

# Examples of Migration Matrices Models and their Performance in Credit Risk Analysis

U. GRZYBOWSKA<sup>a</sup>, M. KARWAŃSKI<sup>a</sup> AND A. ORŁOWSKI<sup>a,b</sup>

<sup>a</sup>Katedra Informatyki SGGW, Nowoursynowska 159, 02-776 Warszawa, Poland

<sup>b</sup>Instytut Fizyki PAN, al. Lotników 32/46, 02-668 Warszawa, Poland

Credit risk models used in banks are based on probability models for occurrence of default. A vast class of the models used in practice (e.g., Credit Metrics) is based on the notion of intensity. In 1997 Jarrow applied Markov chain approach to analyze intensities. The key problem that arises is the selection of appropriate estimators. Within the Markov approach among the most frequently used estimators of a migration matrix are cohort and duration estimators. Migration matrices can also be obtained with help of statistical longitudinal models (GLMM) in which states (rating classes) in discrete time points are regarded as matched pairs. In this paper we compare Markov chain models and GLMM models and the influence of their application on bank portfolio evaluation.

PACS: 02.50. Ga, 87.55.N-, 02.10.Yn

## 1. Introduction

In the late 80-ties measures were undertaken to organize both bank proceedings and methods of evaluating resulting financial risk. The prevailing aim to achieve the sustainable growth keeping up with the assumed financial performance was reached with help of mathematical models. These models allowing correct risk evaluation became an indispensable tool after the financial crisis of 2008.

At the beginning the changes in values of investigated loans were modeled in frames of the Merton theory. In the 90 - ties more attention was put to rating systems measuring credit rating and the process of changing rating in time. For the description of these changes intensity  $\lambda$  was used, namely:

$$P(\text{default} \in [t, t + \Delta t] | H_t) \approx \lambda(t)\Delta t,$$

where  $H_t$  denotes historical information [1].

The easiest way to describe intensities is to consider a finite state Markov space [2-3]. The methods of evaluating intensities were based on the observed migration matrix. There are several methods to estimate migration matrices. The most popular frequency methods are based on directly observed historical performance. The other approach is offered by statistical methods of generalized longitudinal models GLMM.

## 2. Migration matrices and their application

In this paper we try to compare influence of various migration matrices models on risk estimation. At the moment there are four models used in practice: Merton's model, CreditRisk+, CreditMetrics, and CreditPortfolio View. The last two are based on migration matrices.

In developed countries estimators of migration matrices are published by external rating agencies, e.g., S&P [4], Moody's [5-6]. The data is then used by banks to model probability of default. Unfortunately, no such data is available for Poland, mainly due to insufficient number of rated agents. Therefore banks rely on their own internal models. Unfortunately, the notion of default has not yet been defined and the available data concerning migration to default is not reliable as agents try to hide their true financial situation when in danger of default. Moreover internal ratings are constructed with a method called Point-in-Time, i.e., for a given economic situation of the country.

Due to the above remarks we decided to base our research on generated data. Apparently this approach is biased and should not be applied to model evaluation in terms of goodness-of-fit but can be used to discover if the differences between the models are considerable.

## 3. Markov chain approach

We will use the rating consisting of 8 states, proposed in [5], and described in details in [7] and [8]. In literature devoted to the subject of credit rating it is assumed that rating changes can be described with help of a finite homogenous Markov chain with the transition matrix  $P = [p_{ij}]$  based on historical data. The assumption of time homogeneity is usually false, mainly due to external economic factors such as cycles of economy [9–11]. Statistical tests for time invariance (see, e.g., [12]) require a sufficiently large sample for each matrix entry. This cannot be achieved as the off-diagonal events are scarce, hence the elements on the diagonal are estimated with high precision while the entries far away from the

diagonal are estimated with low precision. Special attention should be paid to entries equal to zero, which mean no transition. This is not desirable. Credit migration matrices are diagonally dominant, nonnegative, stochastic, their largest eigenvalue is equal to 1, and they have an absorbing state (default).

There are several methods for estimating migration matrices for count data. The most frequently used are the cohort method and two variants of duration methods – parametric (imposing time homogeneity) and non-parametric (relaxing time homogeneity), known as Nelson-Aalen-Johansen estimator. In this paper we will show that migration matrices can also be obtained with help of longitudinal models (GLMM). The advantage of such an approach is that it allows to incorporate economic indices thus making the estimation more adjust to given economic situation.

### 3.1. Cohort method

Let  $n_{ij}$  be the number of agents rating  $i$  at  $t-1$  and rating  $j$  at  $t$ ,  $\Delta t = 1$ . The estimated transition probability of migrating from  $i$  to  $j$  is:

$$p_{ij} = \frac{n_{ij}}{\sum_j n_{ij}}. \quad (1)$$

Hence the probability of migrating from  $i$  to  $j$  is the proportion of the number  $n_{ij}$  of all agents that migrated from  $i$  to  $j$ , to all agents whose rating was  $i$  at  $t-1$ , i.e.,  $\sum_j n_{ij} = N_i$ .

The cohort method [11, 13, 14], which became the industry standard for count data, has a drawback. Any rating change activity which occurs within the period  $\Delta t$  is ignored.

### 3.2. Migration matrices for $T$ periods

In [15] the following estimator of transition probabilities within  $T$  periods was introduced

$$p_{ij} = \frac{N_{ij}}{N_i} = \frac{\sum_{t=1}^T n_{ij}(t)}{\sum_{t=1}^T N_i(t)}. \quad (2)$$

In the sequel, the above defined estimator will be called Plug-in, and the transition matrix obtained by (2) will be denoted by PI.

### 3.3. Duration methods

Unlike the cohort method, the duration methods use full rating history. If the assumption of time homogeneity is imposed, transition probabilities can be described via a generator or an intensity matrix  $\Lambda$  [14, 16]. Transition matrix  $P(t)$ , for an arbitrary  $t \geq 0$ , can be written as

$$P(t) = \exp(\Lambda(t)) = I + t\Lambda(t) + t^2 \frac{\Lambda^2(t)}{2!} + t^3 \frac{\Lambda^3(t)}{3!} + \dots, \quad (3)$$

$t \geq 0$ , where the entries of  $\Lambda$  satisfy  $\lambda_{ij} \geq 0$  for  $i \neq j$ ,

and  $\lambda_{ii} = -\sum_{j \neq i} \lambda_{ij}$ . The estimate of  $\Lambda$  is given by  $\lambda_{ij} = \frac{n_{ij}(T)}{\int_0^T Y_i(s) ds}$ , where  $Y_i(s)$  is the number of agents rating  $i$  at time  $s$ , and  $n_{ij}(T)$  is the total number of transitions over the period  $T$  from  $i$  to  $j$  for  $i \neq j$ .

For the time-inhomogeneous case, the Nelson-Aalen-Johansen estimator can be used. The transition matrix for the period  $s$  to  $t$  is computed as:  $P(s, t) = \prod_{i=1}^m (I + \Delta \hat{A}(T_i))$ . The off-diagonal elements of  $\Delta \hat{A}$  matrix are given by

$$a_{jk} = \frac{\Delta N_{jk}(T_i)}{Y_j(T_i)}, \quad (4)$$

where  $T_i$  is the jump time in the interval  $(s, t]$ ,  $Y_j(T_i)$  is the number of agents in the rating class  $j$  at a time just prior to  $t$ , and  $\Delta N_{jk}(T_i)$  is the number of transitions from  $j$  to  $k$  in  $T_i$ ,  $m$  is the number of days with at least one mobility [14, 16]. The raw sum is equal to zero.

## 4. Migration matrices and GLM

From the statistical viewpoint the transition probabilities  $p_{ij} = P(Y_t = j | Y_{t-1} = i)$  can be regarded as matched pairs. Therefore it seems reasonable to apply longitudinal GLMM models. These models are generalizations of generalized linear models GLM [18]. In our work we have chosen two types of GLMM models: Marginal Models and Transitional Models.

### 4.1. Marginal GLM

Marginal models belong to a class of generalized linear models, for which both expectation and variance is defined independently. Let the probability distribution  $f(\cdot)$  be in the class of one parameter exponential family [18, 19], let  $\mu_{ij} = b'(\theta_{ij})$  be called the marginal average and let the expectation  $E(Y_{ij} | \mu_{ij})$  for a given marginal average  $\mu_{ij}$  depend on exogenous variables  $X_{ij}$  via a copula  $g(\cdot)$ :  $g(\mu_{ij}) = \eta_{ij} = X_{ij}\beta$ . Let the variance  $Var(Y_{ij})$  depend on the value of marginal average via a given function  $\nu$ :  $Var(Y_{ij}) = \nu(\mu_{ij})\varphi$ . The class of models defined above is called marginal models [18]. Here we have chosen a marginal model consisting of the average and variance described above to produce migration matrix with 8 states and one absorbing in it. The average can be described in terms of an equation with exogenous variables depending on rating classes:

$$\log \{E(Y_{ij})\} = \alpha_i \cdot rating_i^t + \beta_j \cdot rating_j^{t-1} + \theta_{ij} \cdot factor_{ij} + \tau_{ij} \cdot \delta_{ij},$$

where

$$\delta_{ij} = \begin{cases} 1 & \text{for } i \leq j \\ 0 & \text{for } i > j \end{cases}.$$

For the likelihood function the Poisson distribution was taken. We have  $f(Y_{ij} | \mu_{ij}) = \exp[Y_{ij} \log(\mu_{ij}) - \mu_{ij} - \log(\Gamma(Y_{ij} + 1))]$ ,

with the copula function  $g(\cdot) = \log(\cdot)$  and the parameter  $\phi = 1$ .

#### 4.2. Transitional GLM models

In the transitional models all the assumptions concerning generalized linear models are valid except that density function  $f(Y_{ij})$  is conditional and depends on the past as well as on additional external variables  $X_{ij}$  [18, 19]. We assume that the past is given as  $H_{ij} = \{Y_{i1}, Y_{i2}, \dots, Y_{i,j-1}\}$ , then the  $f(Y_{ij} | H_{ij})$  belongs to the class of one parameter exponential family. The conditional average can be represented in the form:  $g(E(Y_{ij} | H_{ij})) = X_{ij}\beta + \sum_{r=1}^s \alpha_r f_r(H_{ij})$ , for some copula  $g(\cdot)$  and transforming functions  $f(\cdot)$ .

For the need of the presented paper and the aim of migration matrices modeling (with 8 states and one absorbing in it), we have chosen equations with identity transforming functions and exogenous variables depending on rating.

The negative binomial distribution was used to construct the likelihood function:

$$f(Y_{ij}) = Y_{ij} \log(k\mu_{ij}) - (Y_{ij}) \log(1 + k\mu_{ij}) + \log\left(\frac{\Gamma(Y_{ij} + 1/k)}{\Gamma(1/k)\Gamma(Y_{ij} + 1)}\right),$$

where  $Var(Y_{ij}) = \mu_{ij} + k\mu_{ij}^2$ ,  $g(\cdot) = \log(\cdot)$ , and  $\phi = 1$ .

### 5. Generators of Markov chains and their application

Whatever estimation method is used the shortest time interval for estimating a migration matrix is one year (one period). The reason is that for shorter periods no enough transitions is observed to get reliable results. In practice, however, one often needs to use migration matrices for shorter periods than one year. This can be done provided that for a transition matrix  $P$  its generator  $Q$  such that  $P = \exp(Q)$  is obtained. Then, the formula

$$P(t) = \exp(tQ) = I + tQ + \frac{(tQ)^2}{2!} + \frac{(tQ)^3}{3!} + \dots \quad (5)$$

gives transition matrix  $P(t)$  for an arbitrary  $t \geq 0$ . In other words finding a generator  $Q$  for a one period matrix  $P$ , i. e., for a discrete time matrix, allows to obtain a continuous time transition matrix [16].

#### Theorem 1 [16]

Let  $P$  be an  $N \times N$  transition matrix and let  $S = \max\{(a-1)^2 + b^2; \text{ where } a + ib, (a, b \in \mathbb{R}) \text{ is an eigenvalue of } P\}$ . Suppose that  $S < 1$ . Then the series

$$\tilde{Q} = (P - I) - \frac{(P - I)^2}{2} + \frac{(P - I)^3}{3} - \frac{(P - I)^4}{4} + \dots \quad (6)$$

converges geometrically quickly and gives rise to an  $N \times N$  matrix  $\tilde{Q}$  such that  $\exp(\tilde{Q}) = P$ .

In our examples eigenvalues were real and the largest was equal to 1. Therefore we were able to find generators of matrices considered in our paper. One has to mention that in [16] and [17] a number of theorems and criterions was given for existence and uniqueness of transition matrix generators.

#### Theorem 2 [16]

Let  $P$  be a transition matrix and let  $p_{ii} > 0.5$  for each  $i$ . Then  $S < 1$ , i.e., the series (6) is convergent.

The off-diagonal entries may be negative. This requires amendment, e.g., defined in [16]. Once we have obtained  $\tilde{Q}$  with possibly some negative off-diagonal entries, we can define a new matrix  $Q$  by setting

$$q_{ij} = \begin{cases} \max(\tilde{q}_{ij}, 0), & \text{for } j \neq i, \\ \tilde{q}_{ii} + \sum_{j \neq i} \min(\tilde{q}_{ij}, 0), & \text{for } i = j. \end{cases} \quad (7)$$

The new matrix will have nonnegative off-diagonal entries, will have row-sums 0 but it will no longer satisfy  $\exp(Q) = P$  exactly. Still after the corrections are made we compute transition matrix  $P$ , for arbitrary  $t \geq 0$  using formula (5).

## 6. Data description and results

The comparison of migration matrices was done based on simulated data obtained by a migration matrix given in Table I [5]; in this simulation 2000 clients were generated with their history for 5 periods. The data was then used to calculate migration matrices\*.

First, one year transition matrices were calculated based on simulated data. These matrices were then used to obtain migration matrices estimated with described methods: cohort for T periods (PI), duration (parametric (D) and non-parametric (AN)), marginal GLM model (M GLM) and transitional GLM model (T GLM). Also the average of one year migration matrices was calculated (AR). Standard errors were calculated for entries of the matrices with boot-strap method [20] (matrices with B at the end) and generators for each matrix were determined with correction introduced in [16]. Note that transition matrices for duration methods are already continuous time matrices. Then transition matrices obtained with generators were calculated (denoted by G) to enable probability estimates in an arbitrary moment. All together 13 matrices were considered.

In order to obtain the one year Aalen-Nelson matrix, first we calculate its generator, then take its logarithm, next we divide the matrix by four and compute the one year AN matrix as the exponent of the result following formulas (5) and (6). The obtained matrix (Table II, III and IV) was then compared with other matrices.

\* Application of real life data requires extra treatment which is beyond the scope of the paper.



TABLE IV

One year Aalen-Nelson migration matrix

	Aaa	Aa	A	Baa	Ba	B	Caa_C	D
Aaa	0.9037	0.0769	0.0174	0.0008	0.0009	0	0	0
Aa	0.0064	0.8949	0.0805	0.0119	0.0041	0.0005	0	0.0013
A	0.0002	0.0125	0.9177	0.0466	0.0171	0.0048	0.0003	0.0005
Baa	0.0001	0.0048	0.0448	0.8724	0.0549	0.0179	0.0012	0.0035
Ba	0.0006	0.0008	0.0061	0.0243	0.8807	0.0606	0.0053	0.0211
B	0.0001	0.0006	0.0018	0.0053	0.0391	0.8657	0.0114	0.0756
Caa_C	0	0.0002	0.0015	0.0061	0.0169	0.0297	0.7801	0.165
D	0	0	0	0	0	0	0	1

The measures defined by (8) and (9) will be denoted by L1 and L2 respectively. We propose also a metric based on maximal difference between entries of matrices

$$M_{\max}(P, \hat{P}) = \max_{ij} |p_{ij} - \hat{p}_{ij}|.$$

In [13] the authors defined a new metric based on SVD, namely on the average of singular values of the so called mobility matrix  $\tilde{P} = P - I$

$$M_{SVD}(P) = \frac{\sum_{i=1}^N \sqrt{\lambda_i(\tilde{P}^T \tilde{P})}}{N}, \tag{10}$$

where  $\lambda_i$  are eigenvalues of  $\tilde{P}^T \tilde{P}$ . The metric based on singular values allows to measure the mobility encapsulated by transition matrix. For matrix comparison we define

$$M_{SVD}(P, \hat{P}) = |M_{SVD}(P) - M_{SVD}(\hat{P})| \tag{11}$$

In the rest of this paper we compare estimated transition matrices and their performance. The first comparison is based on measures between matrices, the second on business performance. To compare matrices we calculate their distance to PI matrix. We calculate also distances between cohort matrices and duration matrices and their generators as well as between GLM matrices and Markov chain matrices.

TABLE V

Comparison of discrete time and continuous time migration matrices

	PI	AR	T GLM
	PI G	AR G	T GLMG
L1	0.0010	0.0011	0.0046
L2	0.0016	0.0017	0.0063
max	0.0027	0.0028	0.0151
SVD	0.0002	0.0003	0.0003

The comparison of distances between matrices and their continuous time counterparts obtained via generators confirms the fact that these matrices almost do not differ (see Table V).

The differences encapsulated in Table VI indicate that it really matters if the matrix was calculated in a straightforward way or by a bootstrap method.

TABLE VI

Comparison of bootstrap generated matrices and ordinary count data matrices

	PI	AN
	PI B	AN B
L1	0.0309	0.0584
L2	0.0454	0.0867
max	0.0742	0.2069
SVD	0.0354	0.1549

TABLE VII

Examples of matrices with large migration differences

	AN	D
	AN B	AN B
L1	0.0371	0.0204
L2	0.0543	0.0361
max	0.2132	0.2492
SVD	0.1795	0.2133

In Table VIII distances calculated between duration matrices: parametric (D) and nonparametric (AN), confirm the conclusion of [13] that these matrices encapsulate similar migration power. The same conclusion holds true for average (AR) and PI matrices, which additionally have very similar entries, so in fact they do not differ much.

TABLE VIII

Examples of matrices with similar migration impact

	AN	D	AR	M GLM
	PI B	AN	PI	PI
L1	0.03	0.0467	0.0021	0.0129
L2	0.0408	0.06	0.0049	0.0146
max	0.0651	0.1215	0.0109	0.0233
SVD	0.0257	0.05	0.0009	0.0003

TABLE IX

Distances between PI matrix and other matrices

	AN B	M GLM	T GLM
	PI B	PI B	PI B
L1	0.0635	0.0358	0.0523
L2	0.1016	0.0467	0.0691
max	0.2127	0.0738	0.1439
SVD	0.1293	0.0357	0.0103

TABLE X

Comparison of Markov chain and GLM matrices

	AN B	AN B	PI B	PI B
	T GLM	M GLM	M GLM	T GLM
L1	0.0886	0.0531	0.0358	0.0523
L2	0.1416	0.0796	0.0467	0.0691
max	0.3566	0.2139	0.0738	0.1439
SVD	0.1190	0.1650	0.0357	0.0103

The comparison of matrices based on measures does not necessarily reflect their performance (no metric can capture the differences in the last column of migration matrix which describes probability of migration to default). Therefore, after calculating generators of matrices, we estimate transition probabilities as the function of time.

For practical application the most important are probabilities of migration to default. In Figs. 1 and 2 probabilities of migration to default from Ba and Caa\_C in the function of time are shown. Apart from the mean value, the confidence intervals were estimated with the bootstrap method. Large values of default probabilities result from the fact that the models are extrapolated to long time horizon. In practice, portfolios from “banking book” are monitored over the horizon up to 10 years.

As can be seen models give different results. Particularly large differences exist for portfolios in the lower rating grade such as the class Caa\_C. The probabilities of migration obtained from models are the input information to calculate the economic capital and thus directly translate into financial results. The differences observed in simulations may result in significant changes (up to 10%) in the costs of maintenance the portfolios. The problem out of the scope of this article is to incorporate the value of migration into liquidity risk models.

It is not only the average value of migration, but also the precision of their estimation. Noticeable differences depend on the probability distributions of default for the simulated data. In practice the situation is more complicated because the distribution of migration rates change over time and depend on many factors. Models based on Markov chain approach do not allow incorporating exogenous variables into the models. The practice however requires considering economic factors in credit risk modeling. This is possible within the GLMM approach.

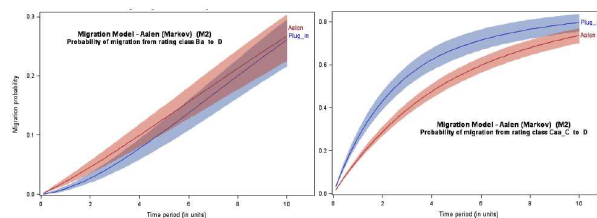


Fig. 1. Comparison of AN and Plug-in models with respect to the probability of migration to default (D) from Ba / Caa\_C.

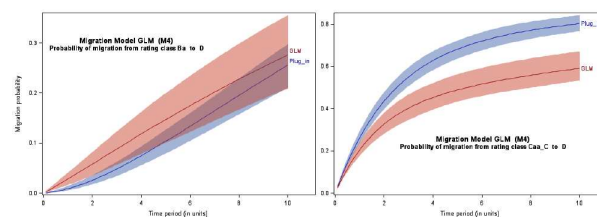


Fig. 2. Comparison of T GLM and Plug-models with respect to the probability of migration to default (D) from Ba / Caa\_C.

For periods of economic stability such as in 2006-2008, the more appropriate were the models of Aalen, in periods when there were rapid changes such as the crisis of 2008/2009 better results could be obtained in GLMM models. Therefore additional studies are required to determine which internal and external factors should be included in the models.

### 8. Summary

Application of migration matrices to risk estimation becomes more and more popular. There are several methods of estimating migration matrices. In our paper we have presented these methods and compared obtained matrices. The comparison of matrices based on distance measures does not reflect their business performance because no metric can capture the differences in the last column of migration matrix which describes probability of migration to default. From the viewpoint of applications much better results can be obtained with measures based on estimation of convergence to default. Such measures involve basic properties of absorbing Markov chains and will be investigated in our next paper.

### References

- [1] R. A Jarrow, D.Lando, S. M.Turnbull, A Markov model for the term structure of credit risk spreads, *Review of Financial Studies*, **10**, 2 (1997).
- [2] W. Feller, *Wstęp do rachunku prawdopodobieństwa*, PWN, Warszawa 1966.
- [3] M. Iosifescu, *Skończone łańcuchy Markowa*, WNT, Warszawa 1987.
- [4] S&P, [www.standardandpoors.com/ratings/en/eu](http://www.standardandpoors.com/ratings/en/eu) .

- [5] Moody's Investors Service, *Historical Default Rates of Corporate Bond Issuers, 1920-1999*, p. 25, Special Comment 2000.
- [6] Moody's, [www.moody.com](http://www.moody.com) .
- [7] Basel Committee on Banking Supervision, *The Internal Ratings-Based Approach. Consultative Document*, 2001, [www.bis.org/publ/bcbcsca05.pdf](http://www.bis.org/publ/bcbcsca05.pdf).
- [8] A. Saunders, *Metody pomiaru ryzyka kredytowego*, Oficyna Ekonomiczna Kraków 2001.
- [9] H. Frydman, J. G. Kallberg, L. D. Kao, Testing the Adequacy of Markov Chain and Mover-Stayer Models as Representations of Credit Behavior, *Operations Research* **33**, 6 (1985).
- [10] H. Frydman, Maximum Likelihood Estimation in the Mover-Stayer Model, *J. American Statistical Association*, **79**, 632 (1984).
- [11] M. T. Jones, *Estimating Markov Transition Matrices Using Proportions Data: An Application to Credit Risk*, IMF Working Paper WP/05/219, <http://www.imf.org/external/pubs/cat/longres.aspx?sk=18387.0> .
- [12] S. Höse, S. Huschens, R. Wania, Rating Migrations in ed. W. Härdle, T. Kleinow, G. Stahl, *Applied Quantitative Finance: Theory and Computational Tools*, Springer, Berlin 2002.
- [13] Y. Jafry, T. Schuermann, Measurement, estimation and comparison of credit migration matrices, *J. Banking and Finance*, **28**, 11 (2004).
- [14] D. Lando, T. M. Skodeberg, Analyzing rating transitions and rating drift with continuous observations, *J. Banking and Finance*, **26**, 423 (2002).
- [15] T. Schuermann, Credit Migration Matrices in: *Encyclopedia Quantitative Risk Analysis & Assessment*, [www.wiley.com//wileychi/risk/](http://www.wiley.com//wileychi/risk/).
- [16] R. B. Israel, J. S. Rosentahl, J. Z. Wei, Finding generators for Markov Chains via empirical transition matrices, with applications to credit rating, *Mathematical Finance*, **11**, 2 (2001).
- [17] S. T. Rachev, S. Trueck, *Rating Based Modeling of Credit Risk Theory and Application of Migration Matrices*, Elsevier Inc, Amsterdam 2009.
- [18] A. Agresti, *Categorical Data Analysis*, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., Hoboken, New Jersey 2002.
- [19] P. Diggle, *Analysis of Longitudinal Data*, Oxford University Press, New York, 2002.
- [20] B. Efron, R. J. Tibshirani, *An Introduction to the Bootstrap*, Chapman & Hall, New York 1993.