L} \Delta \theta_L^{\delta_{dec}^L}. \quad (3.8)$$

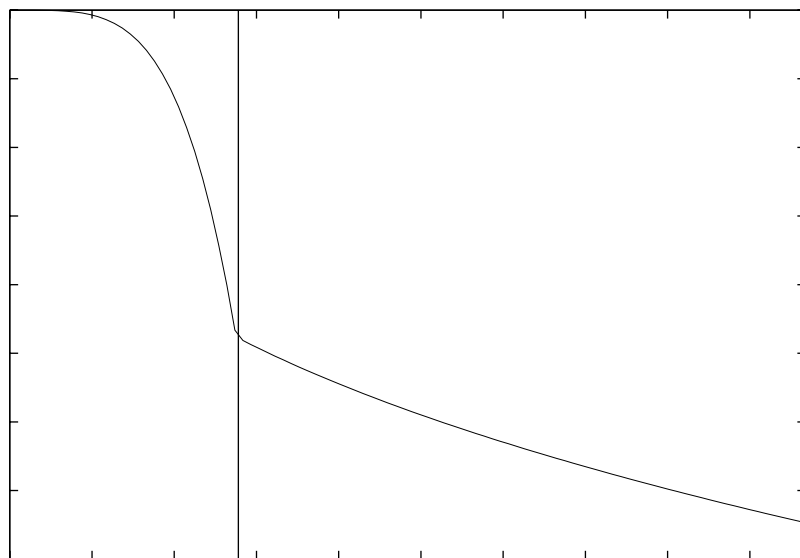
It is described by a *sensitivity/stimulus* framework. The leader for each direction is chosen considering several *potential leaders* (represented by light gray circles in Figure 3.7). An individual  $k$  is defined as a potential leader based on the following indicator function:

$$I_g^k = \begin{cases} 1, & \text{if } d_l \leq d_k \leq d_r \text{ (is in the cone),} \\ & \text{and } 0 < D_k \leq D_{th} \text{ (not too far),} \\ & \text{and } 0 < |\Delta \theta_k| \leq \Delta \theta_{th} \text{ (walking in almost the same direction),} \\ 0, & \text{otherwise,} \end{cases}$$

where  $d_l$  and  $d_r$  represent the bounding left and right directions of the cone in the choice set (defining the region of interest) while  $d_k$  is the direction identifying the position of pedestrian  $k$ .  $D_k$  is the distance between pedestrian  $k$  and the decision maker,  $\Delta \theta_k = \theta_k - \theta_d$  is the difference between the movement direction of pedestrian  $k$  ( $\theta_k$ ) and the angle characterizing direction  $d$ , i.e. the direction identifying the radial cone where individual  $k$  lies ( $\theta_d$ ). The two thresholds  $D_{th}$  and  $\Delta \theta_{th}$  are fixed at the values  $D_{th} = 5D_{max}$ , where  $D_{max}$  is the radius of the choice set, and  $\Delta \theta_{th} = 10$  degrees. This seems to be reasonable and well adapted to pedestrian environment perception. In addition, when looking at the data, this appears to be a good trade off between the computational complexity and the real behavior. We assume an implicit *leader choice* process, executed by the decision maker herself and modeled choosing as leader for each direction, the potential leader at the minimum distance  $D_L = \min_{k \in K}(D_k)$ , illustrated in Figure 3.7 by the darker circle. Once the leader is



(a) Deceleration



(b) Acceleration

Figure 3.6: Impact of free flow acceleration terms on utility (x axis: the speed, y axis: the utility contribution )

identified, we compare her speed. The indicator  $I_{d,acc}^L$  is one if the leader in the cone  $\mathbf{d}$  has been identified with a speed larger than  $v_n$ , and zero otherwise. Similarly,  $I_{d,dec}^L = 1 - I_{d,acc}^L$  is one if the leader in cone  $\mathbf{d}$  has been identified with a speed lower than  $v_n$ , and zero otherwise. Finally, the indicator functions  $I_{v,acc}$  and  $I_{v,dec}$  discriminate between accelerated and decelerated alternatives, as with the free flow acceleration model. The underlying assumption is that faster leaders will have an impact on the acceleration, while slower leaders will have an impact on the deceleration.

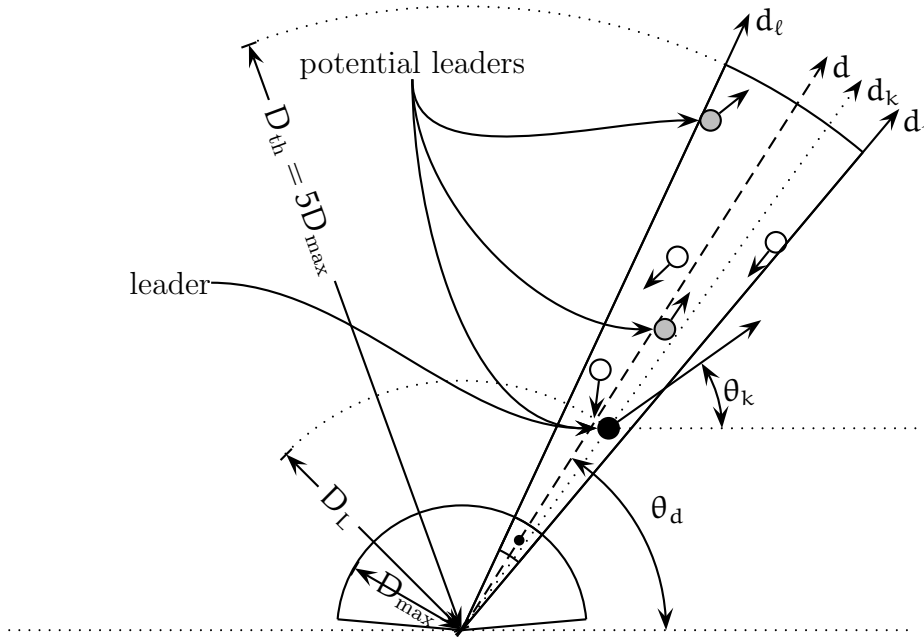


Figure 3.7: Leader and potential leaders in a given cone

For a given leader, sensitivity is described by

$$\text{sensitivity} = \alpha_g^L D_L^{\rho_g^L} \quad (3.9)$$

where  $D_L$  represents the distance between the decision maker and the leader. The parameters  $\alpha_g^L$  and  $\rho_g^L$  have to be estimated and  $\mathbf{g} = \{\mathbf{acc}, \mathbf{dec}\}$  indicates when the leader is accelerating with respect to the decision maker. Both  $\alpha_{acc}^L$  and  $\alpha_{dec}^L$  are expected to be positive while a negative sign is expected for  $\rho_{acc}^L$  and  $\rho_{dec}^L$ .

The decision maker reacts to stimuli coming from the chosen leader. We model the stimulus as a function of the leader's relative speed  $\Delta v_L$  and the leader's relative direction  $\Delta \theta_L$  as follows:

$$\text{stimulus} = \Delta v_L^{\gamma_g^L} \Delta \theta_L^{\delta_g^L} \quad (3.10)$$



with  $\Delta v_L = |v_L - v_n|$ , where  $v_L$  and  $v_n$  are the leader's speed module and the decision maker's speed module, respectively. The variable  $\Delta\theta_L = \theta_L - \theta_d$ , where  $\theta_L$  represents the leader's movement direction and  $\theta_d$  is the angle characterizing direction  $\mathbf{d}$ , as shown in Figure 3.7. Positive signs are expected for both the  $\gamma_{acc}^L$  and  $\gamma_{dec}^L$  parameters, while we expect a negative sign for both the  $\delta_{acc}^L$  and  $\delta_{dec}^L$ . Indeed, we assume the pedestrian to follow a leader who has a behavior that the pedestrian wishes to have. If the pedestrian accelerates, her perception will be heavily impacted by her leader who is strongly accelerating and going in her desired direction. Inversely, if the pedestrian is decelerating, her perception will be significantly impacted by her leader who is strongly decreasing her speed and going in her desired direction. This means that a leader acceleration induces a decision maker's acceleration. A substantially different movement direction in the leader reduces the influence of the latter on the decision maker. Note that in the final specification, the parameter  $\delta_{dec}^L$  appeared not to be significantly different from 0. Therefore, we decided to remove it from the model in the final estimation. The specification (3.8) thus becomes

$$I_{v,acc} I_{d,acc}^L \alpha_{acc}^L D_L^{p_{acc}^L} \Delta v_L^{\gamma_{acc}^L} \Delta \theta_L^{\delta_{acc}^L} + I_{v,dec} I_{d,dec}^L \alpha_{dec}^L D_L^{p_{dec}^L} \Delta v_L^{\gamma_{dec}^L}. \quad (3.11)$$

### 3.4.5 Collision avoidance

This pattern captures the effects of possible collisions on the decision maker's trajectory. For each direction in the choice set, a collider is identified among a set of *potential colliders*. Another individual is selected as a potential collider if she is inside a certain region of interest, *not too far* from the decision maker and walking in the opposite direction. The collider for a radial direction is chosen from the set of potential colliders for that direction as the individual whose walking direction forms the larger angle with the decision maker's walking direction. This pattern is associated with repulsive interactions in the obvious sense that pedestrians change their current direction to avoid collisions with other individuals. The collision avoidance model is given by the following term

$$I_{d,c} \alpha_c e^{p_c D_c} \Delta v_C^{\gamma_c} \Delta \theta_C^{\delta_c}. \quad (3.12)$$

The collider for each direction is chosen considering several *potential colliders*, as shown in Figure 3.8. An individual  $k$  is defined as a potential collider based on the following indicator function:

$$I_C^k = \begin{cases} 1, & \text{if } \mathbf{d}_l \leq \mathbf{d}_k \leq \mathbf{d}_r \text{ (is in the cone),} \\ & \text{and } 0 < D_k \leq D'_{th} \text{ (not too far),} \\ & \text{and } \frac{\pi}{2} \leq |\Delta\theta_k| \leq \pi \text{ (walking in the other direction),} \\ 0, & \text{otherwise,} \end{cases}$$

where  $\mathbf{d}_l$ ,  $\mathbf{d}_r$  and  $\mathbf{d}_k$  are the same as those defined for the leader-follower model.  $D'_k$  is the distance between individual  $k$  and the center of the alternative,  $\Delta\theta_k = \theta_k - \theta_{d_n}$  is

the difference between the movement direction of pedestrian  $k$ ,  $\theta_k$ , and the movement direction of the decision maker,  $\theta_d$ . The value of the distance threshold is now fixed to  $D'_{th} = 10D_{max}$ . We use a larger value compared to the leader-follower model, assuming the collision avoidance behavior to be a longer range interaction, happening also at a lower level of density. We assume an implicit *collider choice* process, which is deterministic and decision-maker specific. Among the set of  $K_d$  potential colliders for direction  $d$ , a collider is chosen in each cone as that individual having  $\Delta\theta_C = \max_{k \in K_d} |\Delta\theta_k|$ . The indicator  $I_{d,C} = 1$  if a collider has been identified, and 0 otherwise. Finally, the collision avoidance term is included in the utility functions of all the alternatives.

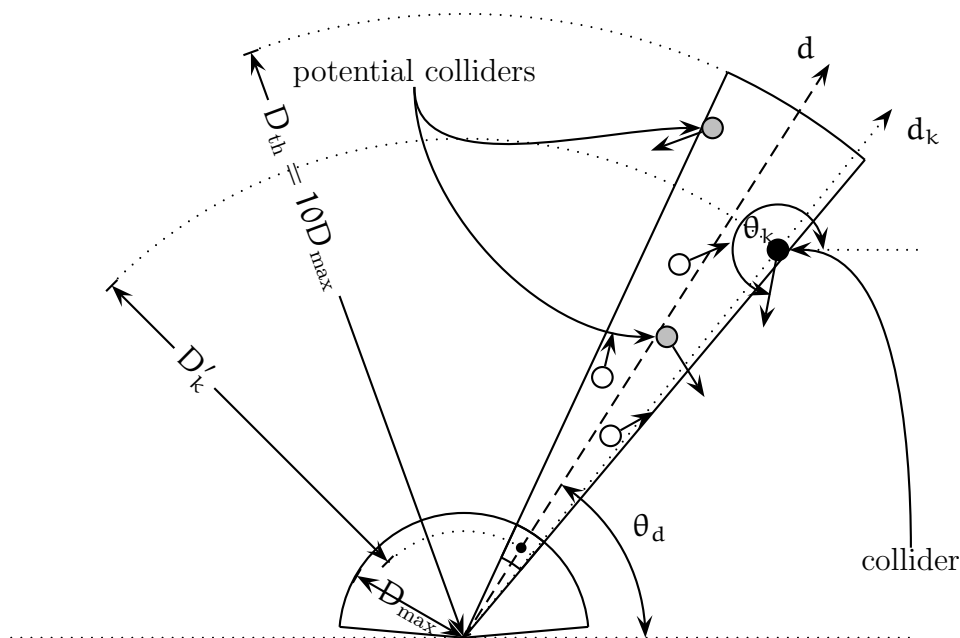


Figure 3.8: Collider and potential colliders in a given cone

We apply a similar sensitivity/stimulus framework, where the sensitivity function is defined as

$$\text{sensitivity} = \alpha_C e^{\rho_C D_C} \quad (3.13)$$

where the parameters  $\alpha_C$  and  $\rho_C$ , that have to be estimated, are both expected to have a negative sign and  $D_C$  is the distance between the collider position and the center of the alternative. The decision maker reacts to stimuli coming from the collider. We model the stimulus as a function of two variables:

$$\text{stimulus} = \Delta v_C^{\gamma_C} \Delta \theta_C^{\delta_C} \quad (3.14)$$

with  $\Delta\theta_C = \theta_C - \theta_{dn}$ , where  $\theta_C$  is the collider movement direction and  $\theta_{dn}$  is the decision maker movement direction, and  $\Delta v_C = v_C + v_n$ , where  $v_C$  is the collider's speed module and  $v_n$  is the decision maker's speed module. The parameters  $\gamma_C$  and  $\delta_C$  have to be estimated and a positive sign is expected for both of them. Individuals walking against the decision maker at higher speeds and in more forward directions (higher  $\Delta\theta_C$ ) generate stronger reactions, weighted by the sensitivity function.

Note that in the final specification, the parameters  $\gamma_C$  and  $\delta_C$  appeared not to be significantly different from 0. Therefore, we decided to remove them from the model in the final estimation. The final specification includes only the sensitivity part (3.13).

### 3.4.6 The error term

We use a cross nested logit (CNL) model (see, among others, Wen and Koppelman, 2001, Bierlaire, 2006, Abbe et al., 2007) specification. Such a model allows flexible correlation structures in the choice set, keeping a closed form solution. The CNL being a Multivariate Extreme Value model (MEV, see McFadden, 1978), the probability of choosing alternative  $i$  within the choice set  $C$  is:

$$P(i|C) = \frac{y_i \frac{\partial G}{\partial y_i}(y_1, \dots, y_J)}{\mu G(y_1, \dots, y_J)} \quad (3.15)$$

where  $J$  is the number of alternatives in  $C$ ,  $y_j = e^{V_j}$  with  $V_j$  the systematic part of the utility described by (3.2) and  $G$  is the following generating function:

$$G(y_1, \dots, y_J) = \sum_{m=1}^M \left( \sum_{j \in C} (\alpha_{jm}^{1/\mu} y_j)^{\mu_m} \right)^{\frac{\mu}{\mu_m}} \quad (3.16)$$

where  $M$  is the number of nests,  $\alpha_{jm} \geq 0, \forall j, m, \sum_{m=1}^M \alpha_{jm} > 0, \forall j, \mu > 0, \mu_m > 0, \forall m$  and  $\mu \leq \mu_m, \forall m$ . This formulation leads to the following expression for the choice probability formula, using  $y_i = e^{V_i}$ :

$$P(i|C) = \sum_{m=1}^M \frac{\left( \sum_{j \in C} \alpha_{jm}^{\mu_m/\mu} y_j^{\mu_m} \right)^{\frac{\mu}{\mu_m}}}{\sum_{n=1}^M \left( \sum_{j \in C} \alpha_{jn}^{\mu_n/\mu} y_j^{\mu_n} \right)^{\frac{\mu}{\mu_n}}} \frac{\alpha_{im}^{\mu_m/\mu} y_i^{\mu_m}}{\sum_{j \in C} \alpha_{jm}^{\mu_m/\mu} y_j^{\mu_m}} \quad (3.17)$$

We assume a correlation structure depending on the speed and direction and we identify five nests: *accelerated*, *constant speed*, *decelerated*, *central* and *not central*. We fix the degrees of membership to the different nests ( $\alpha_{jm}$ ) to the constant value 0.5. We do not want to estimate the membership degrees, as an alternative is supposed to be as correlated with alternatives in its direction cone, than correlated with alternatives belonging to the same speed regime. They are geographic and kinetic correlations. The parameter  $\mu$  is normalized to 1, and the nest parameters  $\mu_m$  are estimated. Note

that the parameter associated with the deceleration nest had been constrained to 1 in the final specification, as it did not appear to be significantly different from that value.

### 3.4.7 Intermediary specification steps

The proposed specification is the result of an intensive modeling process, where many different specifications have been tested. We have gradually refined the specification. We start with a logit with multiple alternatives containing only the *keep direction* and *toward destination* parts. The cross-nested logit structure has been rapidly adopted and kept due to the large fit improvement. An early version of the model has been embedded in a basic simulator in order to check prediction capabilities. They showed abnormal predicted behaviors regarding speeds. There were amplifications of accelerations and decelerations conducting to abnormally high speeds or stops. We decided to account for the *free flow acceleration*, which regulates the speed. The final non-linear specification follows an early piecewise linear specification which helped to find the adapted mathematical formulation. The *leader-follower* and *collision avoidance term* have been lately introduced in the model. We got inspiration from car models, as depicted in Sections 3.4.4 and 3.4.5.

We conclude this section by emphasizing that the above specification ignores heterogeneity in the population. Characteristics such as age, sex, weight, height (among others) probably influence the spatial perception, interpersonal distance and human-human interactions. However, given the nature of the data (trajectories) it is not possible to take them into account in the model. Therefore, a specification with unobserved heterogeneity captured by random coefficient in a panel data setup would have been appropriate. However, the complexity of this specification did not allow us to estimate the model with a sufficiently high number of draws. The proposed model does not explicitly account for the specificity of the cross-walk and the influence of fixed obstacles. These latter are roughly accounted in the model, by declaring unavailable the cells which are filled with obstacles. Regarding the *leader-follower* and *collision avoidance* parts, note that only a single leader or collider is considered for each alternative. The model does not account for the influence of pedestrian groups. This has been not considered due to computational issues.

## 3.5 Data

The data set used to estimate the model consists of pedestrian trajectories manually tracked from video sequences.

It was collected in Sendai, Japan, in August 2000 (see Teknomo et al., 2000, Teknomo, 2002). The video sequence was recorded from the 6th floor of the JTB parking building (around 19 meters above the ground), situated at an important



(a) Japanese scenario

Figure 3.9: A frame from the Japanese video

pedestrian crossing. Two main pedestrian flows cross the street, giving rise to a large number of interactions. A frame extracted from this video is represented in Figure 3.9.

The data set consists of 190 pedestrian trajectories, manually tracked at a rate of 2 processed frames per second, for a total number of 10200 position observations. The mapping between the image plane and the walking plane was performed by Arsenal Research (Bauer, 2007) using a 3D-calibration with the standard DLT algorithm (Shapiro, 1978). The reference system on the walking plane has the origin arbitrarily placed at the bottom left corner of the cross-walk. The  $x$  axis represents the width of the crossing while the  $y$  axis represents the length.

For each frame, the following information for each visible pedestrian was collected: (i) the time  $t$  corresponding to the frame  $f$  (in this case  $t = f/2$ ), (ii) the pedestrian identifier  $n$ , and (iii) the coordinates  $\mathbf{p}_n^f = (x_n^f, y_n^f)$  identifying the location of the pedestrian in the walking plane.

From these raw data, we first derived the current direction and speed of each pedestrian using the current and previous frames, that is

$$\begin{aligned} \mathbf{d}_n &= \mathbf{p}_n^f - \mathbf{p}_n^{f-1}, \\ v_n &= \|\mathbf{d}_n\|/0.5 = 2\|\mathbf{d}_n\|. \end{aligned}$$

In Figure 3.10 we report the speed histogram and in Table 3.1 the speed statistics. High speed (the right tail of the histogram) are associated to persons who are rushing to cross the road. They appear mainly when the pedestrian light gets red. They often appear on the video when already running.

Then, a specific choice set (see Figure 3.4) was constructed for each pedestrian, based on (3.1) where  $t = 1$  sec (that is, 2 frames),  $v = v_n$  for constant speed alternatives,  $v = 0.5v_n$  for decelerated alternatives,  $v = 1.5v_n$  for accelerated alternatives,

Mean	1.31
Standard Error	0.012
Median	1.27
Mode	1.28
Standard Deviation	0.37
Minimum	0.43
Maximum	4.84

---

Note: *standard error* is the estimated standard deviation of the sample mean

Table 3.1: Speed statistics(m/sec)

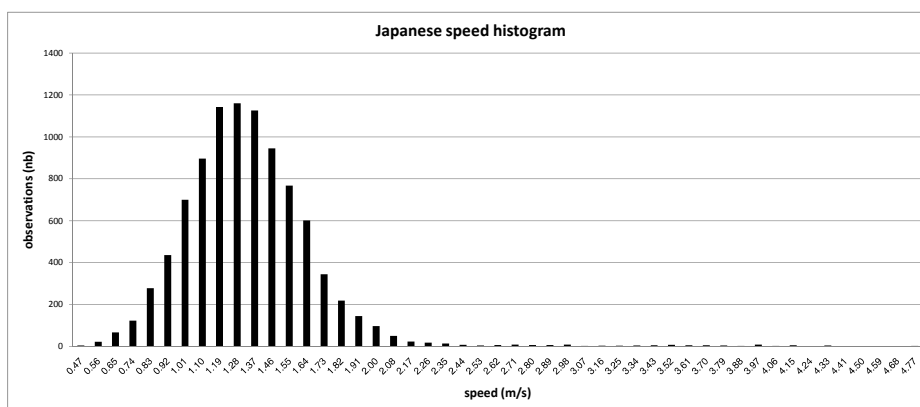


Figure 3.10: Speed histogram

$\mathbf{d} = \mathbf{d}_n$  for alternatives in cone 6 (alt. 6, 17, 28), and  $\mathbf{d} = \text{rot}(\mathbf{d}_n, \zeta)$  is obtained by rotating  $\mathbf{d}_n$  around  $\mathbf{p}_n$  with an angle  $\zeta$  corresponding to the cone, that is

$$\begin{array}{ll} \text{Cone 1: } & \zeta = 72.5^\circ, & \text{Cone 11: } & \zeta = -72.5^\circ, \\ \text{Cone 2: } & \zeta = 50^\circ, & \text{Cone 10: } & \zeta = -50^\circ, \\ \text{Cone 3: } & \zeta = 32.5^\circ, & \text{Cone 9: } & \zeta = -32.5^\circ, \\ \text{Cone 4: } & \zeta = 20^\circ, & \text{Cone 8: } & \zeta = -20^\circ, \\ \text{Cone 5: } & \zeta = 10^\circ, & \text{Cone 7: } & \zeta = -10^\circ. \end{array}$$

For each cell in the choice set, each variable in (3.2) was then computed based on the explanations in Section 3.4. Note that the destination of each individual was defined by her location in the last frame where she is visible. Finally, the chosen alternative has been identified as the cell containing the pedestrian's location after 1 second, that is  $\mathbf{p}_n^{f+2}$ . In the rare instances where  $\mathbf{p}_n^{f+2}$  did not belong to any cell (because of numerical errors due to poor image resolution, or extreme speed variations), the corresponding piece of data was removed from the sample (a total of 919 observations). We represent in Figure 3.11 selected generated choice sets on a given trajectory (representing them all would have been unreadable).

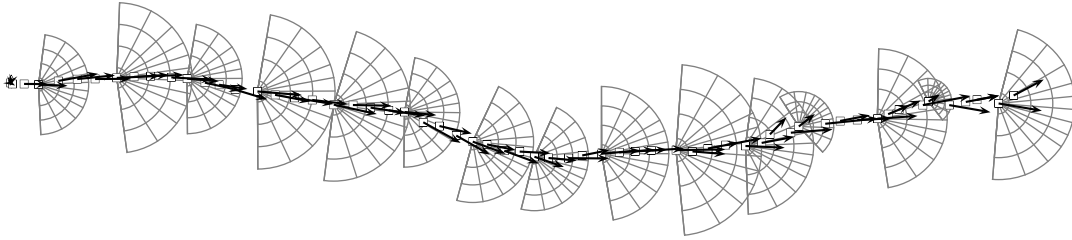


Figure 3.11: Example of one manually tracked trajectory with choice sets

We obtain a total of 9281 observations for 190 pedestrians. In Figure 3.12 we report the frequency of the revealed choices as observed in the data set. The three peaks in the distributions arise on the central alternatives (6, 17, 28), as expected. Note that cells 1, 12, 23 and 33 were never chosen in this sample. A summary of the observations across the nests is detailed in Table 3.2.

Nest	# steps	% of total
acceleration	1065	11.48%
constant speed	7565	81.51%
deceleration	651	7.01%
central	4297	46.30%
not central	4984	53.70%

Table 3.2: Number of chosen steps in each nest for the real data set

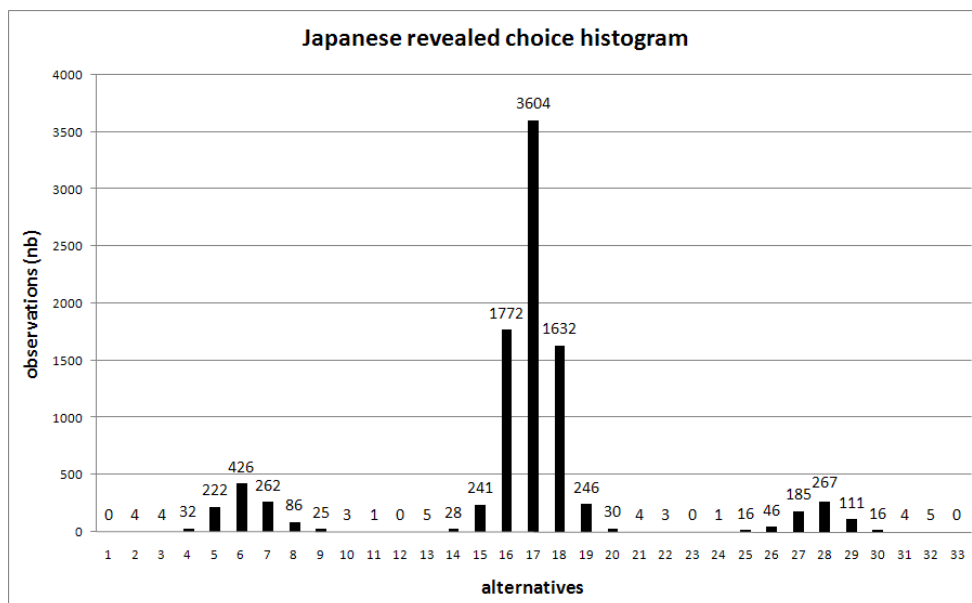


Figure 3.12: Revealed choices histograms

### 3.6 Estimation results

Table 3.3 presents the estimation results. The parameters were estimated using Biogeme (Bierlaire, 2003b, biogeme.epfl.ch).

All estimates have the expected sign. They are coherent with the expectations presented in Section 3.4. Note that the parameter associated with the deceleration nest was clearly insignificant, and fixed to 1. The influence of the extreme tail of the speed distribution (see Figure 3.10) has been checked. It appears to be negligible on the estimation results. Observations associated to these speeds mainly concern the constant speed, because pedestrians are already running when they enter the video (rushing pedestrians).

In addition to the proposed model, we analyze also a simple model, where the utility of each alternative is represented only by an alternative specific constant. This constant-only model perfectly reproduces the observed shares in the sample, with 28 parameters (33 alternatives, minus 4 which are never chosen, minus one constant normalized to 0), but does not capture any causal effect. This is a simplistic, model easy to build in any analysis. With this model, the loglikelihood drops from -13944.74 to -17972.03, illustrating the statistical significance of the proposed specification. Note that a classical likelihood ratio test is not appropriate here, as the hypotheses are not nested. We believe that a more rigorous test is not really necessary given the huge jump in loglikelihood value.

Note that a comparison with a simple model is a common practice in model estimation and validation (see Brogan and Johnson, 2003), but not sufficient as such



Variable name	Coefficient estimate	$t$ test 0	Variable name	Coefficient estimate	$t$ test 0	$t$ test 1
$\beta_{\text{ddir}}$	-0.0793	-24.14	$\rho_{\text{acc}}^{\text{L}}$	-0.465	-1.78	
$\beta_{\text{ddist}}$	-1.52	-11.63	$\gamma_{\text{acc}}^{\text{L}}$	0.552	1.98	
$\beta_{\text{dir\_extreme}}$	-0.0343	-9.71	$\alpha_{\text{dec}}^{\text{L}}$	3.78	5.41	
$\beta_{\text{dir\_side}}$	-0.0553	-22.71	$\rho_{\text{dec}}^{\text{L}}$	-0.654	-6.70	
$\beta_{\text{dir\_central}}$	-0.0320	-13.90	$\gamma_{\text{dec}}^{\text{L}}$	0.658	5.48	
$\beta_{\text{accLS}}$	-4.94	-25.20	$\delta_{\text{acc}}^{\text{L}}$	-0.179	-2.22	
$\beta_{\text{accHS}}$	-7.41	-5.10	$\alpha_{\text{C}}$	-0.00730	-10.84	
$\beta_{\text{dec}}$	-0.0645	-2.46	$\rho_{\text{C}}$	-0.212	-8.38	
$\lambda_{\text{accLS}}$	4.37	20.06	$\mu_{\text{acc}}$	1.66	9.97	3.95
$\lambda_{\text{accHS}}$	0.354	2.02	$\mu_{\text{const}}$	1.45	16.99	5.25
$\lambda_{\text{dec}}$	-2.40	-8.50	$\mu_{\text{central}}$	5.76	2.84	2.34
$\alpha_{\text{acc}}^{\text{L}}$	0.735	1.87	$\mu_{\text{not\_central}}$	1.82	13.12	5.91
Sample size = 9281			Init log-likelihood = -32451			
Nbr of estimated parameters = 24			Final log-likelihood = -13944.74			
$\bar{\rho}^2 = 0.570$			Likelihood ratio test = 37013			

Table 3.3: CNL estimation results for the Japanese data set

to validate the proposed model.

## 3.7 Model validation

Two data sets are used for validation: the Japanese data set used for estimation and described in Section 4.3, and a data set collected in the Netherlands, which was not involved at all in the estimation of the parameters.

In Section 3.7.1, we apply the model on the Japanese data set, and compare the predicted choices with the observed ones. In Section 3.7.2, we test the robustness of the model specification by performing cross-validation, where a subset of the Japanese data set is saved for validation, and the model is estimated on the rest. Finally, in Section 3.7.3, we apply both our model, and a simple constant-only model on the data set collected in the Netherlands.

### 3.7.1 Japanese data set: validation of the model

We first apply our model with the parameters described in Table 3.3 on the Japanese data set, using Biosim (Bierlaire, 2003b). For each observation  $\mathbf{n}$ , we obtain a probability distribution  $P_{\mathbf{n}}(\mathbf{i})$  over the choice set.

Figure 3.13 represents the histogram of the probability value  $P_{\mathbf{n}}(\mathbf{i}_{\mathbf{n}}^*)$  assigned by the model to the chosen alternative  $\mathbf{i}_{\mathbf{n}}^*$  of each observation  $\mathbf{n}$ , along with the hazard

value  $1/33$  (where 33 is the number of alternatives). We consider observations below this threshold as outliers. There are only 7.10% of them. As a comparison, there are 19.90% of outliers with the constant-only model.

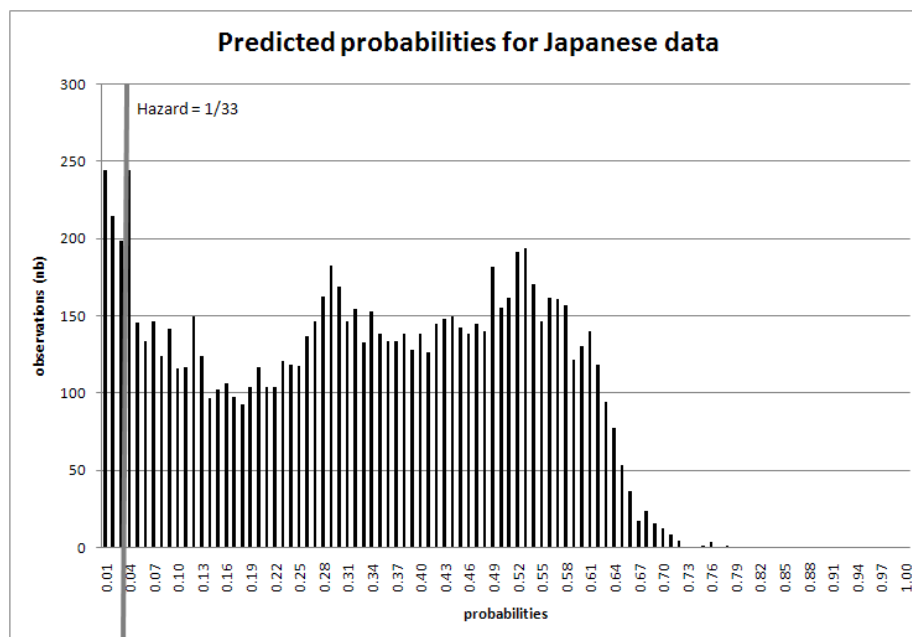


Figure 3.13: Predicted probabilities of the Japanese data

The top part of Figure 3.14 reports, for each  $i$ ,  $\sum_n P_n(i)$ , and the bottom part reports  $\sum_n y_{in}$ , where  $y_{in}$  is 1 if alternative  $i$  is selected for observation  $n$ , 0 otherwise. As expected, the two histograms are similar, indicating no major specification error.

This is confirmed when alternatives are aggregated together, by directions (see Table 3.4) and by speed regimes (see Table 3.5). For a group  $\Gamma$  of alternatives, the quantities

$$\begin{aligned} M_\Gamma &= \sum_n \sum_{i \in \Gamma} P_n(i), \\ R_\Gamma &= \sum_n \sum_{i \in \Gamma} y_{in}, \end{aligned}$$

and

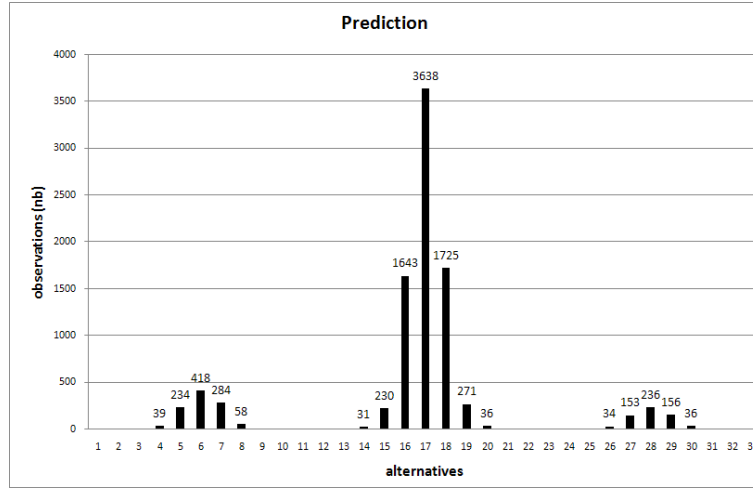
$$(M_\Gamma - R_\Gamma)/R_\Gamma$$

are reported in columns 3, 4 and 5, respectively, of these tables.

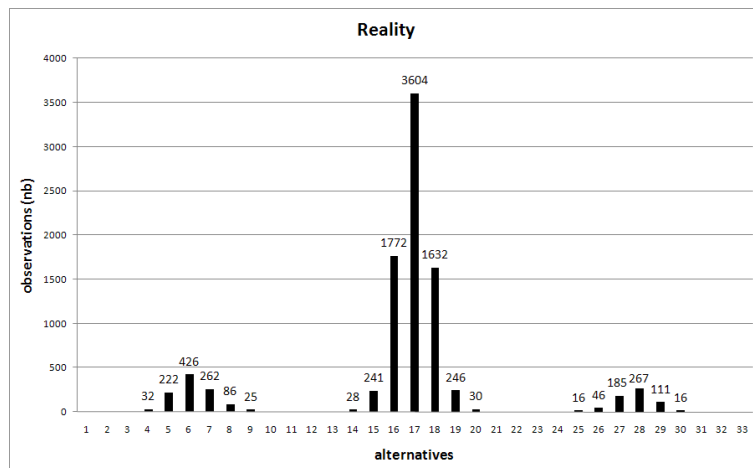
The relative errors showed in Table 3.4 and Table 3.5 are low, except for groups of alternatives with few observations, that is groups corresponding to extreme left and extreme right directions.

### 3.7.2 Japanese data set: validation of the specification

In order to test the proposed specification, we have performed a cross validation done on the Japanese data set. It consists in splitting the data set into 5 subsets, each



(a) Predicted shares



(b) Observed shares

Figure 3.14: Predicted and observed shares for the Japanese data set

Cone	$\Gamma$	$M_\Gamma$	$R_\Gamma$	$(M_\Gamma - R_\Gamma)/R_\Gamma$
Front	5 – 7, 16 – 18, 27 – 29	8486.16	8481	0.0006
Left	3, 4, 14, 15, 25, 26	348.86	367	-0.0494
Right	8, 9, 19, 20, 30, 31	419.29.	407	0.0302
Extreme left	1, 2, 12, 13, 23, 24	12.29	10	0.2292
Extreme right	10, 11, 21, 22, 32, 33	14.39	16	-0.1004

Table 3.4: Predicted ( $M_\Gamma$ ) and observed ( $R_\Gamma$ ) shares for alternatives grouped by directions with the Japanese data set

Area	$\Gamma$	$M_\Gamma$	$R_\Gamma$	$(M_\Gamma - R_\Gamma)/R_\Gamma$
acceleration	1 – 11	1059.85	1065	-0.0048
constant speed	12 – 22	7588.28	7565	0.0031
deceleration	23 – 33	632.87	651	-0.0279

Table 3.5: Predicted and observed shares for alternatives grouped by speed regime with the Japanese data set.

containing 20% of the observations. We perform 5 experiments. For each of them, one of the five subsets is saved for validation purposes, and the model is re-estimated on the remaining 4 subsets. The same procedure has been applied with the constant-only model. The proportion of outliers for each experiment is reported in Table 3.6. We observe that they are consistent with 7.10% (for our model) and 19.90% (for the constant-only model) of outliers obtained with the complete data set, illustrating the robustness of the specification.

Model	Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5
Proposed spec.	8.62%	6.52%	7.44%	7.87%	5.87%
Constant only	20.79%	20.70%	17.13%	19.88%	18.64%

Table 3.6: Summary of the cross-validation performed on the Japanese data set

The above analysis indicates a good specification and performance of the model. However, it is not sufficient to fully validate it. Consequently, we perform now the same analysis on a validation data set, not involved in the estimation of the model.

### 3.7.3 Dutch data set: validation of the model

This data set was collected at Delft University, in the period 2000-2001 (Daamen and Hoogendoorn, 2003b, Daamen and Hoogendoorn, 2003a, Daamen, 2004) where volunteer pedestrians (about 80) were called to perform specific walking tasks in a controlled experimental setup (experiment 4 in Daamen and Hoogendoorn, 2003a): “The walking experiments have been conducted in a large hallway. One rectangular area (10 meters  $\times$  4 meters) was used and taped on the floor. The digital camera was mounted at the ceiling of the hallway, at a height of 10 m, observing an area of approximately 14 m by 12 m.”

For the purposes of our validation procedure we use the subset of the Dutch data set corresponding to a bi-directional flow. This situation is the experimental version of the Japanese data set, which corresponds to a walkway. The subset includes 724 subjects for 47481 observed positions, collected by means of pedestrian tracking techniques on video sequences, at a frequency of 10Hz, that is 10 frames per second. In Figure 3.15 we report one frame from the experimental scenario.

For each frame, we collected for each visible pedestrian the time  $t$  corresponding to the frame  $f$  (in this case  $t = f/10$ ), the pedestrian identifier  $n$ , and the coordinates  $\mathbf{p}_n^f = (x_n^f, y_n^f)$  identifying the location of the pedestrian in the walking plane. From these raw data, we derived the current direction and speed of each pedestrian using the current and previous frames, that is

$$\begin{aligned} \mathbf{d}_n &= \mathbf{p}_n^f - \mathbf{p}_n^{f-1}, \\ v_n &= \|\mathbf{d}_n\|/0.1 = 10\|\mathbf{d}_n\|. \end{aligned}$$

Consistent with the model assumptions, the chosen alternative has been identified as the cell containing the pedestrian's location after 1 second, that is  $\mathbf{p}_n^{f+10}$ .

A summary of the observations across nests is detailed in Table 3.7. Note the very low number of decelerations and accelerations, probably due to the experimental nature of the data.



Figure 3.15: A representative frame from the video sequences used for data collection

Nest	# steps	% of total
acceleration	1273	2.68%
constant speed	45869	96.61%
deceleration	339	0.71%
central	20950	44.12%
not central	26531	55.88%

Table 3.7: Number of chosen steps in each nest for Dutch data

We compare the observed choices for the Japanese and the Dutch data set in Table 3.8 and Figure 3.16. Table 3.8 reports the percentage of observations for cells at the extreme left of the choice set (alts. 1, 2, 12, 13, 23, 24), the left part (alts. 3, 4, 14, 15, 25, 26), the front (alts. 5-7, 16-18, 27-29), the right (alts. 8, 9, 19, 20, 30, 31) and the extreme right (10, 11, 21, 22, 32, 33). Figure 3.16 reports normalized observation, that is, for each alternative  $i$ ,  $\sum_n y_{in}/N$ , where  $y_{in}$  is 1 if

alternative  $i$  is selected for observation  $n$ , 0 otherwise, and  $N$  is the total number of observations. We observe a great similarity in the observed proportions, except for alternatives corresponding to accelerations and decelerations. This suggests that a simple model, with only alternative specific constants, may actually perform well on this data set. The property of this simple model is to reproduce the alternative shares of the estimation data, when used for forecasting at an aggregate level. So a model with constant only and estimated on the Japanese data set should perform well when applied on the Dutch data set. We show below, however that this is not the case.

Data set	Extreme left	Left	Front	Right	Extreme right
Japanese	0.11%	3.95%	91.38%	4.39%	0.17%
Dutch	0.06%	4.40%	91.35%	4.15%	0.04%

Table 3.8: Comparison between Japanese and Dutch data sets for the observations proportions in the direction's cones

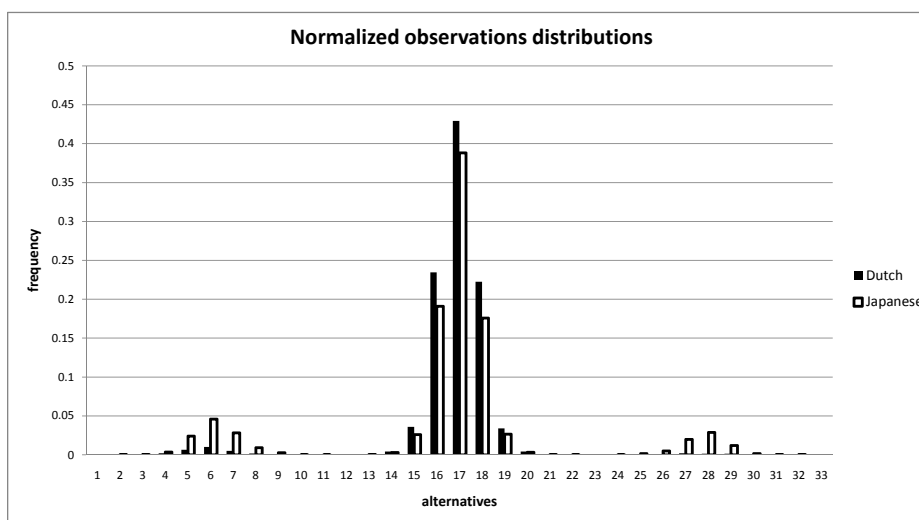


Figure 3.16: Comparison between the Japanese and Dutch normalized observation distributions across the alternatives

We applied our model with the parameters described in Table 3.3 on the Dutch data set, using the Biosim package. For each observation  $n$ , we obtain a probability distribution  $P_n(i)$  over the choice set.

Figure 3.17 represents the histogram of the probabilities  $P_n(i_n^*)$  of the chosen alternatives as predicted by the model, as well as the hazard value  $1/33$  (where 33 is the number of alternatives) illustrating the prediction of a purely random model with equal probabilities. Again, we consider observations below this threshold as outliers. We observe that there are 2.41% of them. This is good news, as it is actually less than for the data set used for parameter estimation. The shape of the

curve, as well as the low number of outliers are signs of the good performance of the model. The shape of the curve is even better in the Dutch case, than in the Japanese case, with higher frequencies for high choice probabilities. When we compare it with predictions obtained with the constant-only model (Figure 3.18), the superior forecasting potential of our model is clear.

The significant superiority of our model over the constant-only model is also illustrated by comparing the proportion of outliers (2.41% vs. 10.31%) or the loglikelihood (-51303.58 vs. -77269.28, as detailed in Table 3.14).

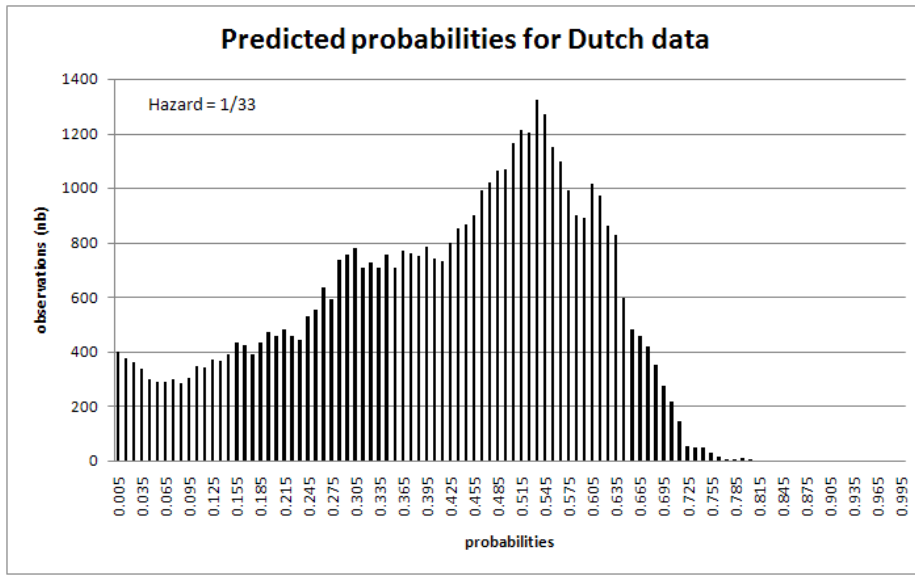


Figure 3.17: Prediction with the proposed model

We now compare the predictions performed by our model with the actual observations. The top part of Figure 3.19 reports the predicted probabilities obtained by sample enumeration, that is, for each  $i$ ,  $\sum_n P_n(i)$ , and the bottom part the observed shares, that is  $\sum_n y_{in}$ . The predictions are very satisfactory, except maybe for decelerations (alternatives 22 to 33) and accelerations (alternatives 1 to 11).

Cone	$\Gamma$	$M_\Gamma$	$R_\Gamma$	$(M_\Gamma - R_\Gamma)/R_\Gamma$
Front	5 – 7, 16 – 18, 27 – 29	43552.36	43374	0.0041
Left	3, 4, 14, 15, 25, 26	1948.77	2089	-0.0671
Right	8, 9, 19, 20, 30, 31	1853.34	1972	-0.0602
Extreme left	1, 2, 12, 13, 23, 24	43.91	27	0.6261
Extreme right	10, 11, 21, 22, 32, 33	82.62	19	3.3485

Table 3.9: Predicted ( $M_\Gamma$ ) and observed ( $R_\Gamma$ ) shares for alternatives grouped by directions with the Dutch data set.

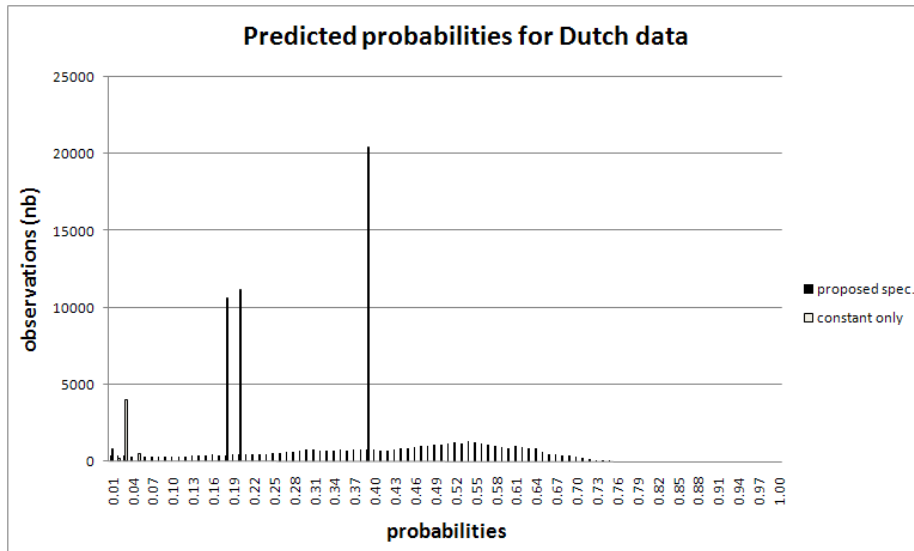
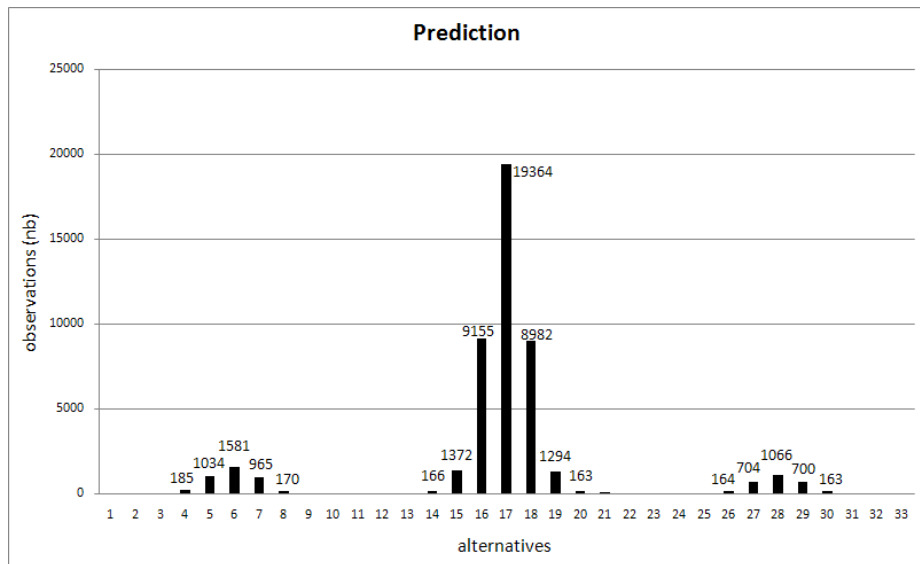


Figure 3.18: Prediction with the constant-only and proposed model

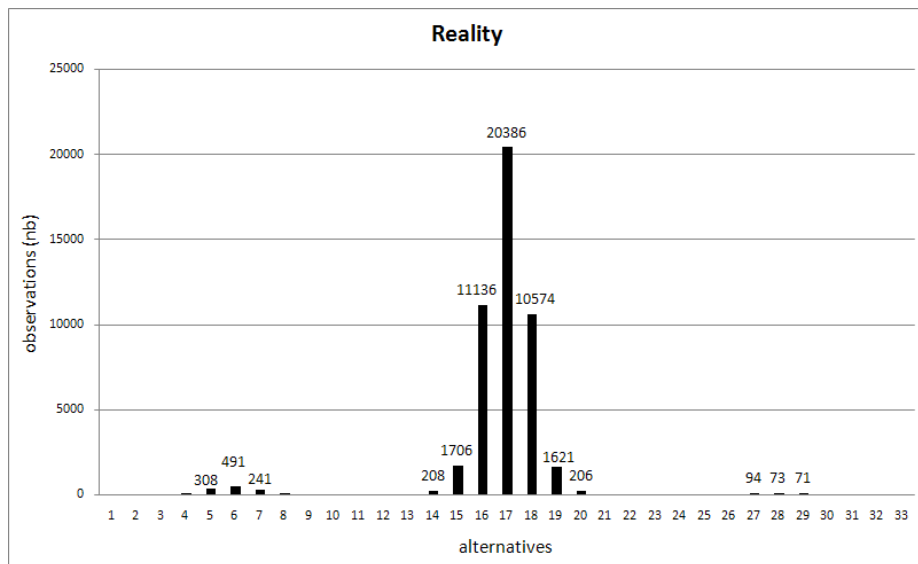
Area	$\Gamma$	$M_\Gamma$	$R_\Gamma$	$(M_\Gamma - R_\Gamma)/R_\Gamma$
acceleration	1 – 11	4022.32	1273	2.1597
constant speed	12 – 22	40581.06	45869	-0.1153
deceleration	23 – 33	2877.62	339	7.4886

Table 3.10: Predicted ( $M_\Gamma$ ) and observed ( $R_\Gamma$ ) shares for alternatives grouped by speed regime with the Dutch data set.





(a) Predicted



(b) Observed

Figure 3.19: Choice histogram predicted by the model against revealed choices in the Dutch data set

We also perform the comparison at a more aggregate level, for groups of cells. Tables 3.9 and 3.10 show a good overall performance of the model. Clearly, the **extreme left** and **extreme right** groups contain too few observations to reach any conclusions. The only bias seems to consist in a systematic over-prediction of accelerations and decelerations. This is consistent with the above-described analysis. The Dutch data set was collected in controlled experimental conditions, which may have introduced a bias in pedestrian behavior, depending on the exact instructions they have received. This assumption is supported by the quasi absence of decelerations in the data set, and by the different shapes of the speed distributions (see Figure 3.20). While the Japanese curve appears to be Gaussian, the Dutch curves contain some non-Gaussian features which are likely the result of the experimental nature of the data. In particular, the support is much narrower, with few high speeds. Note that, in the Japanese case, some pedestrians are running when the traffic light becomes red and cars start moving.

Data Set	Mean speed [m/s]
Dutch (experimental)	1.297
Japanese (real)	1.341

Table 3.11: Average pedestrian speed in the data sets

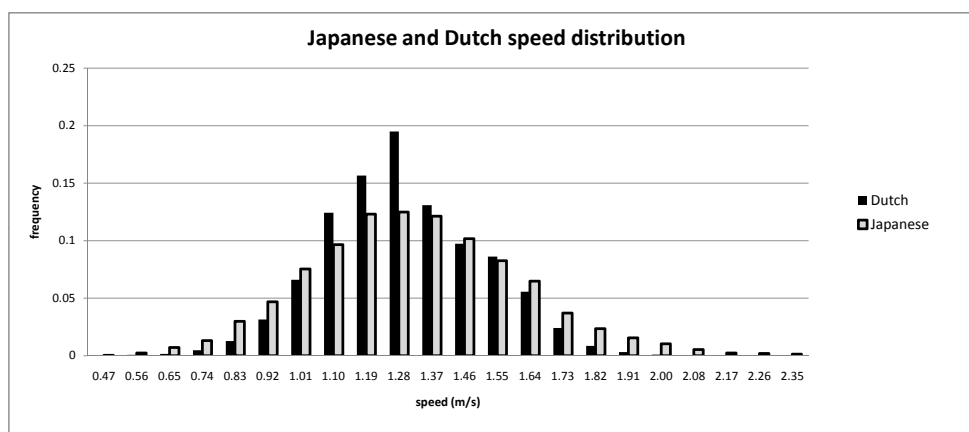


Figure 3.20: Distribution of speed in the two data sets

We now report the same aggregate prediction obtained with the constant-only model in Tables 3.12 and 3.13. The good performance of this simple model at the aggregate level emphasizes the need for the disaggregate validation performed above. Indeed, the relatively good performance of the model is due to the coincidental similarity of proportions of chosen alternatives in the two data sets (see Table 3.8). The detailed analysis presented in Figure 3.18 clearly rejects the simple model, while the aggregate analysis does not.

For the sake of completeness, a constant-only model was calibrated on the Dutch data set, in the same way as for the Japanese. Our model estimated on the Japanese

Cone	$\Gamma$	$M_\Gamma$	$R_\Gamma$	$(M_\Gamma - R_\Gamma)/R_\Gamma$
Front	5 – 7, 16 – 18, 27 – 29	43386.42	43374	0.0003
Left	3, 4, 14, 15, 25, 26	1877.47	2089	-0.1013
Right	8, 9, 19, 20, 30, 31	2082.10	1972	0.0558
Extreme left	1, 2, 12, 13, 23, 24	51.16	27	0.8947
Extreme right	10, 11, 21, 22, 32, 33	81.85	19	3.308

Table 3.12: Predicted ( $M_\Gamma$ ) using the constant-only model and observed ( $R_\Gamma$ ) shares for alternatives grouped by direction with the Dutch data set.

Area	$\Gamma$	$M_\Gamma$	$R_\Gamma$	$(M_\Gamma - R_\Gamma)/R_\Gamma$
acceleration	1 – 11	5448.24	1273	3.2798
constant speed	12 – 22	38700.42	45869	-0.1563
deceleration	23 – 33	3330.34	339	8.824

Table 3.13: Predicted ( $M_\Gamma$ ) using the constant-only model and observed ( $R_\Gamma$ ) shares for alternatives grouped by speed regime with the Dutch data set.

data is better than the constant-only model estimated on the Dutch data, when applied on the Dutch data set, both in terms of log-likelihood (-51303.58 against -71847.69) and prediction (2.41 %, percentage of outliers against 4.33%). We have summarized the various loglikelihood values in Table 3.14, where each column corresponds to a model, and each row to a data set.

Data set	Our model	Constant-only model based on Japanese data	Constant-only model based on Dutch data
Japanese	-13944.74	-17972.03	—
Dutch	-51303.58	-77269.28	-71847.69

Table 3.14: Loglikelihood of each model applied to the two data sets

In summary, we observe that our model applied to the estimation data (Japanese) have few outliers compared to the constant-only model, and reproduces the observed choices well. A forecasting cross-validation based on 80% of the sample illustrates the robustness of the specification. When the model is applied to the validation data (Dutch), we observe few outliers and an excellent probability histogram. Also, it reproduces the observed choices very well, in terms of directions and constant speed. We emphasize that this disaggregate analysis was necessary since the aggregate comparison does not reject the constant-only model.

### 3.8 Contributions

We proposed and validated a discrete choice model for the walking behavior, which lead to several contributions.

Several behavioral patterns have been considered and translated in mathematical formulations. The advantage of such an approach is the modularity. Behavioral patterns can be included in the model without changing the model structure, contrary to physical models (Helbing et al., 2002). The model is estimated on real data, as in the work done by Hoogendoorn et al. (2002). This guarantees the relevance of the modeling assumptions and the quality of the estimates. Contrary to rule-based models (see Blue and Adler, 2001), behavioral assumptions are not only checked at the validation stage, but also at the estimation.

Compared to the previous work of Antonini, Venegas, Bierlaire and Thiran (2006), two supplementary behavioral patterns have been added in the model. They concern the interactions with other pedestrians: *leader-follower* and *collision-avoidance*. They improve significantly the model in terms of fit and prediction.

We proposed and applied a validation methodology. The validation is crucial in order to check the prediction capabilities of the model, before using it in a simulation tool. This has been done in the literature mainly at an aggregate level (Berrou et al. (2007) and Kretz et al. (2008)), with the comparison of simulated and observed flows. In this work, we proposed both an aggregate and a disaggregate validation. The second step of the validation involves a Dutch data set, not used for estimation. In addition of the good validation results, the analysis allowed to compare the Japanese and Dutch data sets. These two data sets have been collected in two different setups: a real context in the Japanese case, and an experience in the Dutch case. Dutch pedestrians are walking at much more at a constant speed, compared to Japanese pedestrians, while the directional behavior is similar.

### 3.9 Conclusions

In this chapter we propose a discrete choice model of pedestrian walking behavior. The short range walking behavior of individuals is modeled, identifying two main patterns: constrained and unconstrained. The constraints are generated by the interactions with other individuals. We describe interactions in terms of leader-follower, and collision avoidance model. These models capture self-organizing effects which are characteristic of crowd behavior, such as lane formation. Inspiration for the mathematical form of these patterns is taken from driver behavior in transportation science, and ideas such as the car following model and lane changing models have been reviewed and re-adapted to the more complex pedestrian case. The difficulties of collecting pedestrian data as well as the limited information conveyed by pure dynamic data sets limit the possibilities in model specification. Important individual effects cannot be captured without the support of socio-economic characteristics. Re-

cent development of pedestrian laboratories (see among others Daamen and Hoogendoorn, 2003a, Nagai et al., 2005, Helbing et al., 2005, Cepolina and Tyler, 2005, Kretz et al., 2006), where controlled experimental conditions are possible, represent an important step in this direction. We use experimental data in a two step validation procedure. First, the model is validated on the same data set used for estimation in order to check for possible specification errors. Second, the model is run on a new data set collected at Delft University under controlled experimental conditions. The proposed validation procedure suggests good stability of the model and good forecasting performance. Few observations are badly predicted, mostly concentrated at the extremes of the choice set. The estimated coefficients are significant and their signs are consistent with our behavioral assumptions. As opposed to other previous models, we can quantify the influence of the relative kinematic characteristics of leaders and colliders on decision-maker behavior. Moreover, such quantitative analysis has been performed using real world pedestrian data.

The validation procedure is rather complete, since it involves several models, including a simple one, and analyzes the results both at an aggregate and a disaggregate level. The next step would be to validate the model within actual tools, such as pedestrian simulators or automatic video tracking systems (Antonini, Venegas, Bierlaire and Thiran, 2006). In particular, a simulation tool would allow for validation at the path level (Brogan and Johnson, 2003) or based on flows and densities (Berrou et al., 2007). Also, the validation of phenomena like spontaneous formation of lanes and queues requires a simulation environment. In such a context, the errors may quickly add up even with a good model, and a mechanism to keep the simulation on track may be necessary.

From a modeling viewpoint, future developments will focus on analyzing more and improving the acceleration and deceleration patterns. In particular, we plan to incorporate some physical and socio-economic characteristics of the pedestrians in the model. Also, we must investigate alternative resolutions of the choice set, with a possible adaptation of the resolution to the circumstances. In particular, the granularity of the most used alternatives (toward the front) may be revised. The current version has been designed to account for the resolution of the available data, and to limit the complexity of the model. Clearly, the behavioral relevance of this resolution should also be analyzed. The influence of the obstacles should be refined in the model. For example, we can adapt the *collision avoidance* pattern in this purpose. Regarding interactions with other pedestrian, the influence of pedestrian groups should also be considered. For the moment, only one single leader and collider are considered for each alternative.

Finally, it is important to emphasize that there is no such thing as a universal walking behavior model, as there are differences in actual walking behavior across circumstances and across cultures (Wiseman, 2007, pp. 262–268). The validation of the model guarantees that the model performs well in similar contexts, and is robust when used for forecasting.

84 3. *THE WALKING BEHAVIOR: MODELING OF PEDESTRIAN MOVEMENTS*

# 4. The judgmental behavior: dynamic facial expression recognition

## 4.1 Introduction

We propose a dynamic facial expression recognition framework based on discrete choice models. We model the choice of a person who has to label a video sequence representing a facial expression. The originality is based on the explicit modeling of causal effects between the facial features and the recognition of the expression. Five models are proposed. The first assumes that only the last frame of the video triggers the choice of the expression. The second model is composed of two parts. The first part captures the evaluation of the facial expression within each frame in the sequence. The second part determines which frame triggers the choice. The third model is an extension of the second model. It assumes that the choice of the expression results from the average of expression perceptions within a group of frames. The fourth and fifth models integrate the panel effect inherent to the estimation data and are respectively based on the first and second models. The models are estimated using videos from the Facial Expressions and Emotions Database (FEED). Labeling data on the videos has been obtained using an internet survey. The prediction capability of the models is studied in order to check their validity, by cross-validation using the estimation data. Estimation results and prediction capabilities of the five models are compared and discussed.

This chapter contains mainly the developments proposed by Robin et al. (2010).

## 4.2 Motivation

Facial expressions are essential to convey emotions and represent a powerful way used by human beings to relate to each other. When developing human machine interfaces, where computers have to take into account human emotions, automatic recognition of facial expressions plays a central role. In addition, The emotion is essential in many choice processes (Lerner and Keltner, 2000, Mellers and McGraw, 2001) and

the facial expression is one of the main indicators of the emotion.

Some coding systems have been proposed to describe facial expressions. Ekman and Friesen (1978) introduced the facial action coding system (FACS). They identified a list of fundamental expressions and associated groups of muscles tenseness or relaxations, called action units (AU) to each basic expression. A FACS expert can recognize AU activated on a face, and then deduct precisely the facial expression mixture. This is now the coding system of reference to characterize facial expressions.

The dynamic facial expression recognition (DFER) refers to the recognition of facial expressions in videos, whereas the static facial expression recognition (SFER) concerns the recognition of facial expressions in images. The DFER is an extension of the SFER. A great deal of research has been conducted in the field. Cohen et al. (2003) have developed an expression classifier based on a Bayesian network. They also propose a new architecture of hidden Markov model (HMM) for automatic segmentation and recognition of human facial expression from video sequences. Pantic and Patras (2006) present a dynamic system capable of recognizing facial AU and expressions, based on a particle filtering method. In this context, Bartlett et al. (2003) use a Support Vector Machine (SVM) classifier. Finally, Fasel and Luetttin (2003) study and compare methods and systems presented in the literature to deal with the DFER. They focus particularly on the robustness in case of environmental changes.

There is a recent interest in quantifying facial expressions in different fields such as robotics, marketing or transportation. In the robotic field, Tojo et al. (2000) have implemented facial and body expressions on a conversational robot. With some experiments, they showed the added value of such a system in the communication between humans and the robot. Miwa et al. (2004) have also developed a humanoid robot able to reproduce human expressions and their associated human hand movements. In the marketing field, Weinberg and Gottwald (1982) have investigated human behavior characterizing impulse purchases. Emotions play a key role and facial expressions appeared to be one of their main indicators. Small and Verrochi (2009) studied how the victim faces displayed on advertisements for charities affect both sympathy and giving.

Measuring user emotions has become an important research topic in transportation behavior analysis. In the car context, it may allow one to adapt the vehicle functionalities to the driver's mood for both well-being and safety reasons. Reimer et al. (2009) develop the concept of "awareness" of the vehicle in order to improve the mobility, performance and safety of older drivers. Information about driver general states, such as respiration, facial expression or concentration, are crucial to correctly apprehend the immediate driver capabilities and adapt the vehicle behavior to it. Moreover, some car manufacturers are currently working on the driver's mood recognition in order to warn the driver about possible dangers generated by other users. This aims at preventing road rage. Currently, the mood recognition is based only on the driver's voice. For routine trips, Abou-Zeid (2009) conducts experiments to measure the travel well-being for both public transportation and car modes. Collected data were employed to estimate mode choice models. Well-being measures are used



as utility indicators, in addition to standard choice indicators.

Contrarily to computer vision algorithms which are calibrated using a ground truth, the proposed models are estimated using behavioral data. Computer vision algorithms can be often considered as a “black box”, as their parameters are difficult to interpret. In our case, a specification is proposed where causal links between facial characteristics and expressions are explicitly modeled. The output of the model is a probability distribution among expressions. We have successfully applied the approach for SFER (Sorci, Antonini, Cruz, Robin, Bierlaire and Thiran, 2010, and Sorci, Robin, Cruz, Bierlaire, Thiran and Antonini, 2010). We propose a logit model, with nine alternatives corresponding to the nine considered expressions. Each utility is a function of measures related to the AU associated to the expression, as defined by the FACS. Sorci, Antonini, Cruz, Robin, Bierlaire and Thiran (2010) have also introduced the concept of expression descriptive units (EDU), that capture interactions between AU. Moreover, some outputs of the computer vision algorithm used to extract measures on facial images, are also included in the utility, in order to account for the global facial perception.

The DFER does not fit into the usual discrete choice applications, so adjustments have to be done. We took inspiration from the work of Choudhury (2007) who uses a dynamic behavioral framework to model car lane changing, and more generally from the framework developed by Ben-Akiva, 2010 for the concept of “planning and action”. Five models are presented in this analysis. Different modeling assumptions have been tested and compared. We first present the behavioral data used to estimate the models. Then the specification of the proposed models and the estimation results are presented. We finally describe the cross-validation and the predictions of the proposed models. In order to ease the understanding of the mentioned acronyms, Table B.1 in Appendix B.1 summarizes them and their definitions.

### 4.3 Data

The data is derived from a set of video sequences from the facial expressions and emotions database (FEED) collected by Wallhoff (2004). They have recorded students watching television. Different types of TV programs are presented to the subjects in order to generate a large spectrum of expressions. The database contains 95 sequences from 18 subjects. The collected videos last between 3 and 6 seconds. In each video, the subject starts with a neutral face (see example in Figure 4.1). Then, at some point the TV program triggers an expression (see example in Figure 4.2).

We have selected 65 videos from 17 subjects. The videos of subject N°17 were removed because of the lack of variability in facial characteristics, and to some discontinuities in the recording. The number of considered videos per subject is shown in Figure 4.3. We have no access to the type of expression that was meant to be triggered during the experiment.

A video is a sequence of images. For each image, numerical data are extracted



Figure 4.1: Snapshot of a FEED database video: neutral face (subject N°2)



Figure 4.2: Snapshot of a FEED database video: expression produced by the TV program (subject N°2)

using an active appearance model (AAM, Cootes et al., 2002). It allows to extract facial distances and angles as well as facial texture information (such as levels of gray) from each image. This technique is based on several principal component analysis (PCA) performed on the image treated as an array of pixel values. The algorithm tracks a facial mask composed of 55 points (see Figure 4.4) used to measure various facial distances and angles. The details of the mask are shown in Figure 4.5(a), as

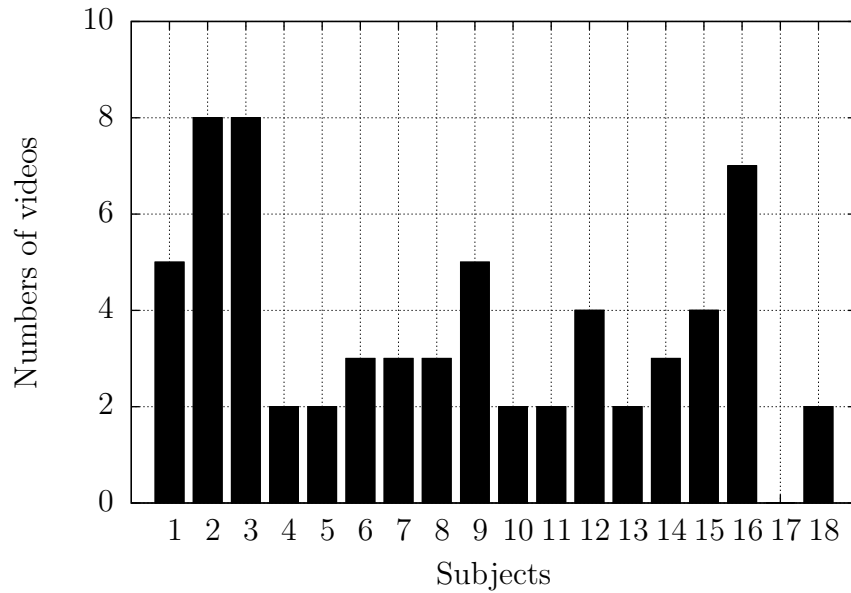


Figure 4.3: Numbers of considered videos per subject

well as the geometrical relationship of the facial measure points (Figure 4.5(b)) and some facial descriptors (Figure 4.5(c)). The correspondences between the measures on the mask displayed in Figure 4.5(b) and the mask presented in Figure 4.5(c), are shown in Table 4.1.

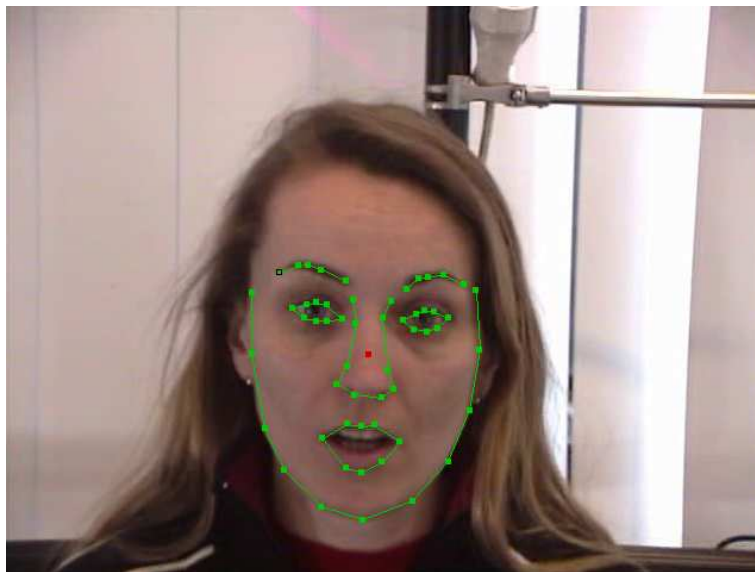


Figure 4.4: Mask tracked by AAM along a video sequence

Different explanatory variables based on the outputs of the AAM, are used to

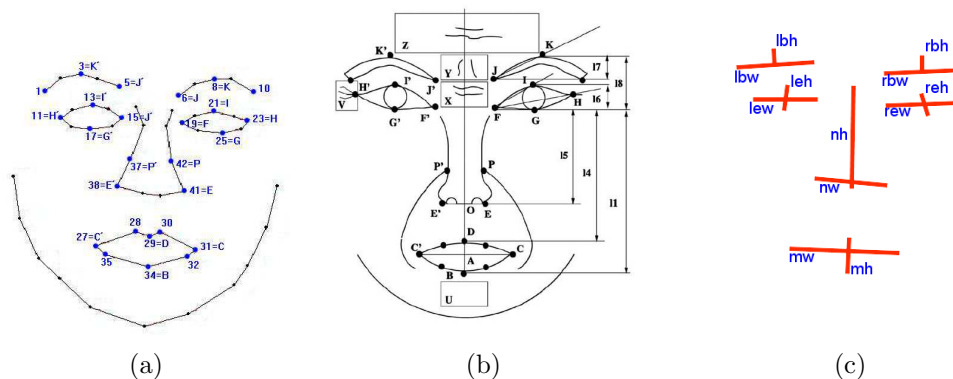


Figure 4.5: a) Facial landmarks (55 points); b) the geometrical relationship of facial feature points; c) some facial descriptors;

reflect the perception of facial expressions. They are coming from the facial action coding system (FACS); they are expression descriptive units (EDU), and also C parameters. We describe them briefly in the following. Note that a complete description of these variables can be found in Sorci, Antonini, Cruz, Robin, Bierlaire and Thiran (2010).

The FACS associates tensenesses and relaxations of muscles to each expression. They call them action units (AU). A sample of AU is presented in Figure 4.6. For example AU 6 is associated to happiness. The details of these associations are presented in Ekman and Friesen (1978). We translate the facial distances and angles extracted from the mask, into AU.



















AU1  Inner Brow Raiser	AU2  Outer Brow Raiser	AU4  Brow Lowerer	AU5  Upper Lid Raiser	AU6  Cheek Raiser	AU7  Lid Tightener
AU9  Nose Wrinkler	AU10  Upper Lip Raiser	AU12  Lip Corner Puller	AU15  Lip Corner Depressor	AU16  Lower Lip Depressor	AU17  Chin Raiser
AU20  Lip Stretcher	AU23  Lip Tightener	AU24  Lip Pressor	AU25  Lips part	AU26  Jaw Drop	AU27  Mouth Stretch

Figure 4.6: Sample of AU

EDU are reported in Table 4.2 and introduced by in Antonini, Sorci, Bierlaire and Thiran (2006). Additionally to the FACS, they account for the interactions between facial descriptors. The first 5 EDU represent, respectively, the eccentricity of eyes, left and right eyebrows, mouth and nose. The EDU from 7 to 9 represent



Figure 4.7: Examples of synthesized faces obtained varying the first  $C$  parameter from the mean face ( $\pm 3\text{std}$ ).

the eyes interactions with mouth and nose, while the 10<sup>th</sup> EDU is the nose-mouth relational unit. The last 4 EDU relate the eyebrows to mouth and nose. The EDU can be intuitively interpreted. For example, in a face displaying a surprise expression, the eyes and the mouth are usually opened and this can be captured by EDU7 ( $\text{eye}_{\text{height}}/\text{mouth}_{\text{height}}$ ).

Another vector  $C$  of values capturing both the facial texture and shape is also generated by the AAM. FACS and EDU provide measures of local facial features but they do not provide a description of a face as a global entity. This information can be obtained considering the appearance vector  $C$  matching the face in the processed image. Figure 4.7 shows the effect of varying the first appearance model parameter, showing changes in identity and expression.

A total of 88 variables capturing distances (number of pixels) and angles (radians), as well as 100 elements of the vector  $C$ , have been generated for each image in each video, which leads to obtain 188 variables per image.

The video is discretized in groups of 25 images, each corresponding to one second of the video, *i.e.* the number of groups of images is equal to the duration in seconds of the video. The features associated with each group of images are the features of the first image of the group. In the following, we use “frame” to refer to what is actually the first image of a group. The features of the 24 remaining images are used to compute variances (see Equation (4.2)).

For a given frame  $t$  and video  $o$ , three sets of variables are introduced:  $\{x_{k,t,o}\}_{k=1,\dots,188}$ ,  $\{y_{k,t,o}\}_{k=1,\dots,188}$ ,  $\{z_{k,t,o}\}_{k=1,\dots,188}$ .  $\{x_{k,t,o}\}_{k=1,\dots,188}$  are the features extracted using the AAM (188 = 88 variables capturing distances + 100 elements of the  $C$  vector).

Frame dynamics is captured by variables  $y_{k,t,o}$ . For each  $x_{k,t,o}$ ,  $k = 1, \dots, 188$ ,  $y_{k,t,o}$  is defined as

$$y_{k,t,o} = x_{k,t,o} - x_{k,t-1,o} \text{ for } t = 2, \dots, T_o, \quad (4.1)$$

where  $T_o$  is the number of frames in the video  $o$ . As each frame corresponds to one second,  $y_{k,t,o}$  can be interpreted as the first derivative of  $x_{k,t,o}$  with respect to time,

approximated by finite differences. It quantifies the level of variation of the facial characteristics between two consecutive frames.

Finally, another set of variables  $z_{k,t,o}$ , is introduced to capture the variation of  $x_{k,t,o}$  within a frame. For each  $x_{k,t,o}$ ,  $k = 1, \dots, 188$ ,  $z_{k,t,o}$  is defined as

$$z_{k,t,o} = \text{Var}(x_{k,t,o}). \quad (4.2)$$

It is the variance of the features calculated over the 25 images preceding the frame  $t$ . It characterizes the short time variations of the facial characteristic  $x_{k,t,o}$ . For logical reasons, we have fixed

$$y_{k,1,o} = z_{k,1,o} = 0 \quad \forall k, o, \quad (4.3)$$

meaning that the derivative and the variance of a variable in the first frame of all videos, is fixed to 0. We have a database of 564 ( $= 188 \times 3$ ) variables for each frame  $t$  in each video  $o$ . The variables have been normalized in the interval  $[-1, 1]$ , in order to harmonize their scale: each variable has been divided by the maximum in absolute value between its observed maximum and minimum over all frames and videos.

An internet survey has been conducted in order to obtain labels of FEED videos. The list of labels is composed of the seven basic expressions described by Keltner (2000): happiness (H), surprise (SU), fear (F), disgust (D), sadness (SA), anger (A), neutral (N). We have also added “Other” (O) and “I don’t know” (DK), to avoid ambiguities in the survey. It is available at <http://transp-or2.epfl.ch/videosurvey/> since august 2008. A screen snapshot is shown at Figure 4.8.

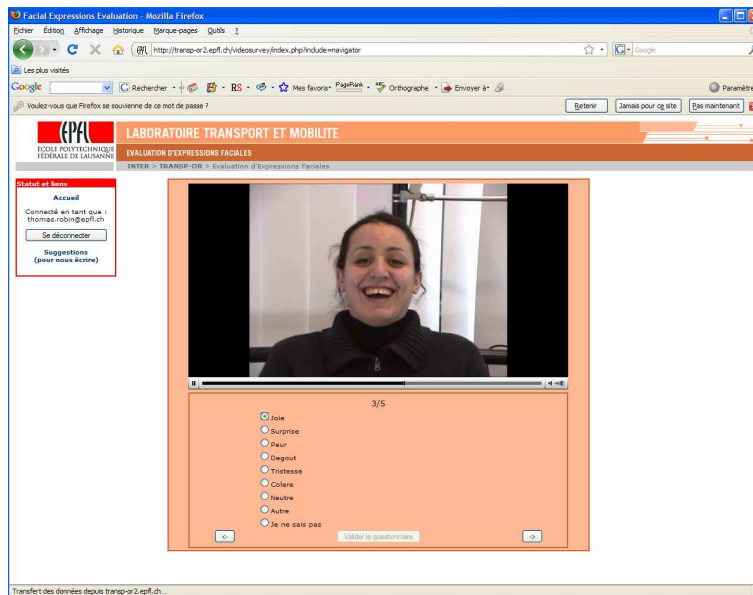


Figure 4.8: Snapshot of the internet survey screen (subject N°15)

For this analysis, we have collected 369 labels from 40 respondents. The repartition of the observations among the expressions is displayed in Figure 4.9.

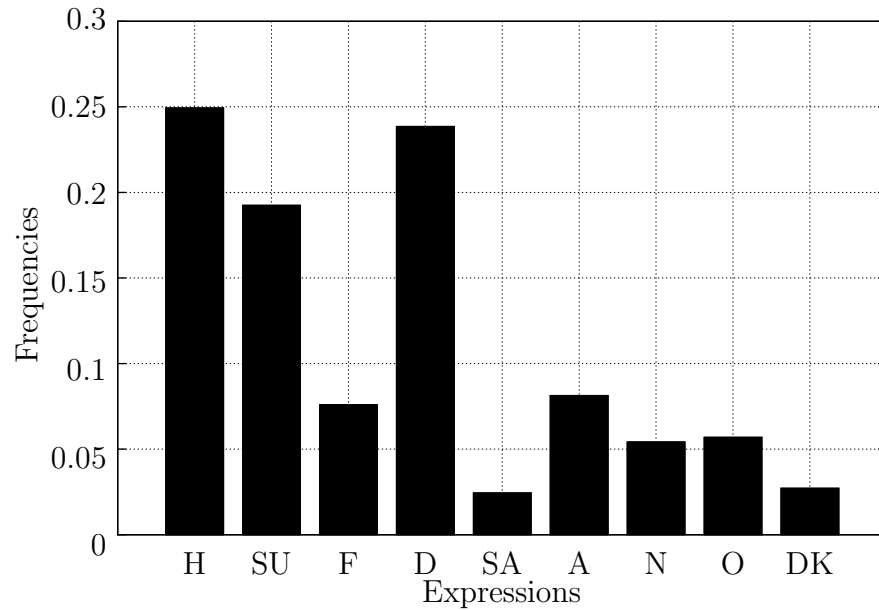


Figure 4.9: Distribution of the collected labels among expressions

FACS Measures	Measures on mask 4.5(a)
$\overline{JJ'}$	Dist(P5, P6)
$\overline{JF}$	Dist(P6, P19)
$\overline{J'F'}$	Dist(P5, P15)
$\overline{KG} \equiv 18$	Dist(P8, P25)
$\overline{K'G'}$	Dist(P3, P17)
$\overline{GI} \equiv 16$	Dist(P25, P21)
$\overline{G'I'}$	Dist(P13, P17)
$\overline{PF}$	Dist(P19, P42)
$\overline{P'F'}$	Dist(P15, P37)
$\overline{FC}$	Dist(P19, P31)
$\overline{F'C'}$	Dist(P15, P27)
$\overline{FD} \equiv 14$	Dist(P25, P29)
$\overline{F'D}$	Dist(P17, P29)
$\overline{OD}$	Dist( $(\frac{P39+P40}{2})$ , P29)
$\overline{OB}$	Dist( $(\frac{39+40}{2})$ , 33)
$\overline{DB}$	Dist(P29, P33)
$\overline{C'C}$	Dist(P27, P31)
$\triangle FHJ$	$\triangle P19P23P6$
$\triangle F'H'J'$	$\triangle P15P11P5$
$\triangle HFI$	$\triangle P23P19P21$
$\triangle H'F'I'$	$\triangle P11P15P13$
$\triangle HGF$	$\triangle P23P25P19$
$\triangle H'G'F'$	$\triangle P15P17P11$

Table 4.1: Correspondences between measures on masks 4.5(b) and 4.5(a)



EDU Measures	Measures definition
EDU1	$\frac{lew+rew}{leh+reh}$
EDU2	$\frac{lbw}{lbh}$
EDU3	$\frac{rbw}{rbh}$
EDU4	$\frac{mw}{mh}$
EDU5	$\frac{nw}{nw}$
EDU6	$\frac{lew}{mw}$
EDU7	$\frac{leh}{mh}$
EDU8	$\frac{leh+reh}{lbh+rbh}$
EDU9	$\frac{lew}{nw}$
EDU10	$\frac{nw}{mw}$
EDU11	$\frac{EDU2}{EDU4}$
EDU12	$\frac{EDU3}{EDU4}$
EDU13	$\frac{EDU2}{EDU10}$
EDU14	$\frac{EDU3}{EDU4}$

Table 4.2: Expressions Descriptive Units

## 4.4 Models specification

We consider a decision-maker who has to label a video sequence by choosing among the list of facial expressions described in Section 4.3 (happiness (H), surprise (SU), fear (F), disgust (D), sadness (SA), anger (A), neutral (N), other (O), not known (DK)). Five models based on different assumptions have been developed. We suppose that the perception of the respondent starts at the first frame of the video. Then, we assume that the respondent updates her perception every second, which corresponds to every frame (see Section 4.3). In the first model we hypothesize that only the last frame of the video influences the observed choice of label. This is the simplest model presented in this analysis because it does not include dynamic aspects and it will be considered as a reference for comparison. This model is called **reduced model**. In the second model, only the most impressive frame is supposed to be influential on the choice of label. It is called **latent model**. In the third model, we hypothesize that it is the average perception of a group of consecutive frames which generates the choice of label. This is called **smoothed model**. Two supplementary models are proposed in order to account for the panel nature of the data, they are based on the first and second models and called **reduced model with panel effect** and **latent model with panel effect**. Note that a **smoothed model with panel effect** is not considered due to its estimation complexity.

The theoretical details and specification of each model are described in Sections 4.4.1, 4.4.2, 4.4.3, 4.4.4 and 4.4.4. They are all extensions of the model proposed by Sorci, Antonini, Cruz, Robin, Bierlaire and Thiran (2010), which is called **static model**. Due to the small number of respondents, their socio-economic characteristics have not been included in the models.

### 4.4.1 The reduced model

We first assume that the perception of the last frame of a video is suggesting the choice of label. The filmed subject starts with a neutral face and evolves toward a certain expression which is triggered by the TV program that she is watching, and the video ends. The subject's face on the last frame should be expressive. The model is a direct application of the **static model** in the last frame.

The model associated to the perception of expressions is denoted by  $P_{M_1}(i|\mathbf{o}, \theta_{M_1})$ . It is the probability for an individual to label the video  $\mathbf{o}$  with the expression  $i$ , given the vector of unknown parameters  $\theta_{M_1}$ . The last frame is supposed to be the only information used by the respondent to label the video  $\mathbf{o}$ . The utility function associated with each expression is defined in Equation (4.4).

$$\begin{aligned}
V_{M_1}(H|T_o, o, \theta_{M_1}) &= ASC_{M_1,H} + \sum_{j=1}^{K_{M_1}} I_{M_1,H,j} \theta_{M_1,j} \sum_{k=1}^{188} I_{M_1,j,k} x_{k,T_o,o} , \\
V_{M_1}(SU|T_o, o, \theta_{M_1}) &= ASC_{M_1,SU} + \sum_{j=1}^{K_{M_1}} I_{M_1,SU,j} \theta_{M_1,j} \sum_{k=1}^{188} I_{M_1,j,k} x_{k,T_o,o} , \\
V_{M_1}(F|T_o, o, \theta_{M_1}) &= ASC_{M_1,F} + \sum_{j=1}^{K_{M_1}} I_{M_1,F,j} \theta_{M_1,j} \sum_{k=1}^{188} I_{M_1,j,k} x_{k,T_o,o} , \\
V_{M_1}(D|T_o, o, \theta_{M_1}) &= ASC_{M_1,D} + \sum_{j=1}^{K_{M_1}} I_{M_1,D,j} \theta_{M_1,j} \sum_{k=1}^{188} I_{M_1,j,k} x_{k,T_o,o} , \\
V_{M_1}(SA|T_o, o, \theta_{M_1}) &= ASC_{M_1,SA} + \sum_{j=1}^{K_{M_1}} I_{M_1,SA,j} \theta_{M_1,j} \sum_{k=1}^{188} I_{M_1,j,k} x_{k,T_o,o} , \\
V_{M_1}(A|T_o, o, \theta_{M_1}) &= ASC_{M_1,A} + \sum_{j=1}^{K_{M_1}} I_{M_1,A,j} \theta_{M_1,j} \sum_{k=1}^{188} I_{M_1,j,k} x_{k,T_o,o} , \\
V_{M_1}(N|T_o, o, \theta_{M_1}) &= 0 , \\
V_{M_1}(O|T_o, o, \theta_{M_1}) &= ASC_{M_1,O} + \sum_{j=1}^{K_{M_1}} I_{M_1,O,j} \theta_{M_1,j} \sum_{k=1}^{188} I_{M_1,j,k} x_{k,T_o,o} , \\
V_{M_1}(DK|T_o, o, \theta_{M_1}) &= ASC_{M_1,DK} , \tag{4.4}
\end{aligned}$$

where  $T_o$  denotes the length of the video  $o$  in seconds, which is also the index of the last frame of the video  $o$ .  $K_{M_1}$  is the total number of parameters associated to facial measurements  $\{x_{k,t,o}\}$  in the **reduced model**.  $I_{M_1,i,j}$  is an indicator equal to 1 if the parameter  $j$  is present in the utility of expression  $i$ , 0 otherwise.  $I_{M_1,j,k}$  is an indicator equal to 1 if the parameter  $j$  is related to the facial measurement  $x_{k,T_o,o}$  collected in the last frame of the video  $o$ , 0 otherwise. We have

$$\sum_{k=1}^{188} I_{M_1,j,k} = 1 \quad \forall j , \tag{4.5}$$

meaning that a parameter  $\theta_{M_1,j}$  is related to only one facial measurement  $x_{k,T_o,o}$ . Each utility contains an alternative specific constant  $ASC_{M_1,i}$  except the neutral, which is taken as the reference, and its utility is fixed to 0. Note that there is no expression specific attributes, as the facial characteristics do not vary across the expressions. The details of the utility specifications are presented in Tables B.2 and B.3. For each parameter  $\theta_{M_1,j}$ , if  $I_{M_1,i,j}$  is equal to 1, there is a “×” in the column of the corresponding expression  $i$ . This notations is used in all Tables in Appendix B.2. If

$I_{M_1,j,k}$  is equal to 1, the relative facial characteristic  $\chi_{k,T_0,o}$  is indicated. The model is a logit, so the probability is

$$P_{M_1}(i|o, \theta_{M_1}) = \frac{e^{V_{M_1}(i|T_0,o,\theta_{M_1})}}{\sum_{j=1}^9 e^{V_{M_1}(j|T_0,o,\theta_{M_1})}}. \quad (4.6)$$

Then the log-likelihood is

$$\mathcal{L}(\theta_{M_1}) = \sum_{o=1}^O \sum_{i=1}^9 w_{i,o} \log(P_{M_1}(i|o, \theta_{M_1})), \quad (4.7)$$

where  $w_{i,o}$  is a weight, corresponding to the number of times the expression  $i$  has been chosen for the video  $o$  in the collected database of annotations (see Section 4.3).

Sorci, Antonini, Cruz, Robin, Bierlaire and Thiran (2010) employed the database proposed by T.Kanade (2000) when collecting behavioral data. The estimated parameters of the static model cannot be used directly in our analysis due to problems of facial position and scale between this database and the FEED (see Section 4.3). The filmed subjects are further from the camera in the FEED, compared to the Cohn-Kanade. Consequently, the model has to be re-estimated. In addition, the specifications of the utilities have been adapted to this analysis because of the lower number of available data. We use 369 observations of labels against 38110 for the work of Sorci, Antonini, Cruz, Robin, Bierlaire and Thiran (2010). This implies the estimation of a lower number of parameters: the utility specifications have been simplified and parameters have been grouped together regarding their sign and interpretability. The proposed model contains 32 parameters against 135 for the **static model**.

#### 4.4.2 The latent model

The assumption supporting this model is that one frame in the video has influenced the observed choice of label, but the analyst does not know which one. The DFER model consists of a combination of two models. The first model quantifies the perception of expressions in a given frame. It is similar to the **reduced model** presented in Section 4.4.1. The second model predicts which frame has influenced the chosen label. It is a latent choice model where the choice set is composed of all frames in the video. The instantaneous perception of expressions and the most influential frame are not observed. Only the final choice of label for the video is observed.

The first model provides the probability for a respondent to choose the expression  $i$  when exposed to the frame  $t$  of the video sequence  $o$ , and is written  $P_{M_2}(i|t, o, \theta_{M_2,1}, \alpha_{M_2})$ . The second model provides the probability for the frame  $t$  of video  $o$  to trigger the choice, and is denoted by  $P_{M_2}(t|o, \theta_{M_2,2})$ . The probability for a respondent to label the video  $o$  with expression  $i$ , is denoted by  $P_{M_2}(i|o, \theta_{M_2}, \alpha_{M_2})$ , which is observable.  $\theta_{M_2,1}$  and  $\theta_{M_2,2}$  are the vectors of unknown parameters to be estimated,

merged into the vector  $\theta_{M_2}$ .  $\alpha_{M_2}$  is a vector of parameters capturing the memory effects, which will be introduced in Equation (4.11), and has to be estimated ( $\alpha_{M_2} = \{\alpha_{M_2,i}\}_{i=H,SU,F,D,SA,A,O}$ ). We obtain

$$P_{M_2}(i|o, \theta_{M_2}, \alpha_{M_2}) = \sum_{t=1}^{T_0} P_{M_2}(i|t, o, \theta_{M_2,1}, \alpha_{M_2}) P_{M_2}(t|o, \theta_{M_2,2}). \quad (4.8)$$

For specifying the model  $P_{M_2}(i|t, o, \theta_{M_2,1}, \alpha_{M_2})$ , we need to define a utility function associated to each expression. We hypothesize that the perception of an expression  $i$  in frame  $t$  depends on the instantaneous perceptions of this expression  $i$  in the frames  $t$  and  $t - 1$ .  $V_{M_2}(i|t, o, \theta_{M_2,1}, \alpha_{M_2,i})$  is a utility reflecting the perception of the expression  $i$  in frame  $t$  for the video  $o$ . We decompose it into two parts. First  $V_{M_2}^s(i|t, o, \theta_{M_2,1})$  concerns the instantaneous perception of the frame  $t$  in the video  $o$ . Second,  $V_{M_2}^s(i|t - 1, o, \theta_{M_2,1})$  concerns the instantaneous perception of the frame  $t - 1$  in the video  $o$ . This is designed to capture the dynamic nature of the decision making process, as illustrated in Figure 4.10. In this figure, the facial measurements  $\{x_{k,t,o}\}$  and  $\{z_{k,t,o}\}$  (introduced in Equation (4.2)) are observed, they are enclosed in rectangles and their influences are represented by plain arrows; whereas the utilities are latent, they are enclosed in ellipses and their influences are marked by dashed arrows.  $\{x_{k,t,o}\}$  and  $\{z_{k,t,o}\}$  influence  $V_{M_2}^s(i|t, o, \theta_{M_2,1})$ , while  $V_{M_2}(i|t, o, \theta_{M_2,1}, \alpha_{M_2,i})$  is only function of  $V_{M_2}^s(i|t, o, \theta_{M_2,1})$  and  $V_{M_2}^s(i|t - 1, o, \theta_{M_2,1})$ .

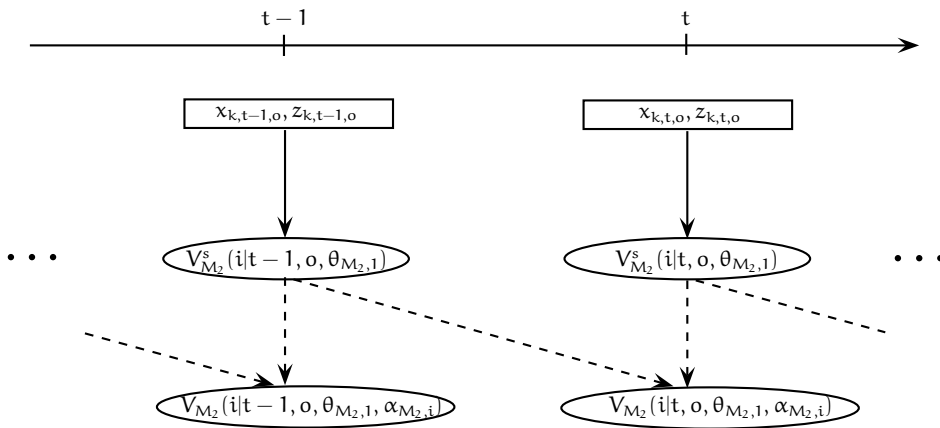


Figure 4.10: The dynamic process of the **latent model**

The specification of  $\{V_{M_2}^s(i|t, o, \theta_{M_2,1})\}$  is presented in Equation (4.9)

$$\begin{aligned}
V_{M_2}^s(H|t, o, \theta_{M_2,1}) &= ASC_{M_2,H} + \sum_{j=1}^{K_{M_2}} I_{M_2,1,H,j} \theta_{M_2,1,j} \sum_{k=1}^{188} I_{M_2,j,k} x_{k,t,o} , \\
V_{M_2}^s(SU|t, o, \theta_{M_2,1}) &= ASC_{M_2,SU} + \sum_{j=1}^{K_{M_2}} I_{M_2,1,SU,j} \theta_{M_2,1,j} \sum_{k=1}^{188} I_{M_2,j,k} x_{k,t,o} \\
&\quad + \sum_{j=1}^{K_{M_2}^z} I_{M_2,SU,j}^z \theta_{M_2,1,j}^z \sum_{k=1}^{188} I_{M_2,j,k}^z z_{k,t,o} , \\
V_{M_2}^s(F|t, o, \theta_{M_2,1}) &= ASC_{M_2,F} + \sum_{j=1}^{K_{M_2}} I_{M_2,F,j} \theta_{M_2,1,j} \sum_{k=1}^{188} I_{M_2,j,k} x_{k,t,o} , \\
V_{M_2}^s(D|t, o, \theta_{M_2,1}) &= ASC_{M_2,D} + \sum_{j=1}^{K_{M_2}} I_{M_2,D,j} \theta_{M_2,1,j} \sum_{k=1}^{188} I_{M_2,j,k} x_{k,t,o} , \\
V_{M_2}^s(SA|t, o, \theta_{M_2,1}) &= ASC_{M_2,SA} + \sum_{j=1}^{K_{M_2}} I_{M_2,SA,j} \theta_{M_2,1,j} \sum_{k=1}^{188} I_{M_2,j,k} x_{k,t,o} , \\
V_{M_2}^s(A|t, o, \theta_{M_2,1}) &= ASC_{M_2,A} + \sum_{j=1}^{K_{M_2}} I_{M_2,A,j} \theta_{M_2,1,j} \sum_{k=1}^{188} I_{M_2,j,k} x_{k,t,o} , \\
V_{M_2}^s(N|t, o, \theta_{M_2,1}) &= 0 , \\
V_{M_2}^s(O|t, o, \theta_{M_2,1}) &= ASC_{M_2,O} + \sum_{j=1}^{K_{M_2}} I_{M_2,O,j} \theta_{M_2,1,j} \sum_{k=1}^{188} I_{M_2,j,k} x_{k,t,o} , \\
V_{M_2}^s(DK|t, o, \theta_{M_2,1}) &= ASC_{M_2,DK} , \tag{4.9}
\end{aligned}$$

where  $K_{M_2}$  is the total number of parameters related to  $\{x_{k,t,o}\}$ .  $K_{M_2}^z$  is the total number of parameters related to  $\{z_{k,t,o}\}$ . The indicators are similar to those introduced in Section 4.4.1.  $I_{M_2,i,j}$  is an indicator equal to 1 if the parameter  $j$  is included in the utility of expression  $i$ , 0 otherwise.  $I_{M_2,j,k}$  is an indicator equal to 1 if the parameter  $j$  is related to the facial measurement  $x_{k,t,o}$  collected in the frame  $t$  of the video  $o$ , 0 otherwise. We have

$$\sum_{k=1}^{188} I_{M_2,j,k} = 1 \quad \forall j , \tag{4.10}$$

meaning that a parameter  $\theta_{M_2,j}$  is related to only one  $x_{k,t,o}$ .  $I_{M_2,SU,j}^z$  and  $I_{M_2,j,k}^z$  have exactly the same role as  $I_{M_2,i,j}$  and  $I_{M_2,j,k}$ , but they concern the parameter  $\theta_{M_2,j}^z$  which is related to  $z_{k,t,o}$ . Each utility contains a constant, except for the neutral expression, whose utility is the reference and is fixed to 0. The presence of  $\{z_{k,t,o}\}$  (short time

variations of facial characteristics) in the surprise utility accounts for the perception of suddenness.  $\{z_{kto}\}$  are better than  $\{y_{k,t,o}\}$  in this case, because they capture faster variations of facial characteristics. This does not lead necessarily to the surprise facial expression, but according to the collected data, fast variations of facial characteristics could be perceived as surprise by respondents.  $\{z_{kto}\}$  have been tested in the *reduced model*, but the associated parameters did not appear to be significant, certainly due to the simplistic assumption about the last frame triggering the expression choice. The detailed specification of  $V_{M_2}^s(i|t, o, \theta_{M_2,1})$  is described in Tables B.4 and B.5. The reading of the tables is exactly the same as for Table B.2 described in Section 4.4.1.

The utility function  $V_{M_2}(i|t, o, \theta_{M_2,1}, \alpha_{M_2,i})$  is supposed to be the sum of  $V_{M_2}^s(i|t, o, \theta_{M_2,1})$  and  $\{V_{M_2}^s(i|t-1, o, \theta_{M_2,1})$  weighted by  $\alpha_{M_2,i}$ , the parameter of memory effect. The specification of  $V_{M_2}(i|t, o, \theta_{M_2,1}, \alpha_{M_2,i})$  is defined in Equation (4.11).

$$\begin{aligned}
V_{M_2}(H|t, o, \theta_{M_2,1}, \alpha_{M_2,H}) &= V_{M_2}^s(H|t, o, \theta_{M_2,1}) \\
&\quad + \alpha_{M_2,H} V_{M_2}^s(H|t-1, o, \theta_{M_2,1}), \\
V_{M_2}(SU|t, o, \theta_{M_2,1}, \alpha_{M_2,SU}) &= V_{M_2}^s(SU|t, o, \theta_{M_2,1}), \\
V_{M_2}(F|t, o, \theta_{M_2,1}, \alpha_{M_2,F}) &= V_{M_2}^s(F|t, o, \theta_{M_2,1}) \\
&\quad + \alpha_{M_2,F} V_{M_2}^s(F|t-1, o, \theta_{M_2,1}), \\
V_{M_2}(D|t, o, \theta_{M_2,1}, \alpha_{M_2,D}) &= V_{M_2}^s(D|t, o, \theta_{M_2,1}), \\
V_{M_2}(SA|t, o, \theta_{M_2,1}, \alpha_{M_2,SA}) &= V_{M_2}^s(SA|t, o, \theta_{M_2,1}) \\
&\quad + \alpha_{M_2,SA} V_{M_2}^s(SA|t, o, \theta_{M_2,1}), \\
V_{M_2}(A|t, o, \theta_{M_2,1}, \alpha_{M_2,A}) &= V_{M_2}^s(A|t, o, \theta_{M_2,1}), \\
V_{M_2}(N|t, o, \theta_{M_2,1}, \alpha_{M_2,N}) &= V_{M_2}^s(N|t, o, \theta_{M_2,1}) = 0, \\
V_{M_2}(O|t, o, \theta_{M_2,1}, \alpha_{M_2,O}) &= V_{M_2}^s(O|t, o, \theta_{M_2,1}) \\
&\quad + \alpha_{M_2,O} V_{M_2}^s(O|t, o, \theta_{M_2,1}), \\
V_{M_2}(DK|t, o, \theta_{M_2,1}, \alpha_{M_2,DK}) &= V_{M_2}^s(DK|t, o, \theta_{M_2,1}). \tag{4.11}
\end{aligned}$$

Note that this is not anymore a linear-in-parameter specification for happiness, fear, sadness and anger, since  $\{\alpha_{i,j}\}$  are estimated. Five memory effect parameters  $\{\alpha_{M_2,i}\}_{i=SU,D,A,N,DK}$  have been fixed to 0 : for neutral because it is the referent alternative, so its utility is fixed to zero; and for ‘‘I don’t know’’ because its utility contains only  $ASC_{M_2,DK}$ , which is invariant across the frames. For surprise, disgust and anger, they do not appeared to be significant in previous specifications of the model (see Section 4.6 and Table B.6).  $\{\alpha_{M_2,i}\}_{i=H,F,SA,O}$  are supposed to be in the interval  $[-1, 1]$  because we hypothesize that the instantaneous perception of expression  $i$  at time  $t$  is more influenced by the instantaneous perception of expression  $i$  at frame  $t$  than at frame  $t-1$ . This dynamic specification has not been tested in the **reduced model**, as in this model we hypothesized that only the last frame of the video was triggering the expression (and not the two last frames). The model for

$P_{M_2}(i|t, o, \theta_{M_2,1}, \alpha_{M_2})$  is a logit model, that is

$$P_{M_2}(i|t, o, \theta_{M_2,1}, \alpha_{M_2}) = \frac{e^{V_{M_2}(i|t, o, \theta_{M_2,1}, \alpha_{M_2, i})}}{\sum_j e^{V_{M_2}(j|t, o, \theta_{M_2,1}, \alpha_{M_2, j})}}. \quad (4.12)$$

The model  $P_{M_2}(t|o, \theta_{M_2,2})$  is also specified as a logit model. Note that we decide to ignore here the potential correlation between error terms of successive frames. A utility  $V_{M_2}(t|o, \theta_{M_2,2})$  is associated to each frame  $t$  in the video  $o$ . The utility depends on variables  $\{y_{k,t,o}\}$  (see Equation (4.1)), and  $\{z_{k,t,o}\}$  (see Equation (4.2)). We define  $V_{M_2}(1|o, \theta_{M_2,2}) = 0$  and, for  $t = 2, \dots, T_o$ ,

$$\begin{aligned} V_{M_2}(t|o, \theta_{M_2,2}) &= \sum_{j=1}^{K_{M_2,2}^y} \theta_{M_2,2,j}^y \sum_{k=1}^{188} I_{M_2,2,j,k}^y y_{k,t,o} \\ &+ \sum_{j=1}^{K_{M_2,2}^z} \theta_{M_2,2,j}^z \sum_{k=1}^{188} I_{M_2,2,j,k}^z z_{k,t,o}, \end{aligned} \quad (4.13)$$

and

$$P_{M_2}(t|o; \theta_{M_2,2}) = \frac{e^{V_{M_2}(t|o, \theta_{M_2,2})}}{\sum_{\ell=1}^{T_o} e^{V_{M_2}(\ell|o, \theta_{M_2,2})}}. \quad (4.14)$$

$K_{M_2,2}^y$  and  $K_{M_2,2}^z$  are numbers of parameters associated to  $\{y_{k,t,o}\}$ , and  $\{z_{k,t,o}\}$  respectively, in the utility related to each frame.  $I_{M_2,2,j,k}^y$  is an indicator equal to 1 if the parameter  $\theta_{M_2,2,j}^y$  is associated to  $y_{k,t,o}$ , 0 otherwise. As for the other indicators, it is related to only one  $y_{k,t,o}$ , we have

$$\sum_{k=1}^{188} I_{M_2,2,j,k}^y = 1 \quad \forall j, \quad (4.15)$$

$I_{M_2,2,j,k}^z$  is similar to  $I_{M_2,2,j,k}^y$ , but is associated to  $z_{k,t,o}$ . The vector of parameters  $\theta_{M_2,2}$  is described in Table B.7. Finally, the log-likelihood function is

$$\begin{aligned} \mathcal{L}(\theta_{M_2}, \alpha_{M_2}) &= \sum_{o=1}^O \sum_{i=1}^9 w_{i,o} \log P_{M_2}(i|o, \theta_{M_2}, \alpha_{M_2}) \\ &= \sum_{o=1}^O \sum_{i=1}^9 w_{i,o} \log \left( \sum_{t=1}^{T_o} P_{M_2}(i|t, o, \theta_{M_2,1}, \alpha_{M_2}) P_{M_2}(t|o, \theta_{M_2,2}) \right). \end{aligned} \quad (4.16)$$

#### 4.4.3 The smoothed model

In this model, we hypothesize that the behavior of the respondent is composed of two consecutive phases, when watching a video. In the first phase, the respondent is



waiting for information, no perception of expressions is influencing the observed choice of label. At a certain point in time, the respondent starts to use the information of the frames to make her choice of label. This consideration of information is continued until the end of the video and constitutes the second phase. The model combines a model related to the perception of expressions and a model which detects the changing of phase. The observed choice of label is supposed to be the average across the frames of the perception of expressions in the second phase. Both models are latent as only the choice of label is observed.

The first model provides the probability for a respondent to choose the expression  $i$  when exposed to frame  $\ell$  of the video sequence  $\mathbf{o}$ , and is written  $P_{M_3}(i|\ell, \mathbf{o}, \theta_{M_3,1})$ . The second model  $P_{M_3}(t|\mathbf{o}, \theta_{M_3,2})$  provides the probability for a respondent to enter in her second phase when being exposed to the frame  $t$ . The probability for a respondent to label the video  $\mathbf{o}$  with expression  $i$ , is denoted by  $P_{M_3}(i|\mathbf{o}, \theta_{M_3})$ , which is observable.  $\theta_{M_3,1}$  and  $\theta_{M_3,2}$  are the vectors of unknown parameters to be estimated within each of the two models, merged into the vector  $\theta_{M_3}$ .  $P_{M_3}(i|\mathbf{o}, \theta_{M_3})$  is the average of  $\{P_{M_3}(i|\ell, \mathbf{o}, \theta_{M_3,1})\}_{\ell=t \dots T_o}$ , weighted by  $P_{M_3,n}(t|\mathbf{o}, \theta_{M_3,2})$ , sum up over  $t = 1 \dots T_o$ . We obtain

$$P_{M_3}(i|\mathbf{o}, \theta_{M_3}) = \sum_{t=1}^{T_o} P_{M_3}(t|\mathbf{o}, \theta_{M_3,2}) \frac{1}{T_o - t + 1} \sum_{\ell=t}^{T_o} P_{M_3}(i|\ell, \mathbf{o}, \theta_{M_3,1}). \quad (4.17)$$

For  $P_{M_3}(i|\mathbf{o}, \theta_{M_3,1})$ , a utility  $V_{M_3}(i|\mathbf{o}, \theta_{M_3,1})$  is associated to each expression  $i$ . The specification of  $\{V_{M_3}(i|\mathbf{o}, \theta_{M_3,1})\}$  is defined in Equation (4.18).

$$\begin{aligned}
V_{M_3}(H|t, o, \theta_{M_3,1}) &= ASC_{M_3,H} + \sum_{j=1}^{K_{M_3}} I_{M_3,1,H,j} \theta_{M_3,1,j} \sum_{k=1}^{188} I_{M_3,j,k} \chi_{k,t,o} , \\
V_{M_3}(SU|t, o, \theta_{M_3,1}) &= ASC_{M_3,SU} + \sum_{j=1}^{K_{M_3}} I_{M_3,1,SU,j} \theta_{M_3,1,j} \sum_{k=1}^{188} I_{M_3,j,k} \chi_{k,t,o} \\
&\quad + \sum_{j=1}^{K_{M_3}^z} I_{M_3,SU,j}^z \theta_{M_3,1,j}^z \sum_{k=1}^{188} I_{M_3,j,k}^z \chi_{k,t,o} , \\
V_{M_3}(F|t, o, \theta_{M_3,1}) &= ASC_{M_3,F} + \sum_{j=1}^{K_{M_3}} I_{M_3,F,j} \theta_{M_3,1,j} \sum_{k=1}^{188} I_{M_3,j,k} \chi_{k,t,o} , \\
V_{M_3}(D|t, o, \theta_{M_3,1}) &= ASC_{M_3,D} + \sum_{j=1}^{K_{M_3}} I_{M_3,D,j} \theta_{M_3,1,j} \sum_{k=1}^{188} I_{M_3,j,k} \chi_{k,t,o} , \\
V_{M_3}(SA|t, o, \theta_{M_3,1}) &= ASC_{M_3,SA} + \sum_{j=1}^{K_{M_3}} I_{M_3,SA,j} \theta_{M_3,1,j} \sum_{k=1}^{188} I_{M_3,j,k} \chi_{k,t,o} , \\
V_{M_3}(A|t, o, \theta_{M_3,1}) &= ASC_{M_3,A} + \sum_{j=1}^{K_{M_3}} I_{M_3,A,j} \theta_{M_3,1,j} \sum_{k=1}^{188} I_{M_3,j,k} \chi_{k,t,o} , \\
V_{M_3}(N|t, o, \theta_{M_3,1}) &= 0 , \\
V_{M_3}(O|t, o, \theta_{M_3,1}) &= ASC_{M_3,O} + \sum_{j=1}^{K_{M_3}} I_{M_3,O,j} \theta_{M_3,1,j} \sum_{k=1}^{188} I_{M_3,j,k} \chi_{k,t,o} , \\
V_{M_3}(O|t, o, \theta_{M_3,1}) &= ASC_{M_3,DK} . \tag{4.18}
\end{aligned}$$

The general description of the utilities is exactly the same as for the utilities in Equation (4.9). The detailed specifications of  $\{V_{M_3}(i|t, o, \theta_{M_3,1})\}$  are presented in Tables B.8 and B.9. Note that a dynamic formulation, as presented in Equation (4.11), has been tested in the expression utilities. It did not appear to be relevant, certainly due to the fact that the dynamics is already accounted for, by the consideration of the two phases. A logit form is postulated for  $P_{M_3}(i|t, o, \theta_{M_3,1})$

$$P_{M_3}(i|t, o, \theta_{M_3,1}) = \frac{e^{V_{M_3}(i|t, o, \theta_{M_3,1})}}{\sum_j e^{V_{M_3}(j|t, o, \theta_{M_3,1})}} . \tag{4.19}$$

The second model  $P_{M_3}(t|o, \theta_{M_3,2})$  captures the change of phases. A utility  $V_{M_3}(t|o, \theta_{M_3,2})$

is associated to each frame  $\mathbf{t}$  in the video  $\mathbf{o}$

$$V_{M_3}(\mathbf{t}|\mathbf{o}, \theta_{M_3,2}) = \sum_{k=1}^{K_{M_3,2}^y} \theta_{M_3,2,k}^y \sum_{k=1}^{188} I_{M_3,2,j,k}^y \mathbf{y}_{k,t,\mathbf{o}}, \quad (4.20)$$

where  $K_{M_3,2}^y$  is the number of parameters associated to this model. The specification of  $V_{M_3}(\mathbf{t}|\mathbf{o}, \theta_{M_3,2})$  is generic.  $I_{M_3,2,j,k}^y$  is an indicator equal to 1 if  $\theta_{M_3,2,k}^y$  is associated to  $\mathbf{y}_{k,t,\mathbf{o}}$ , 0 otherwise.  $\theta_{M_3,2,k}^y$  is linked to only one  $\mathbf{y}_{k,t,\mathbf{o}}$ , we have

$$\sum_{k=1}^{188} I_{M_3,2,j,k}^y = 1 \quad \forall j. \quad (4.21)$$

The model contains only  $\{\mathbf{y}_{k,t,\mathbf{o}}\}$ ,  $\{\mathbf{z}_{k,t,\mathbf{o}}\}$  have been tested but do not appear to be significant.  $\{\mathbf{y}_{k,t,\mathbf{o}}\}$  measure more drastic changes in the face compared to  $\{\mathbf{z}_{k,t,\mathbf{o}}\}$  (see Section 4.6.3). The detailed specifications of the utilities are presented in Table B.10. Finally,  $P_{M_3}(\mathbf{t}|\mathbf{o}, \theta_{M_3,2})$  is a logit model

$$P_{M_3}(\mathbf{t}|\mathbf{o}, \theta_{M_3,2}) = \frac{e^{V_{M_3}(\mathbf{t}|\mathbf{o}, \theta_{M_3,2})}}{\sum_{\ell=1}^{T_0} e^{V_{M_3}(\ell|\mathbf{o}, \theta_{M_3,2})}}, \quad (4.22)$$

and the log-likelihood function is

$$\begin{aligned} \mathcal{L}(\theta_{M_3}) &= \sum_{\mathbf{o}=1}^O \sum_{\mathbf{i}=1}^9 w_{\mathbf{i},\mathbf{o}} \log P_{M_3}(\mathbf{i}|\mathbf{o}, \theta_{M_3}) \\ &= \sum_{\mathbf{o}=1}^O \sum_{\mathbf{i}=1}^9 w_{\mathbf{i},\mathbf{o}} \log \left( \sum_{\mathbf{t}=1}^{T_0} P_{M_3}(\mathbf{t}|\mathbf{o}, \theta_{M_3,2}) \frac{1}{T_0 - \mathbf{t} + 1} \sum_{\mathbf{k}=\mathbf{t}}^{T_0} P_{M_3}(\mathbf{i}|\mathbf{k}, \mathbf{o}, \theta_{M_3,1}) \right). \end{aligned} \quad (4.23)$$

#### 4.4.4 Models with panel effect

The models presented in Sections 4.4.1, 4.4.2 and 4.4.3 do not account for the correlation between labels obtained through the internet survey. In this section, we assume that the labels are correlated through the filmed subject. Other panel structures have been tested (over respondents and videos) but this one appears to be the most relevant. Two models are developed based on the **reduced** and **latent** models.

##### The reduced model with panel effect

This is a direct extension of the **reduced model** presented in Section 4.4.2. The utilities shown in equation 4.4 become

$$\begin{aligned}
V_{M_4}(H|T_o, o, \theta_{M_4}, \varepsilon_{M_4,s}) &= ASC_{M_4,H} + \sum_{j=1}^{K_{M_4}} I_{M_4,H,j} \theta_{M_4,j} \sum_{k=1}^{188} I_{M_4,j,k} \chi_{k,T_o,o} + \sum_{s=1}^{17} I_{o,s} \varepsilon_{M_4,s}, \\
V_{M_4}(SU|T_o, o, \theta_{M_4}, \varepsilon_{M_4,s}) &= ASC_{M_4,SU} + \sum_{j=1}^{K_{M_4}} I_{M_4,SU,j} \theta_{M_4,j} \sum_{k=1}^{188} I_{M_4,j,k} \chi_{k,T_o,o} + \sum_{s=1}^{17} I_{o,s} \varepsilon_{M_4,s}, \\
V_{M_4}(F|T_o, o, \theta_{M_4}, \varepsilon_{M_4,s}) &= ASC_{M_4,F} + \sum_{j=1}^{K_{M_4}} I_{M_4,F,j} \theta_{M_4,j} \sum_{k=1}^{188} I_{M_4,j,k} \chi_{k,T_o,o} + \sum_{s=1}^{17} I_{o,s} \varepsilon_{M_4,s}, \\
V_{M_4}(D|T_o, o, \theta_{M_4}, \varepsilon_{M_4,s}) &= ASC_{M_4,D} + \sum_{j=1}^{K_{M_4}} I_{M_4,D,j} \theta_{M_4,j} \sum_{k=1}^{188} I_{M_4,j,k} \chi_{k,T_o,o} + \sum_{s=1}^{17} I_{o,s} \varepsilon_{M_4,s}, \\
V_{M_4}(SA|T_o, o, \theta_{M_4}, \varepsilon_{M_4,s}) &= ASC_{M_4,SA} + \sum_{j=1}^{K_{M_4}} I_{M_4,SA,j} \theta_{M_4,j} \sum_{k=1}^{188} I_{M_4,j,k} \chi_{k,T_o,o} + \sum_{s=1}^{17} I_{o,s} \varepsilon_{M_4,s}, \\
V_{M_4}(A|T_o, o, \theta_{M_4}, \varepsilon_{M_4,s}) &= ASC_{M_4,A} + \sum_{j=1}^{K_{M_4}} I_{M_4,A,j} \theta_{M_4,j} \sum_{k=1}^{188} I_{M_4,j,k} \chi_{k,T_o,o} + \sum_{s=1}^{17} I_{o,s} \varepsilon_{M_4,s}, \\
V_{M_4}(N|T_o, o, \theta_{M_4}, \varepsilon_{M_4,s}) &= 0, \\
V_{M_4}(O|T_o, o, \theta_{M_4}, \varepsilon_{M_4,s}) &= ASC_{M_4,O} + \sum_{j=1}^{K_{M_4}} I_{M_4,O,j} \theta_{M_4,j} \sum_{k=1}^{188} I_{M_4,j,k} \chi_{k,T_o,o} + \sum_{s=1}^{17} I_{o,s} \varepsilon_{M_4,s}, \\
V_{M_4}(DK|T_o, o, \theta_{M_4}, \varepsilon_{M_4,s}) &= ASC_{M_4,DK} + \sum_{s=1}^{17} I_{o,s} \varepsilon_{M_4,s}, \tag{4.24}
\end{aligned}$$

where  $\varepsilon_{M_4,s}$  is an error term capturing the correlation between observations associated to the filmed subject  $s$ . It is supposed normally distributed,  $\varepsilon_{M_4,s} \sim \mathcal{N}(0, \sigma_{M_4})$ .  $I_{o,s}$  is an indicator equal to 1 if the subject  $s$  appears in video  $o$ , 0 otherwise. The probability of choosing the expression  $i$  is

$$P_{M_4}(i|o, \theta_{M_4}, \varepsilon_{M_4,s}) = \frac{e^{V_{M_4}(i|T_o,o,\theta_{M_4},\varepsilon_{M_4,s})}}{\sum_{j=1}^9 e^{V_{M_4}(j|T_o,o,\theta_{M_4},\varepsilon_{M_4,s})}}. \tag{4.25}$$

Then, for the calculation of the log-likelihood, we have to integrate on  $\varepsilon_{M_4,s}$

$$\mathcal{L}(\theta_{M_4}, \sigma_{M_4}) = \sum_{s=1}^{17} \log \left( \int \left( \prod_{o=1}^O \prod_{i=1}^9 P_{M_4}(i|o, \theta_{M_4}, \varepsilon_{M_4,s})^{w_{i,o} I_{o,s}} \right) f(\varepsilon_{M_4,s}) d\varepsilon_{M_4,s} \right), \tag{4.26}$$

where  $f(\varepsilon_{M_4,s})$  is the probability density function (pdf) of  $\varepsilon_{M_4,s}$ .

### The latent model with panel effect

This model generalizes the model proposed in Section 4.4.2. The utilities introduced in equation 4.11 are reformulated

$$\begin{aligned}
V_{M_5}(H|t, \mathbf{o}, \theta_{M_5,1}, \alpha_{M_5,H}, \varepsilon_{M_5,s}) &= V_{M_5}^s(H|t, \mathbf{o}, \theta_{M_5,1}) + \sum_{s=1}^{17} I_{\mathbf{o},s} \varepsilon_{M_5,s} \\
&\quad + \alpha_{M_5,H} V_{M_5}^s(H|t-1, \mathbf{o}, \theta_{M_5,1}), \\
V_{M_5}(SU|t, \mathbf{o}, \theta_{M_5,1}, \alpha_{M_5,SU}, \varepsilon_{M_5,s}) &= V_{M_5}^s(SU|t, \mathbf{o}, \theta_{M_5,1}) + \sum_{s=1}^{17} I_{\mathbf{o},s} \varepsilon_{M_5,s}, \\
V_{M_5}(F|t, \mathbf{o}, \theta_{M_5,1}, \alpha_{M_5,F}, \varepsilon_{M_5,s}) &= V_{M_5}^s(F|t, \mathbf{o}, \theta_{M_5,1}) + \sum_{s=1}^{17} I_{\mathbf{o},s} \varepsilon_{M_5,s} \\
&\quad + \alpha_{M_5,F} V_{M_5}^s(F|t-1, \mathbf{o}, \theta_{M_5,1}), \\
V_{M_5}(D|t, \mathbf{o}, \theta_{M_5,1}, \alpha_{M_5,D}, \varepsilon_{M_5,s}) &= V_{M_5}^s(D|t, \mathbf{o}, \theta_{M_5,1}) + \sum_{s=1}^{17} I_{\mathbf{o},s} \varepsilon_{M_5,s}, \\
V_{M_5}(SA|t, \mathbf{o}, \theta_{M_5,1}, \alpha_{M_5,SA}, \varepsilon_{M_5,s}) &= V_{M_5}^s(SA|t, \mathbf{o}, \theta_{M_5,1}) + \sum_{s=1}^{17} I_{\mathbf{o},s} \varepsilon_{M_5,s} \\
&\quad + \alpha_{M_5,SA} V_{M_5}^s(SA|t, \mathbf{o}, \theta_{M_5,1}), \\
V_{M_5}(A|t, \mathbf{o}, \theta_{M_5,1}, \alpha_{M_5,A}, \varepsilon_{M_5,s}) &= V_{M_5}^s(A|t, \mathbf{o}, \theta_{M_5,1}) + \sum_{s=1}^{17} I_{\mathbf{o},s} \varepsilon_{M_5,s}, \\
V_{M_5}(N|t, \mathbf{o}, \theta_{M_5,1}, \alpha_{M_5,N}, \varepsilon_{M_5,s}) &= V_{M_5}^s(N|t, \mathbf{o}, \theta_{M_5,1}) = 0, \\
V_{M_5}(O|t, \mathbf{o}, \theta_{M_5,1}, \alpha_{M_5,O}, \varepsilon_{M_5,s}) &= V_{M_5}^s(O|t, \mathbf{o}, \theta_{M_5,1}) + \sum_{s=1}^{17} I_{\mathbf{o},s} \varepsilon_{M_5,s} \\
&\quad + \alpha_{M_5,O} V_{M_5}^s(O|t, \mathbf{o}, \theta_{M_5,1}), \\
V_{M_5}(DK|t, \mathbf{o}, \theta_{M_5,1}, \alpha_{M_5,DK}, \varepsilon_{M_5,s}) &= V_{M_5}^s(DK|t, \mathbf{o}, \theta_{M_5,1}) + \sum_{s=1}^{17} I_{\mathbf{o},s} \varepsilon_{M_5,s} \tag{4.27}
\end{aligned}$$

where  $\varepsilon_{M_5,s}$  is an error term capturing the correlation between observations implicating the same filmed subject  $s$ .  $\varepsilon_{M_5,s}$  is supposed normally distributed,  $\varepsilon_{M_5,s} \sim N(0, \sigma_{M_5})$ . Note that  $\{V_{M_5}^s(i|t, \mathbf{o}, \theta_{M_5,1})\}$  are free of the error components, so there is no double counting of the error terms. The probability of choosing the expression  $i$ , within frame  $t$  of video  $\mathbf{o}$  is

$$P_{M_5}(i|t, \mathbf{o}, \theta_{M_5,1}, \alpha_{M_5}) = \frac{e^{V_{M_5}(i|t, \mathbf{o}, \theta_{M_5,1}, \alpha_{M_5,i}, \varepsilon_{M_5,s})}}{\sum_{j=1}^9 e^{V_{M_5}(j|t, \mathbf{o}, \theta_{M_5,1}, \alpha_{M_5,j}, \varepsilon_{M_5,s})}}, \tag{4.28}$$

and the probability of choosing the expression  $\mathbf{i}$  for video  $\mathbf{o}$  is

$$P_{M_5}(\mathbf{i}|\mathbf{o}, \theta_{M_5}, \alpha_{M_5}, \varepsilon_{M_5,s}) = \sum_{t=1}^{T_o} P_{M_5}(\mathbf{i}|t, \mathbf{o}, \theta_{M_5,1}, \alpha_{M_5}) P_{M_5}(t|\mathbf{o}, \theta_{M_5,2}), \quad (4.29)$$

where  $P_{M_5}(t|\mathbf{o}, \theta_{M_5,2})$  is the influence of the frame  $t$  of video  $\mathbf{o}$  on the choice of label. It is the same than for the **latent model** (equation 4.14). The calculation of the log-likelihood function requires to integrate on  $\varepsilon_{M_5,s}$

$$\mathcal{L}(\theta_{M_5}, \alpha_{M_5}, \sigma_{M_5}) = \sum_{s=1}^{17} \log \left( \int \left( \prod_{o=1}^O \prod_{i=1}^9 P_{M_5}(\mathbf{i}|\mathbf{o}, \theta_{M_5}, \alpha_{M_5}, \varepsilon_{M_5,s})^{w_{i,o} I_{o,s}} \right) f(\varepsilon_{M_5,s}) d\varepsilon_{M_5,s} \right), \quad (4.30)$$

where  $f(\varepsilon_{M_5,s})$  is the pdf of the normal distribution  $N(0, \sigma_{M_5})$ .

## 4.5 Intermediary specification steps

The five models presented in this chapter are the results of a long modeling process, where intermediary models have been generated. We started from the work of Sorci, Antonini, Cruz, Robin, Bierlaire and Thiran (2010) and first tried to apply their model. The results were not satisfactory as the face positions and scales were not the same in the FEED and Cohn-Kanade databases as explained in Section 4.4.1. The model has not only been re-estimated, it has also been adapted due to the small number of observations available in this analysis. The obtained model is the **reduced model**, which is a strong basis for the development of the four other models. In the other proposed models, the model managing the expression perception are extensions of this **reduced model**.

Then, we focus on the **latent model** (Section 4.4.1). In a first step, the model did not include any dynamics (Equation (4.11)). It came after the study of HMM and the dynamic formulation appeared to be meaningful in our work. Note that the accounting for previous probabilities instead of previous utilities has also been tested, but it appeared to be very heavy to manipulate. The incorporation of  $\{z_{k,t,o}\}$  (Equation (4.2)) in the utility of surprise (Equation (4.9)) is consecutive to an analysis of the observed labels. Respondents have tendency to answer “surprise” when they perceive suddenness, and  $\{z_{k,t,o}\}$  are well adapted to reflect it. Several specifications have been tested for the model managing the frame influence. We began with a model giving equal probabilities to all the frames, but the estimated results were not good. Regarding the frame utilities (Equation (4.13)), we started by incorporating only  $\{x_{k,t,o}\}$  (Section 4.3), but parameters were not significant. We continued with models integrating only  $\{y_{k,t,o}\}$  (Equation (4.1)), in order to account for the facial

changes, which made a lot of sense. Finally we refined the model using both  $\{y_{k,t,o}\}$  and  $\{z_{k,t,o}\}$  in order to capture more in details the perception of the changes.

The study of the **latent model** predictions shows an instability of the model in case of several impressive frames in the video, presenting different expressions. The improvement of this model passed by a smoothing of its behavior, that's why we introduced the **smoothed model** (see Section 4.4.3). This new model is relevant for the forecasting, because it is less subject to tiny fluctuations of facial descriptors, which can appear in noisy data. The dynamic formulation of the utilities has also been tested in the **smoothed model**, but it did not appear to improve it. It is certainly due to the fact that the dynamics is already accounted for, with the assumption about the two behavioral phases (see Section 4.4.4). Several utility specifications of the model managing the phase changing have been tested (see Equation (4.20)). Only  $\{z_{k,t,o}\}$  appeared to be significant, certainly due to the fact that facial changes should be drastic for passing from one behavioral phase to the other.

We decided to refine the quality of the estimates for the **reduced** and **latent models** by accounting for the correlation between the observations in the data. It was not considered in the **smoothed** model due to practical difficulties linked to its estimation, induced by the modeling complexity. We obtained the **reduced** and **latent models with panel effect**. We considered sequentially the correlation per respondents, per videos and per filmed subjects. We retained this latter, the reasons are explained in the Section 4.6.5. In this model, we additionally tried several specifications of the error term related to the panel effect. We tried to include *i.i.d.* and homoscedastic error terms in every alternatives (see Equation (4.24) and (4.27)). But the estimation results were equivalent. We kept the final specifications, because in addition of the panel effect, they capture the correlation between all the alternatives, except the neutral. This mimics a nested structure, which as a lot of sense, as the neutral is the default expression.

## 4.6 Estimation of the models

The models are estimated by maximum likelihood (see Equations (4.7), (4.16), (4.23), (4.26) and (4.31)) using the biogeme software (Bierlaire, 2003a and Bierlaire and Fetiariison, 2009). Except for the **reduced models** (with and without panel effect), these models are complex to estimate. The estimation results for the **latent** and **smoothed** models have been also obtained using codes based on biogeme. The estimation of the models with panel effect requires to perform numerical integration. A Monte-Carlo simulation with 1000 draws has been used. General estimation results are presented in Table 4.3.

### 4.6.1 The reduced model

The **Reduced model** is the simplest model because it only accounts for the influence of the last frame on the observed choice of label. The values of the 32 estimated parameters and associated t-tests are presented in Tables B.2 and B.3. Fourteen parameters are related to facial measurements characterizing AU (see Section 4.4.1). The signs are consistent with the work of Sorci, Antonini, Cruz, Robin, Bierlaire and Thiran (2010), and with the FACS (Ekman and Friesen, 1978). The asymmetry of the face is taken into account by associating different parameters to the left and right measurements of a same type.

All parameters related to AU are significantly different from 0 (t-test  $\geq 1.96$ ). This is also the case for the five parameters related to EDU and for the five parameters associated to elements of the vector  $\mathbf{C}$ . Their signs are coherent with the work of Sorci, Antonini, Cruz, Robin, Bierlaire and Thiran (2010).

Some of the eight  $\{\text{ASC}_i\}$  do not appear to be significant, which is a good feature because they are designed to absorb the unobserved perception of respondents.

### 4.6.2 The latent model

For the **latent model**, the values and associated t-tests of the 34 parameters related to the model handling with the expression perception are presented in Tables B.4 and B.5.

Signs and significance of parameters associated to AU, EDU and elements of the vector  $\mathbf{C}$  are correct and consistent with the estimated parameters obtained for the **reduced model**. In addition, the model contains two more parameters. The parameter  $\theta_{M_2,1,22}$  associated to the height of the mouth (“*mouth\_h*”), appears to be significant, while it was not the case for the **reduced model**. This is due to the fact that the **reduced model** accounts only for the perception of the last frame in a video, compared to all the frames here. So the **reduced model** could not be as precisely specified as this model.  $\theta_{M_2,1,1}^z$  is related to the variance of the height of the mouth (“*mouth\_h*”). It is positive meaning that the more the height of the mouth varies during the previous second, the more the surprise will be favored, which is logical.

Four parameters of memory effect ( $\alpha_{M_2,H}$ ,  $\alpha_{M_2,F}$ ,  $\alpha_{M_2,SA}$ ,  $\alpha_{M_2,O}$ ) appear to be significantly different from zero (see Table B.6). They have the same magnitude. Without any constraint, their estimated values are in  $[-1, 1]$  meaning that the present perception is predominant, as expected.

Seven parameters related to the model characterizing the influence of the frames are estimated significantly different from zero (see Table B.7). Six are associated to  $\{\mathbf{y}_{k,t,o}\}$  and one to  $\mathbf{z}_{2,t,o}$ , which is the variance of the distance between eyebrows (“*brow\_dist*”). Their magnitude is larger than for the parameters associated to the model of perception of the expressions. This means that the model is sensitive to small variations of features and tends to produce a sharp probability distribution



among the frames. The signs of the parameters are logical, for example  $\theta_{M_2,2,5}$  is attached to the height of the eyes (“*eye\_h*”) and is negative. This means that the more a subject has the eye closed on a frame, the more the frame has influence on the observed choice of label.

### 4.6.3 The smoothed model

For the **smoothed model**, the model dealing with the perception of the expressions contains 36 parameters (see Tables B.8 and B.9).

Signs and significance of parameters associated to AU, EDU and C parameters are the same than for the **reduced model**. The model contains 4 more parameters.  $\theta_{M_3,1,4}$  and  $\theta_{M_3,1,12}$  are respectively attached to the EDU corresponding to the fraction between the height of the eyebrows and their width (“*RAP\_brow*”), and to the fifth element of the vector C (“*C\_5*”). Both are in the utility of disgust. Compared to the **reduced model**, they appear to be significant due to the fact that we now account for the total number of frames.  $\theta_{M_3,1,1}^z$  and  $\theta_{M_3,1,2}^z$  are respectively related to the variance of the height of the mouth (“*mouth\_h*”) and the variance of the height of the left eye (“*leye\_h*”). They are included in the utility of surprise in order to capture the perception of suddenness. They are positive as expected, meaning that the higher  $z_{1,t,o}$  and  $z_{3,t,o}$  are, the more the surprise is favored, which is logical.

The model designed to detect the first frame of the relevant group of frames contains 8 parameters (see Table B.10). They are all linked with  $\{y_{k,t,o}\}$ . None of the parameters attached to  $\{z_{k,t,o}\}$  appeared to be significant. The perception of the short time variations of facial characteristics is not relevant for activating the second phase of behavior, which seems logical. The change in the facial characteristics should be more drastic, which explains why  $\{y_{k,t,o}\}$  are better adapted. As for the **latent model**, the magnitude of the parameters is larger compared to the model handling with the perception of the expressions. The interpretation remains the same as for the **latent model**.

### 4.6.4 Models with panel effect

Concerning the models with panel effect, the parameters of the **reduced model with panel effect** are shown in Tables B.11 and B.12. The parameters of the **latent model with panel effect** are presented in Tables B.13, B.14, B.15 and B.16. In both cases, the parameter values are respectively rather the same than for **reduced** and **latent** models. Their interpretations remain unchanged. The standard errors  $\sigma_{M_4}$  and  $\sigma_{M_5}$  appear to be significant, so the hypothesis of correlation between labels associated to the same filmed subject is verified.

### 4.6.5 Comparison of the five models

The final log-likelihood is improved between the **reduced** and **latent models**, and the **reduced** and **smoothed models**. The three first models can not be compared using likelihood ratio-tests. We use  $\bar{p}^2$  as a goodness of fit to identify the best model. Looking at Table 4.3 and regarding models without panel effect, the **latent model** appears to be the best model, closely followed by the **smoothed model**. The improvement brought by the dynamic modeling is substantial. This is due to the nature of the videos (see Section 4.3). At the beginning of the videos, the facial expressions are neutral, and then they evolve toward other expressions, so faces are highly expressive on the last frame of the video. This explains why the **reduced models** are working well. Nevertheless, the proposed behavioral hypothesis have sense. The assumption about one single frame triggering the choice seems to be the most relevant (**latent model**), closely followed by the assumption about the two behavioral phases (**smoothed model**). This order seems to be logical as the **latent model** focuses on a “pure” and “strong” perception, which is intuitively the most important, specially in short facial videos. Compared to this model, the **smoothed model** polishes the perceptions in the second behavioral phase. The main advantage of the **smoothed model** is to be less sensitive to data errors, compared to the **latent model**.

Regarding the models with panel effect, the log-likelihood is improved between the **reduced model** and **reduced model with panel effect**, and the **latent model** and **latent model with panel effect**. Out of the five proposed models, the **latent model with panel effect** is the best in terms of fit. The accounting for the correlations between observations related to a same filmed subject, improves significantly the fit. A correlation by respondents has been tested but did not appear to meaningful, the perception of the respondent seems to be homogeneous. This is logical as the respondents are also homogeneous in terms of socio-economic characteristics (they are mainly in Switzerland with an academic background). A correlation per videos has been tested and gave approximately the same results than for the correlation per filmed subjects. It has not been kept because this model was very heavy to manipulate. As there is more videos than filmed subjects, we increased the number of draws which increased dramatically the estimation time, making the cross-validation (see Section 4.7.2) unrealizable.

The magnitude of the parameter values and signs are the same for the five models. For example,  $\theta_{M_1,4}$ ,  $\theta_{M_2,1,4}$ ,  $\theta_{M_3,1,5}$ ,  $\theta_{M_4,4}$  and  $\theta_{M_5,1,4}$  are associated to the opening of the mouth (“RAP\_mouth”), defined as the fraction between the height of the mouth (“*mouth\_h*”) and the width of the mouth (“*mouth\_w*”). They are present in the utilities of surprise and fear. The associated parameters are all positive, showing the stability of the models. Their positive sign is logical because when a person has the mouth opened, the perceived facial expression is more likely to be fear or surprise.

The specifications of the model related to the detection of the most impressive frame in the **latent models** (with and without panel effect), and to the detection of the first frame of the relevant group of frames in the **smoothed model**, are very

	Reduced	Latent	Smoothed	Reduced panel	Latent panel
Nb obs.	369	369	369	369	369
Nb param.	32	45	44	33	46
Null $\mathcal{L}$	-810.78	-810.78	-810.78	-810.78	-810.78
Final $\mathcal{L}$	-475.79	-441.28	-447.67	-470.26	-435.14
$\bar{\rho}^2$	0.374	0.400	0.394	0.379	0.406

Table 4.3: General estimation results

similar. For the **latent models**, it contains parameters associated with both  $\{\mathbf{y}_{k,t,o}\}$  and  $\{\mathbf{z}_{k,t,o}\}$  and for the **smoothed model**, only associated with  $\{\mathbf{y}_{k,t,o}\}$ . For example,  $\mathbf{y}_{2,t,o}$  is present in both models and is related to the height of the mouth (“*mouth\_h*”). Figure 4.11 displays the variation of this feature among frames of a video which are displayed at the top. The sign of the parameters associated to  $\mathbf{y}_{2,t,o}$  ( $\theta_{M_2,2,6}$ ,  $\theta_{M_5,2,6}$  and  $\theta_{M_3,2,8}$ ) is positive for both **latent** and **smoothed models**, which is logical. The higher the difference of mouth height between two consecutive frames, the more important the second frame is. In that special case and regarding only  $\mathbf{y}_{2,t,o}$ , frame 3 seems to be the most important.

In conclusion, the parameters of the models are significant and interpretable. Moreover, the addition of a dynamic part in the models significantly improves the fit. The accounting of the panel effect is successful, as the latent model with panel effect has the best fit.

## 4.7 Prediction capability

The prediction capability is tested in order to ensure the quality of the models. The dataset used in this section is the same as the one used for the estimation (see Section 4.6). We proceed in three steps: the first one consists of comparing the percentages of badly predicted observations for the proposed models. In a second step, the models are validated using the method of cross-validation. In the third step, we study the predictions of the proposed models at a more disaggregated level. This consists of picking a certain video and analyzing the predictions of the models in detail.

### 4.7.1 Aggregate prediction

An observation is considered as badly predicted, if its forecasted choice probability is less than  $\frac{1}{9}$ , which corresponds to the probability predicted by a uniform probability on the number of alternatives. Table 4.4 summarizes the percentages of badly predicted observations per model. The percentages are consistent with the fitting results presented in Section 4.6, which is a good sign. The percentage of badly predicted

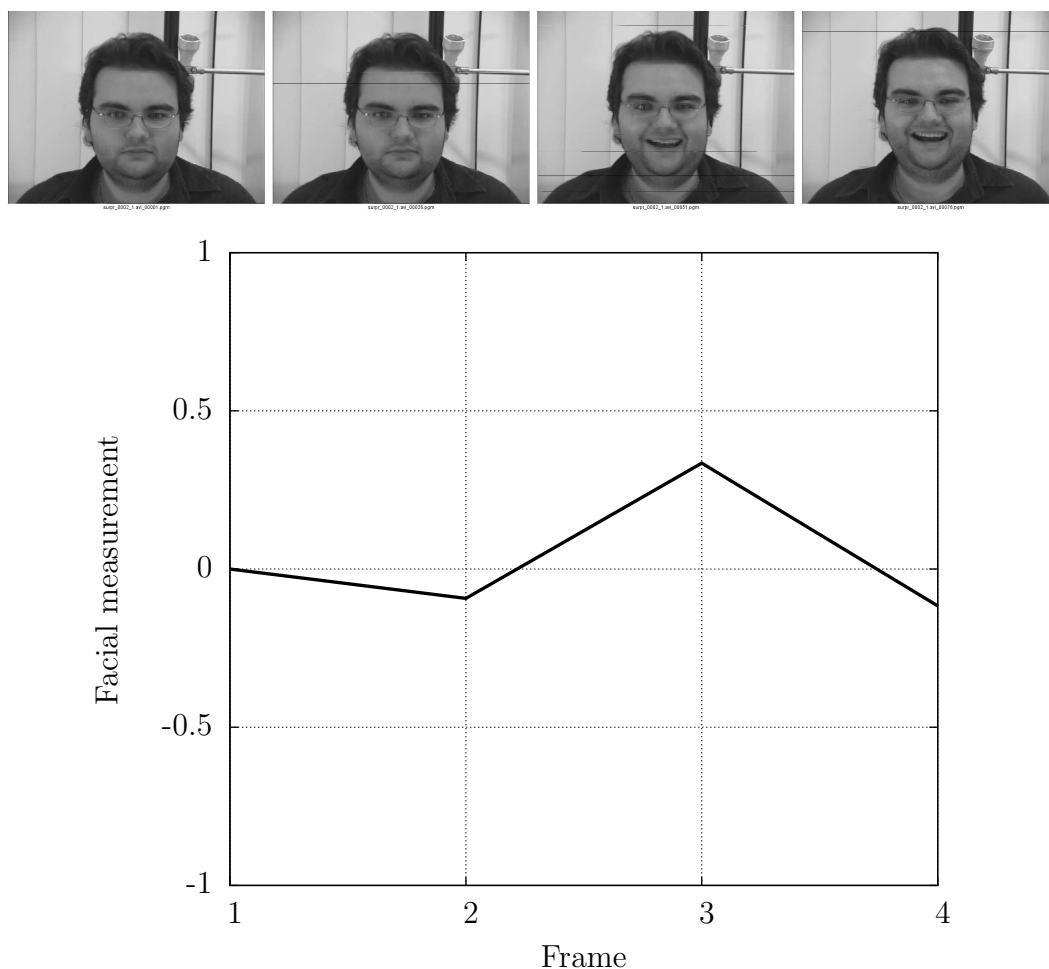


Figure 4.11: Examples of the variations of  $y_{2,t,o}$ , associated to the height of the mouth (“*mouth\_h*”), for a video

observations is already low for the **reduced model**. The improvement brought by the **latent** and **smoothed models** compared to the **reduced model** is minor in terms of prediction. This can be explained by the structure of the considered facial videos. As the “peak” emotion is often observed at the end of the video, there are few observations where the dynamic models could do better. However the **latent model with panel effect** is the best.

The cumulative distributions of the choice probabilities predicted by the models are displayed in Figure 4.12. If the models were perfect, the curves should be flat with a pick for choice probabilities equal to one. This would mean that the models replicate exactly the observed choices of labels. Of course this is not the case. The five curves are close in the “badly predicted” interval (choice probabilities less than  $\frac{1}{9} = 0.11$ ). This is consistent with the results shown in Table 4.4. In the interval  $[0, 0.78]$  the **latent model with panel effect** is the best. In the last interval, it is

Reduced	Latent	Smoothed	Reduced panel	Latent panel
17.89	17.34	15.45	18.43	14.45

Table 4.4: Percentages of badly predicted observations on the estimation data

the **latent model**. This model predicts the highest probabilities (its curve is the last to reach the level of one). The **smoothed model**, is better than the **reduced model** except on  $[0.68, 1]$ . Moreover the **latent model with panel effect** is always better than the **reduced models** (with and without panel effect), which demonstrates the added value of the dynamic modeling.

#### 4.7.2 Cross-validation

The study of the badly predicted observations, described in Section 4.7.1 is done on the estimation data presented in Section 4.3. The finality of the models is to be used on some data not involved in the estimation process, for prediction. Consequently the quality of the model should be tested on some new data, but we do not dispose of such data. In this situation, the cross-validation allows to validate the models. The methodology is inspired from the work of Robin et al. (2009) who successfully cross-validate a model of pedestrian behavior. The dataset is split into an estimation subset and a validation subset. The dataset is randomly split across the videos, in five subsets. Each subset contains twenty percent of the videos. In the data, there are 65 videos, so each subset contains the collected labels related to 13 videos. Four subsets are combined into the estimation dataset. After estimation, the model is applied on the remaining subset. The operation is repeated five times. The percentages of badly predicted observations, calculated over the validation subsets are presented in Table 4.5.

Validation subsets	1	2	3	4	5
<b>Reduced</b>	28.74	26.15	21.31	21.87	28.26
<b>Latent</b>	24.14	13.85	11.48	17.19	21.74
<b>Smoothed</b>	20.69	16.92	18.03	15.63	10.87
<b>Reduced panel</b>	28.73	26.15	22.95	23.43	28.26
<b>Latent panel</b>	28.70	15.38	21.29	*	35.87

Table 4.5: Percentages of badly predicted observations calculated over the validation subsets, obtained when cross-validating the models

Regarding models without panel effect, the two dynamic models (the **latent** and **smoothed models**) are always better than the **reduced model**. In addition, the percentages of badly predicted observations are close from those obtained on the

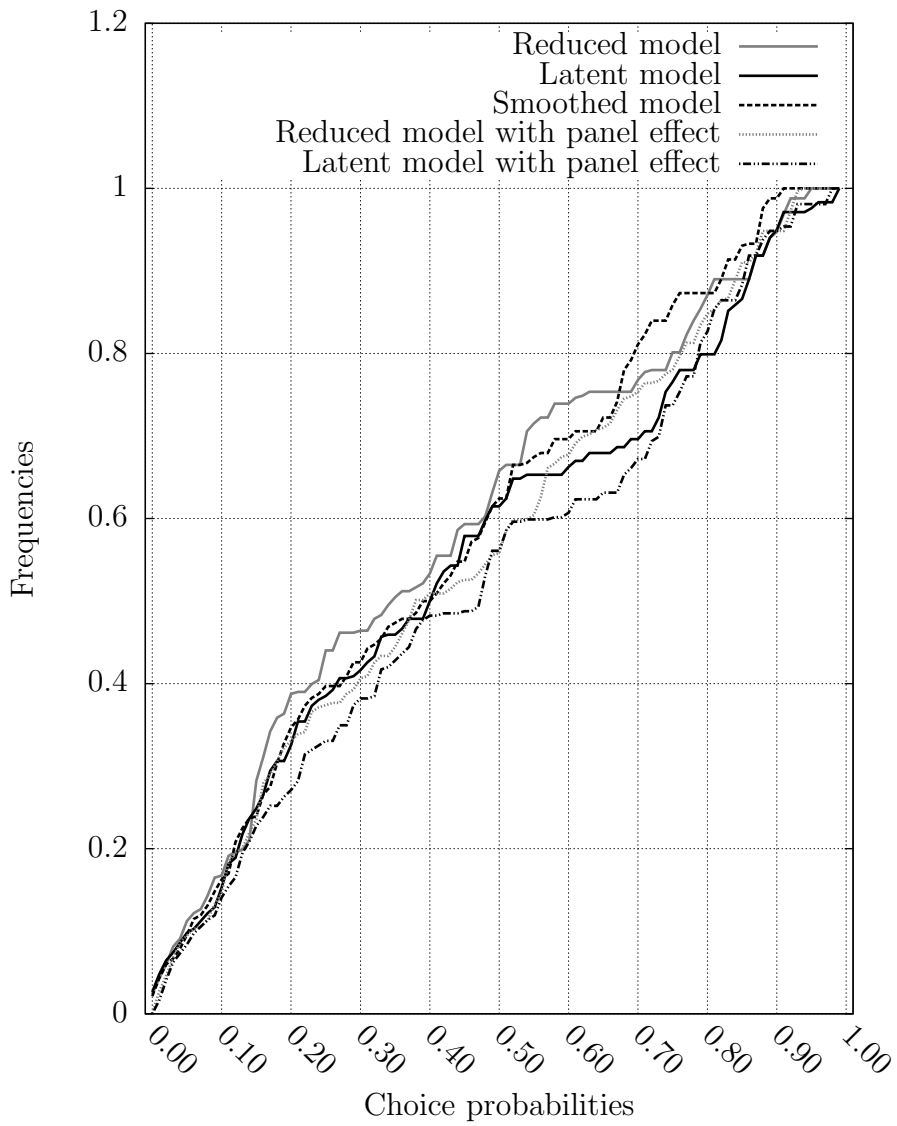


Figure 4.12: Cumulative distributions of the choice probabilities predicted by the five proposed models, on the estimation data

entire estimation data (see Table 4.4) for the **latent** and **smoothed models**, not the **reduced model**. The dynamic models appear to be much more robust than the **reduced model**. This justifies the goodness of the approach and the validity of the dynamic models.

Concerning the models with panel effect, results for the **reduced model with panel effect** are worse than for the **reduced model**. This is also the case for the **latent model with panel effect** compared to the **latent model**. Note that for experience 4, the estimation of the **latent model with panel effect** did not converge (this model is very difficult to estimate due to its complexity). We conclude that the two models with panel effect tends to over fit the data.

### 4.7.3 Disaggregate prediction

We looked at the power of prediction over the estimation dataset, at the aggregate level. The study of a particular video allows to detail precisely the predictions of the five models. The video is the same than the one considered in Figure 4.11. The detailed predictions of the models are shown in Figure 4.13 for the **reduced model**, Figure 4.14 for the **latent model**, Figure 4.15 for the **smoothed model**, Figure 4.16 for the **reduced model with panel effect** and Figure 4.17 for the **latent model with panel effect**. On these figures, each column is associated to a frame, except the extreme right. The first line displays the considered frames. As mentioned in Section 4.3, each frame is the first of a group of images corresponding to one second in a video. The second line concerns the predictions of the model associated to the perception of the expressions. For each frame, the probability distribution among the expressions is presented. The third line shows the influence of the frames. The contributions of the frames sum up to one. For the **reduced models** (with and without panel effect), only the last frame is considered relevant, so the peak is logically on this last frame. For the **latent models** (with and without panel effect), it shows the influence of each frame on the final expression choice. For the **smoothed model**, the peak measures the contribution of the average perception of the following group of frames (until the end of the video), including the frame of the peak. Finally in the extreme right column, you find on the second row the final probability distribution among the expressions, which is predicted by the model, and on the third row, the distribution of the collected labels for the video.

On the first frame of the considered video (see Figure 4.11), the face tends to be neutral, and then evolves toward a different expression. Seven respondents have labeled this video: three gave the label happiness, three gave the label surprise, and one the label anger. Anger does not seem to be appropriate for this video, but it has been kept because there was no proof of mistakes made by the respondent. In addition, the subject on the two first frames of the video could be considered angry. The observed distribution of the collected labels is displayed at the bottom right of the figures. The **reduced model** predicts 65% of happiness, 35% of surprise, and 0% for anger. The prediction seems logical regarding only the facial characteristics

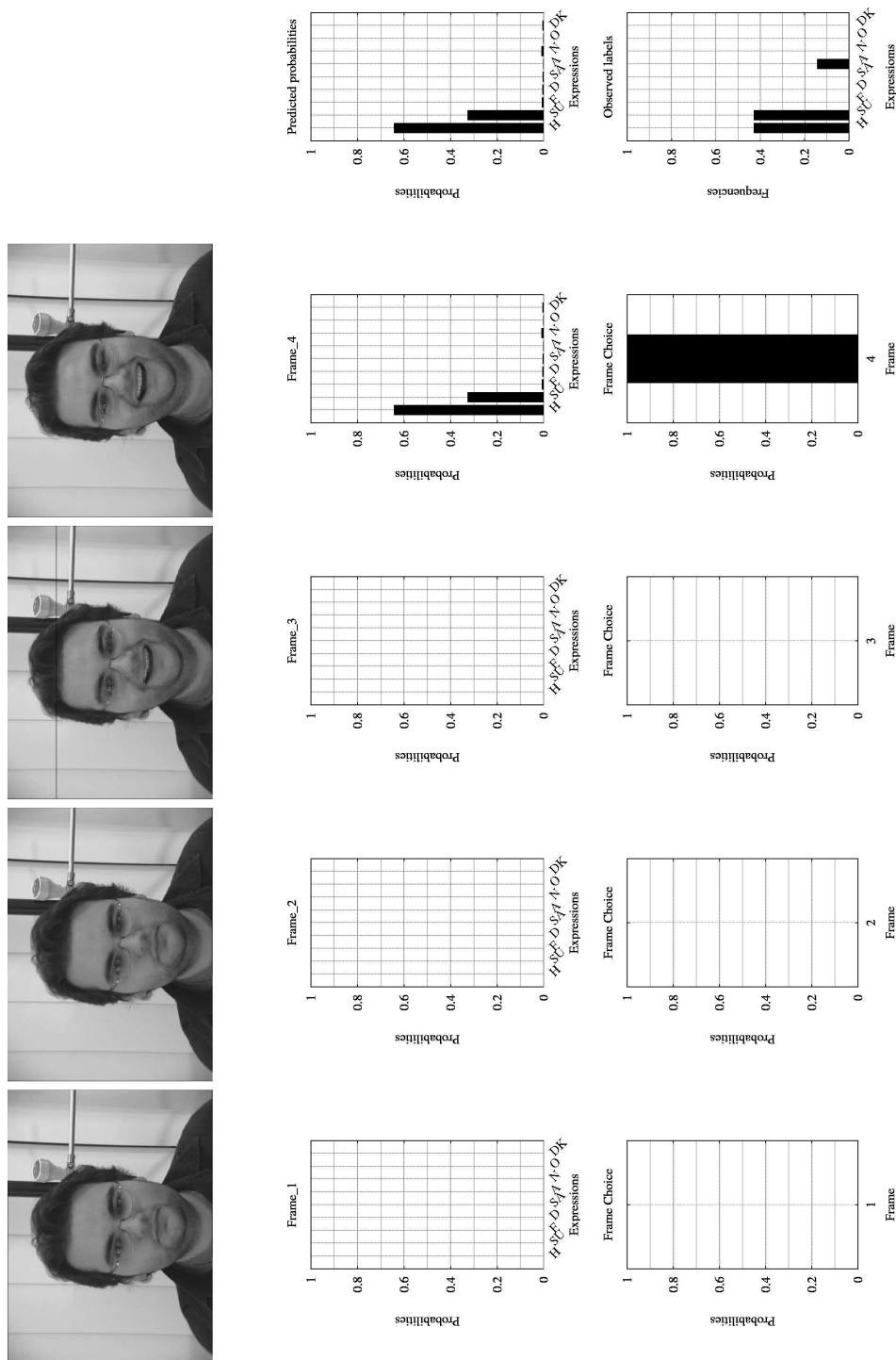


Figure 4.13: Example of a detailed prediction of the **reduced model**



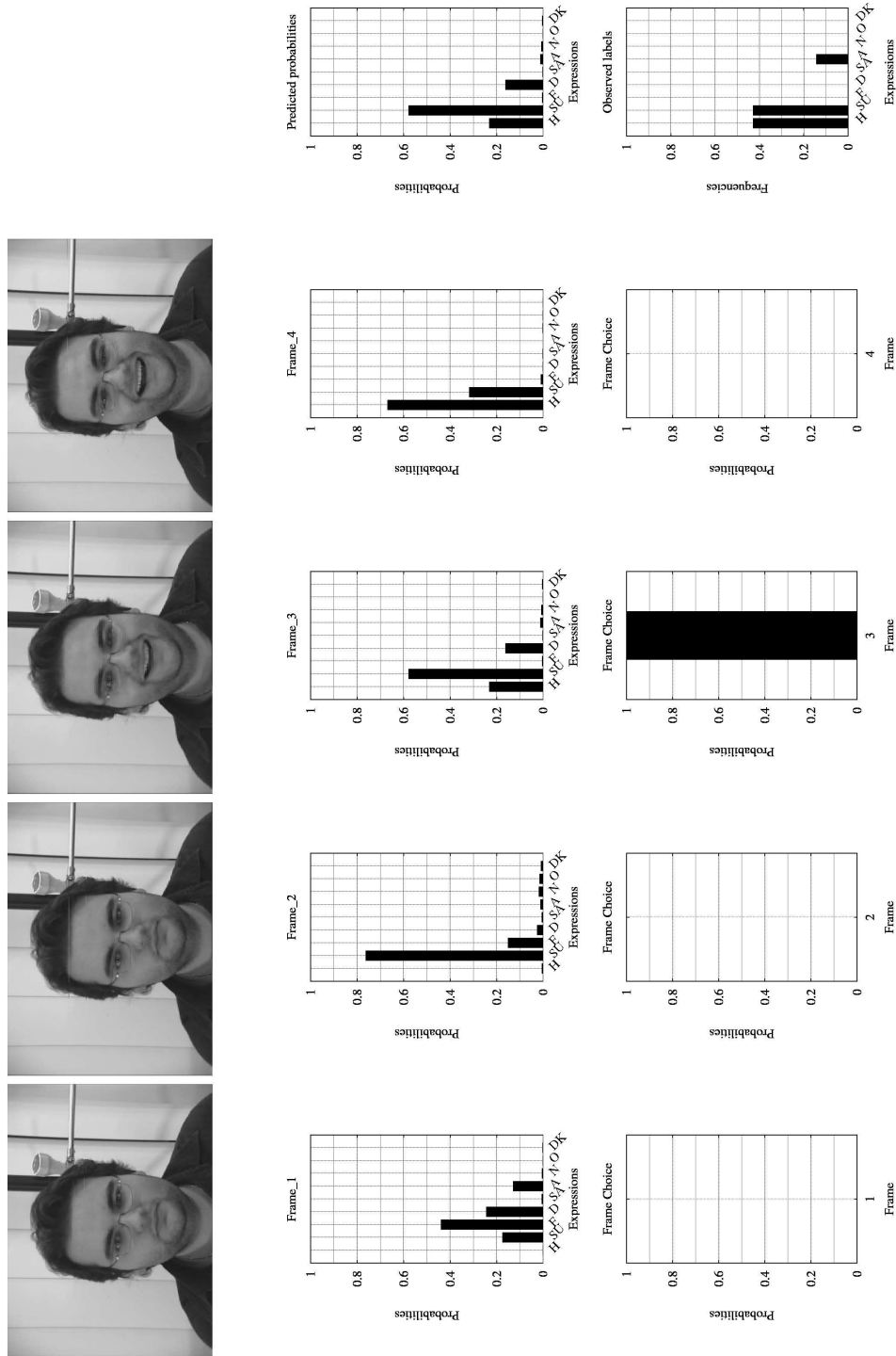


Figure 4.14: Example of detailed prediction of the latent model

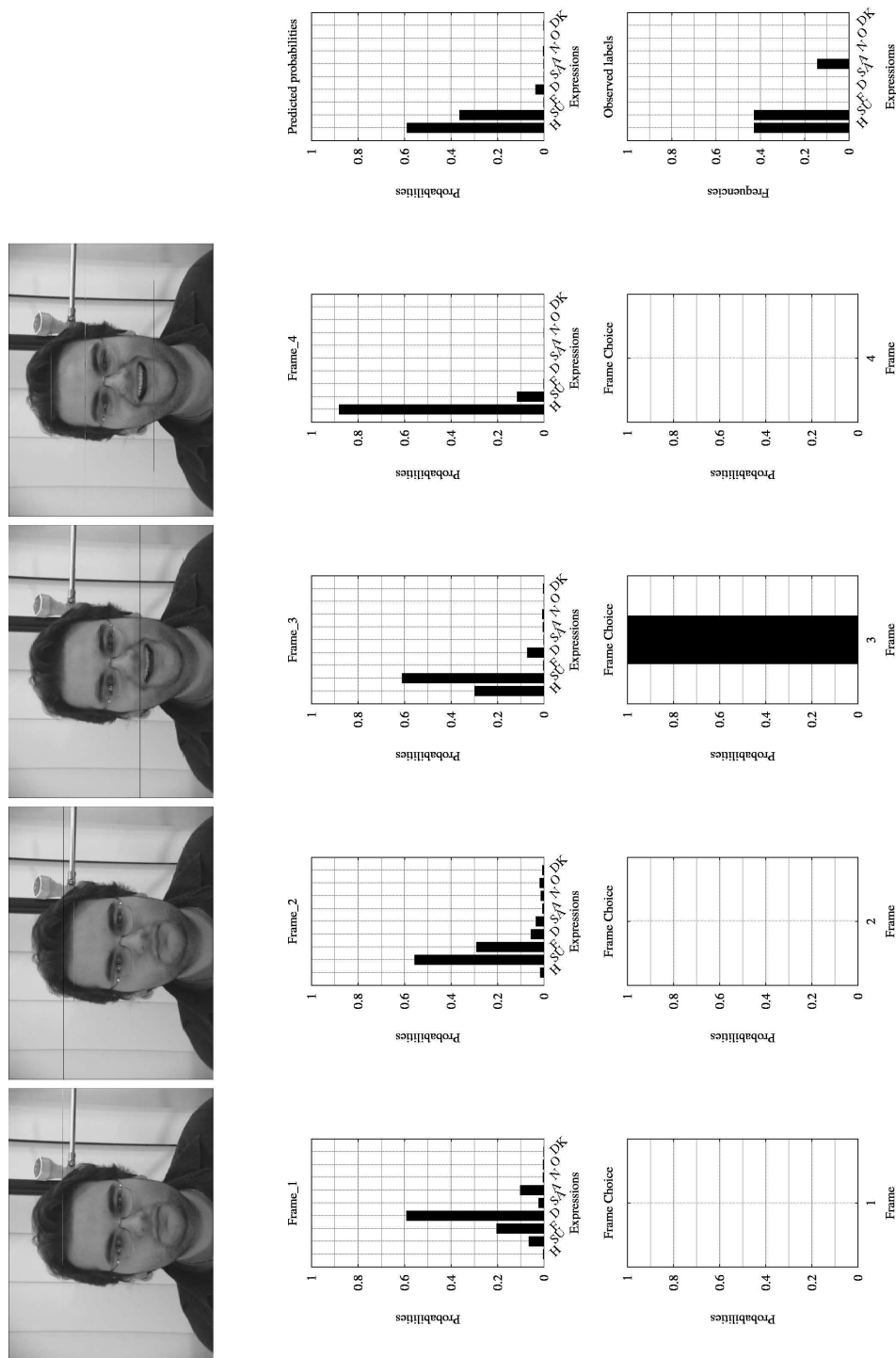


Figure 4.15: Example of detailed prediction of the smoothed model

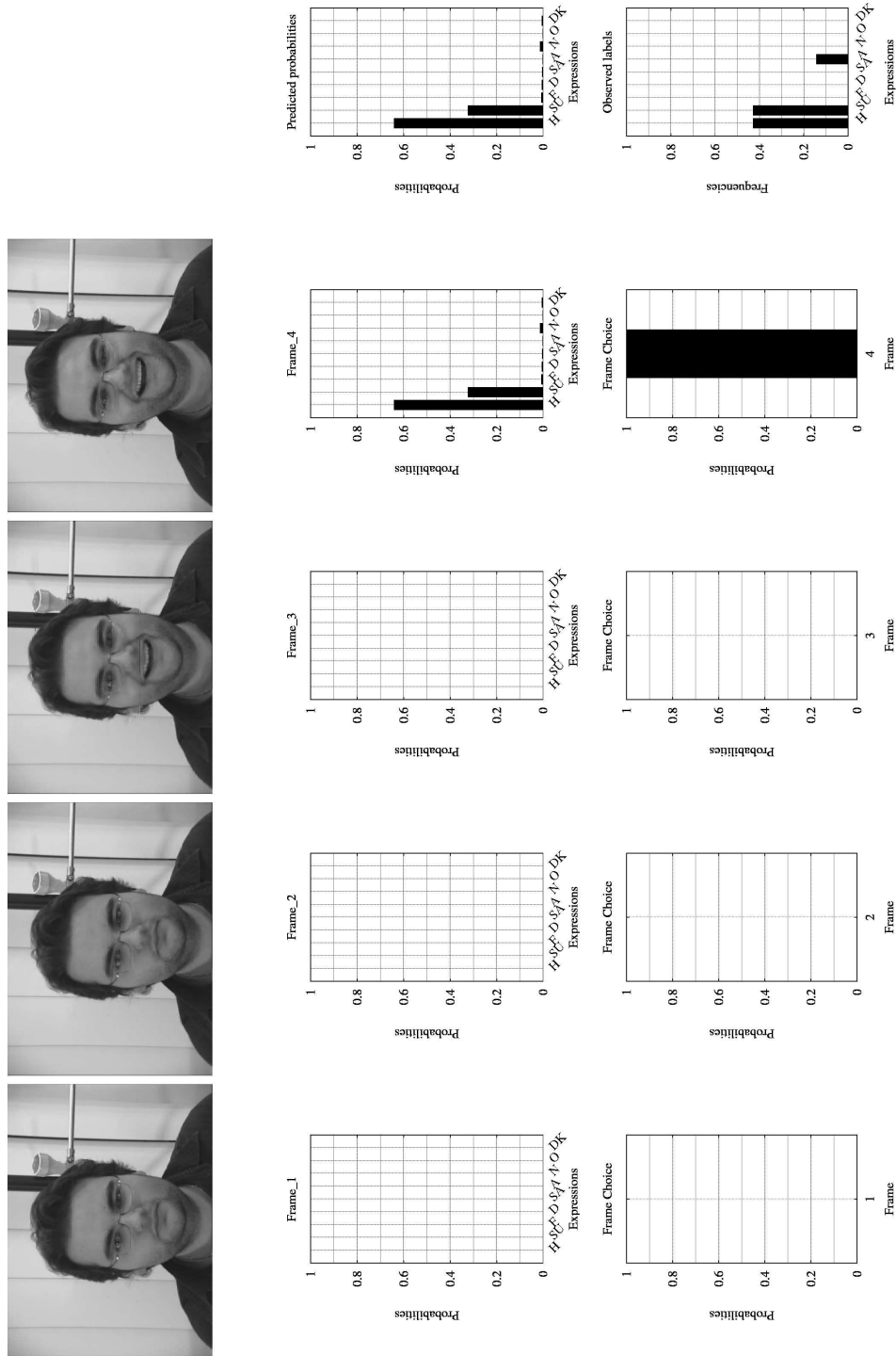


Figure 4.16: Example of detailed prediction of the reduced model with panel effect

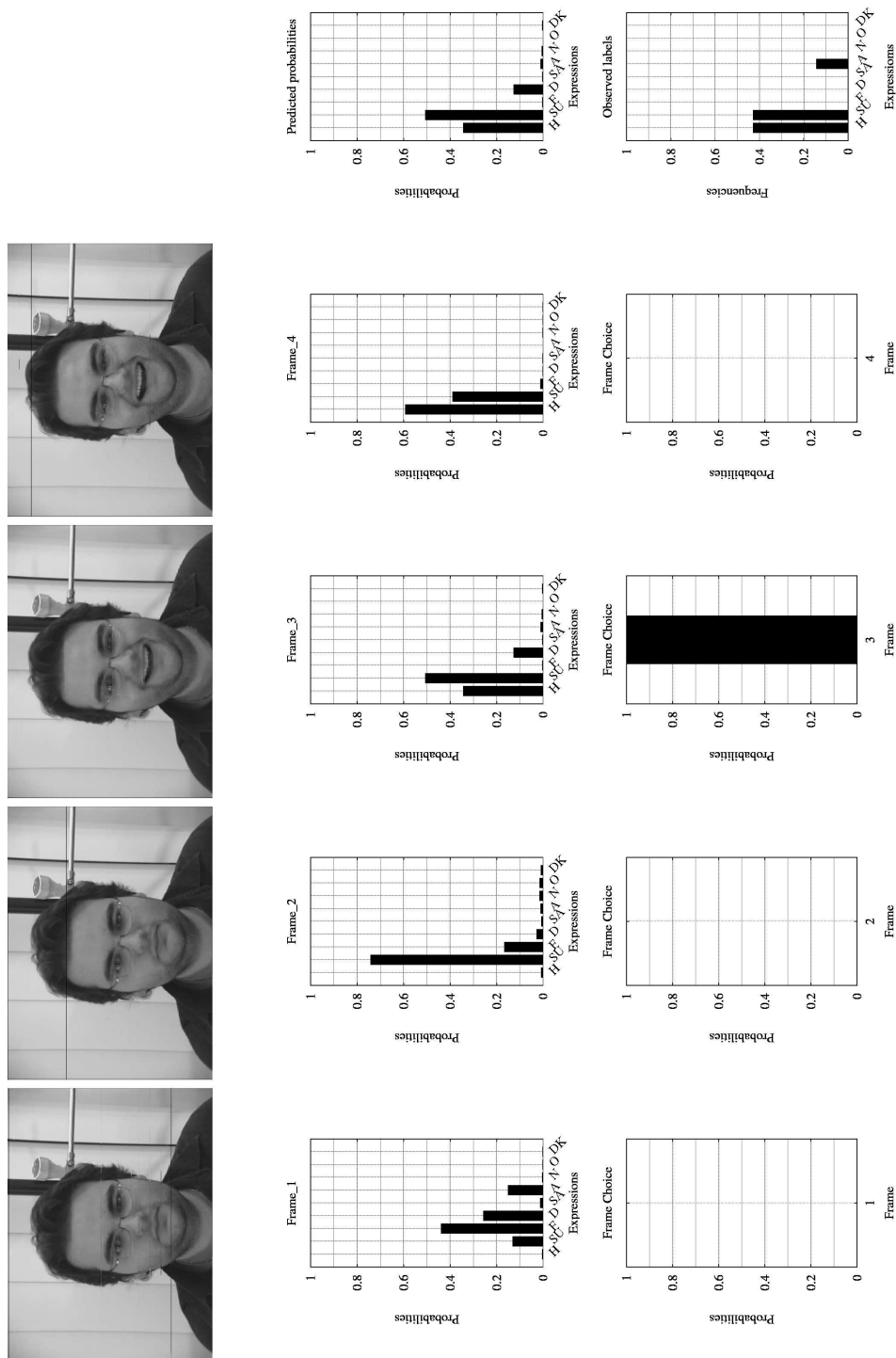


Figure 4.17: Example of detailed prediction of the Latent model with panel effect

in the last frame.

The **latent model** predicts 24% of happiness, 58% of surprise, 18% of disgust and 0% for anger. This is further away from the distribution of the collected labels, compared to the **reduced model**. The model has selected frame 3 as being the most impressive frame, with a probability almost equal to one, so the predictions of the model results only from the perception of this frame. This is logical because the utilities of the frames contain both  $\{y_{k,t,o}\}$  and  $\{z_{k,t,o}\}$  (see Section 4.4.2), and they appear to be very high for frame 3 (see Figure 4.11 for the height of the mouth). For this frame, the predicted probability of surprise is very high. This is logical, because the utility of surprise contains  $\{z_{k,t,o}\}$  (see Equation (4.9)), which account for the perception of suddenness. For this frame, the high probability for happiness is also intuitive due to the facial characteristics. The prediction of disgust does not seem to be appropriate.

The **smoothed model** predicts 58% of happiness, 38% of surprise, 4% of disgust and 0% of anger. The prediction is well adapted to the observed distribution of labels. The model detects frame 3 as being the first frame of the relevant group of frames. As for the **latent model**, this is due to the presence of  $\{y_{k,t,o}\}$  in the utilities of the frames (see Section 4.4.3), and  $\{y_{k,t,o}\}$  are high for this frame (see Figure 4.11). The model handling with the perception of the expressions predicts more surprise than happiness for frame 3, and the contrary for frame 4. This is logical due to the perception of suddenness in frame 3 (see the utility of surprise in Equation (4.18)). The facial characteristics are stabilized in frame 4 and lead to the expression happiness, which is coherent. The final prediction of the model is the average of the perception of expressions among the frames of the relevant group (frames 3 and 4), which explains the balanced share between happiness and surprise.

The results are rather the same for the **reduced model with panel effect** than for the **reduced model**. Regarding the **latent model with panel effect**, it predicts 35% of happiness, 52% of surprise, 13% of disgust and 0% of anger. The model has selected the frame 3 as being the most influential. Even if the results are quite similar compared to those obtained with the **latent model**, they are better because the difference between the predicted probabilities of happiness and surprise are smaller.

The predictions of the five models are explainable. The **smoothed model** seems to be the most interpretable. The **smoothed model** and **latent model with panel effect** predict the closest distributions of probability across the expressions, to the collected labels. The **smoothed model** over predicts happiness and under predicts surprise, contrary to the **latent model with panel effect**.

## 4.8 Contributions

The proposed work overcomes the limitations of the standard approaches in the dynamic facial expression recognition. Standard approaches consist in associating any

two examples with the same facial descriptors to the same expression. One of the main assumption, is that facial expression labels, which are in the data, stand for the true expressions (Cohen et al., 2003, Bartlett et al., 2003). But this assumption does not hold in reality, as people can perceive differently the same expression. Facial expressions are characterized by the ambiguity. In our work, this ambiguity is directly taken into account, as we have adopted a probabilistic approach. Another limitation of the previous approaches is the inability for interpreting the knowledge acquired by the systems. They are often black-boxes, where the interpretations of the links between the inputs (facial descriptors) and the output (expression), are not possible. Due to this black-box nature, it is also impossible to put knowledge in the model in order to improve it. In our proposed work, psychological concepts are translated into mathematical equations, and the maximum likelihood estimation allows to confirm (or infirm) the quality of the model, in terms of interpretation and significance of the parameters. In addition, this allows to learn the behavioral patterns contained in the data. In particular, we have quantified the concepts introduced by Ekman and Friesen (1978).

We generalize the work of Sorci, Antonini, Cruz, Robin, Bierlaire and Thiran (2010), as we worked with facial videos and not with images. More generally, this work is the first attempt for analyzing videos using discrete choice models.

Regarding discrete choice modeling, we have developed models inspired from recent works (Ben-Akiva, 2010). Original formulations have been introduced to capture the dynamics, which can be reused in other analysis. The proposed models are based on different assumptions about video perceptions. The estimation, validation and comparison of the models underlined the relevance of these assumptions. We learned that respondents have tendency to make their expression choices when watching specific frames, and we know how to select these frames with different ways. Our approach explicits and quantifies these psychological concepts.

## 4.9 Conclusions and Perspectives

We propose a new approach of the dynamic facial expression recognition. The estimation of the models is based on labels collected through respondents of an internet survey. The developed models capture up causal effects between facial characteristics and expressions. Statistical tests and model predictions have proved the quality of the models, and the added value of the dynamic formulation (the **latent models** and the **smoothed model** compared to the **reduced models**). In terms of fit, the **latent model with panel effect** is the best. The five models have been cross-validated on the estimation data, the **latent model** and the **smoothed model** appear to be more robust than the **reduced model**. The models with panel effect over fit the data. Consequently they are not worth for forecasting. Finally, some qualitative analysis of the model predictions allow to confirm the modeler's intuition about the facial video. Regarding all the analysis, the **smoothed model** seems to be the more

robust.

As such, the models can be used directly for applications. The major difficulty concerns the computation of the variables. The quality of the considered videos should be very high, in terms of definition and size of the face. The videos of the FEED database are not dedicated to transportation (the stimuli used to generate the facial expressions of the subjects were not necessarily related to the field) but they remain quite general. Some case studies have to be conducted in order to completely prove the model applicability to transportation (Denis, 2009).

In the context of “Aware” vehicles, we think of a system able to manage automatically the interior features of the car, based on the driver’s characteristics, including the facial expression. In case of dedicating the proposed model to this application, a data collection in two steps should be performed. In a first step, we can conduct a survey in a car simulator, by placing the respondents in controlled real driving situations and recording their faces. Then, the respondents are asked to perform actions using the interior car features. In a second step, the collected facial videos are labeled using a survey similar to the proposed internet survey. A model handling the choice of action taking as input the driver’s characteristics and expression can be developed. Then, in a real context, the face of a driver can be monitored with a camera and the proposed model applied.

The proposed models may be also used to analyze travelers satisfaction in public transportation (Friman and Garling, 2001). The facial expression could be used as a measure of satisfaction when conducting transportation surveys. For on-site measures, it is not worth as the facial expressions are most of the time generated by stimuli not related to transportation. The experimental design of the survey should be carefully set, in order to use adapted stimuli.

More generally, for the estimation of hybrid choice models, some indicators of the latent variables are needed. Bolduc and Alvarez-Daziano (2010) propose a hybrid choice model handling with the vehicle choice. In that case, the facial expression of the survey respondent could be used as an indicator of the two latent variables: “Environmental concern” and “Appreciation of new car features”. In addition of the rational behaviors, the latent variables are capturing the emotional states. The facial expression results from a short emotion and it could be used as a proxy of this emotion, or combined with other emotion indicators (questionnaires for example) to reveal it. Practically, in addition to the questionnaires, some well-chosen stimuli have to be shown to the survey respondents (such as short and shocking environmental documentaries, or advertisements of cars having new features), while their faces are recorded. Then, a DFER model is needed for determining the facial expressions. For this application, the facial expression does not enter in the prediction process, but it helps to reinforce the quality of the estimated model.

Finally in the marketing context, MacInnis et al. (1991) studied the ability of individuals to process the brand information from advertisements, the facial expression could enter in the inputs of the model, in addition of eye-tracking data.

Even if this new modeling framework is meaningful, some improvements could be

done. The model has been estimated on a small dataset. More observations would be useful. The number and type of videos is also a critical aspect, feature variabilities are quite low and should be increased. This would allow to have a more complete specification of the utilities. We think of using the specification proposed by Hensher (2010) for the processing of explanatory variables, which is highly relevant due to large amount of information provided on a face. In addition, more complex structures could be tested for the choice models. In the **latent** and **smoothed models**, the model handling with the detection of the most attractive frame and the first frame of the relevant group of frames, can be modified for taking into account the correlation between frames. A cross-nested logit seems to be well adapted to the frame choice, using two nests: “attractive” and “not attractive”. Each frame could belong to the two nests. The membership degrees of the frame to each nest should be defined as a function of their attractivenesses. They can not be generic, as the videos are varying from one observation to the other, as well as the associated frame set. This stands as a research topic in its own. Finally a comparison with a state of the art machine learning method, such as neural networks (NN) or hidden markov models (HMM) would be interesting.



## 5. Conclusion

We synthesize the contents of the thesis and discuss the perspectives of the present works.

### 5.1 Review of the main results

This thesis is based on a collection of papers. Consequently the chapters are rather independent from each other. We synthesize the contents of each chapter and their associated contributions.

In Chapter 2, we presented a model for the investors' behavior. Data provided by a bank are used to characterize the behavioral phenomenon. They are transactions initiated on stocks markets by investors managing six different funds. The modeling of two decisions is considered: the choice of action (buy or sell) and the duration between two actions. An integrated approach has been developed, where the action choice and the duration are modeled together. The action model is a binary logit with latent classes, representing the risk perception. The duration model is a Weibull regression which also accounts for the risk perception. The dynamics of the behavior is modeled explicitly in both models. The perception of the actions is accounted for in the duration model. Explanatory variables are indicators computed by the bank and a market index. New variables are calculated and incorporated in the models to reflect the dynamics.

The models are estimated simultaneously. Parameters are significant and interpretable. Several behavioral mechanisms are underlined. Investors tend to use short-term information when making action choice (daily), whereas they use long-term information for the duration (monthly). The accounting of the dynamics has sense, as investors consider their previous decisions when making current decisions. The specificity of the investors' behavior within each fund appears. The hypothesis about the risk perception is valid in the action and duration models. In *risky* situations, the duration between two consecutive actions is shorter than in *normal* situations, which is logical, as investors adjust more often their portfolio in *risky* situations. Regarding the action choice model, the specificity of the behavior per fund is more important in *risky* than in *normal* situations. This is logical, as individual personalities and emotions are emphasized in panic situations.

Both models are cross-validated on the estimation data. The prediction capabilities are satisfactory for the action model, but limited for the duration model. This is due to the fact that important explanatory variables have not been incorporated in the duration model, such as money flows, as they were not available in the data. Simulation is performed using the action model in order to support its predictive power, and its usefulness in a real context.

In Chapter 3, we focus on the specification, estimation and validation of a discrete choice model capturing the walking behavior. The data are trajectories of pedestrians extracted from videos. The model is a cross nested logit where the choice set is composed of the possible next steps. It is determined by the speed and the direction of the pedestrian. Eleven direction cones and three speed regimes are considered (deceleration, constant speed and acceleration). The combination of the direction cones and the speed regimes lead to thirty-three alternatives. Each alternative belongs to two nests. A nest corresponds to a direction cone or a speed regime. This choice set is individual specific and dynamic as it evolves at each step. The utilities account for the distances and angles of the alternatives toward the destination, the angle with the straight direction, the speed of the pedestrian, and some interactions with the other pedestrians (*leader-follower* and *collision avoidance*).

The model is estimated using trajectories of pedestrians on a cross-walk in Japan. Parameters are interpretable and modeling assumptions are checked. Several interesting behavioral patterns are standing out. Pedestrians have tendency to go straight toward their destinations, without making turns. They tend to keep a stabilized speed. They avoid potential colliders and follow pedestrians going in the same direction, which is logical.

The model is validated using some experimental data collected at the Delft university, which represent bi-directional flows. The validation results show a good prediction accuracy of the model in terms of direction but not for the speed regimes. This is partly explained by the difference of behavior between pedestrians in natural and experimental situations, as well as the difference of culture.

In Chapter 4, we develop some discrete choice models for handling the recognition of dynamic facial expressions. We propose an approach where the ambiguity of the expression perception is accounted for. Data are collected, facial videos of the FEED database are annotated with expressions using an internet survey. Nine labels are proposed: happiness, surprise, fear, disgust, anger, sadness, neutral, other and not known.

Five discrete choice models have been developed, where the choice set is composed of the nine expressions. The **reduced model** is a logit, where the last frame of the video is supposed to trigger the choice of label. The **latent model** is a logit with latent classes, where the most influential frame is supposed to motivate the choice of label. The dynamic process is accounted by a linear combination of the utilities associated to the expressions. The **smoothed model** is an adaption of a logit model

with latent classes, where a group of relevant frames is assumed to generate the choice of label. Two supplementary models are proposed based on the first and second model (**reduced model with panel effect** and **latent model with panel effect**). They account for the correlation between labels obtained besides videos displaying the same filmed subject. Measures are extracted from the faces on the videos using an active appearance model. These measures are used to calculate variables according to the facial action coding system (FACS). They constitute the explanatory variables of the five proposed models.

The models are estimated, and fitted parameters are interpretable. In each model, parameters associated with the expression perception are consistent with the FACS. In the dynamic models (**latent model**, **smoothed model** and **latent model with panel effect**), parameters related to the frame influence have sense, showing that frames with expressive faces are the most influential. The explicit modeling of the dynamic perception is meaningful.

The models are then cross-validated. The prediction capabilities of the **reduced**, **latent** and **smoothed** models are satisfactory, whereas the models with panel effect tend to over-fit the data. Regarding estimation results and predictive accuracy, the **smoothed model** appears to be the best.

We have adopted a general framework for performing analysis in non traditional applications of the behavioral modeling. Three analysis have been successfully conducted in complex and original contexts, characterized by real data. Challenges have been addressed in each tasks of the modeling framework, leading to the generation of several contributions.

Original approaches have been conducted in the three application fields and the different analysis emphasize the added-value of multi-disciplinarity.

Regarding data collection, a data-base of labeled videos has been collected in Chapter 4. An internet survey has been used, which is an emerging way for collecting data. Computer vision techniques were necessary to extract information from facial videos.

Regarding the model specification, behavioral mechanisms are pointed out and quantified, using proper mathematical formulations. In particular, the modeling of the dynamics is considered in Chapters 2 and 4.

For the estimation, the parameters of the developed models are fitted by maximum likelihood. Parameters are significant and their interpretations have been discussed.

Concerning the validation, a complete methodology is proposed and performed in Chapter 3. In Chapters 2 and 4, the prediction capabilities of the models are checked by cross-validation. In each application, the proposed models are operational in real contexts.

Moreover, discrete choice models with latent classes are developed in Chapters 2 and 4. In the literature, these models have been developed for understanding behaviors and rarely for prediction. In this work, both aspects have been studied in details.

## 5.2 Future research

In each chapter, we detailed the perspectives specifically associated to each analysis. In this section, we emphasize the common perspectives. In addition, the global research perspectives of the thesis are developed.

For the three applications, the proposed models should be embedded into simulators. This will allow to conduct case studies in order to illustrate the added-value of the models in decision-aid tools. Regarding the financial behavior, the simulator would give insight for future decisions to fund managers. For the walking behavior, it would be helpful for the urban planning and the design of infrastructures, such as building, stations or public places in general. Concerning the dynamic facial expression recognition, it is particularly relevant in marketing, in transportation in the context of “aware vehicles”, and in any human-machine interfaces.

In the literature, models have been proposed for each application. A comparison between these models and the models developed in this thesis would be interesting. The comparison should be done in terms of understanding and prediction capabilities. This would quantify the added-value of the explicit modeling of behavioral causalities.

In this thesis, no model accounts for socio-economic characteristics of the decision-maker, because they were not available in the data. They obviously impact the behavior. Regarding financial decisions, different styles are known to be adopted by investors, depending on their experience and personality. For the pedestrian modeling, physical characteristics and activities at the destination are relevant. For example, a pedestrian going to work has not the same behavior than a pedestrian who is window-shopping. Concerning the dynamic facial expression recognition, an individual is supposed to label differently a subject having the same ethnic group. In the three applications, the influence of the culture on the decision process is not negligible, and should be further investigated.

For each task of the adopted modeling framework, we underline specific perspectives. Regarding data, in Chapter 4 we have conducted a data collection using an internet survey. More generally, designs of experiments should be adapted to emerging applications. They should be consistent with the new available technologies for collecting data. Moreover, these technologies allow to generate huge amounts of data. New mathematical developments are required to manage and exploit such databases.

Concerning the model specification, a model should be both operational and realistic. Practically, a trade-off has to be found between these two characteristics. In this thesis and for each application, we proposed the most realistic behavioral models which remain operational. The improvement of the modeling techniques would allow to reduce this gap in complex applications. Different aspects should be explored. For example, large choice sets are current in nowadays applications. Rationality of individuals is not always a valid assumption and should be relaxed in some cases. Attitudes are known to be crucial decision factors, so more psychological and sociological concepts should be integrated in behavioral models.

The combination of the new modeling techniques and huge databases complicate the estimation task. The estimation induces the optimization of complex mathematical functions. The development of powerful optimization algorithms is required to guarantee the quality of the fitted parameters. In addition, they should run real-time in order to be usable for applications.

In terms of validation, we have quantified the validity of the proposed models. But, it would be interesting to formally decide if a model is valid or not. In the literature, few statistical tests have been proposed in that purpose. This should be further investigated.



# A. The modeling of investors' decisions

In this appendix, we summarize the estimation results of the five models proposed in Chapter 2.

Parameter	$\nu$	Transform	$t_H$ (day)	$g$	$r$	Value	t-test
$ASC_{B,N}$	1			1,2	N	-1.39	-4.41
$ASC_{B,R}$	1			1,2	R	-2.01	-4.44
$\beta_{B,1}$	Price	Long()	1	1	N	-2.31	-3.15
$\beta_{B,2}$	Price	Perf()	1	2	N,R	6.18	4.19
$\beta_{B,3}$	Price	Short()	1	2	N	-1.02	-3.15
$\beta_{B,4}$	Price	Short()	1	1,2	R	0.916	1.78
$\beta_{B,5}$	Quality	Long()	1	1,2	N	-0.428	-2.57
$\beta_{B,6}$	Quality	Short()	1	1,2	N	0.723	2.03
$\beta_{B,7}$	Sentiment	Long()	1	2	N	0.659	2.78
$\beta_{B,8}$	Technic	Long()	1	1	R	1.06	1.84
$\beta_{B,9}$	Technic	Long()	1	1	N	1.02	2.51
$\beta_{B,10}$	Technic	Long()	1	2	R	-0.805	-2.42
$\beta_{B,11}$	Technic	Short()	1	1	R	1.80	2.45
$\beta_{B,12}$	Technic	Short()	1	1	N	2.78	4.28
$\beta_{B,13}$	Technic	Short()	1	2	R	1.30	3.40
$\beta_{B,14}$	Technic	Short()	1	2	N	-1.47	-2.00
$\beta_{B,15}$	Value	Short()	1	1,2	N	0.625	2.43
$\alpha_N$				1,2	N	-0.222	-5.43
$\alpha_R$				1,2	R	-0.312	-6.77
$\lambda_{B,N}$				1,2	N	-0.0156	3.79
$\lambda_{B,R}$				1,2	R	-0.00946	5.27
$\mu_1$				1,2	N,R	1	
$\mu_2$				1,2	N,R	1.23	5.07

$\mu_3$				1,2	N,R	1.58	4.21
$\mu_4$				1,2	N,R	2.93	3.16
$\mu_5$				1,2	N,R	1.43	2.77
$\mu_6$				1,2	N,R	2.46	3.48

Table A.1: Estimated parameters of the action model ( $\beta$ )

Parameter	$\nu$	Transform	$t_H$ (day)	$g$	Value	t-test
$ASC_{W_A}$	1			1,2	-25.327	
$\omega_{A,1}$	VIX			1	1.37	49.41
$\omega_{A,2}$	VIX			2	1.08	31.39
$\omega_{A,3}$	Sentiment	Sigm()	5	1,2	9.29	5.17

Table A.2: Estimated parameters of the risk model associated to the action model ( $\omega_A$ )

Parameter	$\nu$	Transform	$t_H$ (day)	$g$	$r$	Value	t-test
$ASC_{D,N,1}$	1			1	N	3.54	5.49
$ASC_{D,N,2}$	1			2	N	3.36	10.71
$ASC_{D,R,1}$	1			1	N	0.242	1.54
$ASC_{D,R,2}$	1			2	N	0.398	2.93
$\theta_{D,1}$	Price	Short()	60	1	N	-1.22	-3.47
$\theta_{D,2}$	Price	Sigm()	60	1	N	1.52	1.78
$\theta_{D,3}$	Quality	Long()	60	1	N	-1.08	-3.03
$\theta_{D,4}$	Quality	Short()	60	1	R	0.960	1.80
$\theta_{D,5}$	Quality	Short()	60	2	R	-0.661	-1.54
$\theta_{D,6}$	Quality	Short()	60	1	N	-1.43	-1.96
$\theta_{D,7}$	Sentiment	Long()	60	2	R	-0.716	-3.29
$\theta_{D,8}$	Sentiment	Short()	60	2	R	0.990	3.72
$\theta_{D,9}$	Technic	Long()	60	2	R	1.42	5.38
$\theta_{D,10}$	Technic	Long()	60	1	N	-1.18	-2.36
$\theta_{D,11}$	Technic	Short()	60	1	R	1.79	3.19
$\theta_{D,12}$	Technic	Sigm()	60	2	N	-1.48	-3.91
$\theta_{D,13}$	Value	Short()	60	2	R	1.90	4.84
$\theta_{D,14}$	Value	Short()	60	1	N	2.25	2.60
$\theta_{D,15}$	Value	Sigm()	60	2	N	-0.613	-1.72
$\theta_{D,16}$	VIX	Sigm()	360	2	N	-2.05	-4.99
$\theta_{B,N}$				1,2	N	-0.350	-2.27
$\theta_{B,R}$				1,2	R	-0.261	-2.53
$\alpha_{D,N,1}$				1	N	7.09	4.77



$\alpha_{D,R,1}$				1	R	5.29	3.25
$\alpha_{D,N,2}$				2	R	2.23	1.84
$\eta_D$						0.530	82.26

Table A.3: Estimated parameters of the duration model ( $\theta$ )

Parameter	$\nu$	Transform	$t_H$ (day)	$g$	Value	t-test
$ASC_{W_D,1}$	1			1	-7.48	-1.59
$ASC_{W_D,2}$	1			2	-4.93	-3.76
$\omega_{D,1}$	VIX			1	0.377	2.10
$\omega_{D,2}$	VIX			2	0.263	4.53

Table A.4: Estimated parameters of the risk model associated to the duration model ( $\omega_D$ )



# B. The dynamic facial expression recognition

In this appendix, we summarize the notations and the estimation results of the five models proposed in Chapter 4.

## B.1 Notations

Acronym	Definition
A	Anger
AAM	Active appearance model
ASC	Alternative specific constant
AU	Action unit
D	Disgust
DCM	Discrete choice model
DFER	Dynamic facial expression recognition
DK	I don't know
EDU	Expression descriptive unit
EDU_6	EDU defined as the ratio between the average of the eye width and the mouth width
EDU_8	EDU defined as the ratio between the average of the eyes height and the average of the brow-eyes height
F	Fear
FER	Facial expression recognition
FACS	Facial action coding system
FEED	Facial expression and emotion database
HMM	Hidden markov model
H	Happiness
MEV	Multivariate extreme value
N	Neutral
NN	Neural networks
O	Other

PCA	Principal component analysis
RAP_brow	EDU defined as the ratio between the average of the brow-eyes height and the brow-eyes width
RAP_mouth	EDU defined as the ratio between the height and the width of the mouth
SA	Sadness
SFER	Static facial expression recognition
SU	Surprise
SVM	Support vector machine

Table B.1: Summary of the mentioned acronyms

## B.2 Estimation results

parameter	H	SU	F	D	SA	A	N	O	DK	$x_{k,T_0,o}$	value	t-test 0
$ASC_{M_1,A}$						×				1	0.95	0.28
$ASC_{M_1,D}$				×						1	25.38	7.88
$ASC_{M_1,DK}$									×	1	-0.69	-1.79
$ASC_{M_1,F}$			×							1	0.49	0.19
$ASC_{M_1,H}$	×									1	-3.14	-0.79
$ASC_{M_1,O}$								×		1	6.95	3.20
$ASC_{M_1,SA}$					×					1	10.80	2.54
$ASC_{M_1,SU}$		×								1	-11.27	-5.63

Table B.2: Estimation results of the constants for **reduced model**

parameter	H	SU	F	D	SA	A	N	O	DK	$x_{k,T_0,o}$	value	t-test 0
$\theta_{M_1,1}$				×						EDU_6	-6.52	-3.63
$\theta_{M_1,2}$				×						EDU_8	-4.75	-6.18
$\theta_{M_1,3}$		×				×				RAP_brow	6.70	4.53
$\theta_{M_1,4}$		×	×							RAP_mouth	2.94	2.85
$\theta_{M_1,5}$	×									RAP_mouth	9.36	5.35
$\theta_{M_1,6}$	×									C_1	-16.30	-3.51
$\theta_{M_1,7}$						×				C_2	23.98	3.49
$\theta_{M_1,8}$				×						C_2	26.22	5.16
$\theta_{M_1,9}$	×									C_3	15.34	3.13
$\theta_{M_1,10}$		×								C_3	15.73	3.27
$\theta_{M_1,11}$					×					broweye_l2	153.91	3.17
$\theta_{M_1,12}$		×								broweye_l3	85.58	5.75

parameter	H	SU	F	D	SA	A	N	O	DK	$x_{k,T_0,o}$	value	t-test 0
$\theta_{M_1,13}$		×	×	×	×	×				broweye_r2	-49.81	-4.30
$\theta_{M_1,14}$			×		×					eye_angle_l	58.55	3.43
$\theta_{M_1,15}$					×					eye_brow_angle_l	-140.87	-5.10
$\theta_{M_1,16}$				×						eye_mouth_dist_l2	-69.83	-3.42
$\theta_{M_1,17}$	×				×			×		eye_mouth_dist_l	-36.03	-2.89
$\theta_{M_1,18}$						×				eye_nose_dist_l	245.03	5.05
$\theta_{M_1,19}$			×	×	×			×		eye_nose_dist_l	147.67	4.89
$\theta_{M_1,20}$			×	×	×	×		×		eye_nose_dist_r	-213.93	-6.04
$\theta_{M_1,21}$		×	×							leye_h	20.97	2.09
$\theta_{M_1,22}$					×	×				mouth_nose_dist2	-90.97	-2.15
$\theta_{M_1,23}$	×									mouth_nose_dist	-236.37	-5.65
$\theta_{M_1,24}$	×									mouth_w	188.42	4.90

Table B.3: Estimation results and description of the specification of **reduced model**

parameter	H	SU	F	D	SA	A	N	O	DK	$x_{k,t,o}$	value	t-test 0
$ASC_{M_2,A}$						×				1	-5.86	-1.31
$ASC_{M_2,D}$				×						1	22.73	4.48
$ASC_{M_2,DK}$									×	1	-0.71	-1.83
$ASC_{M_2,F}$			×							1	-4.55	-1.13
$ASC_{M_2,H}$	×									1	3.02	0.22
$ASC_{M_2,O}$								×		1	14.44	4.22
$ASC_{M_2,SA}$					×					1	8.54	1.57
$ASC_{M_2,SU}$		×								1	-25.69	-7.08

Table B.4: Estimation results of the constants for the **latent model**, associated the expression perception model

parameter	H	SU	F	D	SA	A	N	O	DK	$x_{k,t,o}$	value	t-test 0
$\theta_{M_2,1,1}$				×						EDU_6	-6.92	-3.37
$\theta_{M_2,1,2}$				×						EDU_8	-3.92	-5.42
$\theta_{M_2,1,3}$		×				×				RAP_brow	7.84	4.45
$\theta_{M_2,1,4}$		×	×							RAP_mouth	4.93	3.42
$\theta_{M_2,1,5}$	×									RAP_mouth	12.74	2.54
$\theta_{M_2,1,6}$	×									C_1	-38.18	-5.27
$\theta_{M_2,1,7}$						×				C_2	40.99	4.81
$\theta_{M_2,1,8}$				×						C_2	45.77	7.12
$\theta_{M_2,1,9}$	×									C_3	23.96	3.71
$\theta_{M_2,1,10}$		×								C_3	24.46	4.11

parameter	H	SU	F	D	SA	A	N	O	DK	$x_{k,t,o}$	value	t-test 0
$\theta_{M_2,1,11}$					×					broweye_l2	240.75	4.11
$\theta_{M_2,1,12}$		×								broweye_l3	104.09	4.61
$\theta_{M_2,1,13}$		×	×	×	×	×				broweye_r2	-41.76	-2.93
$\theta_{M_2,1,14}$			×		×					eye_angle_l	44.95	2.58
$\theta_{M_2,1,15}$					×					eye_brow_angle_l	-199.01	-6.04
$\theta_{M_2,1,16}$				×						eye_mouth_dist_l2	-73.15	-2.72
$\theta_{M_2,1,17}$	×				×			×		eye_mouth_dist_l	-84.03	-3.83
$\theta_{M_2,1,18}$						×				eye_nose_dist_l	217.99	3.69
$\theta_{M_2,1,19}$			×	×	×			×		eye_nose_dist_l	80.02	2.09
$\theta_{M_2,1,20}$			×	×	×	×		×		eye_nose_dist_r	-211.73	-4.45
$\theta_{M_2,1,21}$		×	×							leye_h	51.35	4.12
$\theta_{M_2,1,22}$	×	×	×	×	×	×				mouth_h	98.27	3.27
$\theta_{M_2,1,23}$					×	×				mouth_nose_dist2	-92.34	-2.04
$\theta_{M_2,1,24}$	×									mouth_nose_dist	-412.5	-5
$\theta_{M_2,1,25}$	×									mouth_w	158.29	2.13
$\theta_{M_2,1,1}^z$										mouth_h, $z_{1,t,o}$	50.21	3.04

Table B.5: Estimation results and description of the specification of the **latent model**, associated to the expression perception model

parameter	value	t-test 0
$\alpha_{M_2,H}$	-0.62	-8.18
$\alpha_{M_2,F}$	-0.33	-2.73
$\alpha_{M_2,SA}$	-0.46	-2.04
$\alpha_{M_2,O}$	-0.70	-2.68

Table B.6: Estimation results of the **latent model**, associated to the memory effects parameters

parameter	$y_{k,t,o}$	value	t-test 0
$\theta_{M_2,2,1}^y$	C_2	-426.75	-1.83
$\theta_{M_2,2,2}^y$	eye_brow_angle	350.53	1.7
$\theta_{M_2,2,3}^y$	mouth_w	407.34	1.76
$\theta_{M_2,2,4}^y$	C_4	463.35	1.75
$\theta_{M_2,2,5}^y$	eye_h	-566.62	-1.79
$\theta_{M_2,2,6}^y$	mouth_h	104.51	1.84
$\theta_{M_2,2,1}^z$	brow_dist, $z_{4,t,o}$	261.65	1.84

parameter	$y_{k,t,o}$	value	t-test 0
-----------	-------------	-------	----------

Table B.7: Estimation results and description of the specification of the **latent model**, associated to the model which detects the most meaningful frame

parameter	H	SU	F	D	SA	A	N	O	DK	$x_{k,t,o}$	value	t-test 0
$ASC_{M_3,A}$						×				1	-7.53	-1.63
$ASC_{M_3,D}$				×						1	20.28	4.03
$ASC_{M_3,DK}$									×	1	-0.69	-1.79
$ASC_{M_3,F}$			×							1	-0.35	-0.09
$ASC_{M_3,H}$	×									1	-7.66	-1.43
$ASC_{M_3,O}$								×		1	12.95	4.38
$ASC_{M_3,SA}$					×					1	4.17	1.04
$ASC_{M_3,SU}$		×								1	-29.15	-7.07

Table B.8: Estimation results of the constants for the **smoothed model**, associated to the expression perception model

parameter	H	SU	F	D	SA	A	N	O	DK	$x_{k,t,o}$	value	t-test 0
$\theta_{M_3,1,1}$				×						EDU_6	-9.19	-3.82
$\theta_{M_3,1,2}$				×						EDU_8	-4.18	-4.09
$\theta_{M_3,1,3}$		×				×				RAP_brow	12.6	5.69
$\theta_{M_3,1,4}$				×						RAP_brow	5.44	2
$\theta_{M_3,1,5}$		×	×							RAP_mouth	2.89	2
$\theta_{M_3,1,6}$	×									RAP_mouth	11.77	4.44
$\theta_{M_3,1,7}$	×									C_1	-23.36	-3.36
$\theta_{M_3,1,8}$						×				C_2	42.46	5.3
$\theta_{M_3,1,9}$				×						C_2	33.98	5.51
$\theta_{M_3,1,10}$	×									C_3	25.82	3.88
$\theta_{M_3,1,11}$		×								C_3	17.61	2.74
$\theta_{M_3,1,12}$				×						C_5	-16.4	-2.5
$\theta_{M_3,1,13}$					×					broweye_l2	149.31	3.15
$\theta_{M_3,1,14}$		×								broweye_l3	128.49	5.76
$\theta_{M_3,1,15}$		×	×	×	×	×				broweye_r2	-61.58	-4.31
$\theta_{M_3,1,16}$			×		×					eye_angle_l	40.99	2.06
$\theta_{M_3,1,17}$					×					eye_brow_angle_l	-126.55	-4.59
$\theta_{M_3,1,18}$				×						eye_mouth_dist_l2	-50.07	-2.13
$\theta_{M_3,1,19}$	×				×			×		eye_mouth_dist_l	-32.09	-2.2
$\theta_{M_3,1,20}$						×				eye_nose_dist_l	163.49	3.75

parameter	H	SU	F	D	SA	A	N	O	DK	$x_{k,t,o}$	value	t-test 0
$\theta_{M_3,1,21}$			×	×	×			×		eye_nose_dist_l	114.66	3.15
$\theta_{M_3,1,22}$			×	×	×	×		×		eye_nose_dist_r	-256.49	-5.39
$\theta_{M_3,1,23}$		×	×							leye_h	52.58	3.73
$\theta_{M_3,1,24}$	×	×	×	×	×	×				mouth_h	90.92	2.96
$\theta_{M_3,1,25}$	×									mouth_nose_dist	-342.14	-6.17
$\theta_{M_3,1,26}$	×									mouth_w	228.81	4.47
$\theta_{M_3,1,1}^z$		×								mouth_h, $z_{1,t,o}$	0.13	4.46
$\theta_{M_3,1,2}^z$		×	×							leye_h, $z_{3,t,o}$	0.04	2.39

Table B.9: Estimation results and description of the specification of the **smoothed model**, associated to the expression perception model

parameter	$y_{k,t,o}$	value	t-test 0
$\theta_{M_3,2,1}^y$	C_1	-234.75	-1.75
$\theta_{M_3,2,2}^y$	eye_brow_angle	548.34	1.76
$\theta_{M_3,2,3}^y$	mouth_w	23.29	1.81
$\theta_{M_3,2,4}^y$	C_2	101.9	1.85
$\theta_{M_3,2,5}^y$	C_3	-221.23	-1.57
$\theta_{M_3,2,6}^y$	C_5	529.64	1.91
$\theta_{M_3,2,7}^y$	eye_h	-122.15	-1.79
$\theta_{M_3,2,8}^y$	mouth_h	119.21	1.88

Table B.10: Estimation results and description of the specification of the **smoothed model**, associated to the model related to the detection of the first frame of the relevant group of frames

parameter	H	SU	F	D	SA	A	N	O	DK	$x_{k,T_o,o}$	value	t-test 0
$ASC_{M_4,A}$						×				1	1.61	0.42
$ASC_{M_4,D}$				×						1	25.40	5.80
$ASC_{M_4,DK}$									×	1	-0.067	-0.10
$ASC_{M_4,F}$			×							1	1.14	0.37
$ASC_{M_4,H}$	×									1	-3.69	-0.94
$ASC_{M_4,O}$								×		1	7.44	2.95
$ASC_{M_4,SA}$					×					1	11.60	3.37
$ASC_{M_4,SU}$		×								1	-9.91	-4.83

Table B.11: Estimation results of the constants for **reduced model with panel effect**



parameter	H	SU	F	D	SA	A	N	O	DK	$x_{k,T_0,o}$	value	t-test 0
$\theta_{M_4,1}$				×						EDU_6	-6.68	-3.27
$\theta_{M_4,2}$				×						EDU_8	-4.57	-3.68
$\theta_{M_4,3}$		×				×				RAP_brow	6.38	4.41
$\theta_{M_4,4}$		×	×							RAP_mouth	2.70	3.33
$\theta_{M_4,5}$	×									RAP_mouth	9.66	5.50
$\theta_{M_4,6}$	×									C_1	-16.70	-2.33
$\theta_{M_4,7}$						×				C_2	22.76	2.80
$\theta_{M_4,8}$				×						C_2	25.20	4.01
$\theta_{M_4,9}$	×									C_3	15.84	2.47
$\theta_{M_4,10}$		×								C_3	15.92	6.03
$\theta_{M_4,11}$					×					broweye_l2	158.76	3.00
$\theta_{M_4,12}$		×								broweye_l3	82.23	5.75
$\theta_{M_4,13}$		×	×	×	×	×				broweye_r2	-52.02	-3.20
$\theta_{M_4,14}$			×		×					eye_angle_l	55.23	3.12
$\theta_{M_4,15}$					×					eye_brow_angle_l	-143.11	-7.56
$\theta_{M_4,16}$				×						eye_mouth_dist_l2	-66.87	-2.49
$\theta_{M_4,17}$	×				×			×		eye_mouth_dist_l	-42.45	-3.40
$\theta_{M_4,18}$						×				eye_nose_dist_l	252.55	5.46
$\theta_{M_4,19}$			×	×	×			×		eye_nose_dist_l	153.93	3.38
$\theta_{M_4,20}$			×	×	×	×		×		eye_nose_dist_r	-214.88	-3.93
$\theta_{M_4,21}$		×	×							leye_h	22.90	1.80
$\theta_{M_4,22}$					×	×				mouth_nose_dist2	-93.02	-2.01
$\theta_{M_4,23}$	×									mouth_nose_dist	-235.84	-3.82
$\theta_{M_4,24}$	×									mouth_w	202.92	4.48
$\sigma$											1.47	4.33

Table B.12: Estimation results and description of the specification of **reduced model with panel effect**

parameter	H	SU	F	D	SA	A	N	O	DK	$x_{k,t,o}$	value	t-test 0
$ASC_{M_5,A}$						×				1	-5.29	-1.44
$ASC_{M_5,D}$				×						1	20.90	4.44
$ASC_{M_5,DK}$									×	1	-0.180	-0.25
$ASC_{M_5,F}$			×							1	-3.30	-0.63
$ASC_{M_5,H}$	×									1	-11.08	-0.96
$ASC_{M_5,O}$								×		1	14.70	3.00
$ASC_{M_5,SA}$					×					1	10.09	1.95
$ASC_{M_5,SU}$		×								1	-22.50	-6.45

parameter	H	SU	F	D	SA	A	N	O	DK	$x_{k,t,o}$	value	t-test 0
-----------	---	----	---	---	----	---	---	---	----	-------------	-------	----------

Table B.13: Estimation results of the constants for the **latent model with panel effect**, associated the expression perception model

parameter	H	SU	F	D	SA	A	N	O	DK	$x_{k,t,o}$	value	t-test 0
$\theta_{M_5,1,1}$				×						EDU_6	-6.10	-4.01
$\theta_{M_5,1,2}$				×						EDU_8	-3.85	-3.65
$\theta_{M_5,1,3}$		×				×				RAP_brow	7.62	3.37
$\theta_{M_5,1,4}$		×	×							RAP_mouth	3.96	2.98
$\theta_{M_5,1,5}$	×									RAP_mouth	17.70	3.27
$\theta_{M_5,1,6}$	×									C_1	-30.40	-4.20
$\theta_{M_5,1,7}$						×				C_2	43.40	5.52
$\theta_{M_5,1,8}$				×						C_2	46.10	5.68
$\theta_{M_5,1,9}$	×									C_3	21.60	3.21
$\theta_{M_5,1,10}$		×								C_3	25.30	3.99
$\theta_{M_5,1,11}$					×					broweye_l2	238.00	4.76
$\theta_{M_5,1,12}$		×								broweye_l3	87.70	4.30
$\theta_{M_5,1,13}$		×	×	×	×	×				broweye_r2	-51.60	-3.14
$\theta_{M_5,1,14}$			×		×					eye_angle_l	39.3	1.80
$\theta_{M_5,1,15}$					×					eye_brow_angle_l	-190.00	-7.84
$\theta_{M_5,1,16}$				×						eye_mouth_dist_l2	-67.8	-1.82
$\theta_{M_5,1,17}$	×				×			×		eye_mouth_dist_l	-84.30	-3.71
$\theta_{M_5,1,18}$						×				eye_nose_dist_l	258.00	3.15
$\theta_{M_5,1,19}$			×	×	×			×		eye_nose_dist_l	106.00	1.59
$\theta_{M_5,1,20}$			×	×	×	×		×		eye_nose_dist_r	-223.00	-3.01
$\theta_{M_5,1,21}$		×	×							leye_h	46.50	3.02
$\theta_{M_5,1,22}$	×	×	×	×	×	×				mouth_h	103.00	2.42
$\theta_{M_5,1,23}$					×	×				mouth_nose_dist2	-121.00	-1.91
$\theta_{M_5,1,24}$	×									mouth_nose_dist	-327.00	-3.00
$\theta_{M_5,1,25}$	×									mouth_w	215.00	4.74
$\theta_{M_5,1,1}^z$										mouth_h, $z_{1,t,o}$	55.20	3.06
$\sigma_{M_5}$											1.20	2.44

Table B.14: Estimation results and description of the specification of the **latent model with panel effect**, associated to the expression perception model

parameter	value	t-test 0
$\alpha_{M_5,H}$	-0.557	-4.29

parameter	value	t-test 0
$\alpha_{M_5,F}$	-0.314	-2.14
$\alpha_{M_5,SA}$	-0.381	-1.31
$\alpha_{M_5,O}$	-0.585	-2.64

Table B.15: Estimation results of the **latent model with panel effect**, associated to the memory effects parameters

parameter	$\mathbf{y}_{k,t,o}$	value	t-test 0
$\theta_{M_5,2,1}^y$	C_2	-506.23	-3.58
$\theta_{M_5,2,2}^y$	eye_brow_angle	311.53	3.93
$\theta_{M_5,2,3}^y$	mouth_w	438.40	3.69
$\theta_{M_5,2,4}^y$	C_4	441.12	3.85
$\theta_{M_5,2,5}^y$	eye_h	-634.03	-3.63
$\theta_{M_5,2,6}^y$	mouth_h	123.99	3.66
$\theta_{M_5,2,1}^z$	brow_dist, $z_{4,t,o}$	295.89	3.76

Table B.16: Estimation results and description of the specification of the **latent model with panel effect**, associated to the model which detects the most meaningful frame



# Bibliography

- Abbe, E., Bierlaire, M. and Toledo, T. (2007). Normalization and correlation of cross-nested logit models, *Transportation Research Part B* **41**(7): 795–808.
- Abou-Zeid, M. (2009). *Measuring and Modeling Travel and Activity Well-Being*, PhD thesis, Massachusetts Institute of Technology.
- Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion, *Journal of Banking & Finance* **26**(7): 1505 – 1518.
- Ahmed, K. I. (1999). *Modeling drivers' acceleration and lane changing behaviors.*, PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.
- Aite (2010). The european equity electronic trading landscape: How deep is your pool?, *Technical report*, Aite group.
- Antonini, G. (2005). *A discrete choice modeling framework for pedestrian walking behavior with application to human tracking in video sequences*, PhD thesis, Ecole Polytechnique Fédérale de Lausanne. 3382.
- Antonini, G. and Bierlaire, M. (2007). A discrete choice framework for acceleration and direction change behaviors in walking pedestrians, in N. Waldau, P. Gattermann, H. Knoflachner and M. Schreckenberg (eds), *Pedestrian and Evacuation Dynamics 2005*, Springer, pp. 145–156. ISBN:978-3-540-47062-5.
- Antonini, G., Bierlaire, M. and Weber, M. (2006). Discrete choice models of pedestrian walking behavior, *Transportation Research Part B* **40**(8): 667–687.
- Antonini, G., Sorci, M., Bierlaire, M. and Thiran, J. (2006). Discrete choice models for static facial expression recognition, in J. Blanc-Talon, W. Philips, D. Popescu and P. Scheunders (eds), *8th International Conference on Advanced Concepts for Intelligent Vision Systems*, Vol. 4179 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, Berlin, pp. 710–721. ISBN: 978-3-540-44630-9.
- Antonini, G., Venegas, S., Bierlaire, M. and Thiran, J.-P. (2006). Behavioral priors for detection and tracking of pedestrians in video sequences, *International Journal of Computer Vision* **69**(2): 159–180.

- Antonini, G., Venegas, S., Thiran, J.-P. and Bierlaire, M. (2004). A discrete choice pedestrian behavior model for pedestrian detection in visual tracking systems, *Proceedings of the Advanced Concepts for Intelligent Vision Systems*, Brussels, Belgium.
- Arthur, W. B., Holland, J. H., LeBaron, B. D., Palmer, R. G. and Tayler, P. (1996). Asset Pricing Under Endogenous Expectations in an Artificial Stock Market, *SSRN eLibrary*.
- Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent measures of risk, *Mathematical finance* **9**(3): 203–228.
- Baker, M. and Wurgler, J. (2007). Investor sentiment in the stock market, *The Journal of Economic Perspectives* **21**(2): 129–151.
- Barberis, N. and Thaler, R. (2003). Chapter 18 a survey of behavioral finance, in M. H. G.M. Constantinides and R. Stulz (eds), *Financial Markets and Asset Pricing*, Vol. 1, Part 2 of *Handbook of the Economics of Finance*, Elsevier, pp. 1053–1128.
- Bartlett, M. S., Littlewort, G., Fasel, I. and Movellan, J. R. (2003). Real time face detection and facial expression recognition: Development and applications to human computer interaction., *Computer Vision and Pattern Recognition Workshop, 2003. CVPRW '03. Conference on*, Vol. 5, pp. 53–53.
- Basak, S. and Cuoco, D. (1998). An equilibrium model with restricted stock market participation, *Review of Financial Studies* **11**(2): 309–341.
- Bauer, D. (2007). Private communication.
- Bauwens, L. and Veredas, D. (2004). The stochastic conditional duration model: a latent variable model for the analysis of financial durations, *Journal of Econometrics* **119**(2): 381–412. Dynamic factor models.  
**URL:** <http://www.sciencedirect.com/science/article/B6VC0-494HPGM-1/2/aca7702ca7b7e4807d312fd243c55f1e>
- Ben-Akiva, M. (2010). Planning and action in a model of choice, in S. Hess and A. Daly (eds), *Choice modelling: the state-of-the-art and the state-of-practice*, Emerald, pp. 19–34.
- Ben-Akiva, M. E. and Lerman, S. R. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press, Cambridge, Ma.
- Ben-Akiva, M., Mcfadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., Bolduc, D., Boersch-Supan, A., Brownstone, D., Bunch, D. S., Daly, A., De Palma, A., Gopinath, D., Karlstrom, A. and Munizaga, M. A. (2002). Hybrid choice models: Progress and challenges, *Marketing Letters* **13**: 163–175.

- Berrou, J. L., Beecham, J., Quaglia, P., Kagarlis, M. A. and Gerodimos, A. (2007). Calibration and validation of the legion simulation model using empirical data, *in* N. Waldau, P. Gattermann, H. Knoflachner and M. Schreckenberger (eds), *Pedestrian and Evacuation Dynamics 2005*, Springer, pp. 167–182.
- Bierlaire, M. (2003a). Biogeme: A free package for the estimation of discrete choice models, Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland.
- Bierlaire, M. (2003b). BIOGEME: a free package for the estimation of discrete choice models, *Proceedings of the 3rd Swiss Transportation Research Conference*, Ascona, Switzerland. [www.strc.ch](http://www.strc.ch).
- Bierlaire, M. (2006). A theoretical analysis of the cross-nested logit model, *Annals of Operations Research* **144**(1): 287–300.
- Bierlaire, M., Antonini, G. and Weber, M. (2003). Behavioral dynamics for pedestrians, *in* K. Axhausen (ed.), *Moving through nets: the physical and social dimensions of travel*, Elsevier, pp. 1–18.
- Bierlaire, M. and Fetiariison, M. (2009). Estimation of discrete choice models: extending biogeme, *Proceedings of the 9th Swiss Transport Research Conference*, Ascona, Switzerland.
- Blue, V. J. and Adler, J. L. (2001). Cellular automata microsimulation for modeling bi-directional pedestrian walkways, *Transportation Research Part B* **35**(3): 293–312.
- Blume, M. (1971). On the assessment of risk, *The journal of finance* **26**(1): 1.
- Bolduc, D. and Alvarez-Daziano, R. (2010). On estimation of hybrid choice models, *in* S. Hess and A. Daly (eds), *Choice modelling: the state-of-the-art and the state-of-practice*, Emerald, pp. 259–288.
- Bollerslev, T., Chou, R. Y. and Kroner, K. F. (1992). Arch modeling in finance : A review of the theory and empirical evidence, *Journal of Econometrics* **52**(1-2): 5 – 59.
- Bollerslev, T. and Mikkelsen, H. O. (1996). Modeling and pricing long memory in stock market volatility, *Journal of Econometrics* **73**(1): 151 – 184.  
**URL:** <http://www.sciencedirect.com/science/article/B6VC0-3VW1TT8-6/2/2ce680f1a1dc3b5fe1e056e3c0faaa2d>
- Borgers, A. and Timmermans, H. (1986). A model of pedestrian route choice and demand for retail facilities within inner-city shopping areas, *Geographical analysis* **18**(2): 115–128.

- Brady, A. T. and Walker, M. B. (1978). Interpersonal distance as a function of situationally induced anxiety, *British Journal of Social and Clinical Psychology* **17**: 127–133.
- Brogan, D. C. and Johnson, N. L. (2003). Realistic human walking paths, *Proceedings of the 16th International Conference on Computer Animation and Social Agents*, IEEE, pp. 94–101.
- Campbell, J. Y. and Shiller, R. J. (2001). Valuation Ratios and the Long-Run Stock Market Outlook: An Update, *SSRN eLibrary* .
- Cepolina, E. and Tyler, N. (2005). Understanding capacity drop for designing pedestrian environments, *Walk21-VI "Everyday Walking Culture", The 6th International Conference on Walking in the 21st Century*, Zurich.
- Chang, T.-J., Meade, N., Beasley, J. and Sharaiha, Y. (2000). Heuristics for cardinality constrained portfolio optimisation, *Computers & Operations Research* **27**: 1271 – 1302.
- Chen, S.-H. and Yeh, C.-H. (2001). Evolving traders and the business school with genetic programming: A new architecture of the agent-based artificial stock market, *Journal of Economic Dynamics and Control* **25**(3-4): 363 – 393.  
**URL:** <http://www.sciencedirect.com/science/article/B6V85-419JHMW-4/2/e00f233b151a01fd8c43e8574e39a524>
- Choudhury, C. F. (2007). *Model Driving Decisions with Latent Plans*, PhD thesis, Massachusetts institute of technology.
- Cohen, I., Sebe, N., Garg, A., Chen, L. S. and Huang, T. S. (2003). Facial expression recognition from video sequences: temporal and static modeling, *Computer Vision and Image Understanding* **91**(1-2): 160 – 187. Special Issue on Face Recognition.
- Collett, P. and Marsh, P. (1974). Patterns of public behavior: Collision avoidance on a pedestrian crossing, *Semiotica* **12**: 281–299.
- Cont, R. and Bouchaud, J.-P. (1997). Herd Behavior and Aggregate Fluctuations in Financial Markets, *SSRN eLibrary* .
- Cootes, T. F., Wheeler, G. V., Walker, K. N. and Taylor, C. J. (2002). View-based active appearance models, *Image and Vision Computing* **20**(9-10): 657 – 664.
- Daamen, W. (2004). *Modelling Passenger Flows in Public Transport Facilities*, PhD thesis, Delft University of Technology, The Netherlands.



- Daamen, W. and Hoogendoorn, S. P. (2003a). Controlled experiments to derive walking behaviour, *European Journal of Transport and Infrastructure Research* **3**(1): 39–59.
- Daamen, W. and Hoogendoorn, S. P. (2003b). Experimental research of pedestrian walking behavior, *Transportation Research Record* **1828**: 20–30.
- Day, R. H. and Huang, W. (1990). Bulls, bears and market sheep, *Journal of Economic Behavior & Organization* **14**(3): 299 – 329.
- De Grauwe, P. and Grimaldi, M. (2004). Bubbles and Crashes in a Behavioural Finance Model, *SSRN eLibrary* .
- de Palma, A., Ben-Akiva, M., Brownstone, D., Holt, C., Magnac, T., McFadden, D., Moffatt, P., Picard, N., Train, K., Wakker, P. and Walker, J. (2008). Risk, uncertainty and discrete choice models, *Marketing Letters* **19**: 269–285. 10.1007/s11002-008-9047-0.
- Dellaert, B. G., Arentze, T. A., Bierlaire, M., Borgers, A. W. and Timmermans, H. J. (1998). Investigating consumers' tendency to combine multiple shopping purposes and destinations, *Journal of Marketing Research* **35**(2): 177–188.
- den Berg, G. J. V. (2001). Chapter 55 duration models: specification, identification and multiple durations, Vol. 5 of *Handbook of Econometrics*, Elsevier, pp. 3381 – 3460.
- Denis, C. (2009). Facial expression recognition project: Collect a database, *Technical report*, Transport and Mobility Laboratory (TRANSP-OR), EPFL, EPFL ENAC INTER TRANSP-OR, Station 18, CH-1015 Lausanne, Switzerland.
- Dentcheva, D. and Ruszczyński, A. (2006). Portfolio optimization with stochastic dominance constraints, *Journal of Banking & Finance* **30**(2): 433 – 451. Risk Management and Optimization in Finance.
- Dosey, M. A. and Meisels, M. (1969). Personal space and self-protection, *Journal of Personality and Social Psychology* **11**(2): 93–97.
- Ekman, P. and Friesen, W. (1978). *Facial action coding system: A technique for the measurement of facial movement*, Consulting Psychologists Press, Palo Alto, California.
- El-Ghaoui, L., Oks, M. and Oustry, F. (2003). Wors-case value-at-risk and robust portfolio optimization: a conic programming approach, *Operation Research* **51**(4): 543 – 556.
- El-Yaniv, R. (1998). Competitive solutions for online financial problems, *ACM Comput. Surv.* **30**: 28–69.

- Embrechts, P. and Schmidli, H. (1994). Modelling of extremal events in insurance and finance, *Mathematical Methods of Operations Research* **39**: 1–34.
- Fasel, B. and Luettin, J. (2003). Automatic facial expression analysis: a survey, *Pattern Recognition* **36**(1): 259 – 275.
- Fermanian, J.-D. and Scaillet, O. (2004). Some Statistical Pitfalls in Copula Modeling for Financial Applications, *SSRN eLibrary* .
- Friman, M. and Garling, T. (2001). Frequency of negative critical incidents and satisfaction with public transport services, *Journal of Retailing and Consumer Services* **8**(2): 105 – 114.
- Frittelli, M. and Gianin, E. R. (2002). Putting order in risk measures, *Journal of Banking & Finance* **26**(7): 1473 – 1486.
- Giles, C. L., Lawrence, S. and Tsoi, A. C. (2001). Noisy time series prediction using recurrent neural networks and grammatical inference, *Machine Learning* **44**: 161–183.
- Greene, W. H. and Hensher, D. A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit, *Transportation Research Part B: Methodological* **37**(8): 681 – 698.
- Hartnett, J. J., Bailey, K. G. and Hartley, C. S. (1974). Body height, position, and sex as determinants of personal space, *Journal of Psychology* **87**: 129–136.
- Helbing, D., Buzna, L., Johansson, A. and Werner, T. (2005). Self-organized pedestrian crowd dynamics: Experiments, simulations and design solutions, *Transportation Science* **39**(1): 1–24.
- Helbing, D., Farkas, I., Molnar, P. and Vicsek, T. (2002). Simulation of pedestrian crowds in normal and evacuation simulations, *in* M. Schreckenberg and S. Sharma (eds), *Pedestrian and Evacuation Dynamics*, Springer, pp. 21–58.
- Helbing, D. and Molnar, P. (1995). Social force model for pedestrian dynamics, *Physical review E* **51**(5): 4282–4286.
- Hensher, D. A. (2010). Attributes processing, heuristics and preference construction in choice analysis, *in* S. Hess and A. Daly (eds), *Choice modelling: the state-of-the-art and the state-of-practice*, Emerald, pp. 35–70.
- Hensher, D. and Jones, S. (2004). Predicting firm financial distress: a mixed logit model, *The accounting review* **79**(4): 1011 – 1038.
- Herman, R. and Rothery, R. W. (1965). Car following and steady-state flow, *Proceedings on 2nd international symposium on the theory of traffic flow*, pp. 1–11.

- Hommel, C. H. (2006). Chapter 23 heterogeneous agent models in economics and finance, Vol. 2 of *Handbook of Computational Economics*, Elsevier, pp. 1109 – 1186.
- Hoogendoorn, S., Bovy, P. and W.Daamen (2002). Microscopic pedestrian wayfinding and dynamics modelling, in M. Schreckenberg and S. Sharma (eds), *Pedestrian and Evacuation Dynamics*, Springer, pp. 123–155.
- Hoogendoorn, S. P. and Daamen, W. (2005). Pedestrian behaviour at bottlenecks, *Transportation Science* **39**(2): 147–159.
- Horowitz, J. J., Duff, D. F. and Stratton, L. O. (1964). Body buffer zone: Exploration of personal space, *Archives of General Psychiatry* **11**(6): 651–656.
- Huang, Z., Chen, H., Hsu, C.-J., Chen, W.-H. and Wu, S. (2004). Credit rating analysis with support vector machines and neural networks: a market comparative study, *Decision Support Systems* **37**(4): 543 – 558. Data mining for financial decision making.
- J.L.Walker (2001). *Extended Discrete Choice Models: Integrated Framework, Flexible Error Structures, and Latent Variables*, PhD thesis, Massachusetts Institute of Technology.
- Johnsen, T. and Melicher, R. W. (1994). Predicting corporate bankruptcy and financial distress: Information value added by multinomial logit models, *Journal of Economics and Business* **46**(4): 269 – 286.
- Keltner, D. Ekman, P. (2000). Facial expression of emotion, *Handbooks of emotions*, M.Lewis & J.M.Havilland, pp. 236–249.
- Kim, K. (2003). Financial time series forecasting using support vector machines, *Neurocomputing* **55**(1-2): 307 – 319. Support Vector Machines.
- Kretz, T., Grünebohm, A., Kaufman, M., Mazur, F. and Schreckenberg, M. (2006). Experimental study of pedestrian counterflow in a corridor, *Journal of Statistical Mechanics: Theory and Experiment* **2006**(10): P10001.  
**URL:** <http://stacks.iop.org/1742-5468/2006/P10001>
- Kretz, T., Hengst, S. and Vortisch, P. (2008). Pedestrian flow at bottlenecks - validation and calibration of vissim's social force model of pedestrian traffic and its empirical foundations.  
**URL:** <http://www.citebase.org/abstract?id=oai:arXiv.org:0805.1788>
- Lamoureux, C. and Lastrapes, W. (1990). Heteroskedasticity in stock return data: volume versus garsh effects, *The journal of finance* **45**(1): 221.

- Lee, G. (1966). A generalization of linear car following theory, *Operations Research* **14**: 595–606.
- Lerner, J. S. and Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgement and choice, *Cognition and Emotion* **14**: 473 – 493.
- Li, T., Jeng, Y. and Chang, S. (2001). Simulating virtual human crowds with a leader-follower model, *Proceedings of Computer Animation Conference*, Computer Graphics Society and IEEE, pp. 93–102.
- Lo, A. W. (2004). The Adaptive Markets Hypothesis: Market Efficiency from an Evolutionary Perspective, *Journal of Portfolio Management*, *Forthcoming* .
- Louviere, J. J., Hensher, D. A. and Swait, J. D. (2000). *Discrete Choice Analysis: Theory and Application to Travel Demand*, Cambridge university press.
- MacInnis, D. J., Moorman, C. and Jaworski, B. J. (1991). Enhancing and measuring consumers' motivation, opportunity, and ability to process brand information from ads, *The Journal of Marketing* **55**(4): 32–53.
- McFadden, D. (1978). Modelling the choice of residential location, in A. Karlquist *et al.* (ed.), *Spatial interaction theory and residential location*, North-Holland, Amsterdam, pp. 75–96.
- McNeil, A. J. and Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach, *Journal of Empirical Finance* **7**(3-4): 271 – 300.
- Mellers, B. A. and McGraw, A. P. (2001). Anticipated emotions as guides to choice, *Current Directions in Psychological Science* **10**(6): 210–214.
- Mikosch, T. and Stărică, C. (2004). Nonstationarities in financial time series, the long-range dependence, and the igarch effects, *Review of Economics and Statistics* **86**(1): 378–390.
- Min, S.-H., Lee, J. and Han, I. (2006). Hybrid genetic algorithms and support vector machines for bankruptcy prediction, *Expert Systems with Applications* **31**(3): 652 – 660.
- Miwa, H., Itoh, K., Matsumoto, M., Zecca, M., Takanobu, S., Rocella, S., Carrozza, P., Dario, A. and A., T. (2004). Effective emotional expressions with emotion expression humanoid robot we-4rii - integration of humanoid robot hand reh-1, *International Conference on Intelligent Robots and Systems*, Vol. 3, pp. 2203–2208.

- Nagai, R., Fukamachi, M. and Nagatani, T. (2005). Experiment and simulation for counterflow of people going on all fours, *Physica A: Statistical Mechanics and its Applications* **358**(2-4): 516–528.
- Newell, G. (1961). Nonlinear effects in the dynamics of car following, *Operations Research* **9**: 209–229.
- Pantic, M. and Patras, I. (2006). Dynamics of facial expression: recognition of facial actions and their temporal segments from face profile image sequences, *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* **36**(2): 433–449.
- Penn, A. and Turner, A. (2002). Space syntax based agent simulation, in M. Schreckenberg and S. Sharma (eds), *Pedestrian and Evacuation Dynamics*, Springer, pp. 99–114.
- Perold, A. (1984). Large-scale portfolio optimization, *Management science* **30**(10): 1143.
- Phillips, J. R. (1979). An exploration of perception of body boundary, personal space, and body size in elderly persons, *Perceptual and Motor Skills* **48**: 299–308.
- Reimer, B., Coughlin, J. and Mehler, B. (2009). Development of a driver aware vehicle for monitoring, managing & motivating older operator behavior, *Technical report*, ITS America.
- Robin, T., Antonini, G., Bierlaire, M. and Cruz, J. (2009). Specification, estimation and validation of a pedestrian walking behavior model, *Transportation Research Part B: Methodological* **43**(1): 36–56.
- Robin, T. and Bierlaire, M. (2011). Modeling the behavior of investors, *Technical report*, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne.
- Robin, T., Bierlaire, M. and Cruz, J. (2010). Dynamic facial expression recognition with a discrete choice model, *Technical Report TRANSP-OR 100423*, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne.
- Rockafellar, R. and Uryasev, S. (2000). Optimization of conditional value-at-risk, *Journal of Risk* pp. 21 – 41.
- Sanders, J. L. (1976). Relationship of personal space to body image boundary definiteness, *Journal of Research in Personality* **10**: 478–481.
- Schadschneider, A. (2002). Cellular automaton approach to pedestrian dynamics — Theory, in M. Schreckenberg and S. Sharma (eds), *Pedestrian and Evacuation Dynamics*, Springer, pp. 75–86.

- Shapiro, R. (1978). Direct linear transformation method for three-dimensional cinematography, *Res. Quart.* **49**: 197–205.
- Shiller, R. J. (1999). Chapter 20 human behavior and the efficiency of the financial system, Vol. 1, Part 3 of *Handbook of Macroeconomics*, Elsevier, pp. 1305 – 1340.
- Shiller, R. J. (2003). From efficient markets theory to behavioral finance, *The Journal of Economic Perspectives* **17**: 83–104(22).
- Shin, K.-S., Lee, T. S. and jung Kim, H. (2005). An application of support vector machines in bankruptcy prediction model, *Expert Systems with Applications* **28**(1): 127 – 135.  
**URL:** <http://www.sciencedirect.com/science/article/B6V03-4DB4VS8-2/2/4d6e679fee1b7982e7ab818c786e51af>
- Small, D. and Verrochi, N. (2009). The face of need: facial emotion expression on charity advertisements, *journal of marketing research* **XLVI**: 777 – 787.
- Sommer, R. (1969). *Personal Space: The behavioral bases of design*, Prentice Hall, Englewood Cliffs, NJ.
- Sorci, M., Antonini, G., Cruz, J., Robin, T., Bierlaire, M. and Thiran, J.-P. (2010). Modelling human perception of static facial expressions, *Image and Vision Computing* **28**(5): 790–806.
- Sorci, M., Robin, T., Cruz, J., Bierlaire, M., Thiran, J.-P. and Antonini, G. (2010). Capturing human perception of facial expressions by discrete choice modelling, in S. Hess and A. Daly (eds), *Choice Modelling: The State-of-the-Art and the State-of-Practice*, Emerald Group Publishing Limited, pp. 101–136. ISBN:978-1-84950-772-1.
- Teknomo, K. (2002). *Microscopic Pedestrian Flow Characteristics: Development of an Image Processing Data Collection and Simulation Model*, PhD thesis, Tohoku University, Japan, Sendai.
- Teknomo, K., Takeyama, Y. and Inamura, H. (2000). Review on microscopic pedestrian simulation model, *Proceedings Japan Society of Civil Engineering Conference*, Morioka, Japan.
- T.Kanade, J.Cohn, Y.-L. (2000). Comprehensive database for facial expression analysis, *Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition (FG'00)*, pp. 46–53.
- Tojo, T., Matsusaka, Y., Ishii, T. and Kobayashi, T. (2000). A conversational robot utilizing facial and body expressions, *Systems, Man, and Cybernetics, 2000 IEEE International Conference on*, Vol. 2, pp. 858–863.

- Toledo, T. (2003). *Integrated Driving Behavior Modeling.*, PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.
- Toledo, T., Koutsopoulos, H. N. and Ben-Akiva, M. (2003). Modeling integrated lane-changing behavior, *Transportation Research Record* **1857**: 30–38.
- Turner, A. (2001). Angular analysis, *In Proceedings 3rd International Symposium on Space Syntax*, pp. 30.1–30.11.
- Venegas, S., Antonini, G., Thiran, J.-P. and Bierlaire, M. (2005). Automatic pedestrian tracking using discrete choice models and image correlation techniques, *in* B. S. and B. H. (eds), *Machine Learning for Multimodal Interaction*, Vol. 3361 of *Lecture Notes in Computer Science*, Springer, pp. 341 – 348. ISBN:978-3540245094.
- Veronesi, P. (1999). Stock market overreactions to bad news in good times: a rational expectations equilibrium model, *Review of Financial Studies* **12**(5): 975–1007.
- Wagner, W. P., Otto, J. and Chung, Q. B. (2002). Knowledge acquisition for expert systems in accounting and financial problem domains, *Knowledge-Based Systems* **15**(8): 439 – 447.
- Wallhoff, F. (2004). Fgnet-facial expression and emotion database, *Technical report*, Technische Universitt Mnchen.  
**URL:** <http://www.mmk.ei.tum.de/waf/fgnet/feedtum.html>
- Webb, J. D. and Weber, M. J. (2003). Influence of sensor abilities on the interpersonal distance of the elderly, *Environment and behavior* **35**(5): 695–711.
- Weinberg, P. and Gottwald, W. (1982). Impulsive consumer buying as a result of emotions, *Journal of Business Research* **10**(1): 43 – 57.  
**URL:** <http://www.sciencedirect.com/science/article/B6V7S-45JWVJH-2W/2/c5ad26cf95e71a37ca1cdb072a7254b>
- Wen, C.-H. and Koppelman, F. S. (2001). The generalized nested logit model, *Transportation Research Part B* **35**(7): 627–641.
- Werner, J. (1985). Equilibrium in economies with incomplete financial markets, *Journal of Economic Theory* **36**(1): 110 – 119.
- Whynes, D., Reedand, G. and Newbold, P. (1996). General practitioners' choice of referral destination: A probit analysis, *Managerial and Decision Economics* **17**(6): 587.
- Wiseman, R. (2007). *Quirkology, The curious science of everyday lives*, Macmillan.
- Zopounidis, C. (1999). Multicriteria decision aid in financial management, *European Journal of Operational Research* **119**(2): 404 – 415.





# CURRICULUM VITAE

## THOMAS ROBIN

### Address

Ch. des Clochetons 41  
1004 LAUSANNE (VD)  
Switzerland  
Phone: +41 21 693 24 35  
E-mail: thomas.robin@epfl.ch  
Homepage: <http://transp-or.epfl.ch/personnal.php?Person=ROBIN>

### Identity

Gender: male  
Date of birth: 6th December 1983  
Place of birth: Chateau-Thierry (France)  
Nationality: French

### Education

- 2007-2011 *PhD Thesis in Mathematics (behavioral modeling)*  
**École Polytechnique Fédérale de Lausanne (EPFL)**
- 2002-2006 *Master*  
*Thesis: Identification du nombre et de l'emplacement de centres de redistribution*  
**École nationale supérieure des techniques industrielles et des mines d'Alès (ENSTIMA), France**
- 2001-2002 *Classe préparatoire physique, chimie et science de l'ingénieur (PCSI)*  
**Lycée Masséna, Nice, France**

### Professional experience

- 2006-2011 **EPFL**  
*Research assistant of Prof. M. Bierlaire (Transport and Mobility Laboratory)*  
*Supervision of semester and diploma projects*
- 2008-2011 **EPFL - Advanced Continuing Education Course**  
*Assistant for the course*  
*"Discrete Choice Analysis: Predicting Demand and Market Shares"*