Proceedings of CONF-APMM 2026 Symposium: Applied Mathematics and Quantitative Modeling
DOI: 10.54254/2753-8818/2026.PU38932

Applications of Markov Chains in Financial Risk Modeling: A Critical Literature Review

Wenbo Li

School of Mathematical and Physical Sciences, Xi'an Jiaotong-Liverpool University, Suzhou, China
wenbo1115@outlook.com

Abstract. Markov chains represent financial risk as movement between credit or market states. This review examines how that representation supports credit migration analysis, default estimation, market risk forecasting, and portfolio decisions. It compares foundational models with subsequent research on rating heterogeneity, switching volatility, and economic conditions. The literature supports Markov models as interpretable tools for connecting current information to future loss distributions, but their usefulness depends on the state definition and stability of transition probabilities. A three-state numerical illustration shows how persistent adverse conditions can amplify cumulative default risk through repeated migration. The review distinguishes physical probabilities used for forecasting from risk-neutral probabilities used for pricing, and explains why a transition matrix alone cannot determine portfolio tail risk. It concludes that Markov models are most credible when combined with economic conditioning, explicit dependence assumptions, parameter uncertainty, and evaluation on observations unavailable during estimation.

Keywords: Markov chains, credit migration, regime switching, financial risk, expected shortfall

1. Introduction

Financial losses depend on the conditions through which an exposure evolves. A borrower can deteriorate before default, while a market can remain volatile for several periods after an initial shock. Markov chains describe these movements through a finite collection of states and transition probabilities. Their appeal in financial mathematics is that repeated uncertainty can be analyzed using matrix operations, while the states retain an economic interpretation.

Two research traditions are central. Hamilton introduced a tractable framework for time series whose parameters change with an unobserved regime [1]. Jarrow, et al. connected credit ratings and default to the term structure of credit spreads [2]. The first tradition primarily describes changing market conditions, the second tracks borrower quality. Both raise the same question: how much information about future risk is retained by the current state?

This paper provides a selective narrative review of these traditions and their implications for risk measurement. This review offers practical implications for financial practitioners to avoid model misuse and improve risk measurement, stress-testing and portfolio risk assessment.

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2026.PU38932

2. Mathematical foundations

Let denote the state at time t, taking values in a finite set. The first-order Markov assumption says that the conditional distribution of the next state depends on the present state alone [3]:

$$\Pr(X_{t+1} = j \mid X_t = i, \ldots, X_0) = \Pr(X_{t+1} = j \mid X_t = i) \tag{1}$$

The transition matrix contains entries p_{ij}. Each entry is nonnegative and each row sums to one. In a time-homogeneous model, these probabilities remain fixed. If is a row vector containing the current state probabilities, the distribution h periods ahead is

$$\pi_{t+h} = \pi_t P^h \tag{2}$$

The matrix power sums the probabilities of every admissible intermediate path. It therefore captures indirect deterioration, such as movement from a strong rating to a weak rating before default. With an absorbing default state D, the entry is the probability of default by horizon h, conditional on starting in state i. It is cumulative, rather than the probability of default only during the final period.

For completely observed transitions at equally spaced dates, a basic maximum-likelihood estimate is $\widehat{p}_{ij} = \frac{N_{ij}}{\sum_k N_{ik}}$, where counts observed transitions from i to j. Infrequent states produce uncertain estimates, and zero observed defaults do not establish zero default risk. Pooling information across comparable borrowers or using Bayesian regularization can stabilize estimation, although the associated assumptions must remain visible.

Continuous-time models use a generator Q, whose off-diagonal entries are transition intensities and whose rows sum to zero. For a homogeneous chain, $P(h) = \exp(hQ)$. Lando and Skødeberg show the value of using continuous rating histories rather than only beginning and ending ratings [4]. Exact transition times retain information about multiple movements within a year. However, an empirical annual transition matrix need not equal the exponential of a valid generator, Israel, Rosenthal, and Wei analyze this embedding problem [5]. Taking an unrestricted matrix logarithm is therefore insufficient.

An observed-state model treats recorded ratings as states. A hidden Markov model instead infers an unobserved condition from returns or other measurements. Its observation equation is as important as its transition matrix. For example, assigning a normal return distribution to each regime can produce volatility clustering, but does not ensure that extreme losses are represented adequately.

Interpretation is also necessary for long-term behavior. An irreducible finite chain has a stationary distribution, and aperiodicity ensures convergence to it. An absorbing credit chain predicts default with probability one, limiting its forecast usefulness for a dynamic bank portfolio, necessitating a finite horizon and defined population.

3. Credit migration and default risk

Credit migration models extend a binary default indicator by retaining intermediate deterioration and recovery. This matters because a downgrade can reduce a bond's market value even when the issuer continues paying. Jarrow, et al. provide an arbitrage-free framework in which rating movements affect defaultable claims [2]. Their contribution is a connection between rating dynamics and

DOI: 10.54254/2753-8818/2026.PU38932

valuation, rather than evidence that one historical matrix is suitable for every borrower or economic environment.

A crucial distinction concerns the probability measure. Historical, or physical, probabilities describe observed migration frequencies and support loss forecasting. Pricing generally requires risk-neutral probabilities calibrated consistently with traded prices and recovery assumptions. These probabilities incorporate compensation for bearing risk. Substituting historical default frequencies directly into a pricing formula can therefore misrepresent spreads even when those frequencies are estimated accurately. Conversely, market-implied default probabilities should not automatically be interpreted as forecasts of actual default frequencies.

The empirical literature questions the stability of a single pooled matrix. Nickell, et al. examine variation with borrower characteristics and business-cycle conditions [6]. Bangia et al. connect migration behavior to economic conditions and use this relationship in credit portfolio stress testing [7]. Together, these studies motivate separating a long-run average transition pattern from a forecast conditional on current conditions. A matrix estimated mainly during expansion can understate the deterioration associated with a sustained downturn.

History also matters within a given rating. Lando et al. identify departures from simple Markov behavior, including effects related to previous rating movements [4]. Frydman and Schuermann address heterogeneity using a mixture of Markov chains [8]. Borrowers with the same observed rating can then have different future migration distributions because their histories imply different latent characteristics. This challenges the interpretation of a rating category as a sufficient summary of credit quality.

Recent work clarifies what a richer state must accomplish. Kalkbrener and Packham model economic conditions and credit ratings jointly as a Markov chain and show that the rating process alone is generally not Markovian [9]. They also formalize differences between point-in-time and through-the-cycle ratings. This supports a specific modeling principle: augmenting the state may explain apparent memory, but Markov behavior of the expanded system does not automatically transfer to an observed component.

For a single exposure with fixed exposure at default and fixed loss given default, expected loss over a horizon equals their product with the cumulative default probability. A portfolio requires more information. Separate borrower matrices determine marginal migration probabilities, but not joint defaults. Common economic states can create dependence, although additional links may remain within sectors or lending networks. Consequently, a model can estimate average loss reasonably while understating extreme portfolio losses. Migration-related changes in market value, collateral, and recovery also require separate valuation assumptions.

Data handling can change the apparent transition process. Withdrawn ratings should not automatically be classified as survival since withdrawal reasons may relate to borrower quality. Similarly, changes in rating definitions can make observations from different periods incomparable. The appropriate estimation sample must specify issuer versus instrument ratings, the observation frequency, and the treatment of exits. These choices determine what population the estimated matrix describes and whether its probabilities can be transferred to another lending book.

4. Market risk and portfolio applications

Market applications usually treat the regime as latent. A low-volatility state and a high-volatility state can have different return distributions and different persistence. Hamilton and Susmel combine changing regimes with autoregressive conditional heteroskedasticity [10]. Haas, et al. develop a tractable switching version of generalized autoregressive conditional heteroskedasticity (GARCH)

DOI: 10.54254/2753-8818/2026.PU38932

models [11]. This combination allows volatility to evolve within each regime as well as change across regimes, avoiding the assumption that every turbulent episode follows one unchanged variance process.

Forecasting requires averaging over uncertainty about tomorrow's state. Let contain filtered regime probabilities based only on information available at time t. The predicted probabilities are $w_{t+1|t} = w_t P$. If is the conditional loss distribution in states, then

$$F_{t+1}(l) = \sum_s w_{t+1|t,s} F_{s,t+1}(l) \tag{3}$$

A loss quantile must be calculated from this mixture distribution. A weighted average of regime-specific quantiles generally gives a different answer. Filtering also differs from smoothing: smoothing uses later observations to infer past regimes. Smoothed classifications are useful for historical interpretation but would introduce future information if used to evaluate a real-time trading or risk forecast.

The research conducted by Ardia et al. favors switching GARCH models for several tail-risk forecasts, particularly for equity returns, while also showing benefits from accounting for parameter uncertainty [12]. The evidence is stronger than an improvement in fitted likelihood because it evaluates forecasts outside the estimation sample. It nevertheless remains conditional on the assets, horizons, distributions, and comparison models examined. The MSGARCH implementation makes such comparisons reproducible and supports both likelihood-based and Bayesian estimation [13].

For a loss variable L, value-at-risk at confidence level alpha is its alpha quantile. Expected shortfall averages the quantiles above that confidence level:

$$ES_\alpha(L) = \frac{1}{1-\alpha} \int_\alpha^1 VaR_u(L)\, du,\ 0 < \alpha < 1 \tag{4}$$

Acerbi and Tasche establish the importance of a careful expected-shortfall definition, especially when the loss distribution has probability masses [14]. Unlike a quantile alone, expected shortfall reflects the severity of tail outcomes. Regime models can affect both measures by changing the probability of entering a stressed state and the loss distribution within that state. A persistent crisis regime with thin-tailed conditional returns may still miss unusually large shocks.

Portfolio allocation adds correlations to the problem. Ang and Bekaert analyze international allocation under regime shifts [15]. Their framework illustrates why diversification should be evaluated conditionally: the joint behavior of assets can change when markets deteriorate. A portfolio constructed from average covariances may therefore provide less protection in adverse conditions than expected. However, recognizing regimes does not establish profitable market timing. Estimated signals can be unstable, and transaction costs can offset gains from frequent rebalancing.

Forecast horizon matters as well. A one-day predictive distribution allows only the next transition, whereas a monthly loss distribution depends on many possible regime paths and successive return shocks. Scaling daily VaR by the square root of time generally lacks justification when regimes and conditional variances evolve. Simulation can instead propagate both the regime and the return process. For nonlinear positions, simulated risk factors must also be passed through a valuation model before the portfolio loss distribution is obtained.

5. An illustrative credit risk calculation

Consider a borrower in one of three states: good credit quality G, vulnerable credit quality V, or default D. Table 1 gives assumed annual transition probabilities for baseline and adverse conditions.

DOI: 10.54254/2753-8818/2026.PU38932

These values are constructed solely to illustrate matrix propagation and are not derived from a rating agency or a historical borrower sample, with default being an absorbing state.

Table 1. Assumed annual credit transition probabilities

Scenario	Current state	To G	To V	To D
Baseline	G	0.94	0.05	0.01
Baseline	V	0.08	0.86	0.06
Baseline	D	0.00	0.00	1.00
Adverse	G	0.88	0.09	0.03
Adverse	V	0.04	0.84	0.12
Adverse	D	0.00	0.00	1.00

Starting from G, the two-year baseline default probability includes three paths: remaining good and then defaulting, becoming vulnerable and then defaulting, or defaulting immediately. Thus,

$$(\mathrm{P}^2)_{\mathrm{GD}} = 0.94 \times 0.01 + 0.05 \times 0.06 + 0.01 = 0.0224 = 2.24\% \tag{5}$$

The corresponding adverse probability is 6.72%. With a constant exposure of USD 1 million and loss given default of 60%, undiscounted two-year expected default loss rises from USD 13,440 to USD 40,320. This calculation excludes changes in exposure, recovery, and market value before default. It isolates the effect of different migration assumptions.

The starting distribution is equally relevant. A book containing vulnerable borrowers would use a weighted combination of the G and V rows, rather than the G row alone. Due to differences in their baseline credit quality and exposure weights, two lenders using the same transition assumptions can therefore obtain different expected losses.

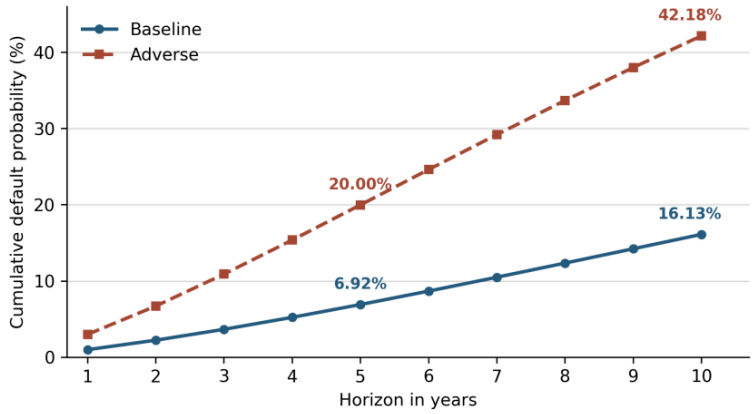


Figure 1. Cumulative default probability from an initially good rating

Figure 1 extends the calculation to ten years by repeatedly multiplying each transition matrix. At five years, the cumulative probabilities are 6.92% and 20.00%. The widening difference arises because the adverse matrix increases direct default risk, accelerates downgrades, and reduces upgrades. One-year comparisons therefore omit part of the effect of persistent deterioration.

DOI: 10.54254/2753-8818/2026.PU38932

The figure is a sensitivity analysis, not empirical validation. If conditions change annually, propagation requires an ordered product of the relevant matrices. If conditions are themselves uncertain, a joint economic-state and rating model is needed. Nor should the plotted cumulative probabilities be added across years, that would count earlier defaults repeatedly. The probability of a new default during a particular year is the increase in cumulative default probability over that year.

6. Conclusion

This study reveals a trade-off between interpretable states and omitted information. A coarse rating system pools dissimilar borrowers, a very detailed system leaves few observations in each category. In a latent model, increasing the number of states can fit unusual observations without identifying stable economic regimes. Model selection should therefore combine statistical criteria with minimum state occupancy, sensitivity to initialization, and an explanation of what each state represents.

Further research should prioritize whether added information improves decisions. Economic covariates can make transition probabilities responsive to conditions, while Bayesian estimation can regularize sparse transitions. Joint credit and market states may reveal dependence that separate models miss. These extensions should be judged by calibration during stressed periods and the stability of resulting risk limits. More flexible models can improve forecasts, but they also make it harder to distinguish economic structure from sample-specific patterns. The strongest contribution would be a reproducible comparison showing when additional complexity materially changes measured risk.

References

[1] Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica, 57*(2), 357–384. https://doi.org/10.2307/1912559

[2] Jarrow, R. A., Lando, D., & Turnbull, S. M. (1997). A Markov model for the term structure of credit risk spreads. *The Review of Financial Studies, 10*(2), 481–523. https://doi.org/10.1093/rfs/10.2.481

[3] Norris, J. R. (1997). *Markov chains*. Cambridge University Press. https://doi.org/10.1017/CBO9780511810633

[4] Lando, D., & Skødeberg, T. M. (2002). Analyzing rating transitions and rating drift with continuous observations. *Journal of Banking & Finance, 26*(2–3), 423–444. https://doi.org/10.1016/S0378-4266(01)00228-X

[5] Israel, R. B., Rosenthal, J. S., & Wei, J. Z. (2001). Finding generators for Markov chains via empirical transition matrices, with applications to credit ratings. *Mathematical Finance, 11*(2), 245–265. https://doi.org/10.1111/1467-9965.00114

[6] Nickell, P., Perraudin, W., & Varotto, S. (2000). Stability of rating transitions. *Journal of Banking & Finance, 24*(1–2), 203–227. https://doi.org/10.1016/S0378-4266(99)00057-6

[7] Bangia, A., Diebold, F. X., Kronimus, A., Schagen, C., & Schuermann, T. (2002). Ratings migration and the business cycle, with application to credit portfolio stress testing. *Journal of Banking & Finance, 26*(2–3), 445–474. https://doi.org/10.1016/S0378-4266(01)00229-1

[8] Frydman, H., & Schuermann, T. (2008). Credit rating dynamics and Markov mixture models. *Journal of Banking & Finance, 32*(6), 1062–1075. https://doi.org/10.1016/j.jbankfin.2007.09.013

[9] Kalkbrener, M., & Packham, N. (2026). A Markov approach to credit rating migration conditional on economic states. *Canadian Journal of Statistics, 54*(1), e70039. https://doi.org/10.1002/cjs.70039

[10] Hamilton, J. D., & Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics, 64*(1–2), 307–333. https://doi.org/10.1016/0304-4076(94)90067-1

[11] Haas, M., Mittnik, S., & Paolella, M. S. (2004). A new approach to Markov-switching GARCH models. *Journal of Financial Econometrics, 2*(4), 493–530. https://doi.org/10.1093/jjfinec/nbh020

[12] Ardia, D., Bluteau, K., Boudt, K., & Catania, L. (2018). Forecasting risk with Markov-switching GARCH models: A large-scale performance study. *International Journal of Forecasting, 34*(4), 733–747. https://doi.org/10.1016/j.ijforecast.2018.05.004

DOI: 10.54254/2753-8818/2026.PU38932

[13] Ardia, D., Bluteau, K., Boudt, K., Catania, L., & Trottier, D.-A. (2019). Markov-switching GARCH models in R: The MSGARCH package. *Journal of Statistical Software, 91*(4), 1–38. https://doi.org/10.18637/jss.v091.i04
[14] Acerbi, C., & Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance, 26*(7), 1487–1503. https://doi.org/10.1016/S0378-4266(02)00283-2
[15] Ang, A., & Bekaert, G. (2002). International asset allocation with regime shifts. *The Review of Financial Studies, 15*(4), 1137–1187. https://doi.org/10.1093/rfs/15.4.1137