dv01 + insights

Roll-Rate Analysis for Marketplace Loan Portfolios Composed of New & Seasoned Loans

ABSTRACT / Roll rate analysis is the prevalent modeling approach in the consumer credit space, but its implication for MPL portfolios has never been studied or proven. In this Insight report, we use data from more than 1.6 million marketplace loans to investigate the applicability and nuances of applying roll rate analysis to MPL portfolios composed of loans of different ages.

July 2017

Wei Wu, Ph.D., CFA
Principal Data Scientist
dv01

sales@dv01.co | 646-854-5258 | 915 Broadway, Suite 802, New York, NY 10010

INTRODUCTION / Accurate default curves are essential to investors, and are a key component in assessing the risk of any consumer loan portfolio.

Investors can use default curves to not only determine the probability of default and expected loss, but as inputs for modeling future cashflows. Doing the latter more precisely estimates the impact of default on the values and timing of cashflows going forward, making it a powerful tool in managing a loan portfolio.

The most common, and easiest, curve to build is a historical default curve for newly originated loans: it is a curve measuring the default probabilities at all future time horizon (ages) through the lifetime of the loans. We calculate the probability at each time period or at each age n by first finding all the loans that were not terminated at time n-1. Then we count how many active loans at time n-1 were charged-off at time n. In other words, we track default statistics since inception for the lifetime of the loan, and use this default rate to predict future time periods.

In practice, however, a loan portfolio most likely has loans of different ages, and the default curves for seasoned loans can be very different from new loans.

ROLL RATE ANALYSIS / When a pool of loans has some seasoning, the method described above is no longer accurate or usable. In order to construct default curves for loans of different ages, you can either:

(1) aggregate historical performance data along one or two variables, or

(2) predict default using statistical models

Using the first method, one would have to construct curves for every loan age, which means dozens of curves if aggregating historical performance. Using the second method, one would have to introduce variables that take into account the effect of age on default curves. Without carefully considering the structure of how age affects the probability of default, adding age into the model may actually make the predictive power of the model worse.

But there is another solution, one that can even be described as a shortcut. In order to take into consideration seasoned loans without having to build a different curve per loan age, it is a common practice in the consumer credit space to use a method called roll rate analysis with "curve anchoring." This means utilizing a single default curve, anchoring the curve to loan age 0, and then adjusting it appropriately for the age of the loan moving forward. (We go into more detail on this method below.)

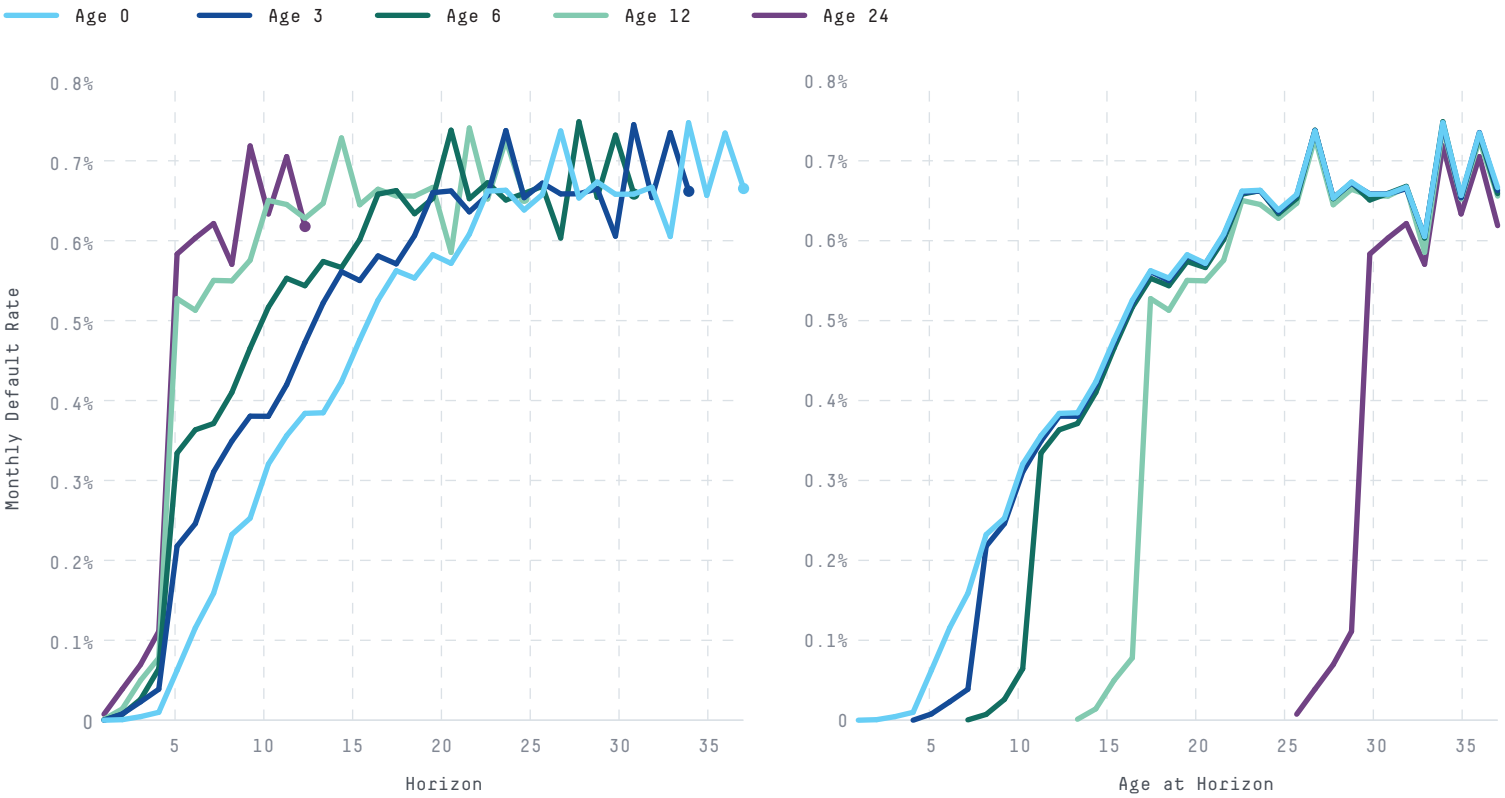
DOES ROLL RATE ANALYSIS APPLY TO MPL? / While roll rate analysis is a common modeling method in the broader consumer credit space, its applicability to unsecured

consumer loans in MPL portfolios has never been studied or proven. Below, we investigate whether traditional roll rate analysis works for MPL, and review nuances unique to this asset class.

DEFAULT CURVE COMPARISON: SEASONED VS. NEWLY ISSUED LOANS / Can modeling practices developed for the broader consumer credit space be applied to MPL? The first question to answer is whether there is a relationship/pattern between the default behavior of seasoned loans and the default behavior of newly originated loans, and whether this pattern is similar to what we see in consumer lending as a whole.

Consider the following: when we calculate the expected default probability of newly issued loans at 18 months out and the expected default probability of a 12-month seasoned loan going forward 6 months, *we count the number of loans surviving at the 17th month and the number of loans charged-off at the 18th month.* This is what's known as the loan age at future time-period, or what we call age at horizon.

FIG 1: Default curves for LC-B Loans with **Current** status at different ages



In order to see whether there is a visible relationship between seasoned and new loans, we plotted historical default curves of current loans (Lending Club, Grade B) at various ages on the left side of *Figure 1*, above. On the right side, we made the x-axis the loan age at the time horizon and shifted the default curves by the loan ages so that the ages at default time for all the curves are aligned.

sales@dv01.co | 646-854-5258 | 915 Broadway, Suite 802, New York, NY 10010

What do you notice?

Default curves of loans of different ages are the shifted age-0 default curve—the default curve anchored at loan origination. In other words, you see the following relationship between the expected default curves of seasoned and newly issued loans:

(1) The front ends of all the default curves are close to zero due to the nature of the transition time from current through delinquent and charged off (3-4 months) in MPL. The front-end values are significantly deviated from the default rates of shifted new loan default curves.

(2) At the long end of the curves, the default rates for seasoned loans converge to the default rate on the new loan curve at the same age as the seasoned loan.

Just like in the broader consumer credit space, there is a close relationship between the default of newly issued loans and seasoned loans. *This pattern suggests that roll rate analysis with curve anchoring at loan origination could be as predictive for MPL portfolios as it is for other consumer credit lending products.*

WHY WOULD ROLL RATE ANALYSIS WORK? / There is an intuitive explanation for the relationship between seasoned and newly issued loans, which can explain why the process of roll rate analysis would be predictive across asset classes.

By definition, the shifted age-0 curves are for seasoned loans of all statuses. At any stage in the lifetime of the pool of loans, there is a very small percentage of loans that are in delinquent status. Delinquent loans have very high probabilities of default in the short term, so even though their percentage is small, their impact on the average short term default rates is significant.

However, in the long term, after the charge-off processes are finished, the pool of loans used to compile the default curves are almost identical because:

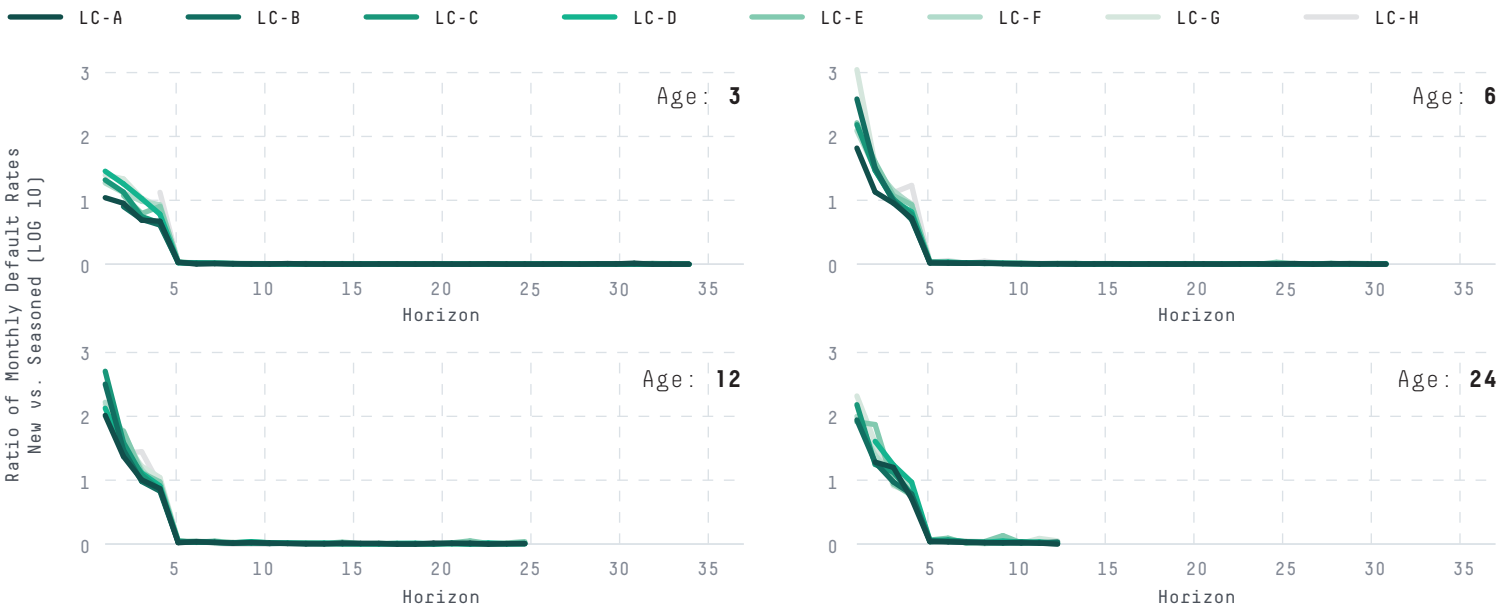
(1) the percent of delinquent loans is very small, and

(2) after the charge-off processes are over, most of the delinquent loans are terminated and gone from the statistics

Therefore, in the long term, the default curves of loans with only current statuses should be very similar to curves for loans with all statuses.

DEFAULT CURVES OF SEASONED VS. NEW LOANS: QUANTITATIVE PATTERN / We can further examine the differences between the default curves of seasoned and newly issued loans quantitatively to see if they resemble consumer credit markets,

FIG 2: Little difference between default curves for **Seasoned** vs. **New** loans for different grades



and therefore support roll rate analysis as an accurate modeling technique.

In *Figure 2*, we graph the ratio of default rates of shifted new loans and seasoned loans in log scale for different loan grades. We can see that the overall shape or profile of the differences is very similar. For all seasoned loans, at the very front end of the time horizon (the first 4-5 months), the default rates of seasoned loans are a couple of orders of magnitude smaller than the shifted default rates curve of newly issued loans. This scaling down makes the short end values of default curves for seasoned loans with current statuses very close to zero.

Beyond the very front end of the time horizon, the default rates for seasoned loans are almost identical to the shifted default curve of new loans. The magnitudes of these differences are similar for different loan ages and loan grades. We find the same shape pattern holds true across loan terms and lending platforms.

This pattern is quite similar to what we see across consumer credit, suggesting that using roll rate analysis is as accurate for MPL portfolios as it is for other consumer credit product portfolios.

ROLL RATE ANALYSIS IS PREDICTIVE FOR MPL PORTFOLIOS / After carefully examining the differences in default curves between seasoned and new loans for different ages, loan grades, loan terms, and even different marketplace lending platforms, we found a common and relatively simple pattern in the curve differences.

We see large modifications in the first few time periods that decrease as time periods increase, and all seasoned curves converge to a shifted new loan curve as the time horizon

passes beyond 5 months.

This closely resembles the pattern we see in the broader consumer credit space, suggesting that the method of roll rate analysis with curve anchoring used in other consumer credit modeling can also be used to build accurate default models for MPL portfolios.

With this insight, the model for constructing default curves for marketplace loans of all ages can be simplified significantly. Instead of building separate models for different loan ages, we can use the curve anchoring method to construct one model for the expected default curve over the lifetime of newly issued loans. We then construct default curves for seasoned loans (with current statuses) of different loan ages by first modeling the patterns of curve difference in default curves of seasoned and new loans. Finally, we apply the curve differences to the age-0 default curve to find default curves for seasoned loans of various ages.

Using this method to construct default curves for seasoned MPL loans not only simplifies the process, but also improves accuracy.

METHODOLOGY / For this study, we mainly used historical data from Lending Club. We used data from July 2007 to December 2016. (However, 99% of the data is 2010 or later.) The data contains more than 1.6 million loan IDs and more than 24 million monthly historical records.

For the platform comparison (mentioned here in brief), we included data from five major marketplace lenders. The data cutoff date was December 2016, but starting date varies for different platforms. The multi-platform dataset contains more than 3 million loan IDs and more than 43 million records.