

INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2004-35

**Stochastic Migration Models
with Application to Corporate Risk**

P. GAGLIARDINI¹
C. GOURIEROUX²

Les documents de travail ne reflètent pas la position de l'INSEE et n'engagent que leurs auteurs.

Working papers do not reflect the position of INSEE but only the views of the authors.

¹ University of Lugano and CREST.

² CREST, CEPREMAP and University of Toronto. (*CREST, 15 Boulevard Gabriel Péri, 92245 Malakoff Cedex, France*). Email : gouriero@ensae.fr

We thank P. Balestra, G. Laroque, E. Renault, T. Schuerman and S. Turnbull for Helpful comments.

STOCHASTIC MIGRATION MODELS WITH APPLICATION TO CORPORATE RISK

P. Gagliardini*, C. Gouriéroux†

First version: June 2004.
This version: January 2005.

*University of Lugano and CREST.

†CREST, CEPREMAP, and University of Toronto.

We thank P. Balestra, G. Laroque, E. Renault, T. Schuerman and S. Turnbull for helpful comments.

Stochastic Migration Models with Application to Corporate Risk

Abstract

In this paper we explain how to use the rating histories provided by the internal scoring systems of banks and by rating agencies in order to predict the future risk of a given borrower or of a set of borrowers. The method is developed following the steps suggested by the Basle Committee. To introduce both migration correlation and non-Markovian serial dependence, we consider rating histories with stochastic transition matrices. We develop the complete methodology to estimate both the number and dynamics of the factors influencing the transitions. Further we explain how to use the stochastic migration model for prediction. As an illustration the ordered Probit model with unobservable dynamic factor is estimated from French data on corporate risk.

Keywords: Migration, Rating, Migration Correlation, Credit Risk, Stochastic Intensity, Autoregressive Gamma Process, Jacobi Process, Ordered Qualitative Model, Kalman Filter, Panel Data.

JEL number: C23, C35, G11.

1 Introduction

In this paper we explain how to use the rating histories provided by the internal scoring systems of banks and by rating agencies in order to predict the future risk of a given borrower, or of a set of borrowers. The method can be applied to both retail credit or corporate loans, and is developed following the steps suggested by the Basle Committee. To highlight the problem, we first describe the usual scoring systems and the approach of the Basle Committee.

1.1 Scoring systems

Several banks, credit institutions or rating agencies have developed scoring systems to predict the future risk of a given borrower. The technical levels of the different scoring systems are rather heterogeneous, but a standard approach consists in predicting the time to failure. More precisely, for a given borrower the probability that the time to failure τ is larger than h is often specified as:

$$P_{i,t}[\tau > h] = \exp \left[-e^{x'_{i,t}b} a(h; \theta) \right], \quad (1)$$

where $x_{i,t}$ are observed covariates, $a(\cdot; \theta)$ is a (baseline) cumulated hazard function, and b, θ are parameters. This specification is the so-called proportional hazard model. The covariates are available in the proprietary data bases hold by banks or rating agencies. For corporations, they can include data on balance sheets, credit histories, corporate bond prices (for large firms only). The score for firm i at date t is: $s_{i,t} = x'_{i,t}\hat{b}$, where \hat{b} is the estimated value of the parameter¹. The number of variables introduced in a score is between 10 and 15 for corporations (including different measures of size, the industrial sector, different financial ratios, ...), up to 50 – 60 for consumer credit. The precise list of variables introduced in a score and the value of \hat{b} are always confidential.

¹The basic score methodology can be improved in various ways, for instance by introducing different scores for different maturities h and by correcting for selectivity bias or competing risks [see Gouieroux, Jasiak (2005) for a complete description of the scoring methodology and for examples of scores].

1.2 The confidentiality restrictions

Before introducing new rules for fixing the capital required to hedge credit risk, the regulator has to account for the existing scoring methodologies described above and for the confidentiality restrictions. The question is currently solved as follows.

First, the regulatory authorities can audit the scoring systems (including the proprietary data bases) and validate the ones, which are sufficiently discriminatory.

Second, the regulator decides the minimal information to be used for the computation of the required capital under observance of the confidentiality restrictions. It has been decided to consider qualitative measures of risk called ratings. These ratings are compatible with a scoring system and are often defined by discretizing the score. The number of rating alternatives has been fixed between 8 and 10.

To summarize, the different existing data bases are described in the table below for corporations.

Balance sheet histories	proprietary data bases
Credit histories	proprietary data bases
Corporate spreads histories (large companies only)	freely available from financial markets
Individual score histories	proprietary data bases
Individual rating histories	can be bought
Aggregate data on rating histories	freely available

1.3 The use of rating histories

In this regulation approach, the rating histories become the basic knowledge. They can be used for different purposes:

- i) to approximate the individual scores $s_{i,t}$ by using the sequence of previous individual ratings. It is known from the scoring practice that such an approximation can be quite accurate, if the score has been well-defined.
- ii) To gather the different individual ratings in order to analyze the risk on a credit portfolio including a (large) number of borrowers.

For both problems correlation matters. More precisely, we have to consider serial correlation to understand the effect of lagged ratings, and cross-sectional correlation to account for the so-called default correlation or, more generally, migration correlation, that is the link between rating upgrades or

downgrades of different firms. The idea behind this paper is that both effects can be appropriately captured by introducing a model explaining the rating transitions with unobservable time varying stochastic factors. Indeed, when these factors are integrated out, we get both dependence of the current rating on all its lagged values and cross-sectional dependence. This type of model is called stochastic migration model.

We want to stress the unobservability of the common time varying factors. Indeed some migration models are currently proposed in the literature with observable factors such as unemployment, business cycle (see references below) or, as in the KMV approach, long term bond yields and equity returns. However, these models cannot be used for prediction purposes without analyzing at the same time the dynamics of the macro-variables which have been introduced. It is not in the spirit of the current regulation approach to also fix the relevant macro-variables and the econometric models to be used for their dynamics.

1.4 Related literature

The literature on credit risk has to be discussed with respect to the regulation approach illustrated above. Thus we have to distinguish:

i) the analysis introducing individual covariates to predict risk. The interest is essentially in determining a score. Recent examples of this literature are Chava, Jarrow (2002), Bharath, Shumway (2004), Duffie, Wang (2004); see also references therein. Very often in academic work the number of covariates is rather small and then the discriminatory power is far behind that of scores currently implemented by some banks for credit granting decisions². For instance Duffie and Wang (2004) construct a score based on a rather small number of individual covariates, that are the distance-to-default, the firm size, the firm earnings performance, the sector earnings performance. To compensate for the lack of microeconomic covariates, they propose to introduce parameters specific to firms, which however create the problem of "incidental parameters" at the estimation level [see e.g. Hsiao et alii (2002)].

ii) The analysis of joint movements of individual risks based on observable macro-variables. This literature is interested in understanding the reasons for joint rating movements and is often related to the business cycle litera-

²However, these papers are useful for academics, who do not have access to the implemented scores.

ture or to the discussion of credit rating philosophies (point in time versus through the cycle). As mentioned above, this type of model is difficult to implement for the prediction of future risk in a credit portfolio. Examples of this literature are Nickel, Perraudin and Variotto (2000), Kavvathas (2001), Bangia et alii (2002), Rosch (2004).

iii) The analysis of joint movements of individual risks based on unobservable factors. This literature focuses on default correlation more than on risk dynamics. Generally the unobservable factors have been (implicitly) assumed serially independent and thus currently available models are not very appropriate for prediction purpose. Examples of this literature are Schonbucher (2000) and Gordy, Heitfield (2001), (2002), while similar ideas underlie the approach of CreditMetrics [Gupton et alii (1997)]. Finally, Duffie and Wang (2004) also introduce in their analysis two time dependent factors: the personal income growth, which is observable, and an unobservable factor (implied by the homogeneous autocorrelation structure of the distance-to-default), assumed serially independent.

1.5 Aim and contribution of this paper

Our paper extends the approach in the latter category iii) to serially dependent unobservable factors, without fixing a priori either the number of factors, or their economic interpretations. The goal is to introduce models which are more appropriate for predicting future risk in a large credit portfolio and, at the same time, are in line with the regulation suggested by the Basle Committee. To this aim, our paper considers a set of Markovian processes with stochastic transition matrices. Basically, this specification extends the standard (doubly) stochastic intensity model introduced by Cox (1972) in the two-state case and used in financial literature for analyzing credit risk and default correlation³. Such a specification is especially appropriate for the joint analysis of rating histories of several corporations, including migration correlation at different terms.

In Section 2 we define the basic model for stochastic migration. In this model the individual qualitative rating histories are independent heterogeneous Markov processes with identical time varying stochastic transition matrices. The underlying process of transition matrices acts as a multivariate

³See Lando (1998), Duffie, Singleton (1999), Duffie, Lando (2001), Gouriéroux, Monfort, Polimenis (2003).

systematic factor which affects all individual histories and creates both the serial dependence between individual ratings and the correlation between histories. We discuss the predictive properties of the model according to the prediction horizon and to the available information set. Different specifications for the dynamics of the stochastic transition matrices are discussed in Section 3. They include the case of i.i.d. transition matrices, factor ordered probit and Gompit model, or reduced form modelling via the Jacobi process. Section 4 focuses on the definition of migration correlation, which is a cornerstone in credit risk analysis. Indeed this notion has not been precisely defined in the previous academic or applied literature; while some estimates are regularly reported by the rating agencies, they are computed "without relying on a specific model driving transitions" [de Servigny, Renault (2002)]. This can explain the following remark in the seminal paper by Lucas (1995) p82: "These historical statistics describe only observed phenomena, not the true underlying correlation relationship". Precise definitions of migration correlations can be provided in the framework of stochastic migration models only. We first discuss the case of i.i.d. transition matrices, which has been (implicitly) assumed in the existing literature on default correlation and migration correlation [see e.g. Bahar, Nagpal (2001) and de Servigny, Renault (2002)]. Then we discuss models with serially dependent factors and highlight their importance to provide non-flat term structures for spreads on credit derivatives. Section 5 is concerned with statistical inference. Since the model with stochastic transition introduced in this paper is a nonlinear factor model for panel data, simulation based estimation methods can be used. However, some specificities of the model deserve a more careful discussion. In particular, the section focuses on i) the consistency of the ML estimator when either the cross-sectional or the time dimension tends to infinity, ii) the problem of default absorbing barrier, iii) the implementation of an approximate linear Kalman filter for large portfolios which greatly simplifies the estimation of factor ordered qualitative model. Finally Section 6 presents an application to the migration data regularly reported by the French central bank. This type of application is especially interesting since the French central bank has a complete internal rating system, but is also strongly linked with the French regulatory authority, that is Commission Bancaire. We display the estimated migration correlations and compare the estimated default correlations with the values currently suggested by the regulator. Then we discuss the relationship between the migration probabilities and the GNP increments (business cycle) in terms of causality analysis. Finally we consider a factor

ordered probit transition model and perform its estimation by an approximated Kalman filter. To our knowledge this is one of the first estimation of such a model [see Feng et alii (2004) for a similar analysis on Standard and Poor's data]. Section 7 concludes.

2 The basic factor model

In this section we introduce a factor model for joint analysis of a large number of qualitative individual histories $(Y_{i,t})$, $i = 1, \dots, n$ with the same known finite state space $\{1, \dots, K\}$. For credit risk applications the individual can be a firm, the states $k = 1, \dots, K$ correspond to the admissible grades such as AAA, AA, ..., D, and a given process $(Y_{i,t})$ to a sequence of individual ratings over time. In particular, grades $k = 1, \dots, K$ are ranked in order of increasing risk, with $k = K$ corresponding to default. The stochastic migration model is defined in Section 2.1 and the homogeneity assumption discussed in Section 2.2. Section 2.3 is concerned by the predictive properties of the model; in particular, the effect of the available information set is carefully discussed.

2.1 Definition

The joint dynamics of individual histories is defined as follows.

Definition 1 *The individual histories satisfy a stochastic migration model if:*

- i) the processes $(Y_{i,t})$, $i = 1, \dots, n$, are independent Markov chains, with identical transition matrices Π_t , when the sequence of transitions (Π_t) is given;*
- ii) the process of transition matrices (Π_t) is a stochastic Markov process.*

In practice the transition matrices are generally written as functions of a small number of factors $\Pi_t = \Pi(Z_t)$, say, satisfying a Markov process. Under a stochastic migration model the whole dependence between individual histories is driven by the common factor Z_t (or Π_t).

The stochastic migration model is a convenient specification to get joint histories featuring cross-individual dependence. Indeed the model might have been defined directly for the joint process of rating histories (Y_t) , where $Y_t = (Y_{1,t}, \dots, Y_{n,t})'$. However, even under the simplifying assumption that the joint process of rating histories (Y_t) is Markov, the associated joint transition matrix would include $K^n (K^n - 1)$ independent transition probabilities

to be estimated, and such an approach is clearly unfeasible in practice. The stochastic migration model is introduced to constrain the transitions and to diminish the number of parameters to be estimated. The latter will include the parameters characterizing the dependence between the transition probabilities Π_t and the factors Z_t , plus the parameters defining the factor dynamics. In particular, the number of parameters does not increase with the number of firms, so that the stochastic migration model is an appropriate framework for the analysis of joint rating migration in large credit portfolios.

The dynamics of rating histories $(Y_{i,t})$ can be analyzed in alternative ways according to the available information.

i) If the past, current and future values of the underlying factors are observed, the processes $(Y_{i,t})$, $i = 1, \dots, n$, are independent Markov chains. They are non stationary, since the transition matrices differ in time.

ii) If the underlying factors are not observed, it is necessary to integrate out the factors (Z_t) [or the transition matrices (Π_t)]. Let us discuss the joint distribution of Y_{t+1} given the lagged ratings $\underline{Y}_t = (Y_t, Y_{t-1}, Y_{t-2}, \dots)$ only. Its transition matrix is characterised by:

$$\begin{aligned} & P(Y_{1,t+1} = k_1^*, \dots, Y_{n,t+1} = k_n^* | \underline{Y}_t) \\ &= E \left[P(Y_{1,t+1} = k_1^*, \dots, Y_{n,t+1} = k_n^* | \underline{Y}_t, (\Pi_t)) | \underline{Y}_t \right] \\ &= E \left[\pi_{k_1 k_1^*, t+1} \dots \pi_{k_n k_n^*, t+1} | \underline{Y}_t \right], \quad \text{where } Y_{1,t} = k_1, \dots, Y_{n,t} = k_n. \end{aligned}$$

We deduce the property below:

Proposition 1 *For a Markov stochastic transition model, the distribution of the joint process (Y_t) is symmetric with respect to individuals, that is invariant by permutation of individual indexes.*

In general the distribution of Y_{t+1} given the past ratings depends on the whole rating history \underline{Y}_t . The distribution of Y_{t+1} given \underline{Y}_t can be summarized by a $K^n \times K^n$ transition matrix from Y_t to Y_{t+1} , whose elements depend on the whole past history \underline{Y}_{t-1} . This transition matrix provides the nonlinear prediction of future rating of any firm i , given the lagged ratings of this firm and of the other ones (see the Introduction). After an appropriate reordering of the states $\{1, \dots, K\}^n$, this transition matrix is given by:

$$P_n = E \left[\overset{n}{\otimes} \Pi(Z_{t+1}) | \underline{Y}_t \right],$$

where $\overset{n}{\otimes}$ denotes n -fold Kronecker product.

iii) Finally both individual histories and factors could be observed up to time t . The available information set becomes $(\underline{Y}_t, \underline{Z}_t)$. Then the joint transition probabilities are derived by integrating out the future factor values only. The transition matrix is given by:

$$Q_n = E \left[\otimes^n \Pi(Z_{t+1}) \mid Z_t \right].$$

2.2 The homogeneity assumption

The property of symmetry in Proposition 1 is a condition of homogeneity of the population of individuals. It implies identical distributions for the individual histories, but also equidependence [see e.g. Frey, McNeil (2001), (2003), Gouriéroux, Monfort (2002)]. The empirical relevance of the homogeneity assumption has to be discussed for the application to credit risk. For this purpose it is necessary to distinguish between retail credit and corporate bonds.

i) Retail credits include consumer credits, such as mortgages, classical consumption credits and revolving credits, as well as over-the-counter credits to small and medium size firms. For such applications the number of borrowers is very large, between 100000 and several millions. The practice for internal rating is to separate the population of borrowers into so-called "homogeneous" classes of risk where the individual risks can be considered as independent, identically distributed within the classes. For consumer credits the number of classes can be rather large (several hundreds), with classes including in general several thousands of individuals. The assumption of identical distributions and cross individual independence can be tested and these tests are the basis for determining the number of classes and their boundaries [see Gouriéroux, Jasiak (2005) for a detailed presentation of the segmentation approach used in the standard score methodology]. The homogeneity assumption considered in this paper extends the usual one in two respects. First it assumes identical dynamics for the individual ratings (not only identical marginal distributions for each $Y_{i,t}$). Second the condition of cross-independence is replaced by a condition of equidependence.

ii) The situation is different when large corporations and associated corporate bonds are considered. It is possible to classify these corporations according to their rating, individual sector, ... and to expect a similar distribution of defaults in the medium run (1 year for S & P, 3 years for Banque de France). Indeed the ratings reported by the rating agencies such as Moody's,

S & P, Fitch are derived according to this criterion. However it is not usual to check that the standard ratings can also be used for classifying defaults at other terms, and that they ensure equidependence. As an example, even if we could accept similar term structures of default (term structures of spreads, respectively) for IBM, General Motors, and Microsoft companies (resp. for the IBM, GM, Microsoft bonds) which have the same rating, the joint probability of IBM and Microsoft defaulting could be much higher than for IBM and GM, for instance. Despite this, the assumption of equidependence adopted in this paper is a valuable paradigm in practice. Indeed, for tractability reason both professional and theoretical literatures often assume independence between defaults within a rating class or, in more recent contributions, a default correlation which is constant in time [see e.g. Lucas (1995), Duffie, Singleton (1999), Schonbucher (2000), Gordy, Heitfield (2002), de Servigny, Renault (2002) and the discussion in Section 4]. The assumption of equidependence is clearly more flexible than the usual conditions of independence or constant default correlation introduced in the previous literature.

2.3 Prediction and information

The transition matrices P_n, Q_n concern migration at horizon 1, that is short run migration. In practice the horizon of interest (for instance the investment or risk management horizon) can be different. In this section we discuss the term structure of migration probabilities, which describes how the rating predictions depend on horizon h .

2.3.1 Prediction formulas

The stochastic migration model can be used to analyze the rating predictions at different horizons. These predictions depend on the available information set, which can include either i) the lagged individual histories only, or ii) the lagged individual histories and the lagged factor values.

i) In the first case the predictive distribution of Y_{t+h} given \underline{Y}_t is given by:

$$P_n(h) = E \left[\overset{n}{\otimes} \Pi(Z_{t+1}) \overset{n}{\otimes} \Pi(Z_{t+2}) \dots \overset{n}{\otimes} \Pi(Z_{t+h}) \mid \underline{Y}_t \right]. \quad (2)$$

This matrix can be rewritten in terms of the transition matrix of the individual chains between t and $t+h$, that is $\Pi(t, t+h) = [\pi_{kl}(t, t+h)]$, where

$\pi_{kl}(t, t+h) = P[Y_{i,t+h} = l \mid Y_{i,t} = k, (Z_t)]$. This transition matrix at horizon h is the product of h transition matrices at horizon 1: $\Pi(t, t+h) = \Pi_{t+1}\Pi_{t+2}\dots\Pi_{t+h}$. Then the predictive distribution of Y_{t+h} given \underline{Y}_t becomes:

$$P_n(h) = E \left[\bigotimes^n \Pi(t, t+h) \mid \underline{Y}_t \right].$$

ii) When the factors are observable up to time t , the predictive distribution of Y_{t+h} given $\underline{Y}_t, \underline{Z}_t$ becomes:

$$\begin{aligned} Q_n(h) &= E \left[\bigotimes^n \Pi(Z_{t+1}) \bigotimes^n \Pi(Z_{t+2}) \dots \bigotimes^n \Pi(Z_{t+h}) \mid Z_t \right] \\ &= E \left[\bigotimes^n \Pi(t, t+h) \mid Z_t \right]. \end{aligned} \quad (3)$$

2.3.2 Factor observability for large population (large portfolios)

In credit risk applications, the cross-sectional dimension n is typically much larger than the time dimension T . In such a situation, it is interesting to study the limiting case $n \rightarrow \infty$ to deduce important implications regarding the observability of the factors driving rating transitions.

Let us consider the historical rating data $(Y_{i,t})$, $i = 1, \dots, n$. These data can be used to compute the migration counts $N_{kl,t}$, $k, l = 1, \dots, K$, $t = 1, \dots, T$, where $N_{kl,t}$ denotes the number of individuals migrating from k to l between $t-1$ and t , and the population structure per rating $N_{k,t}$, $k = 1, \dots, K$, $t = 1, \dots, T$, where $N_{k,t}$ counts the number of individuals in state k at date t . By applying the Law of Large Numbers conditional on factor Z_{t+1} , the transition frequency:

$$\widehat{\pi}_{kl,t+1} = \frac{N_{kl,t+1}}{N_{k,t}},$$

tends to the theoretical transition probability $\pi_{kl,t+1}$ at date $t+1$, if n tends to infinity. Thus, for a large population, at each date t the transition matrix Π_t can be regarded as known, and therefore also the factor Z_t , whenever the mapping $Z \rightarrow \Pi(Z)$ is one-to-one. In practice, such a factor observability for large population has important consequences for prediction purposes. Indeed, when n is large, it is no longer necessary to distinguish between the two information sets considered in Section 2.3.1. More precisely, even if factor Z_t is ex-ante unobservable, the large cross-section of individual transitions can be used to approach its value. Thus, the term structure of migration probabilities can be computed easily according to formula (3). This provides

a simple methodology for predicting the future risk in a large credit portfolio which is in line with the spirit of the current regulation approach.

3 Examples

This section describes different examples of stochastic migration models. We first consider the case of independent, identically distributed transition matrices. This basic framework is important since it underlies the estimates of default correlation which are usually displayed in the literature. The ordered polytomous model for transition matrices is considered in Section 3.2. In this specification, the serial dynamics of transition matrices is introduced by means of a structural latent factor. Finally, in Section 3.3 we consider reduced form models for the dynamics of the stochastic transition matrices, such as the Jacobi process.

3.1 Independent transition matrices

In this section we assume independent, identically distributed transition matrices.

Assumption A.1: *The transition matrices (Π_t) [or factors (Z_t)] are independent, identically distributed (i.i.d.).*

Under Assumption A.1, the joint transition of Y_{t+1} given \underline{Y}_t is given by:

$$P(Y_{1,t+1} = y_{1,t+1}, \dots, Y_{n,t+1} = y_{n,t+1} \mid \underline{Y}_t) = E \left(\prod_{k,l=1}^K \pi_{kl,t+1}^{n_{kl,t+1}} \right),$$

and depends on the individual histories by means of the state indicators for dates t and $t + 1$ only⁴. We deduce the property below.

Proposition 2 *For a stochastic migration model with i.i.d. factors (Z_t) [or (Π_t)], the joint process (Y_t) is an homogeneous Markov process.*

⁴In this formula the expectation is taken with respect to the stochastic transitions $(\pi_{kl,t+1})$, and the current and lagged state indicators are summarized in the observed counts $n_{kl,t+1}$.

The state space of this Markov process is $\{1, \dots, K\}^n$ and, after appropriate ordering of the states, its transition matrix can be written as:

$$P_n = E \left[\otimes^n \Pi(Z_t) \right].$$

The homogeneity of the Markov process implies that $P_n(h) = (P_n)^h$, $\forall h$, that is the term structure of migration intensity is independent of the term. In other words, Assumption A.1 implies a flat term structure of joint migration intensities. This assumption of flat term structure of joint migration intensities has important consequences in terms of credit derivative pricing. As an illustration, let us consider a credit derivative with residual maturity h written on the two firms 1 and 2. The total spread for this derivative can be decomposed as:

$$spread(h) = spread_1(h) + spread_2(h) + spread_{1,2}(h),$$

where $spread_j(h)$ denotes the marginal spread effect corresponding to firm $j = 1, 2$, whereas $spread_{1,2}(h)$ is the component of spread associated with default dependence. Let us assume to simplify that the risk neutral distribution and the historical distribution coincide for default analysis. Then Assumption A.1 will imply a flat term structure not only for the marginal spreads, but also for the joint spread component: $spread_{1,2}(h)$ independent of h . The introduction of serially dependent factors allows for non-flat term structures for each spread component.

To conclude this subsection, we remark that, by a similar argument, any given subset of m rating histories $(Y_{i_1,t}, \dots, Y_{i_m,t})$ is Markov, with a transition matrix which depends on the size m , but not on the specific firms introduced in the set. For instance, for $m = 1$, the individual history $(Y_{i,t})$ of individual i is a Markov process with state space $\{1, \dots, K\}$ and transition matrix $P_1 = E[\Pi(Z_t)]$, independent of the selected individual i . For $m = 2$ the bivariate individual histories $(Y_{i,t}, Y_{j,t})$ of the given pair of individuals (i, j) is a Markov process with state space $\{1, \dots, K\}^2$ and transition matrix $P_2 = E[\Pi(Z_t) \otimes \Pi(Z_t)]$, independent of the selected pair of individuals (i, j) .

3.2 Ordered polytomous model

A specification which is often suggested by the agencies proposing measures of credit risk and also by the Basle Committee is the ordered polytomous

model, see e.g. Gupton, Finger, Bhatia (1997), Crouhy, Galai, Mark (2000), Gordy, Heitfield (2001), Bangia et alii (2002), Albanese et alii (2003), Feng et alii (2004). The idea is to introduce an unobservable quantitative score from which the qualitative ratings are computed. In this subsection we derive a stochastic migration model along these lines by allowing for dynamic unobservable factors in the scores.

More specifically, let us denote by $s_{i,t}$ the underlying quantitative score for corporation i at date t . Let us assume that the conditional distribution of $s_{i,t}$ given the past depends on a factor Z_t (which can be multidimensional) and on the previous rating $Y_{i,t-1}$, and is such that:

$$s_{i,t} = \alpha_k + \beta'_k Z_t + \sigma_k \varepsilon_{i,t}, \quad (4)$$

if $Y_{i,t-1} = k$, where $(\varepsilon_{i,t})$ are iid variables with cdf G , and the common factor (Z_t) is independent of $(\varepsilon_{i,t})$. Thus three parameters are introduced for each initial rating class: α_k represents a level effect, the components of β_k define the sensitivities, whereas σ_k corresponds to a volatility effect.

Let us finally assume that the qualitative rating at date t is deduced by discretizing the underlying quantitative score:

$$Y_{i,t} = l, \text{ iff } c_{l-1} \leq s_{i,t} < c_l, \quad (5)$$

where $c_0 = -\infty < c_1 < \dots < c_{K-1} < c_K = +\infty$ are fixed (unknown) thresholds. This specification is a stochastic migration model, with stochastic transition probabilities given by:

$$\begin{aligned} \pi_{kl,t} &= P[Y_{it} = l \mid Y_{i,t-1} = k, Z_t] \\ &= P\left[c_{l-1} \leq \alpha_k + \beta'_k Z_t + \sigma_k \varepsilon_{it} < c_l \mid Y_{i,t-1} = k, Z_t\right] \\ &= G\left(\frac{c_l - \alpha_k - \beta'_k Z_t}{\sigma_k}\right) - G\left(\frac{c_{l-1} - \alpha_k - \beta'_k Z_t}{\sigma_k}\right), \quad k, l = 1, \dots, K. \end{aligned} \quad (6)$$

We get an ordered polytomous model for each row, with a common latent factor Z_t . When factor Z_t is unobservable, the model has to be completed by specifying the factor dynamics. In particular, when factor Z_t is serially correlated, we get a model with serially dependent transition matrices and non-flat term structures of migration intensities.

The econometric model (4), (5) includes several explanatory variables that are the indicators of the lagged rating class and the cross effects between these indicators and the different factors. The specification focuses on unobservables time dependent factors. It implicitly assumes that individual effects have already been taken into account in the construction of the score. This assumption is coherent with the general approach described in the introduction. Indeed $s_{i,t}$ depends on individual covariates by means of $Y_{i,t-1}$. More precisely we have:

$$s_{i,t} = \sum_{k=1}^K \left(\alpha_k + \beta'_k Z_t + \sigma_k \varepsilon_{it} \right) \mathbb{I}_{c_{k-1} < x'_{i,t-1} \hat{b} \leq c_k},$$

(with the notation of the Introduction). Thus the model implicitly accounts for the lagged individual covariates and the cross-effects with time factors in a nonlinear way⁵.

The parameters of the ordered polytomous model (6) and of the factor dynamics are not identifiable. First, the factor is defined up to an invertible affine transformation; thus we can always assume:

$$\textit{Identifying constraints on the factor dynamics: } E(Z_t) = 0, \quad V(Z_t) = Id.$$

Second, other identifiability problems are due to the partial observability of the quantitative score. The same migration probabilities can be obtained with appropriately combined affine transformations of the quantitative score and of the thresholds. In particular, these transformations have to be the same for each row of the transition matrix, since the thresholds are independent of the initial rating class. Therefore, it is enough to impose the standard identification restrictions for an ordered polytomous model to one row only, the first one, say:

$$\textit{Identifying restrictions for partial observability: } c_1 = 0, \quad \sigma_1^2 = 1.$$

i) Factor probit model. The model reduces to a probit model if the error terms ($\varepsilon_{i,t}$) follow a standard Gaussian distribution. The migration probabilities become:

$$\pi_{kl,t} = \Phi \left(\frac{c_l - \alpha_k - \beta'_k Z_t}{\sigma_k} \right) - \Phi \left(\frac{c_{l-1} - \alpha_k - \beta'_k Z_t}{\sigma_k} \right),$$

⁵For instance the Duffie, Wang (2004) model does not include cross-effects.

where Φ denotes the cdf of the standard normal⁶.

ii) Factor Gompit model. When the error terms are $\varepsilon_{i,t} = \log u_{i,t}$, with u_{it} following an exponential distribution, the cdf G corresponds to a Gompertz distribution: $G(x) = 1 - \exp(-e^x)$. The migration probabilities are given by:

$$\pi_{kl,t} = \exp \left[-\exp \left(\frac{c_{l-1} - \alpha_k - \beta'_k Z_t}{\sigma_k} \right) \right] - \exp \left[-\exp \left(\frac{c_l - \alpha_k - \beta'_k Z_t}{\sigma_k} \right) \right].$$

This model is a multistate extension of the two-state Cox model with stochastic intensity, usually considered for corporate bond pricing [see e.g. Lando (1998)]. Indeed the transitions are induced by an exponential variable crossing a grid of thresholds depending on the starting grade.

The factor probit model is often called structural model in the literature by reference to Merton's model [Merton (1974)], in which the score is the (log) ratio of liability by asset value⁷. Similarly the stochastic intensity model is generally called reduced form model in the literature. The examples above show that this distinction is not relevant. The two approaches are special cases of polytomous ordered qualitative models with simply different assumptions on the distribution of the error term. This explains why we favour the terminology Probit versus Gompit usually employed in microeconometrics.

3.3 Reduced form models

It is also possible to introduce directly a dynamics for the transition matrices without referring to any structural latent variables. The Jacobi process is a continuous time reduced form, which can be appropriate to account for migration in the continuous time pricing models, whereas the logistic autoregression is widely implemented for risk analysis of consumer credit portfolio.

3.3.1 Jacobi specification

Any row of the (stochastic) transition matrix defines a (stochastic) discrete probability distribution on the state space $\{1, \dots, K\}$. Thus, when t varies,

⁶The normality assumption on the latent score is standard in the literature and underlies for instance the implementation of the ordered polytomous model proposed by CreditMetrics.

⁷In standard implemented scores this ratio is only one among several explanatory variables.

we get a stochastic process with values on the set of discrete distributions. The multivariate Jacobi process has been introduced to specify the dynamics of such a stochastic discrete probability distribution [see Gouriéroux, Jasiak (2004)]. A Jacobi specification for the stochastic transitions assumes:

- i) the rows $[\pi_{kl,t}, l = 1, \dots, K, t = 1, \dots, T], k = 1, \dots, K$, are independent stochastic processes;
- ii) any row $[\pi_{kl,t}, l = 1, \dots, K, t = 1, \dots, T]$ corresponds to the discrete time observations of a continuous time multivariate Jacobi process, which satisfies the diffusion system:

$$d\pi_{kl,t} = b_k(\pi_{kl,t} - a_{kl})dt + \sqrt{g_k \pi_{kl,t}} dW_{kl,t} - \pi_{kl,t} \sum_{m=1}^K \sqrt{g_k \pi_{km,t}} dW_{km,t}, \quad l = 1, \dots, K,$$

where $(W_{km,t}), k, m = 1, \dots, K$, are independent Brownian motions, and the parameters satisfy the constraints $b_k < 0, g_k > 0, \sum_{l=1}^K a_{kl} = 1, \forall k, a_{kl} > 0, \forall k, l$.

The drifts of the diffusions suggest that processes $[\pi_{kl,t}, l = 1, \dots, K, t = 1, \dots, T]$ feature a mean-reverting dynamics, with equilibrium levels $a_{kl}, l = 1, \dots, K$, and mean-reverting parameters $b_k, k = 1, \dots, K$. The serial dependence of the stochastic transition matrices (Π_t) is controlled by the mean-reverting parameters b_k . The parameters g_k can be interpreted either as volatility parameters, or as smoothing parameters. In particular, if these parameters tend to infinity, the process $\pi_{kl,t}$ tends to a pure jump process. The restrictions on the parameters ensure that each row is a stationary process, with beta stationary distribution.

3.3.2 Logistic autoregression

Serial dependence can also be directly introduced by considering Gaussian vector autoregressions applied to transformed transition probabilities. For instance, in the two-state case $K = 2$, a logistic autoregression can be introduced for $\pi_{11,t}, \pi_{22,t}$:

$$\begin{aligned} \log \frac{\pi_{11,t}}{1 - \pi_{11,t}} &= c_1 + \varphi_{11} \log \frac{\pi_{11,t-1}}{1 - \pi_{11,t-1}} + \varphi_{12} \log \frac{\pi_{22,t-1}}{1 - \pi_{22,t-1}} + \varepsilon_{1t}, \\ \log \frac{\pi_{22,t}}{1 - \pi_{22,t}} &= c_2 + \varphi_{21} \log \frac{\pi_{11,t-1}}{1 - \pi_{11,t-1}} + \varphi_{22} \log \frac{\pi_{22,t-1}}{1 - \pi_{22,t-1}} + \varepsilon_{2t}, \end{aligned}$$

where $(\varepsilon_{1t}, \varepsilon_{2t})$ is a Gaussian white noise with variance-covariance matrix $\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$. Such a logistic transformation of transition probabilities before introducing the Gaussian autoregression is suggested for instance in the Mc Kinsey methodology.

However the approach by logistic autoregressions is difficult to extend to a larger number of states. Indeed there is no general agreement on a multivariate one-to-one transformation which associates with probabilities $(\pi_1, \dots, \pi_{K-1})$, say, constrained by $0 \leq \pi_k \leq 1$, $k = 1, \dots, K-1$, $0 \leq 1 - \sum_{k=1}^{K-1} \pi_k \leq 1$, an unconstrained vector of \mathbb{R}^{K-1} .

4 Migration correlation

In order to analyse the default risk in a credit portfolio, it is important to take into account carefully the simultaneous rating migrations of different firms in the same direction, such as joint up- or downgrades. The tendency to a common rating migration is called migration correlation [see e.g. Lucas (1995), Bahar, Nagpal (2001), de Servigny, Renault (2002)]. It extends the concept of default correlation, corresponding to the two-state case with default absorbing barrier.

The standard specification with deterministic transition matrices does not feature migration correlation. In our framework, migration correlation is introduced by means of the stochastic transition matrix, which is a common (multivariate) factor across firms. More precisely the basic model (see Section 2) considers the conditional distribution of firm ratings given the sequence of transition matrices and assumes the conditional independence between firms. A dependence between rating dynamics is deduced when the stochastic transition matrices are integrated out, which creates migration dependence.

We first consider the case of i.i.d. transition matrices corresponding to a flat term structure of migration intensity. In Section 4.1 we define precisely the notions of joint bivariate transition and of migration correlation, and explain how they can be displayed in well-chosen matrices. The definitions are illustrated in Section 4.2, where the migration correlations are computed for the Probit and Gompit ordered polytomous models. The definition of migration correlation has to be reconsidered when the transition matrices are serially dependent. This is done in Section 4.3, where the importance of the selected information set is emphasized.

4.1 Definition in the i.i.d. case

Let us consider two firms, whose rating histories are described by the chains $(Y_{i,t})$ and $(Y_{j,t})$, respectively, following a stochastic migration model with i.i.d. transition matrices. From Proposition 2 the bivariate process $(Y_{i,t}, Y_{j,t})$ is still a Markov process, with bivariate joint transition:

$$p_{kk^*,ll^*} = P[Y_{i,t+1} = k^*, Y_{j,t+1} = l^* | Y_{i,t} = k, Y_{j,t} = l] = E[\pi_{kk^*,t}\pi_{ll^*,t}]. \quad (7)$$

It defines a $K^2 \times K^2$ square matrix of joint transition probabilities. Similarly we get:

$$P[Y_{i,t+1} = k^* | Y_{i,t} = k, Y_{j,t} = l] = E[\pi_{kk^*,t}]. \quad (8)$$

Migration correlation is defined in terms of conditional correlation of individual rating indicators⁸:

$$\rho_{kk^*,ll^*} = \text{corr}(\mathbb{I}_{Y_{i,t+1}=k^*}, \mathbb{I}_{Y_{j,t+1}=l^*} | Y_{i,t} = k, Y_{j,t} = l),$$

where $\mathbb{I}_{Y_{i,t+1}=k^*} = 1$, if $Y_{i,t+1} = k^*$, $= 0$, otherwise. The migration correlations can be written in terms of the underlying stochastic transition probabilities:

$$\begin{aligned} \rho_{kk^*,ll^*} &= \frac{\text{cov}(\mathbb{I}_{Y_{i,t+1}=k^*}, \mathbb{I}_{Y_{j,t+1}=l^*} | Y_{i,t} = k, Y_{j,t} = l)}{V(\mathbb{I}_{Y_{i,t+1}=k^*} | Y_{i,t} = k, Y_{j,t} = l)^{1/2} V(\mathbb{I}_{Y_{j,t+1}=l^*} | Y_{i,t} = k, Y_{j,t} = l)^{1/2}} \\ &= \frac{\text{cov}(\pi_{kk^*,t}, \pi_{ll^*,t})}{[E\pi_{kk^*,t}(1 - E\pi_{kk^*,t})]^{1/2} [E\pi_{ll^*,t}(1 - E\pi_{ll^*,t})]^{1/2}}. \end{aligned} \quad (9)$$

Migration correlation between firms i and j depends on their current and future ratings only, not on their names i and j .

There are as many different migration correlations as unordered pairs $(k, k^*), (l, l^*)$, that is $K^2(1 + K^2)/2$. However, these migration correlations are linearly dependent, since they satisfy a set of restrictions such as:

$\sum_{k^*} \rho_{kk^*,ll^*} [E\pi_{kk^*,t}(1 - E\pi_{kk^*,t})]^{1/2} = 0, \forall k, l, l^*$, due to the unit mass restrictions on transition probabilities. In particular they cannot be of the same sign. Among all different migration correlations, some are more appealing for practitioners, especially those which involve migrations of the firms by one rating tick. For instance we can consider correlations between downgrades. If the initial state is (k, l) the correlation is given by:

⁸By a similar approach we can define migration correlations at any horizon h by considering the correlation between the rating indicators $\mathbb{I}_{Y_{i,t+h}=k^*}$ and $\mathbb{I}_{Y_{j,t+h}=l^*}$ conditional on $Y_{i,t} = k, Y_{j,t} = l$.

$$\text{corr} \left(\mathbb{I}_{Y_{i,t+1}=k+1}, \mathbb{I}_{Y_{j,t+1}=l+1} \mid Y_{i,t} = k, Y_{j,t} = l \right), \quad k, l = 1, \dots, K - 1.$$

We get a $(K - 1) \times (K - 1)$ symmetric matrix of down-down migration correlations indexed by the current ratings.

4.2 Migration correlation in a factor i.i.d. ordered qualitative model

Closed form expressions of the migration correlations can be derived for the Probit and Gompit factor ordered qualitative models.

4.2.1 Probit model

When the common factors are i.i.d., the formulas are well-known for the probit model [see e.g. Gordy, Heitfield (2002)] and are available in different documents of the Basle Committee, at least when migration from risk category k to default is considered. Indeed, in the current methodology, the regulator suggests to first consider the latent correlation, which is the correlation between the underlying scores:

$$\rho_k = \text{corr} (s_{i,t}, s_{j,t} \mid Y_{i,t-1} = Y_{j,t-1} = k),$$

conditional on the previous rating category k . Then, the default correlation for rating class k , that is the migration correlation $\rho_{kK,kK}$, is given by:

$$\rho_{kK,kK} = \frac{\int_{-\infty}^{\Phi^{-1}(\pi_k)} \int_{-\infty}^{\Phi^{-1}(\pi_k)} \frac{1}{2\pi\sqrt{1-\rho_k^2}} \exp \left(-\frac{1}{2(1-\rho_k^2)} (x^2 - 2\rho_k xy + y^2) \right) dx dy - \pi_k^2}{\pi_k (1 - \pi_k)} \quad (10)$$

where π_k denotes the expected default probability for category k .

4.2.2 Gompit model

Let us consider the Gompit model introduced in Section 3.2 with a single i.i.d. factor Z_t . An analytic expression for default correlation can be obtained in the special case $\beta_k/\sigma_k = 1$, for any k . Indeed default probabilities are:

$$\pi_{kK,t} = \exp [-\lambda_k \exp (-Z_t)],$$

where $\lambda_k = \exp\left(\frac{c_{K-1} - \alpha_k}{\sigma_k}\right)$. We deduce from (9) the default correlation for two firms in rating classes k and l :

$$\rho_{kK,lK} = \frac{\Psi(\lambda_k + \lambda_l) - \Psi(\lambda_k)\Psi(\lambda_l)}{\sqrt{\Psi(\lambda_k)[1 - \Psi(\lambda_k)]}\sqrt{\Psi(\lambda_l)[1 - \Psi(\lambda_l)]}},$$

where Ψ denotes the Laplace transform of $\exp(-Z_t)$: $\Psi(u) = E[\exp(-u \exp(-Z_t))]$. Equivalently we have:

$$\rho_{kK,lK} = \frac{\Psi[\Psi^{-1}(\pi_k) + \Psi^{-1}(\pi_l)] - \pi_k\pi_l}{\sqrt{\pi_k(1 - \pi_k)}\sqrt{\pi_l(1 - \pi_l)}},$$

where π_k denotes the marginal default probability of risk class k . This value depends on the marginal migration rates and on the factor distribution (by means of the Laplace transform Ψ). Since function $(x, y) \rightarrow \Psi[\Psi^{-1}(x) + \Psi^{-1}(y)]$ corresponds to an Archimedean copula [see e.g. Joe (1997)], a more heterogeneous factor will imply a larger migration correlation.

As mentioned for the probit model, it has been usual among practitioners to compare the value of the migration correlations to the level of a latent correlation corresponding to the model of the underlying score. In this example we have:

$$s_{i,t} = \alpha_k + \beta_k Z_t + \sigma_k \varepsilon_{i,t},$$

and a possible measure of the latent correlation is:

$$\text{corr}_{kl}(s_{i,t}, s_{j,t}) = \frac{\beta_k \beta_l V(Z_t)}{\sqrt{\beta_k^2 V(Z_t) + \sigma_k^2 V(\varepsilon_{i,t})}\sqrt{\beta_l^2 V(Z_t) + \sigma_l^2 V(\varepsilon_{j,t})}}.$$

However a quantitative score is defined up to an increasing transformation. In the extended Cox model it was more natural to consider the transformed score:

$$s_{i,t}^* = \exp(\alpha_k + \beta_k Z_t) u_{i,t}^{\sigma_k},$$

in order to get the crossing of a grid of thresholds by an exponential variate. Therefore another latent correlation can be defined:

$$\text{corr}_{kl}(s_{i,t}^*, s_{j,t}^*) = \frac{\text{cov}(e^{\beta_k Z_t}, e^{\beta_l Z_t})}{\sqrt{V(e^{\beta_k Z_t}) + \frac{V(u_{i,t}^{\sigma_k})}{E(u_{i,t}^{\sigma_k})^2} E(e^{2\beta_k Z_t})}\sqrt{V(e^{\beta_l Z_t}) + \frac{V(u_{j,t}^{\sigma_l})}{E(u_{j,t}^{\sigma_l})^2} E(e^{2\beta_l Z_t})}}.$$

This value can be very different from the value computed directly from the score s , which moreover is not directly observable, and is also very different from the value of the default correlations $\rho_{kK,lK}$. To summarize, the latent quantitative score is defined up to an increasing transformation and there exist as many latent correlations as admissible choices of the transformation. Thus the notion of latent correlation has to be used with care.

4.3 Migration correlations for serially dependent transition matrices

Joint bivariate transition probabilities can also be derived for serially dependent transition matrices. However their expressions depend on the selected information set. They correspond to the matrix P_2 , if the individual histories only are known up to time t , to the matrix Q_2 , if both individual histories and factors are known up to time t ⁹. In particular, for serially dependent transition matrices the joint bivariate process $(Y_{i,t}, Y_{j,t})$, where (i, j) is a given pair of individuals, is no longer Markov. Thus the joint bivariate transitions will not depend on the past through Y_{t-1} only, and in practice will vary with the date t .

Similar remarks apply to the associated migration correlations. They are obtained by a formula similar to (9) but involving expectations conditional on the available information. In particular, migration correlations also depend on the selected information set and vary in time.

It is interesting to reconsider the definitions of migration correlation given above in a static or a dynamic framework in the light of the existing literature. Indeed estimated migration correlations have been displayed in the professional and academic literature [see e.g. Lucas (1995), Bahar, Nagpal (2001), de Servigny, Renault (2002)], but "without relying on a specific model driving transitions" [de Servigny, Renault (2002)]. Since they are considered constant over the whole period of estimation, they correspond intuitively to a flat term structure of migration intensity, that is i.i.d. stochastic transition matrices have been implicitly assumed. At the contrary, in the general framework the migration correlations depend on: i) the initial and final risk categories of the two firms, ii) the horizon, and iii) the date (by means of rating histories or factor values).

⁹These two matrices coincide for large portfolios ($n = \infty$), see Section 2.3.3.

5 Statistical inference

The stochastic migration model is a special case of multifactor model for panel data. The likelihood function or the observable conditional moments involve multidimensional integrals of large dimension, which depend generally on the number of observation dates. In such a framework the standard maximum likelihood or GMM approaches are numerically intractable and can be replaced by simulation approaches as simulated maximum likelihood, simulated method of moments, or indirect inference [see e.g. the surveys by Gouriéroux, Monfort (1995), Gouriéroux, Jasiak (2001)]. However the stochastic migration model presents some specificities and special aspects of statistical inference have to be discussed. For ease of exposition we address these points in a framework with i.i.d. transition matrices, although the main arguments carry on to the general setting. In Section 5.1 we discuss the consistency properties of the Maximum Likelihood (ML) estimator according to the dimension n or T , which tends to infinity. It is explained why a large cross-sectional dimension is not sufficient to get consistency. In Section 5.2 we explain how the asymptotic theory has to be modified when the state space includes an absorbing barrier, such as default, and discuss the limiting case of large homogeneous populations (large portfolios). To conclude, in Section 5.3 we turn to models with serially correlated transition matrices and explain how to estimate in a simple way a factor ordered qualitative model when the cross-sectional dimension is large.

5.1 Consistency of the ML estimator

As noted earlier in Section 3.1, when the transition matrices (Π_t) are i.i.d., the multivariate rating process $Y_t = (Y_{1,t}, \dots, Y_{n,t})'$ is a Markov process. Therefore its distribution is defined by the transition:

$$p(y_{t+1} | y_t; \theta) = E_\theta \left[\prod_{k=1}^K \prod_{l=1}^K \pi_{kl,t+1}^{n_{kl,t+1}} \right], \quad (11)$$

where θ denotes the parameter characterizing the distribution of Π_t . The ML estimator is a solution of the optimization problem:

$$\hat{\theta} = \arg \max_{\theta} \sum_{t=2}^T \log p(y_t | y_{t-1}; \theta). \quad (12)$$

The likelihood function depends on the observations by means of the aggregate migration counts $n_{kl,t}$, $k, l = 1, \dots, K$, $t = 1, \dots, T$, which constitute a sufficient statistics for θ . The likelihood involves multidimensional integrals with dimension less or equal to $K(K - 1)$; this dimension does not depend on the number of observations nT ¹⁰.

If n is fixed and T tends to infinity the general asymptotic theory for Markov chains can be applied [Anderson, Goodman (1957)]. In particular the ML estimator is consistent under regularity conditions including the assumption that the chain is recurrent, that is passes an infinite number of times by any admissible state (when T tends to infinity).

At the opposite, when n tends to infinity and T is fixed, the ML estimator of θ is not consistent¹¹. This feature is easily understood if we consider the case $T = 1$. The sufficient statistics $N_{kl,1}$ can be used to compute the sample transition frequencies at date 1, that is $N_{kl,1}/N_{k,0}$, which tend to $\pi_{kl,1}$, when n tends to infinity (by the Law of Large Numbers applied conditional on the transition matrix Π_1). Therefore the transition matrix Π_1 is perfectly known. However the knowledge of Π_1 , that is a single observation of the sequence of transition matrices, is not sufficient to identify the dynamics of Π_t , that is parameter θ .

In summary the ML estimator is consistent for T tending to infinity, but not for n tending to infinity. The cross-sectional inconsistency of the estimator results from the cross-sectional equidependence, which does not allow the standard mixing conditions for the Law of Large Numbers to be satisfied. The remark on the non-consistency of the cross-sectional ML estimator of parameter θ is also valid when we consider other parameters of interest such as migration correlations (see Section 4), or other estimation methods.

5.2 The problem of absorbing barrier

The consistency of the ML estimator when T tends to infinity is satisfied under the recurrence condition. However, in credit risk applications, there exists an absorbing barrier, namely the default state. Asymptotically all individuals are in default state and the migration parameters associated with

¹⁰This remark is no longer valid when the stochastic transition matrices feature serial dependence.

¹¹See Gagliardini, Gouriéroux (2004) for a more detailed discussion.

a transitory phenomenon generally cannot be identified¹².

A solution to recover the consistency of an estimator of θ for large T is to increase the size of the population in time in order to compensate the defaulted firms. Such a regularly updated population is called "static pool" by Standard & Poor's [Brady, Bos (2002)]. In a simple framework new individuals can be regularly introduced to get a fixed dimension of the population of alive individuals. When an individual i defaults at date t , it is replaced by a new one assigned randomly to a state $k = 1, \dots, K - 1$ according to a distribution $\mu_t = (\mu_{1,t}, \dots, \mu_{K-1,t})$, say. The states occupied by this sequence of individuals define a process \tilde{Y}_t , say, with state space $\{1, \dots, K - 1\}$. Conditional on μ_t, Π_t , the process \tilde{Y}_t is a Markov process with a $(K - 1) \times (K - 1)$ transition matrix $\tilde{\Pi}_t$. The elements of this matrix are:

$$\tilde{\pi}_{kl,t} = \pi_{kl,t} + \pi_{kK,t}\mu_{l,t}, \quad k, l = 1, \dots, K - 1.$$

The theory presented in Section 5.1 can be applied to this transformed process (\tilde{Y}_t) , whenever the matrices (Π_t, μ_t) are assumed i.i.d., with a specified parametric distribution. In particular, in this approach the migration parameters cannot be estimated without estimating jointly the parameters characterizing the process of population renewal.

It is interesting to discuss the limiting case of a very large population of individuals: $n = \infty$. Despite default, the number of individuals alive at any date t is always infinite and some migrations between states can be observed even when t is very large. More precisely, by applying the cross-sectional argument, the transition matrices $\Pi_t, t = 1, \dots, T$ are exactly known, since they are consistently approximated by their sample counterparts. Therefore the ML method can be applied to the observed factor $\Pi_t, t = 1, \dots, T$. The ML estimators will be consistent when $n = \infty$ and T tends to infinity, even if there exists an absorbing barrier. In practice this means that, in the case of an absorbing barrier, the asymptotic bias of the ML estimator computed on a finite population can be diminished by increasing the cross-sectional dimension.

¹²In the special case of cross-sectional independence, neither the absorbing default state nor a finite number of time observations prevent consistency of the ML estimator. Indeed, in such a case the parameters of the transition matrices can be consistently estimated from the cross-section when $n \rightarrow \infty$ [see e.g. Berman, Frydman (1999) for the two-state case].

5.3 Estimation of the factor ordered qualitative model

Let us now explain how to estimate the factor ordered qualitative model introduced in Section 3.2, when the cross-sectional dimension n is large and for instance the factors satisfy a Gaussian VAR process:

$$Z_t = AZ_{t-1} + u_t,$$

where A is a matrix and (u_t) is multivariate standard normal¹³.

From (6) we deduce that:

$$\pi_{kl,t}^* := \sum_{h < l} \pi_{kh,t} = P [Y_{i,t} < l \mid Y_{i,t-1} = k, Z_t] = G \left(\frac{c_{l-1} - \alpha_k - \beta'_k Z_t}{\sigma_k} \right),$$

or equivalently:

$$G^{-1}(\pi_{kl,t}^*) = \frac{c_{l-1} - \alpha_k}{\sigma_k} - \frac{1}{\sigma_k} \beta'_k Z_t.$$

For a large cross-sectional dimension n ($n \rightarrow \infty$), the probability $\pi_{kl,t}^*$ is well-approximated by its cross-sectional sample counterpart $\hat{\pi}_{kl,t}^*$, say; moreover we have $\sqrt{n} (\hat{\pi}_{kl,t}^* - \pi_{kl,t}^*) \xrightarrow{d} N(0, \Omega_{kl,t})$, $t = 1, \dots, T$, and the estimators corresponding to different dates are independent. By applying the δ -method along the lines initially proposed by Berkson (1944) for logit models with repeated observations, we get [see also Amemiya (1976)]:

$$\begin{cases} G^{-1}(\hat{\pi}_{kl,t}^*) \simeq \frac{c_{l-1} - \alpha_k}{\sigma_k} - \frac{1}{\sigma_k} \beta'_k Z_t + \tilde{\Omega}_{kl,t}^{1/2} v_{kl,t}, & \forall k, l, t, \\ Z_t = AZ_{t-1} + u_t, \end{cases}$$

where v_t and u_t are independent multidimensional Gaussian error terms. We get an approximated linear state space model, in which the macro-component corresponds to the transition equation and the microcomponent to the measurement equation. This approximated model can be estimated by a standard linear Kalman filter, under the identification restrictions. This approach provides approximations of the microparameters α , c , β , σ , of the macro-parameters A and of the factor values at each date. From general results on statistical inference for panel models with unobservable dynamic factors [Gouriéroux, Monfort (2004)], it follows that:

¹³For expository purpose and the link with state space representation, we impose $V(u_t) = Id$ instead of $V(Z_t) = Id$ as identifying constraint.

- i) the approximations of the factor values are \sqrt{n} -consistent;
- ii) the estimators of the micro-parameters are \sqrt{nT} -consistent and asymptotically efficient;
- iii) the estimator of the macro-parameters are \sqrt{T} -consistent.

To summarize, in dynamic factor qualitative panel data models, the likelihood function has an intractable form and the standard asymptotic theory does not apply¹⁴, but the large cross-sectional dimension can be useful to introduce simple estimation approaches, based on the possibility to approximate the true transition matrices by their sample counterparts.

6 Application to migration data

6.1 The data set

Migration data are regularly reported by rating agencies as Moody's and Standard and Poor's, or by central banks as the Banque de France [see Foulcher, Gouriéroux, Tiomo (2004) for a comparison of the main rating systems]. The data sets of the agencies concern rather large firms at the international level. The number of rated firms is about 10000 and reliable data are available since 1985. The rating is generally fixed by experts on the basis of information obtained at the moment of bond issuing, for instance.

In contrast, the Banque de France collects yearly the balance sheets of all French firms. The balance sheets are used to construct a quantitative score explaining how default probability at 3 years depend on a set of financial ratios and individual characteristics. Then this score is discretized into rating classes. This data set has several advantages compared to the set of the agencies. First, it concerns about 180000 French firms, which allows to perform some analysis by size or industrial sectors without a too small number of observations. Second the formula for the econometric quantitative score can be followed, as well as the limiting thresholds, which define the rating classes [see Bardos et alii (2004)]. The rating procedure has been stable during the period of observation 1992-2003, which is not necessarily the case for the rating by expertise performed by the agencies, due to the

¹⁴The current literature on credit risk seems to be not sufficiently aware of the nonstandard properties of the estimators for panel data when there is an unobservable factor, or when the parameter size increases with the number of observations [see e.g. Duffie, Wang (2004)].

change of experts or to their different rating behaviours during the phases of the business cycle [see Pender (1992), Blume et alii (1998), Gouriéroux, Jasiak (2005), Chapter 8]. We focus on two economic sectors corresponding to wholesale and retail trade, respectively. They concern about 30000 firms for each economic sector [see Bardos et al (2004)], which are in general of small or medium size.

The Banque de France rating contains 8 risk categories, denoted 0, 1, 2, ... 7. Alternative "0" is for default, whereas alternative "7" represents the lowest risk, that is the usual AAA or Aaa of the rating agencies. The individual rating histories are aggregated to produce the transition matrices between rating classes for different years, categories and time horizons. For instance, a 1-year transition matrix is given in Table 1, for year 2001 and the wholesale sector.

[Table 1: Transition matrix in 2001 for the wholesale sector]

It is immediately seen that this matrix contains an additional column corresponding to the firms, which are not rated (NR) at the end of 2001. The firms are not rated due to missing data, which concern either the total balance sheet, or simply some financial ratios or firm characteristics introduced as explanatory variables in the underlying quantitative score. These missing data are mainly due to a lack of cooperation, which can be voluntary or not. The rate of missing data is rather large (between 10 and 30%) and larger than the rate generally observed for large firms (between 5 and 15%), which are obliged to report regularly some information concerning their balance sheet, especially for bond issuing. As the other rating agencies, the Banque de France does not report the row providing the transition from the NR class to the other risk categories. Therefore this type of matrix has to be transformed into a square transition matrix by assigning the non-rated companies among the other classes. They are usually assigned proportionally, which implicitly assumes the absence of selectivity bias (see all recent applied studies in the list of references for a similar approach). It is seen on Table 1 that the frequency of NR firms has a tendency to increase when the quality of risk diminishes. This fact could be considered either as an additional signal of bad risk, which would create a selectivity bias, or simply it can be due to the fact that providing information to the Banque de France is not a priority when the situation of the firm deteriorates. An idea about the missing row can be obtained from Foulcher et al. (2004) where such a row looks like:

	7	6	5	4	3	2	1	0	NR
NR	2.5	2.5	2.0	2.0	1.4	1.1	0.2	0.5	87.8

This shows that the NR alternative is not an indicator of imminent default, which is in favour of a proportional assignment. This approach is followed in the rest of the section. The adjusted transition matrix corresponding to the matrix in Table 1 is provided in Table 2.

[Table 2: Adjusted transition matrix in 2001 for the wholesale sector]

As usual the adjusted transition matrices contains a lot of very small transition probabilities. The significant elements are essentially around the main principal diagonals, showing that the up- or down-grades are at most of one or two buckets during the year, for firms in a "standard" situation. These matrices are rather different from the transition matrices existing for large firms, for which the ratings are more stable. Typically in the S&P or Moody's data the three main diagonals only have significant elements.

6.2 I.I.D. transition matrices

Let us first consider a model with i.i.d. transition matrices, that is the basic specification underlying the standard measures of migration correlation. As mentioned in Section 3.1, in this framework it is natural to compute the matrix of individual migration $P_1 = E\Pi_t$, and the matrix of joint migration for a pair of firms $P_2 = E(\Pi_t \otimes \Pi_t)$. The theoretical matrices are estimated by their sample counterparts, obtained by averaging on time the associated observed transition frequencies. The estimated matrices \hat{P}_1 and \hat{P}_2 (joint down-grades only) are given in Table 3 and 4 for the wholesale sector:

[Table 3: Individual migration probabilities for the wholesale sector]

[Table 4: Joint down-grade migration probabilities for the wholesale sector]

From these matrices and the results in Section 4.1 we deduce the down-grade migration correlations in the wholesale sector for all pairs of initial states:

[Table 5: Down-grade migration correlations]

Some features of down-grade correlations can be observed, for instance they are generally larger when the two firms are in similar rating classes. More importantly, the displayed migration correlations are rather small, and typically

much smaller than migration correlations reported by de Servigny, Renault (2002) from S&P data with two risk categories only, corresponding to investment and speculative grades. However, these results are difficult to compare since they do not correspond to the same number of risk categories, and it can be expected that the migration correlations will diminish, when the partition becomes thinner. Indeed the correlations are conditional on the available information, that is the chosen segmentation, and generally diminish when the information increases.

Finally, we can compute default correlations. They are reported in Tables 6 and 7 for the wholesale sector and the retail trade sector, respectively.

[Table 6: Default correlations in the wholesale sector]

[Table 7: Default correlations in the retail trade sector]

As for migration correlation, we observe rather small values, clearly much smaller than the values suggested by the Basle Committee. As above this can be due to the partition into rating categories, which is neglected in the basic methodology suggested by the Basle Committee. The disaggregation by rating categories can also have some other consequences for default correlation. For instance it is known that in a large risk the default correlation is necessarily nonnegative [see e.g. Gouriéroux, Monfort (2002) and Frey, McNeil (2001), (2003)]. By disaggregation, we can get subpopulations of smaller size and observe negative default correlations. Anyway it is important to compare the default correlations across the rating classes with the estimated default correlation proposed by the Basle Committee. As mentioned in Section 4.2.1, the regulator suggests a factor probit model, with a latent correlation which is a function of the default probability π_k of the class to which the two firms belong [see the Basle Committee on Banking Supervision (2002)]. The relationship is:

$$\rho_k = 0.24 - 0.12 \frac{1 - \exp(-50\pi_k)}{1 - \exp(-50)}.$$

From (10) we deduce the relationship between the default correlation and the marginal default probability proposed by the Basle Committee. This relationship is displayed in Figure 1 with a zoom on the range of observed default probabilities:

[Insert Figure 1: Default correlation vs default probability by Basle Committee]

It can be directly compared with the relationships estimated on the Banque de France data for the wholesale and retail trade sectors:

[Insert Figure 2: Default correlation vs default probability for two sectors]

The estimated default correlations are systematically much smaller than the values suggested by the regulator, in fact ten times smaller, with direct consequences on the required capital. These values, which are compatible with other recent studies [Feng et alii (2004), Rosch (2004)], are not unrealistic and are not the consequence of the doubly stochastic assumption of the model, as expressed for instance by Schonbucher (2004). In fact, the concepts of default correlation and of contagion are conditional on the information set. Larger the information set, generally smaller the default correlations. In our estimated model the information set includes the rating histories of all firms. Similarly, when the underlying score is based on a larger number of covariates, the default correlations are in general smaller. For instance, the fact that Duffie, Wang (2004) get larger default correlations reflects simply an underlying score based on a single explanatory variable only. Finally we remark that, despite the difference in level, estimated default correlations feature the same type of monotonic dependence with respect to the marginal default probability as suggested by the Basle Committee. However, the slope of the curve has to be adjusted for the economic sector.

6.3 Dynamics of migration probabilities

In Section 6.2 we have analyzed the sequence of migration probabilities under the assumption of i.i.d. transition matrices. This assumption, which is usually adopted in the literature for computing migration correlations (see Section 4), has to be questioned in practice. The aim of this section is to highlight the dynamics of migration probabilities before estimating a dynamic factor model in Section 6.4. In the first subsection we plot the series of up- and down-grade probabilities, and discuss their serial dependence. Then in Section 6.3.2 we consider their relationship with the French business cycle.

6.3.1 The evolution of up- and downgrade probabilities

Let us focus on downgrade and upgrade probabilities involving a migration of at most 2 buckets: $d_{k,t} = \pi_{k,k+1,t} + \pi_{k,k+2,t}$, $u_{k,t} = \pi_{k,k-1,t} + \pi_{k,k-2,t}$, respectively [except for the extreme categories, where the number of buckets

is one]. The time series of downgrade and upgrade probabilities are reported in Figures 3 and 4, respectively.

[Insert Figure 3: Downgrade probabilities]

[Insert Figure 4: Upgrade probabilities]

Each panel corresponds to a rating class at the beginning of the year and provides the dynamics in the wholesale (circles) and retail trade (diamonds) sectors. The first and second order autocorrelations are reported in Table 8 for the different downgrade and upgrade series and for both sectors:

[Insert Table 8: Autocorrelations in the wholesale and in the retail trade sector]

Despite the rather small number of observation dates, it is immediately seen that the serial autocorrelations are rather high. Thus the usual serial independence assumption, which underlies the computation of migration correlations, is not relevant empirically.

6.3.2 Business cycle

It is usual to relate the failure rate with the general state of the economy, that is the so-called business cycle. The Banque de France data base on migration probabilities convey much more information and a better knowledge of the link with business cycles can be expected. Several recent studies have already been performed on the data by Moody's and Standard & Poor's with proxies of the US business cycle [see Nickell, Perraudin, Variotto (2000), Bangia et alii (2002), and Rosch (2004)]. Even if this relationship is not the topic of our paper, it is interesting to give some preliminary elements. In a first step the dynamics of downgrade and upgrade probabilities (see Figures 3-4) can be compared with the evolution of GDP in France in the period 1992-2002, which is provided in Figure 5, first Panel.

[Insert Figure 5: GDP and factor evolutions]

The dynamic linear link between the series can be studied by a causality analysis between the downgrade (resp. upgrade) series and the GDP. The lead and lagged causality measures of downgrade and upgrade probabilities with GDP increment I_t are reported in Tables 9 and 10 for the wholesale and retail trade sector, respectively.

[Insert Table 9: Causality relations in the wholesale sector]

[Insert Table 10: Causality relations in the retail trade sector]

The distribution of causality measures is different for up- and downgrades, and for the different risk categories. The downgrades are generally more reactive to the business cycle than the upgrades. Moreover for the low risk categories the causality from I to d is more important than the causality from d to I . For instance the business cycle clearly affects the downgrades for class 7. But the ordering between both directional measures is reversed in the very risky categories, where the downgrade probabilities provide a leading indicator of the business cycle with a lead between 2 and 3 years.

6.4 Estimation of the factor ordered probit model

Finally we estimate the factor ordered probit model introduced in Section 3.2. We use the approximated linear Kalman filter for large cross-sectional dimension presented in Section 5.3. The estimation is performed for the wholesale sector.

In a first step we compute the transformed series $y_{kl,t} = G^{-1}(\hat{\pi}_{kl,t}^*)$, $\forall k, l$, and perform their principal component analysis, that is the spectral decomposition of matrix $\tilde{Y}\tilde{Y}'$, where the rows of \tilde{Y} are given by $y_{kl,t} - \bar{y}_{kl}$, k, l varying, with $\bar{y}_{kl} = \frac{1}{T} \sum_t y_{kl,t}$. The corresponding eigenvalues are given in decreasing order in the following table:

5.963	2.740	2.166	0.739	0.393	0.314	0.204	0.125	0.051	0
-------	-------	-------	-------	-------	-------	-------	-------	-------	---

Three eigenvalues are much larger than the other ones. The normalized eigenvectors corresponding to the 3 largest eigenvalues are displayed in Figure 5. The pattern of the factor corresponding to the largest eigenvalue is consistent with the evolution of downgrade probabilities reported in Figure 3, for all rating categories except the riskiest one (class 1). Indeed the factor points out an overall decreasing downgrade risk over the sample period, with peaks of downgrade probabilities in 1994 and in 1998-1999¹⁵. The factor corresponding to the second eigenvalue feature a similar pattern, but the peaks occur in 1995-1996 and 1999-2000. In particular the peak in 1995-1996 may be associated with the large downgrade probabilities featured by class 1 (the

¹⁵The sign of the factor has been chosen so that the corresponding estimated coefficients β_k are positive for each class. Thus the larger is the factor value, the larger are the downgrade probabilities.

riskiest class) in those years (see Figure 3). Finally it is important to see how the "business cycle" is related to the three factors. For this purpose the relative change in GDP has been regressed on the constant and the three factors. The regression coefficients are:

$$I_t = 1.970 - 0.303Z_{1,t} - 0.481Z_{2,t} + 0.046Z_{3,t},$$

with $R^2 = 0.19$. This regression analysis and the comparison with the pattern of GDP increments displayed in Figure 5 suggest that the factors corresponding to the two largest eigenvalues are related to the business cycle. Indeed the overall improvement in credit quality in 1992-2001 suggested by the factors is associated with the positive trend in GDP increments over the same period. Moreover the peak in downgrade risk in 1995-1996 may be related to the slow down in GDP increment in these years. However the evolutions of credit cycle and business cycle are not fully parallel [see Feng et alii (2004) for similar findings in US data]. For instance the peak in downgrade risk in 1998-1999 anticipates the slow growth years 2001-2002. This explains the rather poor fit in the regression.

The analysis is completed by applying the linear Kalman filter with 3 factors. This provides the dynamics of the factors and the estimated structural parameters, which are reported in Table 11.

Insert Table 11: Estimated structural parameters

As expected the estimated thresholds c are increasing. Similarly the intercepts α are increasing with respect to the rating index k , which confirms that downgrade risk is higher for the lower rating classes. The β_1 coefficients are higher and more homogeneous, and thus the first factor appears as a general factor. The β_2 coefficients show some opposition between the classes 1 - 4 and the classes 5 - 7, that is between speculative and investment categories. Finally the volatility parameters σ are generally smaller for the riskier rating categories.

Compared to the standard use of the ordered probit model suggested by the Basle Committee, we have followed a more general approach since:

- i) 3 factors have been introduced instead of a single factor as usual;
- ii) the factors have been defined endogenously by a principal component analysis and not chosen a priori;
- iii) the estimation has been performed per economic sector, likely more homogeneous than the whole population.

Nevertheless, it is seen in Figures 6 and 7, which provide some actual and fitted migration probabilities for rating class $k = 1$ (best class) and $k = 4$, respectively, that the fit is not entirely satisfactory (note that the scales are not the same on the different figures).

[Insert Figures 6, 7: Actual and fitted cumulated migration probabilities]

It does not mean that the ordered polytomous model has to be rejected, but there is still some specification errors. Some possible ones are the following:

- i) the population is not sufficiently homogenous;
- ii) factor Z_t can have an instantaneous effect by means of Z_t , but also lagged ones by means of Z_{t-1}, Z_{t-2}, \dots . Such lagged effects likely exist following the causality analysis of Section 6.3;
- iii) the latent distribution G can be different from the Gaussian one, and could feature different tail or skewness behaviours as in the factor Gompit model (see Section 3.2);
- iv) the latent distribution G can depend on the starting rating class. This is likely the main specification error as seen in Figure 7, for rating class 4. Indeed the general patterns of the actual migration probabilities are almost satisfactory, but some of them differ by a drift. This drift can be corrected by means of an appropriate choice of the G function.

These various specification tests are left for further research.

7 Concluding remarks

The stochastic migration model introduced in this paper is a specification which is flexible and appropriate for the joint analysis of rating migration of several firms. We have discussed several properties of the model concerning in particular the prediction of rating transitions, the migration correlations and some special features related to statistical inference. As an illustration, two stochastic migration models have been estimated on the French data set of the Banque de France: a model with i.i.d. stochastic transition matrices and a factor ordered probit specification. The first model underlies the standard measures of migration correlations, but is clearly misspecified, due in particular (but not only) to the effect of the business cycle, which induces serial dependence. The factor ordered probit model is able to account for the dynamics of the migration probabilities. We have performed one of the first estimations of such a model by endogenously selecting the factors, and

we have given some direction of future research for improving this approach. One main finding of this empirical analysis is that estimated migration correlations are much smaller than those suggested by the regulator.

As mentioned in the introduction, the stochastic migration model allows for the joint analysis of the individual rating histories, considered as given. However, it does not give direct information on how to construct the underlying scores. In particular, for large corporations, it does not explain if the score has to be based on fundamentals, such as balance sheets (as in S&P's, Moody's or the French central bank), or on bond and equity prices [as in KMV or in Duffie, Wang (2004)]. The very small values found for the migration correlations simply reflect the quality of the cross-sectional and time information used in the model. Less informative explanatory variables in the underlying scores and more aggregate risk categories will imply a less accurate prediction model, detected by larger migration correlations. The latter in turn induce an increase in the required capital, which naturally compensates for the low accuracy of the model. Instead of imposing unrealistic large values for the default correlation, as currently proposed by the Basle Committee, a better solution is clearly to introduce a coefficient explaining how to pass from the estimated CreditVaR to the required capital, coefficient that can depend on the quality of the rating and of the model. This is exactly the solution previously retained by the Committee to fix the market capital risk as function of the VaR for a portfolio of liquid stocks. Finally, small historical migration correlations do not necessarily imply small risk neutral migration correlations. In other words, even with small historical default correlation, it can be useful to introduce credit derivatives written on several firms.

REFERENCES

Albanese, C., Campolieti, J., Chen, O., and A., Zavidonov (2003): "Credit Barrier Models", University of Toronto DP.

Amemiya, T. (1976): "The Maximum Likelihood, the Minimum Chi-Square, and the Nonlinear Weighted Least Squares Estimators in the General Qualitative Response Model", *JASA*, 71, 347-351.

Anderson, T., and L., Goodman (1957): "Statistical Inference about Markov Chains", *Annals of Math. Statistics*, 28, 89-110.

Bahar, R., and K., Nagpal (2001): "Measuring Default Correlation", *Risk*, March, 129-132.

Bangia, A., Diebold, F., Kronimus, A., Schlagen, C., and T., Schuerman (2002): "Rating Migration and the Business Cycle, with Application to Credit Portfolio Stress Testing", *Journal of Banking and Finance*, 26, 445-474.

Bardos, M., Foulcher, S., and E., Bataille (2004): "Les Scores de la Banque de France: Méthode, Résultats, Applications", Banque de France, Observatoire des Entreprises.

Basle Committee on Banking Supervision (2002), *Quantitative Impact Study 3. Technical Guidance*, Banking for International Settlements.

Berman, S. M., and H., Frydman (1999): "Parametric Estimation of Hazard Functions with Stochastic Covariate Processes", *Sankhya, Series A*, 61, 174-188.

Berkson, J. (1944): "Application of the Logistic Function to Bio-assay", *JASA*, 39, 357-365.

Bharath, S., and T., Shumway (2004): "Forecasting Default with the KMV-Merton Model", Working Paper, University of Michigan.

Blume, M., Liu, F., and C., McKinlay (1998): "The Declining Credit Quality of US Corporate Debt: Myth or Reality?", *Journal of Finance*, 53, 1389-1413.

Brady, B., and R., Bos (2002): "Record Default in 2001: the Result of Poor Credit Quality and a Weak Economy", Standard and Poor's, February.

Chava, S., and R., Jarrow (2002): "Bankruptcy Prediction with Industry Effects", Working Paper, Cornell University.

Cox, D. R. (1972): "The Statistical Analysis of Dependences in Point Processes", in *Symposium on Point Processes*, ed. by P. A. W. Lewis. New York, John Wiley, 55-66.

Crouhy, M., Galai, D., and R. Mark (2000): "A Comparative Analysis of Current Credit Risk Models", *Journal of Banking and Finance*, 29, 59-117.

De Servigny, A., and O., Renault (2002): "Default Correlation: Empirical Evidence", Standard and Poor's, September.

Duffie, D., and D., Lando (2001): "Term Structure of Credit Spreads with Incomplete Accounting Information", *Econometrica*, 69, 633-664.

Duffie, D., and K., Singleton (1999): "Modeling the Term Structure of Defaultable Bonds", *The Review of Financial Studies*, 12, 687-720.

Duffie, D., and K., Wang (2004): "Multi-Period Corporate Failure Prediction with Stochastic Covariates", Working Paper.

Feng, D., Gouriéroux, C., and J., Jasiak (2004): "The Ordered Qualitative Model for Credit Rating Transitions", CREST DP.

Foulcher, S., Gouriéroux, C., and A., Tiomo (2004): "The Term Structure of Defaults and Ratings", *Insurance and Risk Management*, .

Frey, R., and A., McNeil (2001): "Modelling Dependent Defaults", Working Paper, ETH Zurich.

Frey, A., and A., McNeil (2003): "Dependent Defaults in Models of Portfolio Credit Risk", *Journal of Risk*, 6, 59-92.

Gagliardini, P., and C., Gouriéroux (2004): "Migration Correlation: Definition and Consistent Estimation", forthcoming *Journal of Banking and Finance*.

Gordy, J., and E., Heitfield (2001): "Of Moody's and Merton: a Structural Model of Bond Rating Transitions", Working Paper, Federal Reserve.

Gordy, J., and E., Heitfield (2002): "Estimating Default Correlation from Short Panels of Credit Rating Performance Data", Working Paper, Federal Reserve Board.

Gouriéroux, C., and J., Jasiak (2001): *Financial Econometrics*, Princeton University Press.

Gouriéroux, C., and J., Jasiak (2004): "Jacobi Process with Application to Switching Regimes", forthcoming *Journal of Econometrics*.

Gouriéroux, C., and J., Jasiak (2005): *Microeconometrics for Credit, Insurance, and Marketing*, forthcoming Princeton University Press.

- Gouriéroux, C., and A., Monfort (1995): *Simulation Based Econometric Methods*, Louvain, CORE Lecture Series, Oxford University Press.
- Gouriéroux, C., and A., Monfort (2002): "Equidependence in Qualitative and Duration Models with Application to Credit Risk", CREST DP.
- Gouriéroux, C., and A., Monfort (2004): "Error-in-Factor Models", CREST DP.
- Gouriéroux, C., Monfort, A., and V., Polimenis (2003): "Affine Models for Credit Risk", CREST DP.
- Gupton, G., Finger, C., and M., Bahtia (1997): "Creditmetrics-technical document", Technical Report, the RiskMetrics Group.
- Hsiao, C., Pesaran, H., and K., Tahmiscioglu (2002): "Maximum Likelihood Estimation of Fixed Effects Dynamic Panel Data Models Covering Short Time Periods", *Journal of Econometrics*, 109, 107-150.
- Joe, H. (1997): *Multivariate Models and Dependence Concepts*, Monographs on Statistics and Applied Probability, 73, Chapman & Hall.
- Kavvathas, D. (2001): "Estimating Credit Rating Transition Probabilities for Corporate Bonds", Working Paper.
- Lando, D. (1998): "On Cox Processes and Credit Risky Securities", *Review of Derivative Research*, 2, 99-120.
- Lucas, D. (1995): "Default Correlation and Credit Analysis", *Journal of Fixed Income*, March, 76-87.
- Merton, R. (1974): "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates", *Journal of Finance*, 20, 449-470.
- Nickell, P., Perraudin, W., and S. Variotto (2000): "Stability of Rating Transitions", *Journal of Banking and Finance*, 24, 203-227.
- Pender, K. (1992): "Demystifying the Ratings Game", *The San Francisco Chronicle*, February, 17.
- Rosch, D. (2004): "An Empirical Comparison of Default Risk Forecasts from Alternative Credit Rating Philosophies", forthcoming in *International Journal of Forecasting*.
- Schonbucher, P. (2000): "Factor Models for Portfolio Credit Risk", University of Bonn DP.
- Schonbucher, P. (2004): "Frailty Models, Contagion and Information Effects", Working Paper, ETH Zurich.

Table 1
1-year transition matrix

	7	6	5	4	3	2	1	0	NR
7	0.7155	0.0979	0.0241	0.0113	0.0031	0.0002	0.0002	0.0002	0.1474
6	0.1226	0.5977	0.1031	0.0321	0.0187	0.0022	0.0008	0.0011	0.1216
5	0.0169	0.2544	0.4232	0.1145	0.0477	0.0079	0.0025	0.0026	0.1305
4	0.0085	0.0600	0.2519	0.3517	0.1223	0.0297	0.0109	0.0070	0.1579
3	0.0011	0.0413	0.0924	0.2450	0.2934	0.0755	0.0339	0.0200	0.1974
2	0.0000	0.0114	0.0509	0.1500	0.2500	0.1640	0.0842	0.0307	0.2588
1	0.0000	0.0076	0.0317	0.0544	0.1903	0.1224	0.1813	0.0650	0.3474
0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000

Table 1: 1-year transition matrix for the wholesale sector in year 2001.

Table 2
Adjusted 1-year transition matrix

	7	6	5	4	3	2	1	0
7	0.8392	0.1148	0.0282	0.0133	0.0036	0.0003	0.0003	0.0003
6	0.1396	0.6804	0.1174	0.0366	0.0213	0.0025	0.0009	0.0013
5	0.0194	0.2925	0.4867	0.1316	0.0549	0.0090	0.0029	0.0030
4	0.0101	0.0713	0.2991	0.4177	0.1452	0.0352	0.0130	0.0084
3	0.0014	0.0514	0.1152	0.3053	0.3656	0.0940	0.0422	0.0249
2	0.0000	0.0154	0.0686	0.2024	0.3373	0.2213	0.1136	0.0414
1	0.0000	0.0116	0.0486	0.0833	0.2917	0.1875	0.2778	0.0995
0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000

Table 2: Adjusted 1-year transition probabilities for the wholesale sector in 2001.

Table 3
Average 1-year transition matrix

	7	6	5	4	3	2	1	0
7	0.7968	0.1491	0.0317	0.0165	0.0040	0.0007	0.0003	0.0009
6	0.1227	0.6476	0.1458	0.0515	0.0272	0.0030	0.0010	0.0012
5	0.0162	0.2388	0.4935	0.1627	0.0703	0.0125	0.0032	0.0028
4	0.0084	0.0597	0.2587	0.4371	0.1795	0.0373	0.0124	0.0069
3	0.0026	0.0344	0.1002	0.2777	0.4095	0.1136	0.0430	0.0190
2	0.0017	0.0129	0.0494	0.1642	0.3420	0.2524	0.1338	0.0436
1	0.0015	0.0102	0.0276	0.0937	0.2432	0.2037	0.3312	0.0889
0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000

Table 3: Estimated average 1-year transition matrix for the wholesale sector.

Table 4
Joint downgrade probabilities

	7	6	5	4	3	2	1	0
7	0.0225	0.0219	0.0244	0.0269	0.0170	0.0199	0.0132	—
6	0.0219	0.0214	0.0239	0.0263	0.0167	0.0196	0.0130	—
5	0.0244	0.0239	0.0267	0.0295	0.0187	0.0220	0.0145	—
4	0.0269	0.0263	0.0295	0.0326	0.0207	0.0243	0.0161	—
3	0.0170	0.0167	0.0187	0.0207	0.0131	0.0155	0.0102	—
2	0.0199	0.0196	0.0220	0.0243	0.0155	0.0183	0.0120	—
1	0.0132	0.0130	0.0145	0.0161	0.0102	0.0120	0.0081	—
0	—	—	—	—	—	—	—	—

Table 4: Joint downgrade probabilities for two firms in the wholesale sector. Row and column numbers denote the initial rating class of the two firms.

Table 5
Downgrade correlations

	7	6	5	4	3	2	1	0
7	0.0021	0.0013	0.0011	0.0008	0.0006	0.0000	-0.0004	-
6	0.0013	0.0011	0.0010	0.0011	0.0009	0.0007	0.0003	-
5	0.0011	0.0010	0.0016	0.0019	0.0017	0.0016	0.0003	-
4	0.0008	0.0011	0.0019	0.0027	0.0024	0.0026	0.0010	-
3	0.0006	0.0009	0.0017	0.0024	0.0024	0.0025	0.0011	-
2	0.0000	0.0007	0.0016	0.0026	0.0025	0.0034	0.0011	-
1	-0.004	0.0003	0.0003	0.0010	0.0011	0.0011	0.0031	-
0	-	-	-	-	-	-	-	-

Table 5: Downgrade correlations for two firms in the wholesale sector. The row and column numbers denote the initial rating classes of the two firms.

Table 6
Default correlation in the wholesale sector

	7	6	5	4	3	2	1	0
7	0.0004	0.0002	0.0002	0.0000	0.0000	0.0000	0.0000	-
6	0.0002	0.0004	0.0002	0.0002	0.0002	0.0000	0.0000	-
5	0.0002	0.0002	0.0003	0.0001	0.0002	-0.0001	0.0003	-
4	0.0000	0.0002	0.0001	0.0003	0.0002	0.0002	0.0005	-
3	0.0000	0.0002	0.0002	0.0002	0.0006	-0.0003	0.0009	-
2	0.0000	0.0000	-0.0001	0.0002	-0.0003	0.0006	0.0000	-
1	0.0000	0.0000	0.0003	0.0005	0.0009	0.0000	0.0031	-
0	-	-	-	-	-	-	-	-

Table 6: Default correlations for two firms in the wholesale sector. The row and column numbers denote the initial rating classes of the two firms.

Table 7
Default correlation in the retail trade sector

	7	6	5	4	3	2	1	0
7	0.0009	0.0000	0.0003	0.0003	0.0009	0.0007	0.0005	—
6	0.0000	0.0001	0.0000	0.0001	0.0000	0.0000	0.0001	—
5	0.0003	0.0000	0.0003	0.0000	0.0004	0.0003	0.0003	—
4	0.0003	0.0001	0.0000	0.0007	0.0006	0.0004	0.0006	—
3	0.0009	0.0000	0.0004	0.0006	0.0013	0.0010	0.0011	—
2	0.0007	0.0000	0.0003	0.0004	0.0010	0.0014	0.0007	—
1	0.0005	0.0001	0.0003	0.0006	0.0011	0.0007	0.0014	—
0	—	—	—	—	—	—	—	—

Table 7: Default correlations for two firms in the retail trade sector. The row and column numbers denote the initial rating classes of the two firms.

Table 8
Autocorrelations

wholesale		7	6	5	4	3	2	1
down	order 1	0.0430	0.7902	0.6907	0.4373	0.2541	-0.2316	0.2516
	order 2	0.3182	0.3571	0.6878	0.5267	0.4373	0.1686	-0.3434
up	order 1	-	0.0551	0.3146	0.3815	0.2800	0.0283	0.2059
	order 2	-	-0.2284	0.0627	0.1954	0.4626	0.5549	0.0974

retail trade		7	6	5	4	3	2	1
down	order 1	0.0521	0.7573	0.4433	0.2096	0.4497	0.7131	-0.2307
	order 2	-0.1677	0.6729	0.3902	0.3687	0.5290	0.3226	-0.1481
up	order 1	-	0.4506	0.0959	0.0730	0.4863	0.1014	0.6707
	order 2	-	0.3620	0.3880	0.3867	0.4573	0.2222	0.3436

Table 8: Autocorrelations of order 1 and 2 for downgrade and upgrade probabilities in the wholesale sector (upper Panel), respectively in the retail trade sector (lower Panel).

Table 9
Causality relations in the wholesale sector

	7	6	5	4	3	2	1
$C_{I \rightarrow d}$	10.51	5.62	5.79	5.79	7.27	4.68	0.47
$C_{d \rightarrow I}(1)$	1.18	0.44	0.08	0.47	1.47	3.25*	1.02
$C_{d \rightarrow I}(2)$	1.54	1.86	0.02	0.14	0.52	2.43	3.10*
$C_{d \rightarrow I}(3)$	2.39	5.12	2.74*	3.51*	3.07*	2.12	5.79
$C_{I \rightarrow u}$	—	2.27	8.97	6.72	7.96	7.69	0.44
$C_{u \rightarrow I}$	—	0.34	1.38	1.38	1.52	2.36	0.36

Table 9: Causality relations in the wholesale sector. The causality measures are multiplied by T . Under the null hypothesis of no linear link, these standardized statistics are asymptotically $\chi^2(1)$ -distributed. Bold entries (resp. entries with an asterisk) correspond to significant values at the 5% (10%, respectively) level. For the causality directions $d \rightarrow I$ we provide measures at horizon $h = 1, 2, 3$.

Table 10
Causality relations in the retail trade sector

	7	6	5	4	3	2	1
$C_{I \rightarrow d}$	29.28	2.74	2.70	4.08	2.11	0.24	0.01
$C_{d \rightarrow I}(1)$	1.75	0.11	0.20	0.81	0.01	1.50	0.98
$C_{d \rightarrow I}(2)$	1.82	0.23	0.17	0.26	0.00	0.80	6.54
$C_{d \rightarrow I}(3)$	2.58	1.06	2.49	1.97	2.51	2.44	2.06
$C_{I \rightarrow u}$	—	8.01	4.43	3.42	1.39	0.82	0.42
$C_{u \rightarrow I}$	—	1.58	1.18	1.23	0.00	0.62	0.23

Table 10: Causality relations in the retail trade sector. The causality measures are multiplied by T . Under the null hypothesis of no linear link, these standardized statistics are asymptotically $\chi^2(1)$ -distributed. Bold entries correspond to significant values at the 5% level. For the causality directions $d \rightarrow I$ we provide measures at horizon $h = 1, 2, 3$.

Table 11
Estimated structural parameters

$c_1 = 0$	$c_2 = 0.6611$	$c_3 = 0.9470$	$c_4 = 1.1776$	$c_5 = 1.3714$
$c_6 = 1.4680$	$c_7 = 1.5662$			

$\alpha_1 = -0.4410$	$\beta_{1,1} = 0.0210$	$\beta_{2,1} = 0.0389$	$\beta_{3,1} = 0.0097$	$\sigma_1 = 1$
$\alpha_2 = 0.6843$	$\beta_{1,2} = 0.0020$	$\beta_{2,2} = 0.0091$	$\beta_{3,2} = 0.0065$	$\sigma_2 = 0.4067$
$\alpha_3 = 1.0026$	$\beta_{1,3} = 0.0086$	$\beta_{2,3} = 0.0066$	$\beta_{3,3} = 0.0049$	$\sigma_3 = 0.2621$
$\alpha_4 = 1.1589$	$\beta_{1,4} = 0.0066$	$\beta_{2,4} = 0.0073$	$\beta_{3,4} = 0.0048$	$\sigma_4 = 0.2316$
$\alpha_5 = 1.2902$	$\beta_{1,5} = 0.0280$	$\beta_{2,5} = -0.0010$	$\beta_{3,5} = 0.0033$	$\sigma_5 = 0.2039$
$\alpha_6 = 1.3884$	$\beta_{1,6} = 0.0252$	$\beta_{2,6} = 0.0050$	$\beta_{3,6} = 0.0002$	$\sigma_6 = 0.1655$
$\alpha_7 = 1.4766$	$\beta_{1,7} = 0.0134$	$\beta_{2,7} = 0.0025$	$\beta_{3,7} = 0.0157$	$\sigma_7 = 0.1712$

$A_{11} = 0.2538$	$A_{12} = -0.0087$	$A_{13} = -0.1887$
$A_{21} = -0.7646$	$A_{22} = -0.0832$	$A_{23} = -0.2578$
$A_{31} = -0.0675$	$A_{32} = -0.1186$	$A_{33} = 0.0324$

Table 11: Estimated structural parameters for the factor ordered probit model. Thresholds c , intercepts α , slope coefficients β for the three factors, volatilities σ and autoregressive coefficients A are displayed.

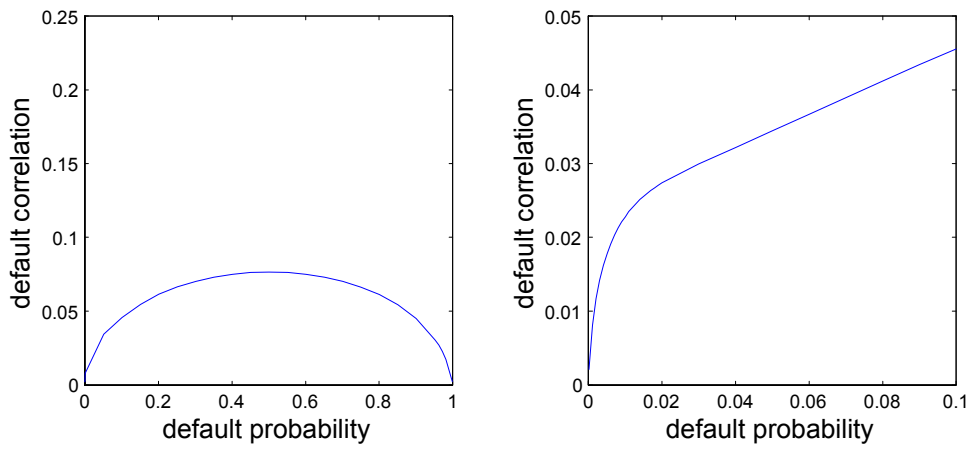


Figure 1: Default correlation as a function of default probability by Basle Committee.

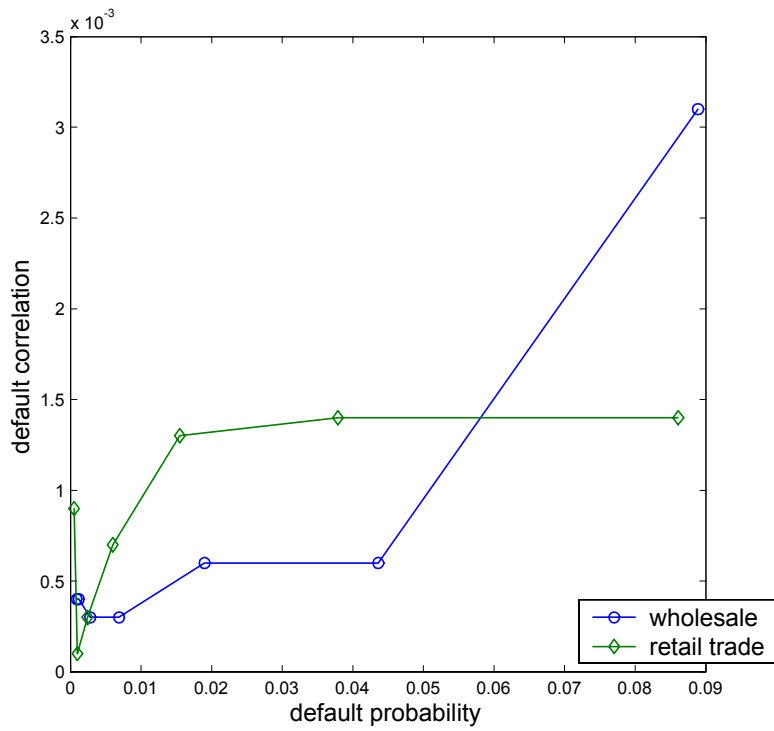


Figure 2: Default correlation as a function of default probability across the different rating classes for the wholesale and the retail trade sectors.

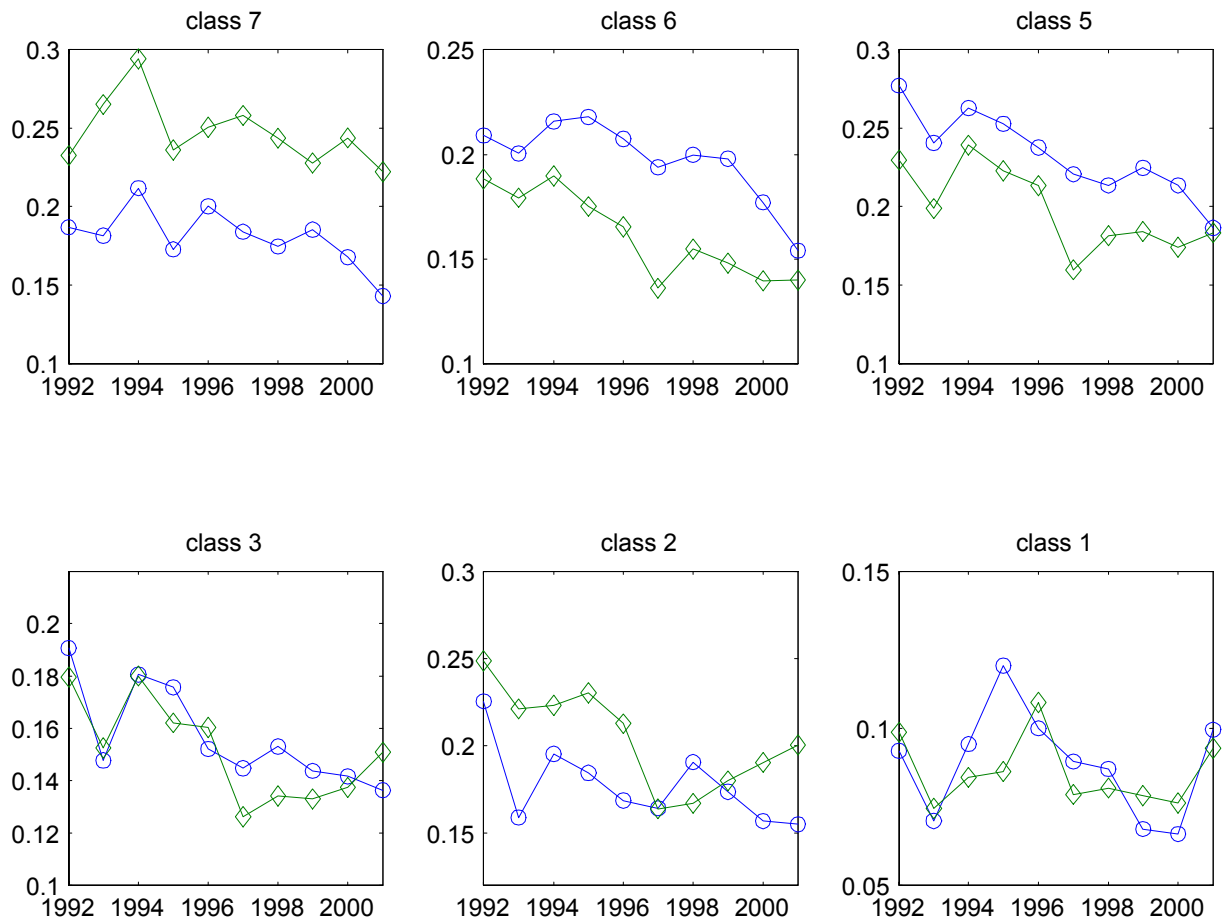


Figure 3: Probabilities of a downgrade of at most 2 buckets for the different rating classes. The circles (resp. the diamonds) correspond to the wholesale (retail trade) sector.

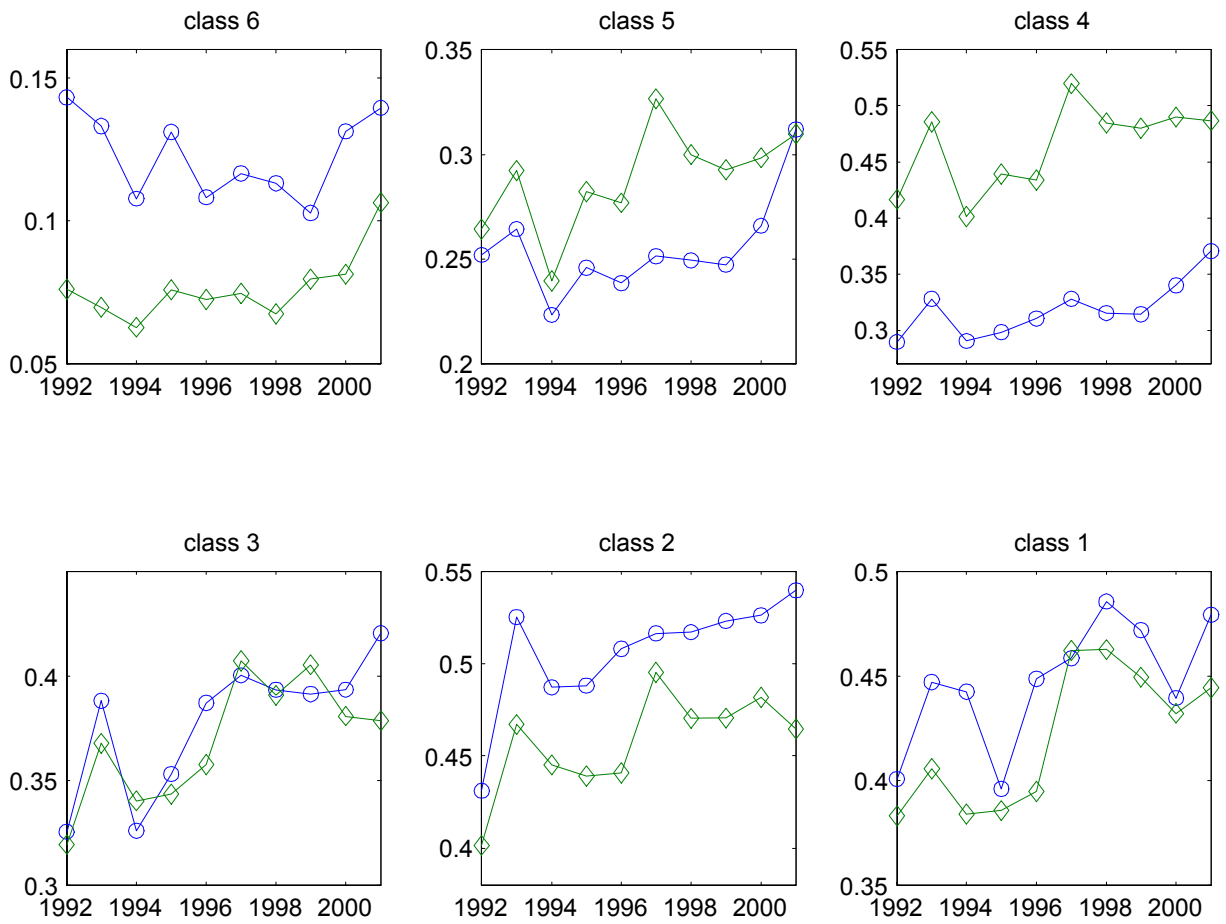


Figure 4: Probabilities of an upgrade of at most 2 buckets for the different rating classes. The circles (resp. the diamonds) correspond to the wholesale (retail trade) sector.

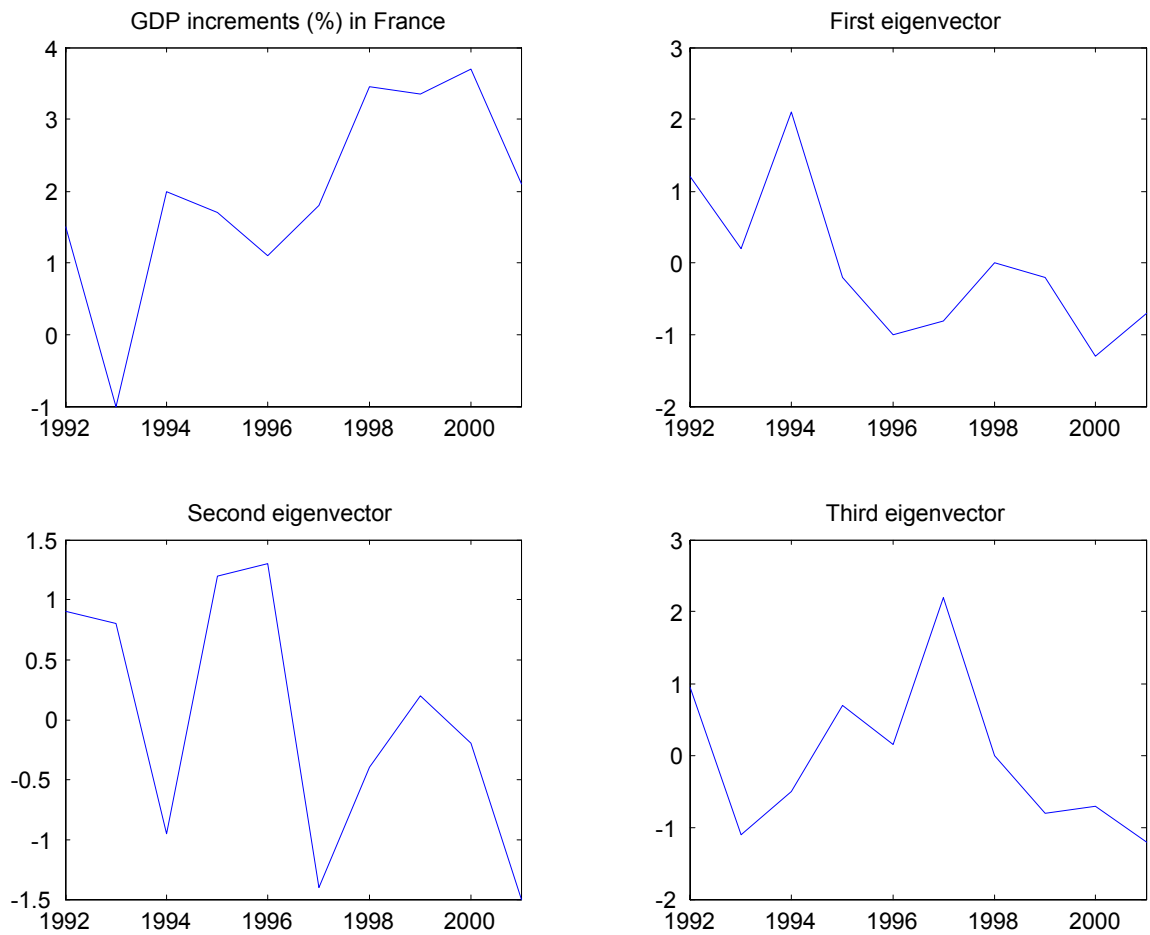


Figure 5: The upper left Panel displays the GDP percentage increments in France in the period 1992-2001. The other Panels display the time evolution of the factors corresponding to the three largest eigenvalues.

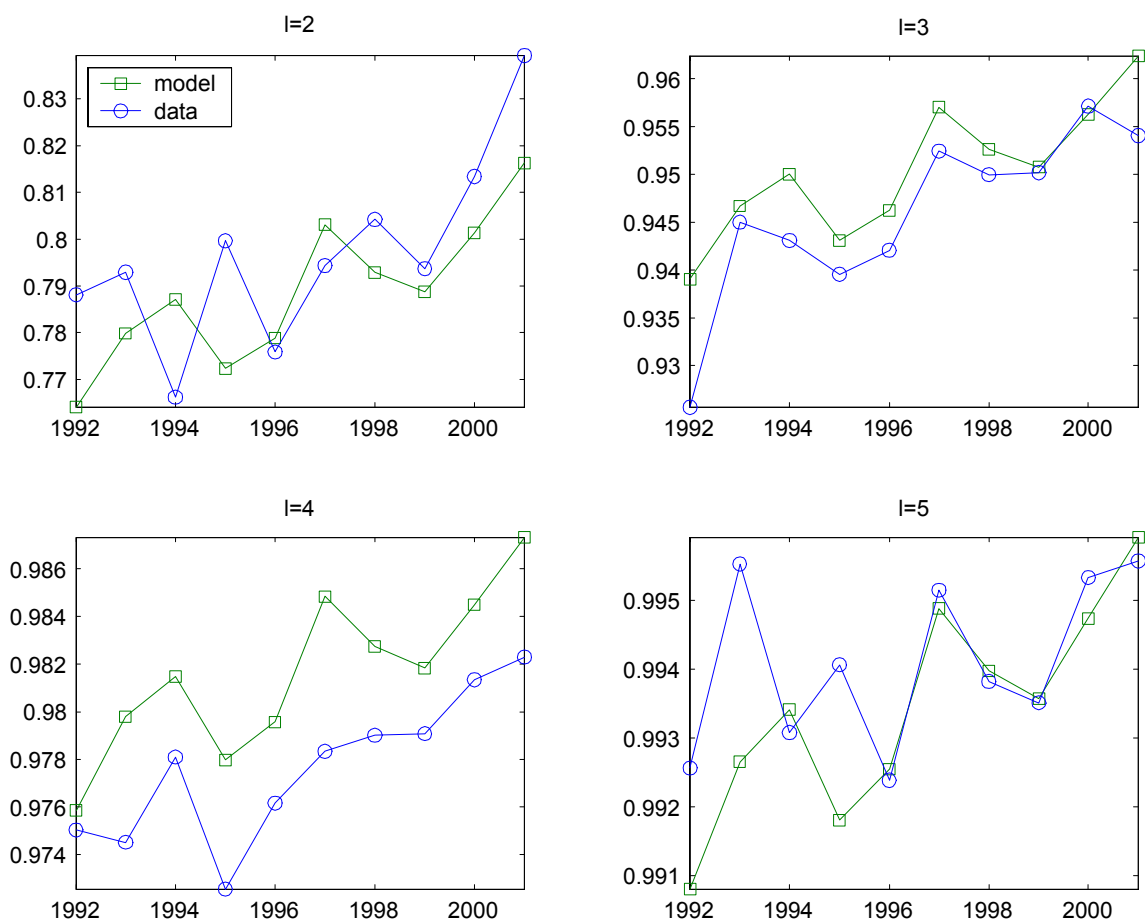


Figure 6: Cumulated probabilities $\pi_{kl,t}^*$ for rating class $k = 1$ and different indices l in the wholesale sector: circles correspond to observed probabilities and squares to fitted probabilities.

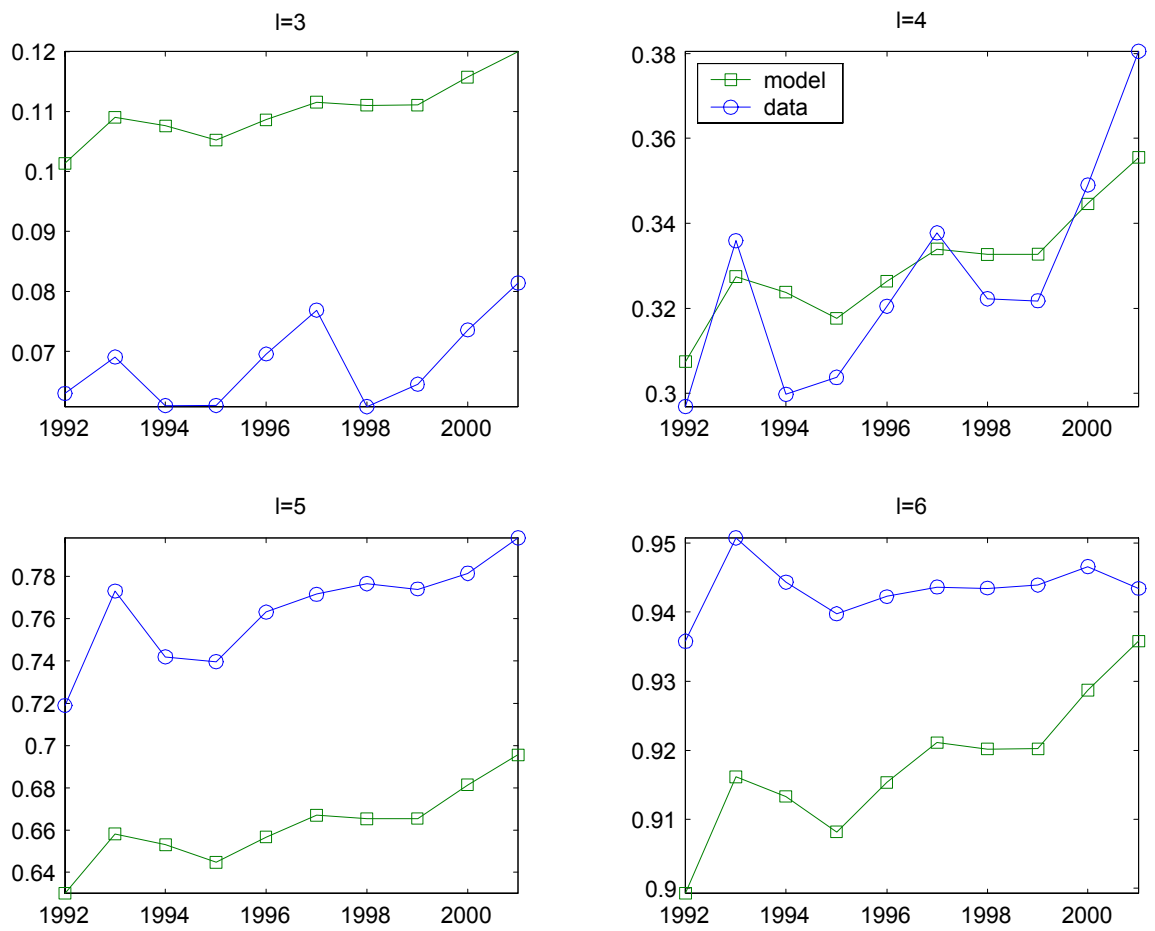


Figure 7: Cumulated probabilities $\pi_{kl,t}^*$ for rating class $k = 4$ and different indices l in the wholesale sector: circles correspond to observed probabilities and squares to fitted probabilities.