

Contents

Introduction	1
Data overview and model inputs	2
Modeling	2
Evaluation	4
References	6

PitchBook | LCD Default Predictor Data and Methodology

PitchBook is a Morningstar company providing the most comprehensive, most accurate, and hard-to-find data for professionals doing business in the private markets.

Introduction

The PitchBook | LCD Default Predictor delivers a monthly forward-looking estimate of the aggregate default rate based on the constituents in the Morningstar LSTA US Leveraged Loan Index. Leveraging PitchBook | LCD's comprehensive loan-level dataset, which provides current pricing for thousands of loans in the index, the model employs a logistic regression framework to estimate the aggregate default rate (defined as the number of companies that will formally default) over the next six months.

PitchBook | LCD offers extensive coverage of the leveraged loan market, regularly surfaced through data and analytics related to the Morningstar LSTA US Leveraged Loan Index. The long history of loan-level pricing data that underpins the index creates a strong foundation for understanding how market signals correspond to future defaults.

This document describes the methodology behind the PitchBook | LCD Default Predictor, including the data and inputs used to train the model, the estimation process, and the transformations applied to produce the final default rate estimate.

Questions or feedback regarding this methodology may be directed to support@pitchbook.com.



Data overview and model inputs

The PitchBook | LCD Default Predictor is built on a loan-level panel dataset maintained by PitchBook | LCD. This dataset provides monthly snapshots of individual leveraged loans in the Morningstar LSTA US Leveraged Loan Index going back to 1997 and includes loan pricing and default-event information. While the loan data extends back to 1997, data quality improves meaningfully from 2005 onward; therefore, the modeling process described below begins in January 2005.

The following inputs, which capture the level and recent trajectory of loan prices, are used to train a logistic regression model to estimate whether a loan will default within the next six months.

PitchBook | LCD Default Predictor model inputs

Input	Definition
Loan price	Bid price of loan
Average index price	Cross-sectional mean loan price at the observed date
Price percentage change over 1 month	1-month price percentage change per loan
Price percentage change over 2 months	2-month price percentage change per loan
Price percentage change over 3 months	3-month price percentage change per loan

Sources: PitchBook | LCD, Morningstar

Modeling

The PitchBook | LCD Default Predictor is a logistic regression model that generates loan-level default probabilities six months forward. These probabilities are then aggregated into a composite default rate and calibrated to align with the historical distribution of observed defaults. The result is a forward-looking default rate estimate comparable to the reported Morningstar LSTA US Leveraged Loan Index default rate, a trailing 12-month (TTM) metric based on the count of issuers that defaulted.

Preprocessing and training pipeline

The modeling pipeline applies standard scaling to the five price-related inputs before fitting a logistic regression model, with default within the next six months as the dependent variable.

To avoid look-ahead bias, the model uses an expanding-window training procedure. Beginning in January 2010, after an initial five-year training window starting in January 2005, the model is refit each month on all data available up to that point, generating a probability of default over the next six months for each loan in the index. Therefore, every prediction is made using only information that would have been available at the time, and the entire prediction history is out-of-sample by construction.



Aggregation into a composite default rate

Default probabilities for each loan are first aggregated by issuer. Assuming that a default on any single loan i causes all of the loans for issuer j to be treated as defaulted, the issuer-level default probability is taken as the maximum loan-level default probability \widehat{P} among that issuer's loans:

$$P(\text{issuer } j \text{ defaults})_t = \max_{i \in j} \widehat{p}_{i,t}$$

Taking the maximum reflects the view that an issuer's default risk is driven by its most distressed obligation. The issuer-level probabilities are then summed across all M_t issuers to estimate the expected number of issuer defaults $E[D]$ at time t :

$$E[D]_t = \sum_{j=1}^{M_t} P(\text{issuer } j \text{ defaults})_t$$

Constructing a TTM estimate

The model's expected issuer defaults over the next six months are combined with observed defaults over the prior six months to produce a forward-looking estimate expressed in TTM terms, which is the same convention used by the reported index default rate. Specifically, with the most recent month end defined as $t = 0$, the projected TTM default rate six months forward is calculated as:

$$DR_{t_6} = \frac{D_{[t-5,t]} + E[D_{(t,t+6)}]}{\sum_{j=1}^M d_{j,t-6} = 0}$$

Calibration

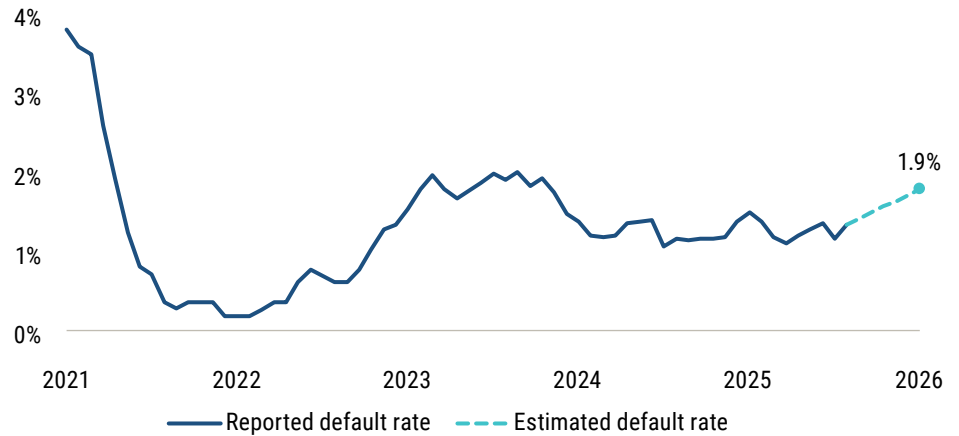
A calibration step adjusts the raw estimated default rate series to match the historical distribution of actual default rates. This step is applied because loan prices tend to be more volatile than realized defaults, which can cause the raw estimated rate to fluctuate far more than the actual series during periods of market stress.

The calibration applies a linear transformation in logit space, ensuring that the calibrated estimate remains bounded between zero and one. The transformation is parameterized to match the mean and standard deviation of the observed default rate distribution over an expanding historical window.^{1,2}

The calibrated default rate estimate can then be compared with the reported Morningstar LSTA US Leveraged Loan Index default rate. Because the model produces a single-point estimate six months forward rather than a monthly path, we display a straight-line extension from the most recent reported value to the current estimate.



Reported Morningstar LSTA US Leveraged Loan Index default rate and estimated default rate in next six months



Sources: PitchBook | LCD, Morningstar • Geography: US • As of May 31, 2026
Note: Default rates are calculated based on the count of issuers.

Evaluation

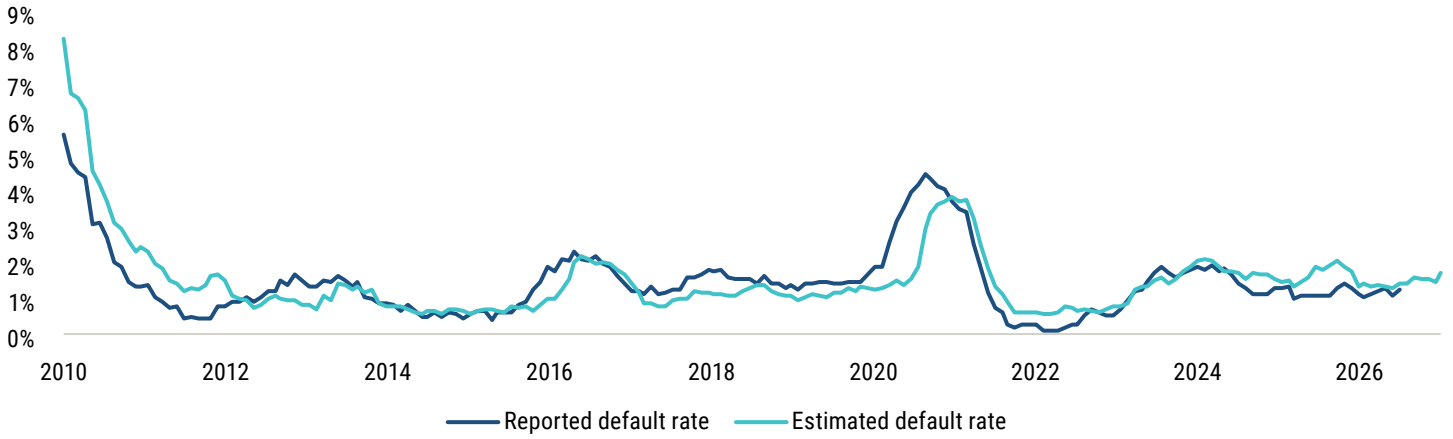
The model is evaluated by comparing the estimated default rate series against the reported index default rate over the full out-of-sample history. Because every prediction is generated using only data available at that point in time, no separate held-out test set is needed.

The model demonstrates a meaningful ability to track both the level and movement of the Morningstar LSTA US Leveraged Loan Index default rate. On average, the level estimate falls within roughly 40 to 50 basis points of the actual default rate. In terms of month-over-month changes, the model correctly anticipates the direction of movement approximately three out of five times (63.9%), a statistically significant result. A correlation of 67% and an R-squared of 44.9% indicate that the model captures a meaningful portion of the variation in actual default rate movements.

Overall, while the model is not intended to pinpoint exact default rate levels, it serves as a valuable forward-looking gauge of aggregate default trends, offering useful directional context within broader credit market analysis. Because the model relies on market-based inputs, it is not designed to anticipate sudden exogenous shocks, such as the COVID-19 pandemic, that are not yet reflected in loan prices.



Reported versus estimated Morningstar LSTA US Leveraged Loan Index default rate



Sources: PitchBook | LCD, Morningstar • Geography: US • As of May 31, 2026
Note: Default rates are calculated based on the count of issuers.



References

- 1: A minimum of 24 months of history is required before moment-matching begins; prior to that threshold, the raw uncalibrated estimate is used in order to preserve a longer time series.
- 2: Because the moment-matching is performed in logit space, the calibrated series will not match the target mean and standard deviation exactly when mapped back to rates space due to Jensen's inequality.