Asset Backed Analytics in DROP

v3.65 17 July 2018

Market Place Lending Credit Model Methodology

Overview of Credit Model Methodology

1. Risk Estimation in MPL Platforms: DROP (2015) has introduced version 2.0 of the Credit Model for estimating prepayment and default risk on a pool of US unsecured consumer loans issued by a select set of marketplace lending ("MPL") platforms.
2. Valuation, Prepay, and Default Drivers: This extends the version 1.0 Credit Model published previously in August 2015. In both cases the goal is to enable the market participants understand the drivers behind valuation, prepayment, and default risk across their holdings in a transparent and robust manner.
3. Usage of Markov Logit Models: A main objective is to apply the rigor and forward looking tools of mortgage analytics, namely the Multinomial Logit and the Markov Chain approach to modeling the MPL collateral.
4. Simulation of Prepay and Default: The second objective is to have the ability apply simulations to prepayment and default rates, and ultimately the cash flows.
5. Usage of Statistical Learning Techniques: The final objective is to use machine learning techniques such as regularization and cross-validation to improve the robustness of the modeling.

Credit Methodology – Purpose and Introduction

1. MPL Growth/Institutional Investors Participation: The dramatic growth in Marketplace Lending (MPL) has been paired with the transition of the consumer and the SME credit to the

capital markets. Institutional investors are increasingly funding credit risk in whole loan, structured credit, or warehouse formats.

2. MPL Risk and Analytics Tools: A new set of analytics and risk pricing tools are necessary to enable institutional capital efficiently access, price, and exchange risk. A core component of risk pricing is independent 3rd party credit models that forecast cash flows on pools of loans.
3. Improvement in Liquidity and Transparency: By releasing a new credit model, the aim is to improve transparency and standards in MPL by providing risk management tools promoting independent pricing and whole loan and ABS market liquidity.
4. Valuation and Risk MPL Common Language: Independent credit models increase activity by allowing market participants to speak a common language in reference to valuation and risk.
5. Independent 3rd Party Credit Models: Several sources have described on various occasions the robust scaling and growth in the US and global marketplace lending sector. Large institutional investors require a 3rd party credit model to understand their credit risk exposures, and to participate in size.
6. Short Duration, High Yield Risk: Further, credit spreads have generally tightened since the 2008 credit crisis in a backdrop of quantitative easing, healthy global economy, and stringent regulation. As a result, investors see MPL offering an attractive short-duration, high-yield credit risk compared to the alternatives.
7. Engagement in the MPL Space: Investors have been engaged in the market through a variety of activities including, but not limited to:
    a. Investing in marketplace lending platform equity
    b. Facilitating funding for the platforms via a provision of credit facilities
    c. Investing in MPL securitizations
    d. Directly lending to borrowers.
8. Proliferation among MPL Asset Classes: In addition, there has been a proliferation of marketplace lenders across a multitude of asset classes including consumer, purchase finance, education finance, real estate, merchant cash advance, and small businesses.
9. Asset Pricing and Regulation Obligation: As a consequence, institutional investors, diversified financial services firms, and funding providers require an independent 3rd party to help them price their holdings or satisfy other regulatory obligations.

10. MPL Valuation Standards Methodology Enhancement: Finally it is imperative that 3rd parties commit to improving and enhancing their standards and methodologies to serve what is a rapidly growing and evolving market. The 2.0 Credit Model is an attempt at enhancements over several key areas foe projections of cash flows on historical data above 1.0.
11. Additional Loan Specific Risk Factors: 2.0 addresses the need to incorporate additional loan specific measures of risk. Version 1.0 segments the loan by 6 factors:
    a. The originator
    b. Loan credit quality grade as provided by the originator
    c. Origination vintage
    d. Loan term
    e. Loan status
    f. Loan age
12. Incorporating Macro-economic Driving Factors: 2.0 provides a structure for incorporating macro-economic factors that drive the estimates of default and prepayment.
13. Reducing Dependence on Recent Issuance: Reducing the reliance on the most recently issued set of loans to drive expectation of prepay and default is another objective (1.0 applied the default and prepay experience of the most recently issued cohort corresponding to the risk factors above).
14. Usage of Statistical Learning Techniques: Version 2.0 applies advanced statistical and machine learning techniques to develop a predictive model for prepay and default. Thus Credit Model Version 2.0 aims to address and overcome the shortcomings from the Model 1.0 above, and bring rigorous techniques to an expanding asset class with a growing investor base.

Scope of the 2.0 Model

1. Loans Originated by Lending Club: For the purposes of demonstration, PeerIQ (2015) illustrate the construction and performance of Model Version 2.0 on public data from loans originated by Lending Club ("LC") with reporting months from 1 January 2010 to 1 July 2015.
2. The Data Model: PeerIQ's data model is proprietary and unified in its methodology for cleaning, enriching, and housing data across all MPL originators, and gets expanded as additional asset classes and originators are on-boarded.
3. Similarity with Lending Club Model: Although DROP does tailor the Model to specific data classes and originators, the methodology and the model structure is substantially similar to the Lending Club model.

Data Model Construction Rules

1. Amendments to Originator Generated Payments: DROP has made specific amendments (or transformations) on the raw originator-generated payments and balances for loans.
2. Consistency and Accuracy across Cohorts: These changes have been made in a rule-based fashion based on conversations with marketplace lenders to ensure the calculation of the cohort payments and balances in a consistent and accurate manner.
3. Reconciliation between Borrowers and Originators: Many of the above rules help reconcile between a borrower snapshot file (the borrower file) and a cumulative payments file (the payments file) published by the originators.

Loan Data Quality Rules

1. Inaccurate Originator Loan Level Record: Some originators publish inaccurate records for loans that have previously charged or paid off. These records are excluded from the cleaned data set and calculations.
2. Identification and Removal of Duplicates: If the loan has more than one record with the same originator loan ID, loan month, month on book, and outstanding principal BOP balance, it is assumed that the subsequent records are duplicates and that they must be removed.
3. Combine Payments for a Given Month: If the loan as multiple payments for a given month, then these must be combined to form a single payment. The formula applied is: combine all rows where count of loan month, originator loan ID > 1.
4. Entry for Maximum Loan Month: Each loan that has a non-zero EOP balance for a month prior to the maximum loan month should have a record for the maximum loan month. For example, if the maximum loan month on file is February 2015, and the loan is current in January 2015 but does not have a February 2015 record, a record will need to be created.
5. Loan Month Issue Date Consistency: Each loan ID should have a record where the loan month equals the issue date. Further, a record must exist for each loan between the issue date and the current file date, charge off date, or fully paid date, whichever is earlier.
6. Loan Age/Days Past Due: Days Past Due value should never be negative. The expected loan age is calculated as the loan month minus the issue date.
7. Charge Off/Fully Paid Fields:
    a. All instances where the

$$CO\ Amount\ != 0$$

    should have the charge off flag set to 1.
    b. All loans that have a status of fully paid should have an end-of-period balance > 0.
    c. BOP principal minus principal received and charge-off amount should equal to 0.
8. Principal, Interest, and Fee Payments:
    a. Interest paid for the given loan on a given month should equal between the borrower and the originator
    b. Amount paid should equal the sum of principal, interest, and fees paid

c. All principal payments, interest payments, and fee payments should be positive.

9. Field Unchanged through the Loan Life: The following fields should remain unchanged and populated through the loan life: loan purpose, loan interest rate, loan grade, loan term, loan state, original principal, issuance date.
10. Consistency of the Recovery Fields: All recoveries should be positive, and should be recorded at the month the loan charges off.

Lending Club Loan Level Data

1. Lending Club Loan Types Considered: As a starting point for the demonstration of the modeling approach, DROP uses the loan level public data from Lending Club. As such, the loan products considered are fixed rate, fixed term, fully amortizing 36 month and 60 month loans issued by Lending Club.
2. Number and Size of Loans: In all, PeerIQ (2015) uses over 9 million loan months of Lending Club data in constructing the model. The table below contains high level descriptive statistics for select items from the dataset.
3. Descriptive Statistics for LC Data: Source: PeerIQ Research

Field	Mean	Standard Deviation	Minimum	Maximum	First Quartile	Median	Third Quartile
Age (Months on Balance)	10.1	8.9	0.0	60.0	3.0	8.0	15.0
Vintage	February 2013	NA	February 2007	February 2015	April 2012	March 2013	February 2014
Original	$14,254	$8,254	$500	$35,000	$8,000	$12,000	$20,000

Principal							
Monthly Gross Income	$6,066	$4,599	$250	$725,549	$3,750	$5,167	$7,333
Term (Months)	42.7	10.8	36.0	60.0	36.0	36.0	60.0
Coupon	13.7%	4.3%	5.3%	29.0%	10.6%	13.5%	16.3%
FICO Origination	699	31	612	847	677	692	717
DTI (ex-mortgage)	16.6%	7.8%	0.0%	39.0%	11.0%	16.0%	22.0%
Total Borrower Accounts	25	11	1	162	16	23	31
Revolving Utilization Rate	57%	24%	0%	892%	40%	58%	75%
Inquiries in Last 6 Months	0.9	1.2	0.0	33.0	0.0	0.0	1.0
DQ Accounts in Last 2 Years	0.3	0.8	0.0	39.0	0.0	0.0	0.0
Months since Last	34	22	0	188	16	31	50

DQ							
Months since Last Public Record	76	29	0	129	55	79	102
Total Open Credit Lines	11	5	0	90	8	10	14

4. Period of LC Loan Origination: Overall, the sample used contains 245,243 distinct loans originated between February 2007 and February 2015. Average loan size is a little over $14,000, varying between $500 and $35,000 for Lending Club.
5. Variation Among the Underwriting Parameters: There is also considerable variation among other under-writing information, including DTI and revolving debt utilization rates, for example.
6. Defaults/Prepayments by Origination Year: Ultimately the goal is derive insight into termination events (defaults and prepays) from loan pools, and the table below lists some simple summaries of defaults and prepayments by origination year.
7. LC Prepay and Default Exits: Source: PeerIQ Research.

Origination Year	**Origination Volume ($mm)**	**Prepays ($mm)**	**Cumulative Defaults ($mm)**	**Cumulative Defaults to Date (%)**
2010	132	26	12	8.82%
2011	262	62	27	10.49%
2012	718	164	75	10.39%
2013	1,982	405	135	6.83%

2014	3,504	387	88	2.51%

8. Peaking of Defaults by Vintage: The table above shows that the defaults in the current cycle have peaked for the 2011 and the 2012 vintages.

Loan Credit Model Implementation

1. Discrete Loan Level Status Codes: The approach to modeling starts with the observation that at a given point in time, a loan has several status codes, or 'state's: current, delinquent, fully prepaid, or charged off.
2. Explicit Transition Between Loan States: The key conceptual idea is to model transitions between the various states explicitly. For example, a loan may move from a state of 'current' to '30-day delinquent', or move further down the delinquency queue (e.g., move from 30 day to 60 day delinquent). Alternatively, a loan may 'cure' and move from a status of current to delinquent.
3. Exit From the Transition Graph: The only exits for a loan from this network are a transition to prepay or default, which are of course the end statuses we are most interested in. The graph table below illustrates the transition network.
4. Directed State Transition Network Graph: Source: PeerIQ Research.

State	**C**	**P**	**D3**	**D6**	**D6+**	**D**
C	Y	Y	Y	N	N	Y
P	N	Y	N	N	N	N
D3	Y	Y	Y	Y	N	N

D6	N	N	Y	Y	Y	Y
D6+	N	N	N	Y	Y	Y
D	N	N	N	N	N	Y

5. State Transition Network Graph Annotation: Y's indicate directionality of the possible transitions. Note that prepay and default here are "absorbing" states, i.e., states from which exits are not possible. For modeling purposes, and due to data considerations, the 90-180 day delinquent states have been combined into the D6+ category.
   a. C => Current
   b. P => Prepay
   c. D3 => 30 day Delinquent
   d. D6 => 60 day Delinquent
   e. D6+ => 90 and more days Delinquent
   f. D => Default
6. Cross-State Transition Probabilities: Mathematically the probabilities of these moves can be represented in a transition matrix thereby describing the propensity of the borrower to move from one state to another.
7. State -> State Transition Setup:

$$P_{i,j}(t) = \begin{bmatrix} P_{C,C}(t) & \cdots & P_{C,D}(t) \\ \vdots & \ddots & \vdots \\ P_{6+,C}(t) & \cdots & P_{6+,D}(t) \end{bmatrix}$$

Each entry in the matrix represents the probability of the borrower moving from the row state to the column state in a particular month. For example, $P_{C,C}(t)$ represents the probability of the borrower moving from current to current (that is, staying current) over month t, and so on.

8. Transition Probabilities from MNL Frameworks: The next step is to decide on the parameterization of the entries in the transition matrix, i.e., how does one model the transition probabilities $P_{i,j}(t)$ (commonly referred to as the 'roll rate's?).
9. Transition Matrix in Markov Process: The MNL framework has features prominently in the mortgage modeling literature in modeling prepay and default. Such a modeling setup is known as the *Markov* process, which is practice means that each monthly observation of a loan's transitions is independent of any prior observations.
10. MNL Formulation in Logit Framework: Under the logit assumption inherent in MNL, the transition probabilities assume the following form:

$$P_{i,j}(t) = \frac{e^{\overrightarrow{\beta_0}+\overrightarrow{\beta_1}\cdot\overrightarrow{X_{\iota,J}}(t)}}{1+e^{\overrightarrow{\beta_0}+\overrightarrow{\beta_1}\cdot\overrightarrow{X_{\iota,J}}(t)}}$$

11. Loan Variables as Regressor Factors: Here $\overrightarrow{X_{\iota,J}}$ represents the matrix of regressors (or predictors) observed at a particular time (such as consumer credit variables, loan age, cohort, for example, and macro-economic variables such as inherent rates). Note that we are not restricted to using the same set of predictors for every transition.
12. Normalization Under the MNL Framework: The parametrization above implies that the probabilities $P_{i,j}(t)$ will lie between zero and one, as they should. Further, since the probabilities across a given row should sum up to one, there will always be a status s for which the probabilities are determined as

$$P_{i,s}(t) = 1 - \sum_j P_{i,j}(t) = \frac{1}{1+\sum_k e^{\overrightarrow{\beta_0}+\overrightarrow{\beta_1}\cdot\overrightarrow{X_{\iota,k}}(t)}}$$

13. Reduction of the Transition Probabilities: In estimating the model, one does not generally estimate all the probabilities from one state to another, as this would make the models unnecessarily complex, especially if the transitions are rare.

14. Current to Charge-Off Transition: While there are various reasons a loan can theoretically transition from current to charge-off (skipping intermediate statuses such as delinquency), such as due to the death of the borrower, these tend to be rare empirically.
15. Current-to-Default Transition Likelihood: This, it is quite intuitive that the probability of going from a state of current to a state of charge-off in a month should be quiet low. As shown in PeerIQ (2015) this monthly transition rate can be seen to be at most 0.03% for 60 month loans. Therefore this probability is not estimated.
16. Estimation of the Sparse Matrix: Thus, a sparse matrix, which is a subset of all the possible transitions – is estimated. The grid below gives the transitions that are estimated in the model, and those that are not.
17. Sparse Transition Matrix and Determinants: Source: PeerIQ Research.

States	**C**	**D3**	**D6**	**D6+**	**D**	**P**
C	Y	Y	N	N	N	Y
D3	Y	Y	Y	N	Y	Y
D6	N	Y	Y	Y	N	N
D6+	N	N	Y	Y	Y	N

Clearly the number of transitions to be estimated will vary depending on the originator and the observed frequency.

18. Transition Probabilities Using Separate Models: Since there is no restriction to using the same set of regressors for each probability, it is instructive to think of each of the entries estimated in the sparse matrix as separate models.

DROP Credit Model Selection Methodology

1. Cross Validation Based Model Selection: DROP applies a rigorous out-of-sample based testing procedure to estimate each of the models discussed above. The procedure for the model selection is described as follows.
2. Candidate Variables for Model Selection: DROP has identified a set of candidate variables for selection. Those variables and their data types are listed in the table below.
3. Candidate Variables and their Types: Source: PeerIQ Research.

Variable	Implementation Type
Age (Months Remaining)	Evenly Spaced Linear Splines
Term (36 or 60 months)	Categorical Variable
Interaction Variable (Age Spline * Term)	Linear Spline * Categorical Variable
DTI	Continuous Variable
FICO at Origination	Categorical Variable
Vintage (Origination Year)	Categorical Variable
Seasonality (Month of Year)	Categorical Variable
Original Loan Size Bucket	Categorical Variable
Coupon Stack Bucket	Categorical Variable
Loan Purpose	Categorical Variable
Employment Length	Categorical Variable
Inquiries in the Past 6 Months	Untransformed
Monthly Gross Income	Untransformed
Total Outstanding Accounts	Untransformed

Revolving Credit Utilization	Untransformed
Delinquent Accounts in last 2 Years	Untransformed
Total Open Credit Lines	Untransformed

4. Model Variables from Candidate Regressors: Altogether the combinations of the categorical and the numerical variables generate 86 candidate regressors from which the best choice for each transition probability is optimized. The details of the procedure for model variable selection follow.
5. Likelihood Estimation for Traditional MNL: Variable selection was achieved via a regularized logistic regression procedure. In traditional logistic regression one typically optimizes for the values of $\overrightarrow{\beta_0}$ and $\overrightarrow{\beta_1}$ in the equation above. Applying maximum likelihood across borrowers, the formulation for the transition probability between states (for e.g., borrowers moving from current to prepay) follows, as shown below.
6. Example Current -> Prepay MLE Setup:

$$Max\left[L\left(\overrightarrow{\beta_0}, \overrightarrow{\beta_1}\right)\right] = \sum_{m=1}^{N} \log P_{c,p,m}$$

where $P_{c,p,m}$ represents the probability of the m^{th} borrower moving from current to prepay (dropping time subscripts for convenience) with the probabilities given from

$$P_{i,j}(t) = \frac{e^{\overrightarrow{\beta_0} + \overrightarrow{\beta_1} \cdot \overrightarrow{X_{i,j}}(t)}}{1 + e^{\overrightarrow{\beta_0} + \overrightarrow{\beta_1} \cdot \overrightarrow{X_{i,j}}(t)}}$$

7. Penalty Based Regularized Logistic Regression: Regularized logistic regression is similar to the equation above, but it imposes a penalty on $\overrightarrow{\beta_0}$ and $\overrightarrow{\beta_1}$ to ensure that the coefficients that

are not predictive of the transition probability are not unnecessarily added (and thus eliminated from the regression).

8. Current Prepay Penalized MLE Setup:

$$Max\left[L\left(\overrightarrow{\beta_0}, \overrightarrow{\beta_1}\right)\right] = \sum_{m=1}^{N} \log P_{c,p,m} + \alpha \sum_{j=1}^{p} \left|\beta_j\right|$$

where β_j represents the set of predictors in $\overrightarrow{\beta_1}$, α represents the parameter that determines the strength of the penalty function.

9. Optimal Penalty Loading Selection Algorithm: The penalty loading parameter α is selected using the following step:
    a. Select a possible range of values for α
    b. For each value of α in the range above, train the logistic regression model on a training dataset, and test the predictive power of the model on a separate test dataset
    c. The test data is defined by sampling every third observation to minimize the risk of over-fitting the data
    d. Choose the value of α which results in the 'best' performance on the testing dataset
10. Classifier Performance Using Sample AUC: The AUC is a well-known measure of predicting accuracy of a classification technique (Receiver Operating Characteristic (Wiki)). Thus, the 'best' in the algorithm above is decided by looking at that value of α that gives the best out-of-sample AUC, e.g., on classifying borrowers who move from current to prepay in a given month.

Empirical Results – Regressor Contribution Weights

1. C -> P Predictor Loading Strength Order:

Strong Positive	**Strong Inverse**	**Weak/Neutral**
a. Age_k0.0	a. Revolving_utilization	a. Coupon_20+
b. Total_accounts	b. Total_open_credit_lines	b. Cohort_2014
c. Inq_6m	c. Age_k0.0_term_60	c. Cohort_2013
	d. Age_k18.0	d. Fico_750_800
	e. Dti_ex_mortgage	e. Lp_debt_consolidation
		f. Coupon_15-20
		g. Cohort_2015
		h. Mo_3
		i. Mo_5
		j. Mo_2
		k. Mo_4
		l. Size_15_20
		m. Mo_10
		n. El_10+years
		o. El_<1year
		p. Mo_8
		q. Mo_9
		r. Mo_11
		s. Size_gte20
		t. Fico_650_700
		u. Lp_small_business
		v. Coupon_<10
		w. Cohort_2011
		x. Cohort_2009
		y. Cohort_2008
		z. Cohort_2010
		aa. El_n/a

		bb. Mo_12 cc. Term_60 dd. Age_k12.0

2. <u>C -> D3 Predictor Loading Strength Order</u>:

Strong Positive	**Strong Inverse**	**Weak/Neutral**
a. Age_k0.0 b. Inq_6m	a. Age_k6.0 b. Monthly_gross_income c. Age_k12.0	a. Coupon_20+ b. Coupon_15-20 c. Dti_ex_mortgage d. Dq_accounts_past_2_years e. Lp_small_business f. Lp_educational g. Cohort_2008 h. Lp_moving i. El_n/a j. Mo_10 k. Lp_other l. Fico_650_700 m. Mo_7 n. Lp_medical o. El_<1year p. Mo_6 q. Mo_11 r. Size_5_10 s. Mo_9 t. El_10+years u. Lp_major_purchase v. Lp_debt_consolidation

		w. Mo_3 x. Size_lte5 y. Cohort_2015 z. Mo_2 aa. Mo_5 bb. Fico_750_800 cc. Lp_wedding dd. Cohort_2014 ee. Total_accounts ff. Fico_800_850 gg. Term_60 hh. Lp_home_improvement ii. Age_k24.0 jj. Cohort_2012 kk. Cohort_2013 ll. Age_k24.0term60 mm. Coupon_<10 nn. Total_open_credit_lines oo. Revolving_utilization

3. D3 -> D3 Predictor Loading Strength Order:

Strong Positive	**Strong Inverse**	**Weak/Neutral**
a. Age_k0.0	a. Cohort_2015	a. Coupon_15-20
b. Dq_accounts_past_2_years	b. Cohort_2014	b. Size_15_20
c. Cohort_2009	c. Age_k24.0term_60	c. Fico_650_700
d. Cohort_2008	d. Cohort_2013	d. El_4years
e. Cohort_2010	e. Size_lte5	e. Mo_7
f. Revolving_utilization	f. Total_open_credit_lines	f. Mo_10

g. Size_gte20	g. Age_k18.0term_60	g. Mo_8
h. Coupon_20+	h. Cohort_2012	h. Fico_750_800
	i. Term_60	i. Mo_12
	j. Size_5_10	j. El_3years
	k. El_7years	k. Mo_9
		l. Mo_5
		m. Coupon_<10
		n. Mo_11
		o. Lp_credit_card
		p. El_5years
		q. El_6years
		r. El_8years

4. <u>D3 -> C Predictor Loading Strength Order</u>:

Strong Positive	**Strong Inverse**	**Weak/Neutral**
a. Monthly_gross_income	a. Total_open_credit_lines	a. Revolving_utilization
b. Dq_accounts_past_2_years	b. Cohort_2014	b. Age_k12.0term_60
	c. Cohort_2015	c. Mo_3
	d. Dti_ex_mortgage	d. Mo_4
	e. Cohort_2013	e. Lp_moving
		f. Fice_650_700
		g. Size_lte5
		h. Lp_car
		i. Cohort_2008
		j. Lp_medical
		k. Cohort_2010
		l. Size_5_10
		m. Coupon_<10

		n. Size_gte20 o. Cohort_2009 p. Lp_other q. Age_k12.0 r. Lp_debt_consolidation s. El_7years t. Mo_5 u. El_5years v. El_8years w. Size_15_20 x. El_9years y. Mo_6 z. Inq_6m aa. El_10+years bb. Mo_12 cc. Mo_7 dd. Term_60 ee. Cohort_2012 ff. Age_k30.0 gg. Mo_11 hh. El_n/a

5. D3 -> P Predictor Loading Strength Order:

Strong Positive	**Strong Inverse**	**Weak/Neutral**
a. Total_accounts b. Age_k6.0 c. Monthly_gross_income d. Fico_800_850	a. Age_k30.0 b. Age_k0.0term_60 c. Dti_ex_mortgage d. Lp_car	a. Size_gte20 b. Mo_10 c. Lp_credit_card d. Mo_6

e. Cohort_2009	e. Mo_8	e. Mo_7
f. El_10+years	f. Lp_small_business	f. Cohort_2014
g. El_8years	g. Fico_700_750	g. Cohort_2013
h. El_5years	h. Size_lte5	h. Size_15_20
i. Cohort_2015		i. Cohort_2010
j. El_6years		j. Term_60

6. <u>D3 -> D Predictor Loading Strength Order</u>:

Strong Positive	**Strong Inverse**	**Weak/Neutral**
a. Age_k0.0	a. Coupon_15_20	a. Size_5_10
b. Total_accounts	b. El_3years	b. Cohort_2012
c. Age_k6.0	c. Coupon_20+	c. Mo_5
d. Dti_ex_mortgage		d. Mo_6
e. Size_lte5		e. El_7years
f. Cohort_2013		f. Fico_700_750
g. Cohort_2014		g. Term_60
h. El_10+years		h. El_<1year
		i. Mo_12
		j. Mo_2
		k. Size_15_20
		l. Size_gte20
		m. El_4years

7. <u>D6 -> D6 Predictor Loading Strength Order</u>:

Strong Positive	**Strong Inverse**	**Weak/Neutral**
a. Age_k0.0	a. Cohort_2014	a. Cohort_2009

b. Age_k6.0 c. Revolving_utilization d. Coupon_20+ e. Total_accounts f. El_4years	b. Cohort_2015 c. Cohort_2013 d. Cohort_2012 e. Cohort_2011 f. Age_k24.0term_60 g. Age_k30.0term_60 h. Size_lte5 i. Lp_major_purchase j. Term_60	b. Coupon_15-20 c. El_<1year d. Fico_750_800 e. Mo_5 f. Size_gte20 g. Lp_home_improvement h. Mo_2 i. Lp_other j. Lp_debt_consolidation k. Mo_3 l. Mo_4 m. El_10+years n. Age_k18.0term_60 o. Fico_700_750 p. Lp_credit_card q. Mo_10 r. Mo_12 s. Mo_11 t. Size_15_20 u. El_7years v. El_9years w. Mo_7 x. El_6years y. Size_5_10 z. El_8years aa. Dti_ex_mortgage

8. <u>D6 -> D6+ Predictor Loading Strength Order</u>:

Strong Positive	**Strong Inverse**	**Weak/Neutral**
a. Age_k30.0term_60	a. Age_k30.0	a. Cohort_2012
b. Age_k24.0term_60	b. Total_accounts	b. Coupon_20+
c. Age_k18.0term_60	c. Age_k0.0	c. Mo_11
d. Total_open_credit_lines	d. Cohort_2008	d. El_9years
	e. Age_k6.0	e. Cohort_2015
		f. Size_gte20
		g. Mo_7
		h. Fico_650_700
		i. El_4years
		j. Cohort_2013
		k. Mo_8
		l. Lp_credit_card
		m. Age_k36.0term_60
		n. Mo_12
		o. Mo_2
		p. Coupon_<10
		q. Monthly_gross_income
		r. El_<1year
		s. Mo_6
		t. Fico_750_800
		u. Mo_9
		v. Term_60
		w. El_10+years
		x. Cohort_2009
		y. El_n/a
		z. Cohort_2010
		aa. Mo_4
		bb. Age_k24.0

		cc. Size_5_10 dd. Revolving_utilization ee. Size_lte5 ff. Mo_3

9. D6+ -> D6+ Predictor Loading Strength Order:

Strong Positive	Strong Inverse	Weak/Neutral
a. Age_k12.0	a. Age_k0.0	a. Mo_11
b. Cohort_2008	b. Cohort_2013	b. Mo_9
c. Cohort_2009	c. Cohort_2014	c. Revolving_utilization
	d. Cohort_2012	d. Coupon_20+
	e. Cohort_2011	e. Fico_650_700
		f. Size_gte20
		g. Mo_7
		h. Coupon_15-20
		i. Age_k18.0
		j. Mo_12
		k. Lp_debt_consolidation
		l. El_3years
		m. El_6years
		n. Mo_6
		o. Size_5_10
		p. Lp_other
		q. Mo_4
		r. Mo_10
		s. Mo_3
		t. Size_lte5
		u. Age_k0.0term_60

10. D6+ -> D Predictor Loading Strength Order:

Strong Positive	Strong Inverse	Weak/Neutral
a. Cohort_2012	a. Age_k6.0	a. Lp_car
b. Cohort_2013	b. Age_k12.0	b. Size_15_20
c. Cohort_2014	c. Dq_accounts_past_2_years	c. Dti_ex_mortgage
d. Mo_7	d. Cohort_2008	d. Mo_10
e. Term_60	e. Cohort_2009	e. El_6years
f. Age_k24.0term_60	f. Total_accounts	f. El_9years
g. Mo_12	g. Mo_2	g. El_n/a
h. Mo_6		h. Mo_5
i. Cohort_2011		i. El_5years
		j. El_7years
		k. El_3years
		l. Coupon_15_20
		m. Lp_debt_consolidation
		n. Mo_9
		o. El_10+years
		p. Mo_8
		q. Mo_3
		r. Size_lte5
		s. Coupon_<10
		t. Cohort_2010

Empirical Analysis of Seasoning Effects

1. Default and Prepay over Time: Perhaps one of the most important predictors of a loan's prepayment and default rate is just the passage of time. The empirical exhibits of the conditional prepayment C-P probabilities over age (months on balance) by PeerIQ (2015) illustrate that monthly prepayment probabilities generally increase with loan seasoning.
2. Time Impact on the Delinquency: Conversely, monthly rolls to delinquency (and ultimately to default) generally fall as a loan seasons. Conditional on the fact that a loan has not defaulted to a certain point in time, it is more generally likely to prepay and not default in the future.
3. Actual vs. Fitted C-P Probabilities: PeerIQ (2015) find that the fitted C-P probabilities correspond quiet well to the actual C-P probabilities indicating that the age splines are quiet explanatory (however less so at later ages where there is some noise due to the thinning of the sample size). Nonetheless the general uptrend of prepayment by loan age is evident in both the actual and the fitted C-P probabilities.
4. Hump in the Delinquency Curve: Conversely, rolls into delinquency peak early in the life of a loan (within the first 15 months of age), followed by a gradual decline through the life of the loan. The result is particularly true for the 60 month loans which (due to their longer term) season meaningfully towards the back end.

Analysis of the Vintage/Cohort Effects

1. Vintage/Cohort Impact on Loan: Vintage and cohort effects are designed to capture the difference in prepayment and default experience of loans identical in all respects except for the fact that they were originated at different times.
2. Impact on Prepay/Delinquency Rolls: While the monotonicity of the prepayment rates are likely due to the seasoning effects described above (for e.g., the relative lower prepayment probability of the younger 2015 vintage loans), the peak in fitted delinquency rolls around the 2012 vintage for the 60 month loans and the2013 vintage for the loan 36 month loans is both interesting and consistent with the prepay/default exist in LC data seen before.

Analysis of Empirical Seasonality Effects

1. Usage of Origination DTI Metric: Once again the DTI metrics used in the model are at origination. Again, the results are quite intuitive and fit the data well.
2. DTI Impact on Prepay/Delinquency: Borrowers with higher DTI tend to prepay slower and tend to roll to delinquency at a higher rate, albeit not with the same effect as that observed on prepayment.

CPR And CDR Curve Estimation

1. Definition of CPR and CDR: The fit probabilities (and therefore the sparse transition matrices) obtained from the above are used to derive CPR and CDR curves. CPR and CDR curves are defined as

$$CDR(t) = 1 - \left[1 - \frac{Default(t)}{BeginBalance(t)}\right]^{12}$$

and

$$CPR(t) = 1 - \left[1 - \frac{Prepay(t)}{BeginBalance(t)}\right]^{12}$$

2. CPR/CDR for Loan Pools: CPR and CDR therefore represent an annualized measure of prepay and default for every dollar of principal outstanding at the start of a period. Clearly, in order to estimate the CPR and CDR, we need to project the cash flows for a given pool of loans.

3. Projecting Each Loan in Pool: A first step to projecting the cash flows is to project the status of each loan in the pool at all future points in time. This is achieved by using the transition matrix

$$P_{i,j}(t) = \begin{bmatrix} P_{c,c}(t) & \cdots & P_{c,d}(t) \\ \vdots & \ddots & \vdots \\ P_{6+,c}(t) & \cdots & P_{6+,d}(t) \end{bmatrix}$$

4. Using MNL Loan Transition Probabilities: Probabilities are estimated at all points in time, as per the earlier section. Suppose, for example, that at a given time t the loan status is current. Aw: The multinomial distribution for where the loan can transition to at time $t + 1$ is used, and it is simply by the first row of the transition matrix.
5. Loan Target Status Random Draw: Therefore a random draw is made from this distribution to project the status of the loan at time $t + 1$. If, in that draw, the loan ended up in a status of 30-day delinquent, a simulation is done using the second row of $P_{i,j}(t + 1)$ and so on, until the earliest of one of maturity, default, or prepay.
6. Projection of Loan Cash Flows: After having obtained the future status of all the loans in the pool, we can compute the cash flows appropriate to the product under consideration (in this case the 36 month vs. the 60 month fixed rate loans).
7. Loan State Dependent Cash Flow: For example, in the case of Lending Club, one can continue to apply the monthly fixed payment on the loan if the loan is in current status, and accrue interest if the loan becomes delinquent, or discontinue further payments altogether if the loan voluntarily prepays or defaults.
8. CPR/CDR from Cash Flows: Such logic allows the computation of the loan balances from period to period, and ultimately the CPR and the CDR above.
9. CPR and CDR Simulation: As a final example, we project the loan status, derive cash flows, and compute CPR and CDR for the entire population of outstanding Lending Club loans – outstanding as of January 2013 – and examine the results. DROP generates the results of one simulation on the portfolio containing 14,000 Lending Club loans as at 1 January 2013.
10. Actual vs. Simulated CPR/CDR: As evidenced from the above, the projection for the CPR and the CDR agrees with the realized values for much of the projection period. For CDR,

there is some volatility towards the latter end of the projection period where the sample thins out, and where the number of defaults (as a rare transition) can have a meaningful impact.

11. Multi-Path CPR/CDR Simulation: In addition to such 'single path' projections of CPR and CDR on a portfolio, because the calibration produces a *distribution* of transitions, one can generate multiple paths for the CPR and CDR on a loan portfolio. Using the model DROP simulated 10 paths for the LC portfolio resulting in a distribution of CPR and CDR for the portfolio.

Credit Model Future Enhancements

1. Inclusion of Macro-economic Regressors: For the purposes of keeping the analysis simple, DROP's model excludes macro-economic regressors from the analysis. However economic variables can be integrated quiet easily into the structure of the model.
2. Stochastic Simulation of Market Variables: In future versions we plan to add in carefully selected market variables such as interest rates or unemployment, which can be used as scenario variables, or stochastically simulated via parameters calibrated from the prices of traded instruments (e.g., interest rate options).

References

- PeerIQ (2015): PeerIQ Analytics Credit Model Methodology, Version 2.0.
- Receiver Operating Characteristic (Wiki).