Numerical estimates of risk factors contingent on credit ratings

T. Gärtner · S. Kaniovski · Y. Kaniovski

Received: 2 March 2020 / Accepted: 4 June 2021 / Published online: 24 June 2021 © The Author(s) 2021

Abstract
Assuming a favorable or an adverse outcome for every combination of a credit class and an industry sector, a binary string, termed as a macroeconomic scenario, is considered. Given historical transition counts and a model for dependence among credit-rating migrations, a probability is assigned to each of the scenarios by maximizing a likelihood function. Applications of this distribution in financial risk analysis are suggested.

Two classifications are considered: 7 non-default credit classes with 6 industry sectors and 2 non-default credit classes with 12 industry sectors. We propose a heuristic algorithm for solving the corresponding maximization problems of combinatorial complexity. Probabilities and correlations characterizing riskiness of random events involving several industry sectors and credit classes are reported.

Keywords Macroeconomic scenario · Combinatorial complexity · Maximum likelihood · Random search · Genetic algorithm · Penalty

Mathematics Subject Classification 90C30 · 90C90

This research was supported by the Free University of Bozen - Bolzano. We thank the two referees for their constructive critique and valuable suggestions.

Y. Kaniovski
YKaniovskyi@unibz.it

T. Gärtner
Timon.Gaertner@provincia.bz.it

S. Kaniovski
Serguei.Kaniovski@wifo.ac.at

1 Provincial Statistics Institute (ASTA T), Autonomous Province of Bolzano-South Tyrol, Via Canonico Michael Gamper 1, 39100 Bolzano, Italy

2 Austrian Institute for Economic Research (WIFO), Arsenal 20, 1030 Vienna, Austria

3 Faculty of Economics and Management, Free University of Bozen-Bolzano, Piazza Università 1, 39100 Bolzano, Italy
1 Motivation

Intuitively, migrations towards more secure credit classes should be more frequent during economic upturns and less frequent during downturns, while migrations towards riskier states should exhibit the opposite pattern. With this observation in mind, we analyze statistically the variety of economic conditions affecting different industries and credit classes. Given the actually observed over a period of time credit-rating migrations, we identify the most likely distribution $D$ of upturns and downturns for this period. We illustrate how the distribution can be used in financial risk analysis.

The two phases of a business cycle—a downturn or, more specifically, a contraction and an upturn or an expansion—, affecting all debtors in the economy, render their credit-rating migrations dependent. This mechanism of dependence is implemented in several models. See among others Bangia et al. (2002), McNeil and Wendin (2007), Frydman and Schuermann (2008), Stefanescu et al. (2009), Xing et al. (2012). In the same vein, Fei et al. (2012) consider three phases: an expansion, a mild recession and a severe recession. The GDP dynamics is necessary for identifying the corresponding time periods. Knowing them, the transition matrices, if discrete time is considered, or the respective infinitesimal generator matrices, in a continuous time setting, are estimated for each of the phases. In sum, the dependent credit-rating migrations are modelled relying on the conventional business cycle theory or its extension to more than two phases of a business cycle. As a consequence, the migration probabilities are not industry specific. Therefore, neither are intensities of credit-rating events considered, nor their interdependence. In this paper we propose a novel model, where the intensities as well as the interdependencies can be industry specific.

Consider a pool of debtors. Assume that there are $M$ non-default credit classes and $S$ industry sectors. Let every debtor be completely characterized by a couple $m \in \{1, 2, \ldots, M\}$ and $s \in \{1, 2, \ldots, S\}$. We postulate, first, that the economic conditions affecting a debtor can be favorable or adverse and, second, that the conditions are specific for every combination of a credit class and an industry sector. Then, keeping 1 for the favorable outcome and 0 for the adverse one, a state of the economy, or a macroeconomic scenario, is encoded with a binary string having $MS$ positions. There are $2^{MS}$ such strings. For a practically interesting choice of $M$ and $S$, this can be a huge number. The conventional setting, assuming just two phases of a business cycle for the whole economy, has neither the sectoral classification nor the riskiness one. Therefore, $M$ and $S$ takes value 1 and there are just two strings in this case, 1 and 0.

Given a model for the probabilistic dependence among individual credit-rating migrations, a discrete time dynamic of the credit rating can be specified and the corresponding likelihood function can be written down. We formulate the model through the conditional on a macroeconomic scenario transition probabilities. Then the likelihood function depends upon a distribution $D$ defined on the $2^{MS}$ binary strings. Given the actually observed migrations, the maximum likelihood principle allows to identify the most likely $D$. Should the support of this distribution contain just two sample points—the binary string formed by 1 repeated $MS$ times and the binary string with 0 at all $MS$ positions—this setting would reduce to the conventional model with just two phases of a business cycle. In fact, all industries and credit classes would be affected by the same economic conditions in this case. Otherwise, if there are more than two
sample points in the support, analyzing the distribution $D$ and its moments can, on the one hand, shed light on hidden, because the corresponding market outcomes are at most partly observed, dependencies among credit events in different credit classes and industry sectors and, on the other hand, imply more specific and precise, because the sectoral classification is taken into account, estimates of the risk associated with a particular portfolio. In sum, the classical two-phases regime can emerge for some input data. However, the estimates given below show that the support of $D$ contains typically more than two sample points. Consequently, a fine-grained differentiation of the phases of a business cycle is possible. We present several new numerical estimates of financial risk factors based on this more general view of a business cycle.

Within our approach, the sectoral classification is not the only one possible. In particular, the model can be reformulated classifying debtors according to their capitalization level. Such a specification is typical for some macroeconomic risks factor models. See Aretz et al. (2010), Boons (2016), Fama and French (2017), Cooper and Maio (2019), Goncalves et al. (2020) among others. On the one hand, they employ more particular than the phases of a business cycle macroeconomic indicators—exchange rate, employment, inflation rate, GDP, etc.—, on the other hand, their predictions address the whole economy rather than a credit class. In other words, conceptually such models are richer, but, ceteris paribus, our analysis is more fine-grained.

For testing the approach, we use the same S&P’s dataset as Boreiko et al. (2017). See https://doi.org/10.1371/journal.pone.0175911.s001. Two couples of $M$ and $S$ are considered: $M = 7$ with $S = 6$ and $M = 2$ with $S = 12$. Since $2^{42}$ and, respectively, $2^{24}$ macroeconomic scenarios are involved, the number of unknown parameters greatly exceeds the number of available observations. A heuristic algorithm is suggested for solving the corresponding likelihood maximization problems of combinatorial complexity.

A numerical technique for evaluating the risk factors contingent on the dependence among credit-rating migrations is the main contribution of the paper. The key element is an estimate of the hidden distribution that shapes the migrations.

In the next session we specify the probabilistic dependence among individual migrations, the corresponding likelihood function and the non-linear programming problem for estimating the unknown parameters. A heuristic algorithm for solving this problem with a combinatorial number of unknowns is described in Sect. 3. Several new risk estimates are suggested in Sect. 4. A detailed characterization of the input data and the heuristic solutions is presented in Sect. 5. Section 6 contains the main results of the paper—several risk estimates based on the distribution $D$. Section 7 concludes. Auxiliary technicalities are given in Appendix 1. Appendix 2 contains some statistical characteristics of the input data.

## 2 Dependence model and optimization problem

Modeling credit-rating migrations as trajectories of time-homogeneous Markov chains has a long tradition in financial risk analysis. It is implemented in CreditMetrics, a toolkit for understanding and managing credit risk. See Gupton et al. (1997). In the model considered next, a time-homogeneous discrete-time Markov chain is as well one
of the two components governing credit-rating migrations. This component is termed as idiosyncratic because, should it be the only driver of the credit rating migrations, they would not depend on each other and the resulting dynamic would be Markovian.

Let us formulate precisely the assumptions regarding the credit-rating dynamics. Consider a portfolio involving debtors classified into \( S \geq 1 \) industry sectors. There are \( M + 1 \) levels of creditworthiness. Non-default credit classes are numbered in descending order of creditworthiness: the most secure assets are labelled by 1, while the next to default credit class is indexed by \( M \). Defaulted debtors receive the index \( M + 1 \). They will never return to business.

Denote by \( \{0, 1\}^{MS} \) the set of binary strings (or vectors) \( V \) with \( MS \) positions (or coordinates). They encode all possible macroeconomic scenarios in the economy. We need a rule for assigning coordinates of a binary vector \( V \) to industries and credit classes. To this end, let the coordinate \( VM(s−1)+m \) characterize the economic conditions affecting the credit class \( m \) of the industry \( s \). That is, the industries occupy blocks of \( M \) coordinates each. The blocks are numbered in ascending order of \( s \). Within a block, the credit classes are listed in ascending order of \( m \).

Let \( X(t) \) be the credit rating of a debtor at time \( t = 1, 2, \ldots \). The rating randomly changes in time, becoming \( X(t+1) \) at \( t + 1 \), while the assignment to an industry remains the same. We postulate that \( X(t+1) \) is a weighted sum of an idiosyncratic \( \xi^t \) and a common \( \eta^t \) component:

\[
X(t+1) = \delta^t \xi^t + (1 - \delta^t)\eta^t. \tag{1}
\]

First, we characterize distributions of the random variables \( \delta^t, \xi^t \) and \( \eta^t \). Let the debtor belong to an industry \( s \) and let \( X(t) = m \) for some \( m \leq M \). Then:

- \( \delta^t \) stands for a Bernoulli random variable whose probability of success equals \( Q_{m,s} \). The \( M \times S \) matrix \( Q \) with entries \( Q_{m,s} \) is not known. It has to be estimated.
- Migrations in industry \( s \) that cannot be attributed, at least directly, to market mechanisms are characterized by an \( M \times (M + 1) \) migration matrix \( P^{(s)} \). It is known. Its entry \( P^{(s)}_{i,j} \) equals the probability that a debtor moves from a credit class \( i \) to a credit class \( j \) in one time instant. The distribution of \( \xi^t \) is:

\[
P\{\xi^t = j\} = P^{(s)}_{m,j}, \quad j = 1, 2, \ldots, M + 1.
\]

- The conditional on \( V_{M(s−1)+m} \) distribution of \( \eta^t \) reads:

\[
P^{(s)}_{m,j}(1) = \begin{cases} 
P^{(s)}_{m,j} & \text{if } j < m, \\
P^{(s)}_{m,m} & \text{if } j = m, \\
0 & \text{if } j > m;
\end{cases}
\]

\[
P^{(s)}_{m,j}(0) = \begin{cases} 
1 - P^{(s)}_{m,j} & \text{if } j > m, \\
(1 - \Delta_{m,s})P^{(s)}_{m,m} & \text{if } j = m, \\
0 & \text{if } j < m.
\end{cases}
\]
Here, $P_{m}^{(s)} = P_{m,1}^{(s)} + P_{m,2}^{(s)} + \ldots + \Delta_{m,s} P_{m,m}^{(s)}$. The $M \times S$ matrix $\Delta$ with elements $\Delta_{m,s} \in [0,1]$ is not known. It has to be estimated.

Then the conditional on a macroeconomic scenario $V$ distribution of $X(t + 1)$ can be written down in the following way:

$$Q_{m,s} P_{m,j}^{(s)} + (1 - Q_{m,s}) P_{m,j}^{(s)} (V_{M(s-1)+m}), \, j = 1, 2, \ldots, M + 1. \quad (2)$$

Note that, defining the conditional migration probabilities, our model does not require identification of the time instants corresponding to the phases of a business cycle. Consequently, no GDP dynamics is necessary to estimate the probabilities. There are several possibilities to relate the trend of credit rating migrations to the phases of a business cycle. In the simplest case, it is required that the probabilities of migrating towards more secure credit classes are higher for economic upturns than for downturns. Incorporating such linear inequality constraints into a likelihood maximization problem, two migration matrices, one for upturns and the other for downturns, are estimated as well as the probability of an upturn. For survey analysis, this approach was used in Hölzl et al. (2019).

Second, let us characterize the dependence among the random variables involved in (1). For a fixed time instant $t$:

- for every debtor, the random variables $\delta^t$, $\xi^t$ and $\eta^t$ are independent;
- the random variables $\delta^t$, $\xi^t$ are independent across debtors, while the random variables $\eta^t$ are conditionally, given a macroeconomic scenario, independent across debtors;
- $\delta^t$, $\xi^t$ and $\eta^t$ do not depend upon their realizations in the past.

Considering a period of observation from $t = 1$ to $t = T$, these assumptions imply the following likelihood function:

$$L(D, Q, \Delta) = \prod_{t=1}^{T} \sum_{V \in \{0,1\}^{MS}} D(V) \prod_{s=1}^{S} \prod_{m=1}^{M} \prod_{m+1}^{M+1} F(s, V, m_1, m_2, Q, \Delta) I'(s,m_1,m_2),$$

where

$$F(s,V,m_1,m_2,Q,\Delta) = \begin{cases} 
 Q_{m_1,s} + (1 - Q_{m_1,s}) \frac{\Delta_{m_1,s}}{p_{m_1}^{(s)}} & \text{if } m_1 = m_2, \, V_{M(s-1)+m_1} = 1; \\
 Q_{m_1,s} + (1 - Q_{m_1,s}) \frac{1 - \Delta_{m_1,s}}{1 - p_{m_1}^{(s)}} & \text{if } m_1 = m_2, \, V_{M(s-1)+m_1} = 0; \\
 Q_{m_1,s} + \frac{1 - Q_{m_1,s}}{p_{m_1}^{(s)}} & \text{if } m_1 > m_2, \, V_{M(s-1)+m_1} = 1; \\
 Q_{m_1,s} + \frac{1 - Q_{m_1,s}}{1 - p_{m_1}^{(s)}} & \text{if } m_1 < m_2, \, V_{M(s-1)+m_1} = 0; \\
 Q_{m_1,s} & \text{otherwise.}
\end{cases}$$

In the formula for $L$ we ignore the multiplier

$$\prod_{t=1}^{T} \prod_{s=1}^{S} \prod_{m_1=1}^{M} \prod_{m_2=1}^{M+1} D_{m_1,m_2}^{(s)} I'(s,m_1,m_2)$$
that does not contain the unknown parameters.

Transition counts \( I_t(s, m_1, m_2) \) for the period of observation is the only input required. By \( I_t(s, m_1, m_2) \) we denote the number of debtors in the industry \( s \) that migrated from the credit class \( m_1 \) to the credit class \( m_2 \) in the period \( t \). (The industry-specific Markovian migration matrices \( P^{(s)} \) are estimated from these counts as time averages.)

All the unknowns have to satisfy the box constraints: \( D(V) \in [0, 1] \), \( Q_{m,s} \in [0, 1] \) and \( \Delta_{m,s} \in [0, 1] \). Since \( D \) is a probability distribution, the following equality has to be satisfied:

\[
\sum_{V \in \{0, 1\}^{MS}} D(V) = 1. \tag{3}
\]

The value \( Q_{m,s} \) determines the impact of the idiosyncratic component in (1). In particular, \( Q_{m,s} = 1 \) implies that migrations of debtors belonging to the credit class \( m \) and the industry \( s \) are independent on migrations of other debtors. The existing models assume that \( \Delta_{m,s} = 1 \) for all possible \( m \) and \( s \). See Kaniovski and Pflug (2007), Wozabal and Hochreiter (2012) or Boreiko et al. (2017). Suggesting more general formulas for the conditional probabilities and estimating \( \Delta_{m,s} \), we attempt to test empirically this assumption. All other things being equal, a larger \( \Delta_{m,s} \) implies smaller probabilities \( P^{(s)}(m,j)(1), j < m \), and larger probabilities \( P^{(s)}(m,j)(0), j > m \). Consequently, the known models may overestimate the default rates \( P^{(s)}(m,M+1)(0) \).

The conditional distribution of \( X_n(t+1) \) defined by (2) typically deviates from the \( m \)th row of the corresponding \( P^{(s)} \). To guarantee that the unconditional distribution of \( X_n(t+1) \) in (1) equals \( P^{(s)}(j), j = 1, 2, \ldots, M + 1 \), the following equalities have to be satisfied:

\[
\sum_{V \in \{0, 1\}^{MS}} V_{M(s-1)+m} D(V) - \Delta_{m,s} P^{(s)}_{m,m} = \tilde{P}_{m}^{(s)}, \quad m=1, 2, \ldots, M, \ s = 1, 2, \ldots, S. \tag{4}
\]

Here, \( \tilde{P}_{m}^{(s)} = P^{(s)}_{m,1} + P^{(s)}_{m,2} + \ldots + P^{(s)}_{m,m-1} \).

The natural logarithm of \( L \) is maximized subject to the linear equality constraints (3) and (4). There are \( 2MS + 2^{MS} \) parameters to estimate: \( MS \) probabilities \( Q_{m,s} \), \( MS \) probabilities \( \Delta_{m,s} \) and \( 2^{MS} \) probabilities \( D(V) \).

Having a distribution \( D \), probabilities of fine-grained macroeconomic events associated with the phases of a business cycle can be estimated. For example, probabilities of adverse for a combination of credit-classes and industries events. Since a time period in the past is considered, frequencies is a more precise denomination for such values. Numerical characteristics other than probabilities can be evaluated as well. In particular, correlations between the economic factors affecting different industries and credit classes.

In what concerns credit risk, our approach, accounting for the sectoral dimension, implies a more specific and, consequently, precise analysis of credit events. In fact, the
Numerical estimates of risk factors contingent on credit ratings

Assessing how likely is a contagion in the financial sector, probabilities of adverse events affecting several credit classes can be used. Unlike the majority of models addressing systemic risk, our estimates do not require any information regarding interconnections in the banking sector. See Glasserman and Young (2016) for a comprehensive review of systemic risk models.

Let us show that, under natural assumptions regarding the migration matrices \( P^{(s)} \), the feasible set defined by the linear relations (3) and (4) is not empty. Since the probabilities \( Q_{m,s} \) are not involved in these relations, it is enough to indicate a distribution \( D \) and probabilities \( \bar{p}_{m,s} \) satisfying (3) and (4).

**Proposition** If \( p \leq \bar{p} \), then the feasible set defined by (3) and (4) is not empty.

**Proof** Consider the binary vectors \((1, 1, \ldots, 1)\) and \((0, 0, \ldots, 0)\). Let \( D((1, 1, \ldots, 1)) = \tilde{P} \) and let \( D((0, 0, \ldots, 0)) = 1 - \tilde{P} \). Then equality (3) holds true, while constraints (4) read:

\[
\tilde{P} - \Delta_{m,s} P_{m,m}^{(s)} = \tilde{P}_{m}^{(s)}, \quad m = 1, 2, \ldots, M, \quad s = 1, 2, \ldots, S,
\]

or

\[
\Delta_{m,s} P_{m,m}^{(s)} = \tilde{P} - \tilde{P}_{m}^{(s)}, \quad m = 1, 2, \ldots, M, \quad s = 1, 2, \ldots, S. \tag{5}
\]

If \( p \leq \tilde{P} \leq \bar{p} \), then \( 0 \leq \tilde{P} - \tilde{P}_{m}^{(s)} \leq P_{m,m}^{(s)} \) implying that there are always \( \Delta_{m,s} \in [0, 1] \) satisfying equalities (5). In other words, the feasible set is not empty. \( \square \)

Typically, every probability \( P_{m,m}^{(s)} \) to retain the current credit rating \( m \) does not fall below \( \frac{1}{2} \). Then \( \bar{p} \geq \frac{1}{2} \). If \( P_{m,m}^{(s)} \geq \frac{1}{2} \), then \( \tilde{P}_{m}^{(s)} \) does not exceed \( \frac{1}{2} \). Therefore, \( p \leq \frac{1}{2} \).

In sum, \( p \leq \bar{p} \). This argument implies, since \( P_{m,m}^{(s)} > \frac{1}{2} \) for all combinations of \( m \) and \( s \) in the migration matrices quoted in Appendix 2, that the assumptions of Proposition hold true for our input data.

Since quarterly \( P_{m,m}^{(s)} \) typically do not fall below, all other things being equal, their annual counterparts, verification of assumptions of Proposition is easier dealing with quarter-by-quarter transition counts.

Even if concavity of \( \ln L \) cannot be established, its derivatives are known and, apart from the box constraints, there are just linear equality constraints. If \( 2^{MS} \) is not too large, a standard solver can be used for finding a solution. Restarting the algorithm for different initial points, the solution can be improved.

This course of action was implemented by Boreiko et al. (2017) for \( M = 2 \) and \( S = 6 \). The Interior Point (IP) method was used. (Postulating that all \( \Delta_{m,s} \) are equal to 1, the model in Boreiko et al. (2017) is slightly simpler.) On the one hand, this approach looks like a classical optimization technique. In fact, the classification in Gilli and Schumann (2012) requires for such a technique “at least well-behaved objective functions” so that a kind of gradient descent can be implemented. See p. 130. On the

\[ \text{Springer} \]
other hand, a heuristic—restarting the gradient descent from different initial points—is used because concavity of the likelihood function cannot be guaranteed. (As a matter of fact, all of the restarts implied the same solution.) In sum, a combination of a classical technique and a heuristic is applied even in this simplest case.

A more detailed analysis of credit-rating migrations considers two non-default credit classes as well, but the number of industries is doubled. Therefore, the total number of unknowns becomes $24 + 2^{12}$. This level of computational complexity requires special algorithms. We suggest and test one of such methods.

Maximizing $\ln L$, there are two challenges. First, the number $2^{MS}$ of unknown probabilities $D(V)$ is typically combinatorial for a practically interesting choice of $M$ and $S$. Second, the estimated distribution $D$ can be nested in too many sample points. In such a case, it is difficult to analyze the corresponding macroeconomic scenarios.

Confronting with a large $MS$, the sample space $\{0, 1\}^{MS}$ can be split into smaller parts so that the likelihood maximization problem could be solved for each of them with a standard algorithm. Since there can be too many such subsets, instead of analyzing all of them, a random sample can be used. Passing from one of the subsets to the next one, we need a rule for identifying the sample points that are retained. In sum, there are two optimization processes: a continuous space search for inputs $Q, \Delta$ and $D$ maximizing the likelihood function defined on a subset of $\{0, 1\}^{MS}$ and a discrete space search for a subset of $\{0, 1\}^{MS}$ with a greater maximum likelihood value. According to the classification given in Gilli and Winker (2009), this is a combination of a local search method and a constructive method. See p. 87. A heuristic algorithm based on these principles is described next.

### 3 Heuristics

The heuristic search for a better solution relies on several concepts. Let us explain and motivate them.

We begin with characterizing the subsets that can be used in a partition of the sample space. We call $V \subseteq \{0, 1\}^{MS}$ a suitable set, if there exists a probability distribution $D_V$ such that the support of $D_V$ belongs to $V$ and the relations

$$\sum_{V \in V} V_{M(s-1)+i} D_V(V) - \Delta_{i,s} P_{i,i}^{(s)} = F_i^{(s)}, \quad i = 1, 2, \ldots, M, \quad s = 1, 2, \ldots, S,$$

(6)

hold true for some $\Delta_{i,s} \in [0, 1]$. Note that every extension $V'$ of a suitable set $V$, that is a subset of $\{0, 1\}^{MS}$ containing $V$, is suitable as well. In fact, let

$$D_{V'}(V) = \begin{cases} D_V(V) & \text{if } V \in V, \\ 0 & \text{if } V \in V' \setminus V. \end{cases}$$

Keeping the values $\Delta_{i,s}$ unchanged, relations (6) for $V'$ will hold true with this distribution $D_{V'}$. 

Springer
Next, let us describe a continuous space optimization problem for identifying optimal inputs $Q$, $\Delta$ and $D$ corresponding to a given suitable set. Consider a suitable set $V$, the likelihood function

$$L_V(D, Q, \Delta) = \prod_{t=1}^{T} \sum_{V \in V} D(V) \prod_{s=1}^{S} \prod_{m_1=1}^{M} \prod_{m_2=1}^{M+1} F(s, V, m_1, m_2, Q, \Delta) I^{(s,m_1,m_2)},$$

and the constraints

$$\sum_{V \in V} V_{M(s-1)+i} D(V) - \Delta_{i,s} P_{i,i}^{(s)} = \tilde{P}_{i}^{(s)}, \quad i = 1, 2, \ldots, M, \quad s = 1, 2, \ldots, S,$$

$$\sum_{V \in V} D(V) = 1. \quad (7)$$

$$\sum_{V \in V} D(V) = 1. \quad (8)$$

That is, our analysis is restricted now to the distributions $D$ that are nested in $V$. Ceteris paribus, the values of $L$ and $L_V$ coincide for such a $D$. Since $V$ is a suitable set, the feasible set defined by the linear equations (7) and (8) is not empty. In fact, the distribution $D_V$ involved in the definition of $V$, see relations (6), satisfies these constraints.

If the cardinality of $V$ is low enough, maximizing $\ln L_V$ subject to the linear constraints (7) and (8), the corresponding optimal distribution $D_V^*$ can be estimated. A standard solver can be used, say the IP method. Since all unknowns are probabilities, the box constraints for $Q_i$, $\Delta_i$ and $D(V)$ are necessary as well. Restarting the algorithm from different initial points, the solution can be improved.

If $D_V^*(V)$ is sufficiently small, the macroeconomic scenario encoded by $V$ is not (statistically) significant and, therefore, it can be ignored passing from $V$ to its extension $V'$. A heuristic for retaining significant elementary outcomes is described next.

For a practical numerical search algorithm in $\{0, 1\}^{MS}$, consider a threshold $\epsilon \in (0, 1)$ and the $\epsilon$-support $V_\epsilon$ of $D_V^*$:

$$V_\epsilon = \{ V \in V : D_V^*(V) > \epsilon \}.$$

Note that $\sum_{V \in V_\epsilon} D_V^*(V) \uparrow 1$ as $\epsilon$ decreases and $D_V^*$ satisfies constraints (7) and (8). Therefore, $V_\epsilon$ is a suitable set for all sufficiently small $\epsilon$. In fact, the linear constraints (6) and (8), where $V$ and $D$ are substituted with $V_\epsilon$ and $D_V^*$, can be satisfied with any precision.

If $\epsilon$ decreases, the number of binary strings in $V_\epsilon$ can increase. In fact, if the strings were equally probable, the total number of them would grow as $\epsilon^{-1}$. A non-negligible $D_V^*(V \setminus V_\epsilon)$ can be another effect of a sparse $D_V^*$. In particular, $D_V^*(V \setminus V_\epsilon) > \epsilon$. Keeping the threshold $\epsilon$ constant, the following heuristic can be used to obtain a more concentrated than $D_V^*$ optimal distribution. As a measure of concentration, we require that the total probability of the significant macroeconomic scenarios exceeds $1 - \epsilon$. 

© Springer
To this end, consider

$$
\bar{L}_V(D, Q, \Delta) = \ln L_V(D, Q, \Delta) + R \sum_{V \in \mathcal{V}} D^2(V).
$$

We interpret $R \geq 0$ as a penalty parameter. It determines the relative importance of the two components in $\bar{L}_V$. Let us maximize $\bar{L}_V$ subject to the linear constraints (7) and (8). For a sufficiently large $R$, we expect that the optimal distribution $\bar{D}_V$ obtained in this way should be more concentrated than $D^*_V$. (The latter corresponds to $R = 0$.)

As a heuristic argument supporting this intuition, observe that the function of $N$ variables

$$
\sum_{k=1}^{N} x_k^2
$$

attains its maximum value 1 under the constraints

$$
x_k \geq 0, \ k = 1, 2, \ldots, N, \quad \text{and} \quad \sum_{k=1}^{N} x_k = 1,
$$

if one of the addends equals 1. (Its minimum value obtains if all of them are equal.) Adding such a term to $\ln L$, we expect, given the box constraints and (8), to get a more concentrated than $D^*_V$ optimal distribution $\bar{D}_V$. Since the weight of the second term in $\bar{L}_V$ increases in $R$, we expect the same for $\bar{D}_V(\bar{V}_\epsilon)$. Therefore, $\bar{D}_V(\bar{V}_\epsilon) > 1 - \epsilon$ for all sufficiently large $R$. By $\bar{V}_\epsilon$ we denote the $\epsilon$-support of $\bar{D}_V$.

Now the rule for identifying the significant macroeconomic scenarios can be formulated more precisely. Fix a maximum mismatch in the linear constraints (7) and (8). Denote this value by $\epsilon$. The set $\bar{V}_\epsilon$ such that $\bar{D}_V(\bar{V}_\epsilon) > 1 - \epsilon$ is a suitable set. In fact, substituting $\bar{V}_\epsilon$ and $\bar{D}_V$ instead of $V$ and $D_V$ into the linear constraints (7) and (8), the maximum error will not exceed $\epsilon$. Therefore, for this numerical precision, $\bar{V}_\epsilon$ is a suitable set.

The structure of our discrete space search is as follows:

1. At the beginning, a suitable set $V$ with a low cardinality has to be identified and a threshold value $\epsilon$ has to be chosen.
2. Solve the optimization problem for specifying $V_\epsilon$. Decrease $\epsilon$ if $V_\epsilon$ is not suitable. Keep $\epsilon$ as large as possible. If $V_\epsilon$ contains too many elements or/and $D_V(V \setminus V_\epsilon) \geq \epsilon$, consider the corresponding $\bar{L}_V$ and identify $\bar{V}_\epsilon$. Increasing $R$, achieve that $\bar{D}_V(\bar{V}_\epsilon) > 1 - \epsilon$. Keep $R$ as small as possible.
3. Better solutions (in terms of the likelihood value) can be found by extending $\bar{V}_\epsilon$ to a larger set. Since any set containing a suitable set is also suitable, the extension preserves suitability. Let $V' = \bar{V}_\epsilon \cup \bar{V}$, where $\bar{V}$ is a subset of $\{0, 1\}^{M \times S} \setminus V$. This is a suitable set, an extension of $\bar{V}_\epsilon$. $\bar{V}$ should not be too large so that the optimization problem with $V'$ instead of $V$ could be solved.
4. We repeat steps 1.–3. with $V$ replaced by $V'$ and keep on doing extensions.
There are two clarifications. First, we can always argue about an extension of $\hat{V}_\epsilon$. In fact, $V_\epsilon = \hat{V}_\epsilon$ for $R = 0$. Second, since $V' \supseteq \hat{V}_\epsilon$, the extension cannot lead to a smaller maximum value of $L$.

If the assumptions of Proposition hold true, every set $V \subseteq \{0, 1\}^{MS}$ containing the string with 1 at all positions and the string with 0 at all positions is suitable. Consequently, the discrete space search can always be initiated. We suggest a particular extension process exploiting the interpretation of a binary string as a macroeconomic scenario.

Let us describe this particular scheme for generating extensions. Initially, we assume that every industry is affected by the same economic conditions. The corresponding binary $MS$-string is termed as a block-structure. In a block-structure, all binary $M$-substrings allocated to different industries coincide. There are $2^M$ block-structures. Containing the macroeconomic scenario favorable for all industries and credit classes as well as the macroeconomic scenario adverse for them all, the set $V$ of all block-structures is suitable.

Solve the optimization problem for identifying $D^*_V$. This is not a hard task. In fact, just $2MS + 2^M$ unknowns are involved. Let $V^k$, $k = 1, 2, \ldots, K$, be the block-structures whose probabilities $D^*_V(V^k)$ exceed $\epsilon$. It is convenient to number them in descending order of the probabilities. In order to reduce $K$, the modified objective function $\hat{L}_V$ can be used instead of $L_V$. Denote by $v^k$ the binary $M$-substring such that $V^k$ is formed by $S$ copies of $v^k$. We call $v^k$ a block.

A scheme for generating new macroeconomic scenarios as mutations of the block-structures $V^k$, $k = 1, 2, \ldots, K$, is motivated by elementary facts from genetics. In particular, substituting one block $v^k$ in $V^k$ by a block $v^i$, $i \neq k$, we get a mutant with a single mutation of the block-structure $V^k$. Conceptually, this is a macroeconomic scenario where all industries, except for one, are affected by the economic conditions encoded with $v^k$, while the remaining industry is affected by the economic conditions summarized in $v^i$. Trying all possible $V^k$, $k = 1, \ldots, K$, industries $s = 1, \ldots, S$ or, equivalently, positions of the mutated block, and its type $v^i$, $i = 1, \ldots, K$, $i \neq k$, all mutants $V^1$ with a single mutation will be obtained. Similarly $V^n$, can be defined for $n > 1$. That is, $V^n$ contains all possible mutants with $n$ mutations. The mutations can coincide: defining $V^n$ for $n > 1$, a block $v^i$ can appear more than one time in a mutant stemming from the block-structure $V^k$. Then $V^n \cap V^{n'} = \emptyset$, if $n, n' < S/2$, $n \neq n'$. Defining the sets $V^n$ for $n \geq S/2$, some of the mutants, originating from different block-structures, can be listed several times. Such repetitions have to be avoided. For example, if $S$ is an even number, a mutant containing $S/2$ blocks $v^k$ and $S/2$ blocks $v^i$ can be regarded as originating from the block-structure $V^k$ as well as originating from the block-structure $V^i$. Such mutants have to be listed in $V^{S/2}$ just one time.

There are $K^S$ possibilities for allocating $K$ building blocks among $S$ positions. If $K < 2^M$, this number is smaller than $2^{MS}$. Consequently, restricting the analysis to all possible combinations of blocks, on the one hand, reduces the scope of macroeconomic scenarios, but, on the other hand, the corresponding solution should not be necessarily optimal even if all of the combinations were considered. A smaller $K$ implies a stronger reduction of the search space.
Instead of dealing with all binary vectors from $\mathbb{V}^n$, a random sample can be used. Generating a binary vector from $\mathbb{V}^n$, for the first mutation there are $S$ equally probable positions, for the second—the remaining $S - 1$, etc. The simplest way for choosing mutations, is to sample them with equal probabilities $\frac{1}{K}$. The resulting distribution of mutants is referred to as uniform.

The S&P’s dataset used for estimating unknown parameters contains 103,723 transition counts. See https://doi.org/10.1371/journal.pone.0175911.s001. We consider two classifications: $M = 7$ with $S = 6$ and $M = 2$ with $S = 12$. The corresponding numbers of unknowns are $84 + 2^{42}$ and $48 + 2^{24}$. They are huge and, more importantly, these numbers greatly exceed 103,723—the number of available observations. Common sense suggests that no classical approach can be used. Looking for a way out, we observe that there are $2^{42} (2^{24})$ probabilities $D(V)$ among the unknowns. Dealing with a real life problem, unlikely outcomes can be ignored. A lower bound $\epsilon$ for probability of a significant outcome implies that at most $\epsilon^{-1}$ such outcomes have to be considered. In particular, if $\epsilon = 10^{-3}$ then at most 1000 macroeconomic scenarios are statistically significant. This number is approximately 100 times smaller than 103,723. Therefore, taking into account that the upper bound of 1000 corresponds to a practically impossible situation when all of the significant outcomes are equally probable, the task of estimating $D$ does not look hopeless. In sum, a classical solution is not possible given the input, instead a conceptually plausible approximation for $D$ and, consequently, for the hidden dependence structure among industries and credit classes can be suggested. Motivating this approach, we rely on the principles formulated in Gilli and Winker (2009): “Often the term ‘heuristic’ is linked to algorithms mimicking some behavior found in nature, ... a heuristic should be able to provide high-quality (stochastic) approximation to the global optimum at least when the amount of computational resources spent on a single run of the algorithm ... is increased.” See p. 83.

Searching for a plausible approximation, we first hypothesize that all industries are affected by the same economic conditions. This assumption corresponds to the standard business cycle theory. A new element is a fine-grained classification of creditworthiness. Under such a restriction on the class of admissible macroeconomic scenarios, we identify the significant outcomes, referred to as block-structures. Since no penalty is necessary and the restarts of the IP method from different initial points result in the same likelihood value, all criteria of Gilli and Schumann (2012) for a classical optimization technique are satisfied. See p. 130. Therefore, this is a classical solution. Next we consider a richer set of admissible macroeconomic scenarios—block structures and their mutants having just one mutation. It is a further extension of the standard business cycle theory because industries can be affected by different economic conditions. This extension is the simplest possible since at most one variation per macroeconomic scenario is allowed. For all choices of $M$ and $S$ considered in this paper, the number of unknowns allows, given our computational resources, to use the IP method. Since a penalty was necessary, an approximation to the solution was obtained for this extension. Next we turn to macroeconomic scenarios where at most two of the industries can be affected by different economic conditions. Conceptually, this is a further generalization of the standard business cycle theory. We list all mutants with exactly two mutations. Given our computational resources, the
corresponding number of unknowns for the couple $M = 7$ and $S = 6$ is too high. 
Thus, we are forced to split this set into equally large subsets. First, we consider one 
of the subsets and the significant binary strings with at most one mutation. The corre-
sponding continuous space maximization problem is solved with the IP method. The 
significant binary strings identified at this step are considered together with the next 
subset, etc. Having tested all of the subsets, we obtain a heuristic approximation to the 
most likely, given the input, distribution on the binary strings with at most two muta-
tions of the block-structures. This is just an approximation. In fact, non-zero penalties 
were used in several extensions and, more importantly, we cannot guarantee that the 
greatest likelihood value achieved by the sequential analysis of the subsets coincides 
with the maximum value attained on the union of them, should the corresponding 
maximization be done. If $M = 2$ and $S = 12$, all binary strings with two mutations 
are used for a single extension. Further extensions are done in the same way with 
randomly generated mutants having $3, 4, \ldots, S - 1$ mutations. No binary string may 
be considered more than one time. All necessary details are given in Appendix 1.

As a standard measure of performance for the heuristic, we present in Appendix 1 
the percentage of increase of the likelihood value. Also, since for $M = 2$ and $S = 6$ 
a classical solution exists, we compare it in Appendix 1 with the heuristic solution. 
This is an instructive example illustrating the particularities of both the likelihood 
maximization problem and the heuristic algorithm.

The discrete space search described above mimics biological evolution: first, only 
macroeconomic scenarios more probable (or more frequent, because we refer to 
a period in the past) than a threshold are retained to the next round of selection 
(where they compete with the new scenarios of the extension) and, second, more 
complex macroeconomic scenarios obtain as mutations of the original forms—the 
block-structures. The minimum viable population argument is a particular justifica-
tion for the selection rule. Business cycles is another natural point of reference: we 
begin with the standard setting and attempt to arrive at a more fine-grained view. Some 
of the reported solutions are exact, but typically an approximation is found. The quality 
of a solution increases as the computation efforts do. In particular, widening the scope 
of admissible macroeconomic scenarios, a more realistic view of a business cycle can 
be obtained. However, it requires solving more complicated optimization problems. 
The random search in the space of mutants with $n \geq 3$ mutations exhibits, accord-
ing to our experience, path dependence. Moreover, the shape of $L$ depends upon the 
transition counts available. For this reason, any estimate of the empirical distribution 
of the maximum likelihood value would be input-specific as well. Therefore, such an 
empirical convergence analysis, even if it is recommended in the literature on heuristic 
methods in econometrics, see among others Gilli and Winker (2009) or Gilli and Schu-
mann (2012), cannot be used. A bootstrap procedure can be suggested instead. Then, 
for a set of estimated parameters, new credit-rating migrations are generated according 
to formula (1). Applying the algorithm, this input is transformed by into a new set 
of estimates. If they match reasonably well the original values, we conclude that the 
algorithm works correctly. For a simple example of such a convergence analysis see 
Boreiko et al. (2016).

Given $D$, $Q$ and $\Delta$, some risk estimates are suggested in the next section.
4 Methodology

The following quantitative characteristics of risk can be evaluated:

- Probability

\[
\pi_{m_1, m_2, \ldots, m_L}^{s_1, s_2, \ldots, s_L}(I_{s_1}^{m_1}, I_{s_2}^{m_2}, \ldots, I_{s_L}^{m_L}) = \sum_{V \in \{0, 1\}^{MS}} D(V) \prod_{i=1}^{L} \max[I_{s_i}^{m_i}, V_{M(s_i-1)+m_i}, (1 - I_{s_i}^{m_i})(1 - V_{M(s_i-1)+m_i})]
\]

of a favorable, if \(I_{s_i}^{m_i} = 1\), (an adverse, if \(I_{s_i}^{m_i} = 0\),) for the industry \(s_i\) and the credit class \(m_i\) macroeconomic scenario, \(i = 1, 2, \ldots, L\). Both indexes can repeat. That is, this probability can concern more than one credit class from the same industry or/and several industry sectors having the same creditworthiness.

- Correlation

\[
C_{m_1, m_2}^{s_1, s_2}(I_{s_1}^{m_1}, I_{s_2}^{m_2}) = \frac{\pi_{m_1, m_2}^{s_1, s_2}(I_{s_1}^{m_1}, I_{s_2}^{m_2}) - \pi_{s_1}^{m_1}(I_{s_1}^{m_1})\pi_{s_2}^{m_2}(I_{s_2}^{m_2})}{\sqrt{\pi_{s_1}^{m_1}(I_{s_1}^{m_1})\pi_{s_2}^{m_2}(I_{s_2}^{m_2})(1 - \pi_{s_1}^{m_1}(I_{s_1}^{m_1}))\pi_{s_2}^{m_2}(I_{s_2}^{m_2})(1 - \pi_{s_2}^{m_2}(I_{s_2}^{m_2}))}}
\]

between the indicators of the following macroeconomic outcomes: if \(I_{s_1}^{m_1} = 1\) \((I_{s_1}^{m_1} = 0)\), one of them is favorable (adverse) for the credit class \(i_1\) of the industry \(s_1\), the other one is favorable (adverse) for the credit class \(i_2\) of the industry \(s_2\) if \(I_{s_2}^{m_2} = 1\) \((I_{s_2}^{m_2} = 0)\).

- Percentage of variation

\[
v^{(s)}_{m, j}(I_s^m) = 100 \times \left( \frac{Q_{m,s} P_{m,j}^{(s)} + (1 - Q_{m,s}) P_{m,j}^{(s)}(I_s^m)}{P_{m,j}^{(s)}} - 1 \right),
\]

of the conditional probability \(Q_{m,s} P_{m,j}^{(s)} + (1 - Q_{m,s}) P_{m,j}^{(s)}(I_s^m)\) against its unconditional counterpart \(P_{m,j}^{(s)}\). (The paragraph containing formula (2) introduces this conditional distribution.) In particular, \(v^{(s)}_{m, M+1}(I_s^m)\) measures the relative increase or the relative decrease of the default probability in the credit class \(m\) of the industry \(s\) due to the economic conditions.

Let us characterize the input data and the corresponding heuristic solutions.

5 Inputs and heuristic solutions

If \(M = 7\), the S&P’s credit classes AAA, AA, A, BBB, BB, B and C are numbered by 1, 2, \ldots, 7. The following 6 industry sectors are considered: 1—agriculture, mining...
Numerical estimates of risk factors contingent on credit ratings

The case of two non-default credit classes, index 1(2) is assigned to the investment-grade (non-investment-grade) debtors. The investment-grade debtors occupy the S&P’s ratings from AAA to BBB. The non-investment-grade debtors are those whose creditworthiness is lower: BB, B or C. There are 12 industry sectors: 1—industry sectors are considered: aero, auto, capital goods, metal; 2—consumer, service; 3—energy, natural resources; 4—financial institutions; 5—forest and building products, homebuilders; 6—health care, chemicals; 7—high technology, computers, office equipment; 8—insurance, real estate; 9—leisure time, media; 10—telecommunications; 11—transportation; 12—utilities. Nagpal and Bahar (2001) estimated default correlations for these 12 industries and two non-default credit classes.

5.1 Input

We use annual transition counts $I_t(s, m_1, m_2)$ covering the period 1991–2015: $t = 1$ corresponds to 1991 and $T = 25$ corresponds to 2015. These transition counts are available at https://doi.org/10.1371/journal.pone.0175911.s001. Distribution of the counts among credit classes and industries is characterized in Appendix 2.

5.2 Solution

All continuous space optimization problems were solved with the IP method. The significance threshold equals $10^{-3}$ everywhere.

Consider first the couple, where $M = 7$ and $S = 6$. There are $K = 14$ block-structures whose probabilities exceed $10^{-3}$. Total probability assigned to the remaining block-structures does not exceed 0.0506%. Numbering the significant $V^k$ in descending order of their probabilities, the corresponding blocks $v^k$ are listed in Table 1. Searching for the heuristic solution, mutants with up to 5 mutations were considered. (The most important technicalities regarding this search are given in Appendix 1.) Table 2 contains the 22 significant macroeconomic scenarios and their probabilities. Interestingly enough, the block $v^{14}$ does not appear in these scenarios. Total probability of the remaining sample points falls below 0.0223%. In particular, probability of the significant macroeconomic scenario numbered by $n = 1$ is 0.2661. It can be thought of as a mutant with four mutations of the block-structure $V^1$. In fact, the block $v^1$ is assigned to industries 1 and 5, while the remaining industries received all different blocks: $v^{13}, v^8, v^{11}$ and $v^4$. Since this macroeconomic scenario allocates the block $v^8$ to the industry 3, economic conditions are favorable for the transportation, technology and utility firms rated at AAA, AA, BBB and BB, while the debtors in this industry sector rated at A, B and C are affected by adverse economic conditions. In fact, the row of Table 1 allocated to $v^8$ contains 1 at the positions 1, 2, 4 and 5, while 0 is assigned to the remaining cells. The significant scenarios listed in Table 2 contain up to five mutations. (In this case all blocks in the corresponding row are different.) The minimum number of mutations in the macroeconomic scenarios quoted in Table 2
is three. Therefore, the simpler outcomes containing zero, one or two mutations were wiped out in the course or calculations. The content of Tables 1 and 2 suggests that the significant macroeconomic scenarios do not exhibit the pattern assumed within the classical business cycle theory.

A referee, whose insight was crucial for streamlining this paper, observed that some of the blocks in Table 1 encode scenarios that look counterintuitive: being favorable for lower rated debtors they are adverse for higher rated ones. For example, the scenario encoded by $v^1$ is adverse for debtors rated at $A$, but it is favorable for those rated at $BBB$. He suggested to consider only “monotone” blocks. In particular, dealing with 7 non-default credit classes, there are 8 such blocks: $1111111, 1111110, 1111100, \ldots, 1000000, 0000000$. This proposal is promising because, on the one hand, the counterintuitive outcomes are excluded and, on the other hand, the total number of unknowns can be substantially reduced. In fact, dealing with $M$ non-default credit classes and $S$ industries, the total number of sample points will be $(M + 1)^S$ instead of $2^{MS}$. We are going to test the corresponding model numerically.

The following matrices $Q$ and $\Delta$ were estimated:

$$Q = \begin{pmatrix} 0.3237 & 0.9775 & 0.6230 & 0.0000 & 1.0000 & 0.8279 \\ 0.4579 & 0.9134 & 0.8652 & 0.6744 & 0.5809 & 0.9789 \\ 0.8580 & 1.0000 & 0.8559 & 0.9035 & 0.6219 & 1.0000 \\ 0.9185 & 0.9471 & 0.8280 & 0.9772 & 0.9482 & 1.0000 \\ 0.8026 & 0.7764 & 0.8761 & 0.9611 & 0.6946 & 1.0000 \\ 0.6013 & 0.6452 & 0.4498 & 0.6297 & 0.5575 & 0.8064 \\ 0.7337 & 0.6797 & 0.2978 & 0.7654 & 0.3793 & 0.7509 \end{pmatrix}.$$ 

$$\Delta = \begin{pmatrix} 0.8610 & 0.9984 & 0.8824 & 0.8930 & 0.9313 & 0.9166 \\ 0.7492 & 0.3822 & 0.8716 & 0.7724 & 0.7834 & 0.9754 \\ 0.0234 & 0.3558 & 0.0900 & 0.1353 & 0.0000 & 0.0135 \\ 0.7835 & 0.7617 & 0.9445 & 0.4770 & 1.0000 & 0.6968 \\ 0.1311 & 0.1262 & 0.2871 & 0.0263 & 0.0092 & 0.1328 \\ 0.6996 & 0.6091 & 0.6935 & 0.6234 & 0.5572 & 0.7495 \\ 0.6530 & 0.5243 & 0.5325 & 0.9610 & 0.4042 & 0.2946 \end{pmatrix}.$$ 

If $M = 2$ and $S = 12$, all four block-structures are significant macroeconomic scenarios. In Table 3 blocks are listed in descending order of probabilities assigned to the corresponding block-structures. The main facts regarding the heuristic search in this case are summarized in Appendix 1. Table 4 contains the significant macroeconomic scenarios and their probabilities. Total probability assigned to the remaining sample points does not exceed 0.0506%. The matrices $Q$ and $\Delta$ read:

$$Q = \begin{pmatrix} 0.9950 & 1.0000 & 0.9654 & 1.0000 & 1.0000 & 0.9296 & 1.0000 & 0.5724 & 1.0000 & 0.7797 & 0.8134 & 0.5285 \\ 0.5965 & 0.6376 & 0.9647 & 0.8608 & 0.5975 & 0.6929 & 0.3688 & 0.2531 & 0.5018 & 0.4219 & 0.6318 & 0.6870 \end{pmatrix}.$$ 

$$\Delta = \begin{pmatrix} 0.7279 & 1.0000 & 0.2829 & 0.8447 & 0.5333 & 0.7441 & 0.8806 & 0.6872 & 0.6435 & 0.7686 & 0.7520 & 0.7620 \\ 0.7223 & 0.6923 & 0.9515 & 0.8551 & 0.7194 & 0.6607 & 0.6443 & 0.6715 & 0.6410 & 0.8039 & 0.6745 & 0.7841 \end{pmatrix}.$$ 

In the considered so far sectoral models of dependent credit rating migrations, all entries of $\Delta$ are assumed to be equal to 1. See Boreiko et al. (2017), for example.
estimates for \( \Delta \) imply that this assumption cannot be justified empirically. Also, since a larger \( \Delta_{m,s} \) causes larger conditional probabilities \( P^{(s)}_{m,j}(0), j > m \), the default rates reported in Boreiko et al. (2017) are overestimated.

Having \( Q, \Delta \) and \( D \), transition paths mimicking the actually observed historical migrations can be simulated. Using the Monte-Carlo method, losses generated by a portfolio can be estimated. Rather than moving in this direction, we concentrate in the next section on the risk characteristics given by formulas involving the estimates of \( Q, \Delta \) and \( D \).

6 Estimates

If \( Q_{m,s} = 1 \), macroeconomic factors do not affect migrations in the credit class \( m \) of the industry \( s \), so such combinations are not considered further. As a curious fact regarding the combination of \( M = 7 \) with \( S = 6 \), note that \( Q_{1,5} = 1 \) can be a quantitative argument supporting the “too big to fail” theory in the financial sector. In particular, an evidence of implicit too-big-to-fail bail-out guarantee policies of the regulatory authorities. If \( M = 2 \) and \( S = 12 \), \( Q_{1,4} = 1 \) can be interpreted in the same way. In sum, the financial sector debtors belonging to the most secure credit classes seem to enjoy a special treatment.

Let us analyze first the results for the combination of \( M = 7 \) with \( S = 6 \). Given in Table 5 percentages regarding the financial sector imply that the credit classes AA, BBB, B and C exhibit under adverse conditions much stronger increase of default rate than A and BB. (Since \( Q_{1,5} = 1 \), the variation equals zero for the credit class AAA. See the formula for \( v^{(5)}_{1,8}(0) \) in Sect. 4.) The corresponding conditional default probabilities are quoted in Table 6. For the four credit classes exhibiting the highest increase
Table 2  0.001-support of $D$ for $S = 6$ and $M = 7$

| Industry sector | $s = 1$ | $s = 2$ | $s = 3$ | $s = 4$ | $s = 5$ | $s = 6$ | Probability |
|-----------------|---------|---------|---------|---------|---------|---------|-------------|
| $v^1$           | $v^8$   | $v^{11}$| $v^1$   | $v^4$   |         |         | 0.2661      |
| $v^2$           | $v^9$   | $v^7$   | $v^{13}$| $v^2$   | $v^{11}$|         | 0.1941      |
| $v^3$           | $v^4$   | $v^1$   | $v^3$   | $v^{10}$| $v^7$   |         | 0.1124      |
| $v^4$           | $v^9$   | $v^1$   | $v^1$   | $v^9$   | $v^3$   | $v^8$   | 0.0533      |
| $v^{13}$        | $v^8$   | $v^{11}$| $v^{12}$| $v^6$   | $v^2$   |         | 0.0404      |
| $v^7$           | $v^3$   | $v^9$   | $v^1$   | $v^5$   | $v^8$   |         | 0.0395      |
| $v^9$           | $v^1$   | $v^2$   | $v^{11}$| $v^3$   | $v^3$   |         | 0.0294      |
| $v^6$           | $v^7$   | $v^9$   | $v^4$   | $v^{12}$| $v^3$   | $v^2$   | 0.0262      |
| $v^{11}$        | $v^{12}$| $v^9$   | $v^3$   | $v^9$   | $v^{11}$|         | 0.0256      |
| $v^5$           | $v^9$   | $v^4$   | $v^{11}$| $v^3$   | $v^2$   |         | 0.0192      |
| $v^{11}$        | $v^{12}$| $v^4$   | $v^6$   | $v^3$   | $v^8$   |         | 0.0192      |
| $v^{11}$        | $v^4$   | $v^4$   | $v^4$   | $v^7$   | $v^8$   |         | 0.0191      |
| $v^7$           | $v^4$   | $v^2$   | $v^9$   | $v^{13}$| $v^1$   |         | 0.0162      |
| $v^7$           | $v^2$   | $v^{13}$| $v^7$   | $v^7$   | $v^{12}$|         | 0.0147      |
| $v^9$           | $v^3$   | $v^{13}$| $v^{13}$| $v^3$   | $v^{12}$|         | 0.0086      |
| $v^{11}$        | $v^5$   | $v^4$   | $v^{12}$| $v^{11}$| $v^{11}$|         | 0.0078      |
| $v^5$           | $v^1$   | $v^{13}$| $v^1$   | $v^7$   | $v^{11}$|         | 0.0075      |
| $v^{11}$        | $v^8$   | $v^{13}$| $v^1$   | $v^5$   | $v^{13}$|         | 0.0052      |
| $v^{11}$        | $v^5$   | $v^2$   | $v^{12}$| $v^3$   | $v^{13}$|         | 0.0046      |
| $v^{11}$        | $v^7$   | $v^2$   | $v^9$   | $v^{13}$| $v^3$   |         | 0.0032      |

Table 3 Blocks, $S = 12$ and $M = 2$

| Block | Credit class | Investment grade | Non-investment grade |
|-------|--------------|------------------|----------------------|
| $v^1$ |              | 1                | 1                    |
| $v^2$ |              | 0                | 1                    |
| $v^3$ |              | 1                | 0                    |
| $v^4$ |              | 0                | 0                    |

of default rate, frequencies of adverse periods are given in Table 7. Explaining the outlier estimated for the credit class BBB, note that BBB is the lowest investment grade creditworthiness level. Therefore, a migration towards a creditworthiness level below BBB means a quality break. As a consequence, a downgrading within the investment grade credit classes could be an easier decision for the rating agency than a downgrading to a junk level. In sum, a low frequency of adverse periods or, equivalently, a high frequency of favorable years corresponds to the role of the credit class BBB as a last
resort for investment grade debtors. Analyzing the respective entries of $Q$, provides with a quantitative argument supporting such an explanation. In fact, $Q_{2.5} = 0.5809$ and $Q_{3.5} = 0.6219$ imply that credit rating migrations of the debtors rated at A and at AA are strongly affected by market forces, while $Q_{4.5} = 0.9482$, the second largest entry of $Q$ estimated for the financial sector, causes almost idiosyncratic migrations of the debtors rated at BBB. Of course, a list of the downgraded investment grade debtors, rather than anonymized data, would be necessary for a more comprehensive explanation.

Arguing about how likely is a contagion, as a factor of systemic risk, in the European financial networks, Glasserman and Young (2015) “consider the possibility that the failure of a bank causes the next two largest banks to default”. See p. 396. Depending upon the country, the largest bank can be an investment grade debtor, like Deutsche Bank in Germany, as well as a non-investment grade debtor, like Alpha Bank in Greece. The frequencies quoted in Table 8 can be an input for such an analysis. We consider three out of the four credit classes with the highest increase of default probabilities.
under adverse conditions: A, B and C. (Given the nearly idiosyncrasy of migrations in BBB indicated above, triples involving this credit-class have to be analyzed separately.) The corresponding indexes are: \( m_1 = 2, m_2 = 6 \) and \( m_3 = 7 \). Not all possible triples are included. In fact, the argument of Glasserman and Young can hardly address a group of three AA rated banks. Comparing the values quoted in Tables 7 and 8, we conclude that the frequency of an adverse outcome for a triple of credit classes coincides with the frequency of an adverse outcome for the most creditworthy entity involved in the triple. Given that the creditworthiness is typically positively related to the size, the pattern exhibited by the frequencies in Table 8 supports indirectly the argument of Glasserman and Young. (For a less approximative analysis, a dataset, where the counts regarding banks are separated from the counts characterizing the remaining financial institutions, is necessary.)

Table 9 characterizes the increase of default rate under adverse conditions in the (non-investment grade) credit class B for six industries. For a comparison, Table 10 quotes analogous values for the non-investment grade debtors classified into twelve industry sectors.

Tables 11 and 12 contain correlations between indicators of favorable outcomes, a measure of dependence between the corresponding couples of a credit class and an industry sector. These dependencies are not directly observable, but they affect occurrence of credit events involving these couples. The values quoted in Table 11 concern all seven non-default credit classes of two industry sectors: correlations regarding agriculture, mining and construction are quoted below the main diagonal, while those for transportation, technology and utility age given above the main diagonal. Every diagonal element equals one as the correlation coefficient of an indicator with itself. Table 12 contains correlations characterizing couples of an investment grade credit debtor and a non-investment grade debtor classified into twelve industry sectors. Only combinations of industry sectors corresponding to the entries of \( Q \) that fall below 1 are included. (Remember, if \( Q_{m,s} = 1 \), macroeconomic factors do not affect migrations in the credit class \( m \) and industry \( s \).) In a portfolio, strongly positively correlated assets
can provoke cascades of defaults. The simulations reported in Kaniovski and Pflug (2007) illustrate this possibility. To the contrary, combining strongly negatively correlated assets can imply smaller losses. To this end, the cells with correlations given in italic mark the combinations that should be avoided, while the cells with correlations given in bold correspond to the couples of credit classes that can mitigate losses in the corresponding industries. The cutoff levels are ±0.5.

7 Conclusion

A numerical technique is suggested for estimating hidden and observable risk characteristics. It exploits a probability distribution on macroeconomic scenarios. A scenario characterizes the conditions affecting every combination of a credit class and an industry sector. The conditions can be favorable or adverse. Given historical migrations, the maximum likelihood principle is used to estimate the distribution. As a result, a fine-grained extension of the standard business cycle theory emerges. Since through-the-cycle ratings were used, this fine-grained structure is rather unexpected. See Kiff et al. (2013) for an analysis of the through-the-cycle rating approach. We guess that, applying our technique to point-in-time ratings, even a more interesting pattern can emerge. Dealing with such an input, constraints (4) can be too restrictive.

Transition counts is the only required input. For practically interesting combinations of the number of non-default credit classes $M$ and the number of industry sectors $S$,
the number of unknown parameters is combinatorial and it greatly exceeds the number of observations. Therefore, the corresponding non-linear estimation problem cannot be solved with a classical method. Instead, a heuristic algorithm was suggested. It entails a local search method and a constructive method. Performing the local search, a set of macroeconomic scenarios is fixed. This is a continuous space optimization problem. All results presented in this paper were obtained with the IP method as a continuous space solver. The constructive search in the space of macroeconomic scenarios resembles a genetic algorithm (GA). Unlike in a classical GA, recombination is not used and an allele entails more than one digit in our case. Instead of dealing with all mutants, a random sample can be used in some instances.

Two combinations of \( M \) and \( S \) are considered: \( M = 7 \) with \( S = 6 \) and \( M = 2 \) with \( S = 12 \). The corresponding numbers of unknowns are \( 84 + 2^{42} \) and \( 48 + 2^{24} \).

In Hölzl et al. (2019) a similar numerical technique was applied for analyzing survey data.

Funding Open access funding provided by Libera Università di Bolzano within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Appendix 1

The threshold for identifying the significant macroeconomic scenarios is 0.001. Solving a continuous space optimization problem, up to five different initial points were tried in several cases. In all instances the restarts did not affect the solution.

First, the set \( V^0_{0.001} \) of significant block-structures is extended with \( V^1 \). If \( M = 7 \) and \( S = 6 \), there are 14 binary strings in \( V^0_{0.001} \) and 6552 binary strings in \( V^1 \). The respective numbers equal 4 and 1728 for the combination of \( M = 2 \) with \( S = 12 \). The continuous space optimization problems for this extension, containing 6650 and, respectively, 1780 unknowns, were solved with the IP method. Since \( R = 10 \) was used for both combinations of \( M \) and \( S \), these are approximations rather than exact solutions. Denote by \( \tilde{V}^1_{0.001} \) the corresponding set of all statistically significant macroeconomic scenarios.

Second, \( \tilde{V}^1_{0.001} \) is extended using \( V^2 \). For \( M = 2 \) and \( S = 12 \) there are 1584 strings in \( V^2 \). The corresponding continuous space optimization problem was solved with the IP method. Since \( R = 20 \) was used, an approximation was obtained. If \( M = 7 \) and \( S = 6 \), there are 35490 strings in \( V^2 \). For our computational resources, the corresponding continuous space optimization problem is too hard. Therefore, we had to split \( V^2 \) into 10 subsets, each containing 3549 elements. They were used one after another for extensions. Two values of \( R \) were applied: 20 and 50. At the end, an
approximation to the solution for the sample space limited to binary strings with at most two mutations was obtained.

Third, the set of significant macroeconomic scenarios $\vec{V}_{0,001}^2$ with at most two mutations is extended using a random sample from $V^3$. In the sample, binary strings may not repeat. If $M = 7$ and $S = 6$, the sample contains 35000 strings. This set comprise 10 subsets, each formed by uniformly, that is with the probabilities specified in Sect. 3, sampled 250 mutants of a block-structure. The subsets are used one after another for extensions. The penalties term with $R = 20$ and $R = 50$ was applied in some of the extensions. At the end, an approximation to the solution for the sample space limited to binary strings with at most three mutations was obtained. If $M = 2$ and $S = 12$, sampling uniformly 350 mutants for each block-structure, 1400 binary strings were generated. They were used for a single extension. Since $R = 20$ was applied, an approximation to the solution for the sample space limited to binary strings with at most three mutations was obtained.

Thereafter extensions employed random samples from $V^n$, $n = 4, 5, \ldots, S - 1$. Always 250 mutants were sampled uniformly for each block-structure and these subsets were used sequentially for extensions if $M = 7$ and $S = 6$, while, sampling uniformly 350 mutants per block-structure, a single bulk of 1400 binary strings was used if $M = 2$ and $S = 12$. With such numbers of mutants, the total number of binary strings used for extensions remains approximately equal to the cardinality of $V^2$. In several cases, the value of $R$ was 10, 20 or 50 dealing with $S = 6$ and $M = 7$, while in the case of $S = 12$ and $M = 2$ the three values, 10, 20 and 30, of $R$ were applied.

Table 13 characterizes the improvement of the maximal likelihood value $L^H$ achieved by the heuristic algorithm. As the benchmark the maximal value $L^0$ attained on the block-structures is used. A smaller relative increase of the likelihood value for a larger $MS$ is consistent with a smaller ratio of the number of binary strings tested by the heuristic algorithm to the total number $2^{MS}$ of mutants. We observed that the values $Q_{i,s}$ stabilize relatively early in the course of calculations. Testing mutants with at most three mutations is typically sufficient to stabilize the third digit after the decimal point. To the contrary, the entries of $\Delta$ keep adjusting to the variations of $D$.

Since for $S = 6$ and $M = 2$ there are only $2^{12} = 4096$ binary strings involved, the corresponding likelihood maximization problem can be solved exactly with the IP method. We consider this case in order to illustrate efficiency of the heuristics and some of the complications discussed above. The threshold 0.001 remains. This threshold implies 64 significant macroeconomic scenarios according in the exact solution. Total probability of insignificant scenarios amounts to 72.59%. On the one hand, there are too many, for a conceptual analysis, significant macroeconomic scenarios. On the other hand, the total probability of all insignificant scenarios is too large.

The heuristic solutions was evaluated according to the scheme explained above. The four building blocks and their probabilities are given in Table 14.

### Table 13: Improvement of the likelihood value

| $S$ and $M$ | $\frac{L^H - L^0}{L^0} \times 10^6$ | $\frac{\ln L^H - \ln L^0}{\ln L^0} \times 100$ |
|------------|---------------------------------|---------------------------------|
| $S = 6$ and $M = 7$ | 1.1522 $\times 10^6$% | 0.6883% |
| $S = 12$ and $M = 2$ | 44241% | 1.0030% |
Table 14 Blocks and their probabilities, $S = 6$ and $M = 2$

| Block | Credit class | Investment grade | Non-investment grade | Probability |
|-------|--------------|-------------------|----------------------|-------------|
| $v^1$ | 1            | 0                 | 0.4235               |
| $v^2$ | 0            | 1                 | 0.3051               |
| $v^3$ | 1            | 1                 | 0.2448               |
| $v^4$ | 0            | 0                 | 0.0265               |

Table 15 0.001-support of heuristic solution, $M = 2$ and $S = 6$

| $n \backslash$ Industry sector $s$ | 1   | 2   | 3   | 4   | 5   | 6   | Probability |
|----------------------------------|-----|-----|-----|-----|-----|-----|-------------|
| 1                               | $v^1$ | $v^4$ | $v^1$ | $v^4$ | $v^3$ | $v^1$ | 0.5563 |
| 2                               | $v^2$ | $v^2$ | $v^2$ | $v^4$ | $v^2$ | $v^2$ | 0.2330 |
| 3                               | $v^4$ | $v^1$ | $v^3$ | $v^4$ | $v^1$ | $v^2$ | 0.0812 |
| 4                               | $v^3$ | $v^2$ | $v^3$ | $v^2$ | $v^1$ | $v^2$ | 0.0756 |
| 5                               | $v^3$ | $v^1$ | $v^3$ | $v^1$ | $v^1$ | $v^2$ | 0.0216 |
| 6                               | $v^3$ | $v^1$ | $v^4$ | $v^4$ | $v^1$ | $v^2$ | 0.0120 |
| 7                               | $v^3$ | $v^1$ | $v^3$ | $v^1$ | $v^4$ | $v^1$ | 0.0108 |
| 8                               | $v^3$ | $v^1$ | $v^4$ | $v^3$ | $v^4$ | $v^3$ | 0.0077 |
| 9                               | $v^3$ | $v^4$ | $v^1$ | $v^1$ | $v^4$ | $v^4$ | 0.0013 |

Table 16 Variation of $L^H$ for $S = 6$ and $M = 2$

| $\frac{L^H - L^0}{L^0}$ 100 | $\frac{L^H - L^*}{L^*}$ 100 | $\frac{\ln L^H - \ln L^0}{\ln L^0}$ 100 | $\frac{\ln L^H - \ln L^*}{\ln L^*}$ 100 |
|------------------------------|------------------------------|---------------------------------------------|---------------------------------------------|
| $6.8783 \times 10^{20}\%$    | $-0.9347\%$                  | $7.1386\%$                                  | $-0.0014\%$                                |

Since there are 360 binary strings in $V^2$ in this case, sampling 90 mutants for every block-structure, random samples from $V^n$, $n = 3, 4, 5$ were generated, each containing 360 binary strings. Using penalty values 5 and 10, the approximate distribution $D$ given in Table 15 was estimated. Probability assigned to all insignificant macroeconomic scenarios was 0.05%. Table 16 characterizes the relative standing of $L^H$ against $L^0$ as well as against the maximum value $L^*$ found with the exact algorithm. The heuristic solution appears to be just slightly worse than the exact one, being more convenient for a conceptual analysis of the corresponding macroeconomic scenarios.

Appendix 2

Distribution of migration counts among credit classes and industry sectors is characterized by Tables 17 and 18. The cell belonging to row $s$ and column $m$ contains the number of counts available for the industry sector $s$ and the credit class $m$. 

 Springer
Table 17 \( \sum_{t=1}^{T} \sum_{m_1=1}^{M+1} I^t(s, m, m_1) \) for \( M = 2 \) and \( S = 12 \)

| \( m \) | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 | Total |
|------|----|----|----|----|----|----|----|----|----|----|----|----|------|
| 1    | 3391 | 5242 | 2073 | 17518 | 1298 | 2370 | 2053 | 27447 | 1264 | 1661 | 2136 | 9246 | 75699 |
| 2    | 3129 | 4096 | 1661 | 1647 | 1296 | 1776 | 1920 | 7246 | 2316 | 1026 | 1021 | 890 | 28024 |
| Total| 6520 | 9338 | 3734 | 19165 | 2594 | 4146 | 3973 | 34693 | 3580 | 2687 | 3157 | 10136 | 103723 |

Table 18 \( \sum_{t=1}^{T} \sum_{m_1=1}^{M+1} I^t(s, m, m_1) \) for \( S = 6 \) and \( M = 7 \)

| \( s \) | 1  | 2  | 3  | 4  | 5  | 6  | 7  | Total |
|------|----|----|----|----|----|----|----|------|
| 1    | 118 | 263 | 719 | 1531 | 1125 | 1068 | 128 | 4952 |
| 2    | 315 | 1159 | 3774 | 4483 | 3655 | 3501 | 331 | 17218 |
| 3    | 369 | 1634 | 5362 | 6232 | 1861 | 1840 | 259 | 17557 |
| 4    | 39 | 225 | 805 | 1278 | 1004 | 1126 | 64 | 4541 |
| 5    | 2760 | 11186 | 17978 | 13142 | 5730 | 2691 | 543 | 54030 |
| 6    | 52 | 367 | 607 | 1301 | 1316 | 1642 | 140 | 5425 |
| Total| 3653 | 14834 | 29245 | 27967 | 14691 | 11868 | 1465 | 103723 |

The following migration matrices were estimated for \( S = 6 \) and \( M = 7 \):

\[
P^{(1)} = \begin{pmatrix}
0.9492 & 0.0508 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0114 & 0.9202 & 0.0646 & 0.0038 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0167 & 0.9110 & 0.0723 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0007 & 0.0033 & 0.0202 & 0.9406 & 0.0320 & 0.0007 & 0.0020 & 0.0007 \\
0.0000 & 0.0000 & 0.0000 & 0.0507 & 0.8640 & 0.0782 & 0.0027 & 0.0044 \\
0.0000 & 0.0000 & 0.0000 & 0.0047 & 0.0496 & 0.8558 & 0.0552 & 0.0346 \\
0.0000 & 0.0000 & 0.0000 & 0.0078 & 0.0000 & 0.1797 & 0.5234 & 0.2891 \\
\end{pmatrix},
\]

\[
P^{(2)} = \begin{pmatrix}
0.9238 & 0.0667 & 0.0095 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0017 & 0.9241 & 0.0733 & 0.0009 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0101 & 0.9367 & 0.0509 & 0.0024 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0002 & 0.0272 & 0.9275 & 0.0379 & 0.0060 & 0.0004 & 0.0007 \\
0.0000 & 0.0003 & 0.0005 & 0.0487 & 0.8627 & 0.0777 & 0.0022 & 0.0079 \\
0.0000 & 0.0000 & 0.0009 & 0.0014 & 0.0583 & 0.8538 & 0.0531 & 0.0326 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0060 & 0.1148 & 0.5347 & 0.3444 \\
\end{pmatrix},
\]

\[
P^{(3)} = \begin{pmatrix}
0.9133 & 0.0813 & 0.0027 & 0.0027 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0031 & 0.8727 & 0.1138 & 0.0086 & 0.0006 & 0.0012 & 0.0000 & 0.0000 \\
0.0000 & 0.0084 & 0.9215 & 0.0660 & 0.0026 & 0.0004 & 0.0006 & 0.0006 \\
0.0000 & 0.0003 & 0.0345 & 0.9387 & 0.0209 & 0.0029 & 0.0006 & 0.0021 \\
0.0011 & 0.0000 & 0.0054 & 0.0752 & 0.8280 & 0.0742 & 0.0064 & 0.0097 \\
0.0000 & 0.0000 & 0.0060 & 0.0022 & 0.0804 & 0.7984 & 0.0636 & 0.0495 \\
0.0000 & 0.0000 & 0.0039 & 0.0039 & 0.0077 & 0.1081 & 0.4865 & 0.3900 \\
\end{pmatrix},
\]

\[
P^{(4)} = \begin{pmatrix}
0.9487 & 0.0513 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.9378 & 0.0622 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0037 & 0.9019 & 0.0944 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.0282 & 0.9124 & 0.0540 & 0.0047 & 0.0000 & 0.0008 \\
0.0010 & 0.0000 & 0.0020 & 0.0488 & 0.8516 & 0.0906 & 0.0020 & 0.0040 \\
0.0000 & 0.0000 & 0.0000 & 0.0018 & 0.0533 & 0.8810 & 0.0409 & 0.0231 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0938 & 0.4688 & 0.4375 \\
\end{pmatrix},
\]
For $S = 12$ and $M = 2$ the transition matrices are:

$$
\begin{align*}
P^{(5)} & = \\
& = \begin{pmatrix}
0.8851 & 0.1080 & 0.0047 & 0.0007 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0014 \\
0.0066 & 0.9057 & 0.0826 & 0.0038 & 0.0002 & 0.0007 & 0.0003 & 0.0002 \\
0.0016 & 0.0503 & 0.8932 & 0.0506 & 0.0018 & 0.0008 & 0.0004 & 0.0014 \\
0.0023 & 0.0088 & 0.0829 & 0.8361 & 0.0595 & 0.0078 & 0.0011 & 0.0015 \\
0.0009 & 0.0077 & 0.0209 & 0.1878 & 0.7003 & 0.0679 & 0.0084 & 0.0061 \\
0.0030 & 0.0052 & 0.0111 & 0.0364 & 0.1903 & 0.6815 & 0.0465 & 0.0260 \\
0.0037 & 0.0000 & 0.0018 & 0.0037 & 0.0387 & 0.1510 & 0.6759 & 0.1252 \\
0.0038 & 0.0577 & 0.0192 & 0.0192 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0136 & 0.9401 & 0.0300 & 0.0163 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0132 & 0.9110 & 0.0741 & 0.0000 & 0.0016 & 0.0000 & 0.0000 \\
0.0008 & 0.0000 & 0.0292 & 0.9185 & 0.0430 & 0.0077 & 0.0000 & 0.0008 \\
0.0000 & 0.0000 & 0.0000 & 0.0403 & 0.8754 & 0.0752 & 0.0023 & 0.0068 \\
0.0000 & 0.0000 & 0.0006 & 0.0300 & 0.0621 & 0.8660 & 0.0420 & 0.0262 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0286 & 0.1500 & 0.5357 & 0.2857
\end{pmatrix},
\end{align*}
$$

$$
P^{(6)} = \\
= \begin{pmatrix}
0.9752 & 0.0245 & 0.0003 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.9762 & 0.0235 & 0.0004 \\
0.0291 & 0.9338 & 0.0371 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0229 & 0.9424 & 0.0347 \\
0.9928 & 0.0063 & 0.0010 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.9646 & 0.0339 & 0.0015 \\
0.0522 & 0.9101 & 0.0376 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0309 & 0.9352 & 0.0340 \\
0.9771 & 0.0224 & 0.0005 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.9677 & 0.0311 & 0.0012 \\
0.0250 & 0.9240 & 0.0510 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.1813 & 0.8033 & 0.0153 \\
0.9789 & 0.0181 & 0.0030 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.9841 & 0.0154 & 0.0005 \\
0.0263 & 0.9006 & 0.0731 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0264 & 0.9354 & 0.0382 \\
0.9875 & 0.0115 & 0.0011 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.9753 & 0.0130 & 0.0000 \\
0.0111 & 0.0300 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0229 & 0.9506 & 0.0265 \\
\end{pmatrix}.
$$

References

Aretz K, Bartram SM, Pope PF (2010) Macroeconomic risks and characteristic-based factor models. J Bank Finance 34:1383–1399

Bangia A, Diebold FX, Kronimus A, Schagen C, Schuermann T (2002) Ratings migration and the business cycle, with applications to credit portfolio stress testing. J Bank Finance 26(2/3):445–474

Boons MF (2016) State variables, macroeconomic activity and the cross-section of individual stocks. J Financ Econ 119:489–511

Boreiko DV, Kaniovski SY, Kaniovski YM, Pflug GC (2017) Traces of business cycles in credit-rating migrations. PLoS ONE 2017:1–29. https://doi.org/10.1371/journal.pone.0175911

Boreiko D, Kaniovski Y, Pflug GC (2016) Modeling dependent credit rating transitions: a comparison of coupling schemes and empirical evidence. Cent Eur J Oper Res 24(4):989–1007

Cooper I, Maio P (2019) Asset growth, profitability, and investment opportunities. Manag Sci 65:3988–4010

Fama EF, French KR (2017) International tests of a five-factor asset pricing model. J Financ Econ 123:441–463

Fei F, Fuertes A-M, Kalotychou E (2012) Credit rating migration risk and business cycles. J Bus Finance Account 39(1–2):229–263

Frydman H, Schuermann T (2008) Credit rating dynamics and Markov mixture models. J Bank Finance 32(6):1062–1075

Gilli M, Schumann E (2012) Heuristic optimisation in financial modelling. Ann Oper Res 193(1):129–158

Gilli M, Winker P (2009) Heuristic optimization methods in econometrics. In: Belsley DA, Kontoghiorghes E (eds) Handbook of computational econometrics. Wiley, New York

Glasserman P, Young HP (2016) Contagion in financial networks. J Econ Lit 54(3):779–831

Glasserman P, Young HP (2015) How likely is contagion in financial networks? J Bank Finance 50:383–399

Goncalves AS, Xue C, Zhang L (2020) Aggregation, capital heterogeneity, and the investment CAPM. Rev Financ Stud 33:2728–2771
Gupton GM, Finger ChC, Bhatia M (1997) CreditMetrics—Technical Document. J.P. Morgan Inc., New York
Hölzl W, Kaniiovski S, Kaniiovski Y (2019) Exploring the dynamics of business survey data using Markov models. Comput Manag Sci 16(4):621–649
Kaniiovski YM, Pflug GCh (2007) Risk assessment for credit portfolios: a coupled Markov chain model. J Bank Finance 31(8):2303–2323
Kiff J, Kisser M, Schumacher L (2013) Rating through-the-cycle: what does the concept imply for rating stability and accuracy? IMF Working Paper, WP 13/64
McNeil A, Wendin J (2007) Bayesian inference for generalized linear mixed models of portfolio credit risk. J Empir Finance 14(2):131–149
Nagpal K, Bahar R (2001) Measuring default correlation. Risk 14(3):129–132
Stefanescu C, Tunary R, Turnbull S (2009) The credit rating process and estimation of transition probabilities: a Bayesian approach. J Empir Finance 16(2):216–234
Wozabal D, Hochreiter R (2012) A coupled Markov chain approach to credit risk modeling. J Econ Dyn Control 36(3):403–415
Xing H, Sun N, Chen Y (2012) Credit rating dynamics in the presence of unknown structural breaks. J Bank Finance 36(1):78–89

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.