Risk and risk management in the credit card industry

FLORENTIN BUTARU (OCC)
QINGQING CHEN (OCC)
BRIAN CLARK (OCC)
SANMAY DAS (WASHINGTON UNIVERSITY)
ANDREW LO (MIT)
AKHTAR SIDDIQUE (OCC)

JANUARY 29, 2015

*This views expressed in this paper and presentation are those of the authors alone and do not necessarily reflect those of the U.S. Treasury or Office of the Comptroller of the Currency.

Key Takeaways

- Machine learning models (with six banks' credit card data) out-of-sample and out-of-time forecasts of credit-card-holder delinquencies and defaults outperform the traditional logistic regressions.

- Analysis and comparison of risk management practices across different banks along with comparison of the drivers of delinquency at different institutions show:
  - Substantial heterogeneity in risk factors, sensitivities, and predictability of delinquency across banks
  - In particular, macroeconomic variables affect different institutions differently.

Key Takeaways: 2

- One model fit, in terms of risk factors or sensitivities to risk factors, would not work very well across all institutions.

- Predictability of delinquency and the quality (i.e. effectiveness) of risk management varies across institutions

  - Leveraging information from one bank may be useful for analyzing the risk and performance for another bank.

- Policy implications that bank specific supervision and capital requirements would be more appropriate.

Consumer Credit Modeling

- Two broad schools of modeling:
  - Traditional statistical models (logistic regression)
    - Segmentation analysis, more relevant and predictive attributes
      - Mester, 1997; Morrison, 2004;Fensterstock, 2005
    - Tend to be easier to interpret and thus useful in practice and may be more useful over long time periods
  - Machine learning and data mining method
    - Decision tree models (Davis, Edelman, & Gammerman, 1992), neural network models (Desai, Crook, & Overstreet, 1996; Malhotra & Malhotra, 2002; West, 2000), support vector machines (Huang, Chen, Wang, 2007)
    - Machine learning is most useful for large-scale complex problems, such as modeling human behavior
- Why use ML for modeling consumer credit card defaults?
  - ML works well when classification is the key model output and is likely to work over short time horizons
  - Banks can actively manage credit cards to manage their exposure
    - Line cuts
    - Freeze accounts

Concerns with Data Mining

- Machine learning (a.k.a. data mining) is widely used in many scientific fields, but is often viewed with skepticism in economics research because:
  - A lack of theory
  - Overfitting the data (in-sample)
  - Difficult to interpret the results

- We address the above issues by:
  - Testing out of sample
  - Sensitivity analysis
  - Analyzing the decision trees

Data Sources

- Credit Card Data:
  - Detailed account-level characteristics (days past due, balance, utilization, FICO, behavioral score, etc.)
    - Cannot link the accounts across individuals
  - Collected by the OCC
  - Entire credit card portfolios of nine (six) large U.S. banks
  - Monthly data starting January 2008
- Attribute Data:
  - Detailed borrower characteristics from credit bureau (linked by account)
  - Quarterly staring in 2009 and ending 2013.
- Macroeconomic Variables:
  - Collected from various sources (linked by ZIP code associated with the account)
  - Employment data, HPI, average wages, average hours, etc.
- In total, over 20TB of raw data
  - Takes weeks to process

Sample Selection

- Start with the full dataset from all six banks
- Create a panel dataset that tracks accounts over time
- Simple random sample the data
  - Retain every n^{th} account where $(100/n) = s\%$ sample size
  - Sample before merging attribute and bureau data (merge has roughly a 70% success rate)
  - Each month, we retain s% of the new accounts so our sample size mimics the bank's portfolio size
- Since we are forecasting default, we drop all accounts currently in default (e.g., 90DPD+, chargeoffs, etc.)
- Retain everything else

Variable Selection

- Roughly based on Glennon, et al. (2009)
  - Raw data with 186 raw data items (106 account-level items and 80 credit-bureau items)
  - Including account level variables, credit bureau variables, and macroeconomic variables
  - Compile into 87 attributes/features
  - These are the results we will focus on today

- Variable selection is carried out in two ways:
  1. Linear best first selection (forward, backward, and bi-directional)
     1. Very few attributes are retained regardless of the method
  2. Decision tree analysis from the full models and retain the top 20 variables from each bank (more on this later)

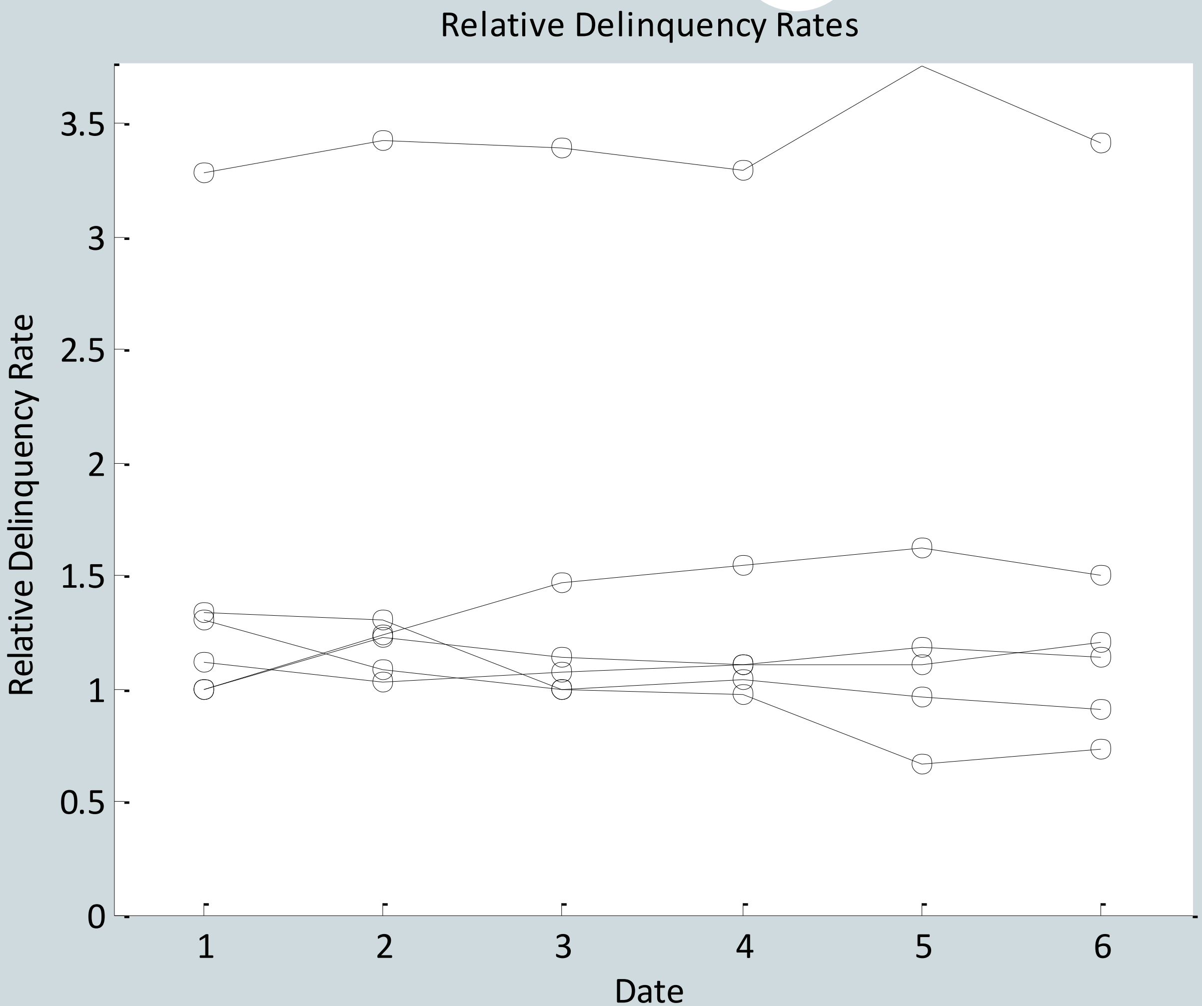
Variables: Tradelines and attributes

Account Level Features:	Credit Bureau Features:	Macroeconomic Features:
Cycle end balance	Flag if greater than 0 accounts 90 days past due	Unemployment rate
Refreshed credit score	Flag if greater than 0 accounts 60 days past due	Unemployment rate (3 mo. chg.)
Behavioral score	Flag if greater than 0 accounts 30 days past due	Unemployment rate (12 mo. chg.)
Current credit limit	Flag if greater than 0 bank cards 60 days past due	Number of total nonfarm (NSA)
Line frozen flag (0,1)	Flag if greater than 0 retail cards 60 days past due	Number of total nonfarm (NSA) (3 mo. chg.)
Line decrease in current mo. flag (0,1)	Flag if total limit on all bank cards greater than zero	Number of total nonfarm (NSA) (12 mo. chg.)
Line increase in current mo. flag (0,1)	Flag if total limit on all retail cards greater than zero	Total private (NSA) (3 mo. chg.)
Actual payment / minimum payment	Flag if greater than 0 accounts opened in the past year	Total private (NSA) (12 mo. chg.)
Days past due	Total number of accounts	Avg. weekly hours worked (private) (3 mo. chg.)
Purchase volume / credit limit	Total balance on all accounts / total limit	Avg. weekly hours worked (private) (12 mo. chg.)
Cash advance volume / credit limit	Total non-mortgage balance / total limit	Avg. hourly wage (private) (3 mo. chg.)
Balance transfer volume / credit limit	Total number of accounts 60+ days past due	Avg. hourly wage (private) (12 mo. chg.)
Flag is the card is securitized	Total number of bank card accounts	Avg. weekly hours worked (trade and transportation) (3 mo. c
chg. in securitization status (1 mo.)	Utilizatiion of all bank card accounts	Avg. weekly hours worked (trade and transportation) (12 mo.
Percent chg. in credit limit (lagged 1 mo.)	Number of accounts 30+ days past due	Avg. hourly wage (trade and transportation) (3 mo. chg.)
Percent chg. in credit limit current 1 mo.)	Number of accounts 60+ days past due	Avg. hourly wage (trade and transportation) (12 mo. chg.)
Total fees	Number of accounts 90+ days past due	Avg. weekly hours worked (leisure) (3 mo. chg.)
Workout program flag	Number of accounts under wage garnishment	Avg. weekly hours worked (leisure) (12 mo. chg.)
Line frozen flag (1 mo. lag)	Number of accounts in collection	Avg. hourly wage (leisure) (3 mo. chg.)
Line frozen flag (current mo.)	Number of accounts in charge off status	Avg. hourly wage (leisure) (12 mo. chg.)
Product type	Total balance on all 60+ days past due accounts	House price index
3 mo. chg. in credit score	Total number of acocunts	House price index (3 mo. chg.)
6 mo. chg. in credit score	Total credit limit to number of open bank cards	House price index (12 mo. chg.)
3 mo. chg. in behavioral score	Total credit limit to number of open retail accounts	
6 mo. chg. in behavioral score	Total number of accounts opened in the past year	
mo.ly utilization	Total balance of all revolving accounts / total balance on all accounts	
1 mo. chg. in mo.ly utilization	Flag if total balance over limit on all open bank cards = 0%	
3 mo. chg. in mo.ly utilization	Flag if total balance over limit on all open bank cards = 100%	
6 mo. chg. in mo.ly utilization	Flag if total balance over limit on all open bank cards > 100%	
Cycle utilization		
1 mo. chg. in cycle utilization		
3 mo. chg. in cycle utilization		
Account exceeded the limit in past 3 mo.s (0,1)		
Payment equal minimum payment in past 3 mo.s (0,1)		
6 mo. chg. in cycle utilization		

Sample Description

- Six large U.S. banks [From the original nine]
  - 2.5% sample for large banks and upto 40% of the smaller banks.
- Yields between ~90K and ~1MM observations each period per bank
- Portfolio size varies over time (mostly grow, some decline) and the sample size is representative of the true portfolio
- Substantial heterogeneity in the time series and cross sectional distribution of delinquency rates

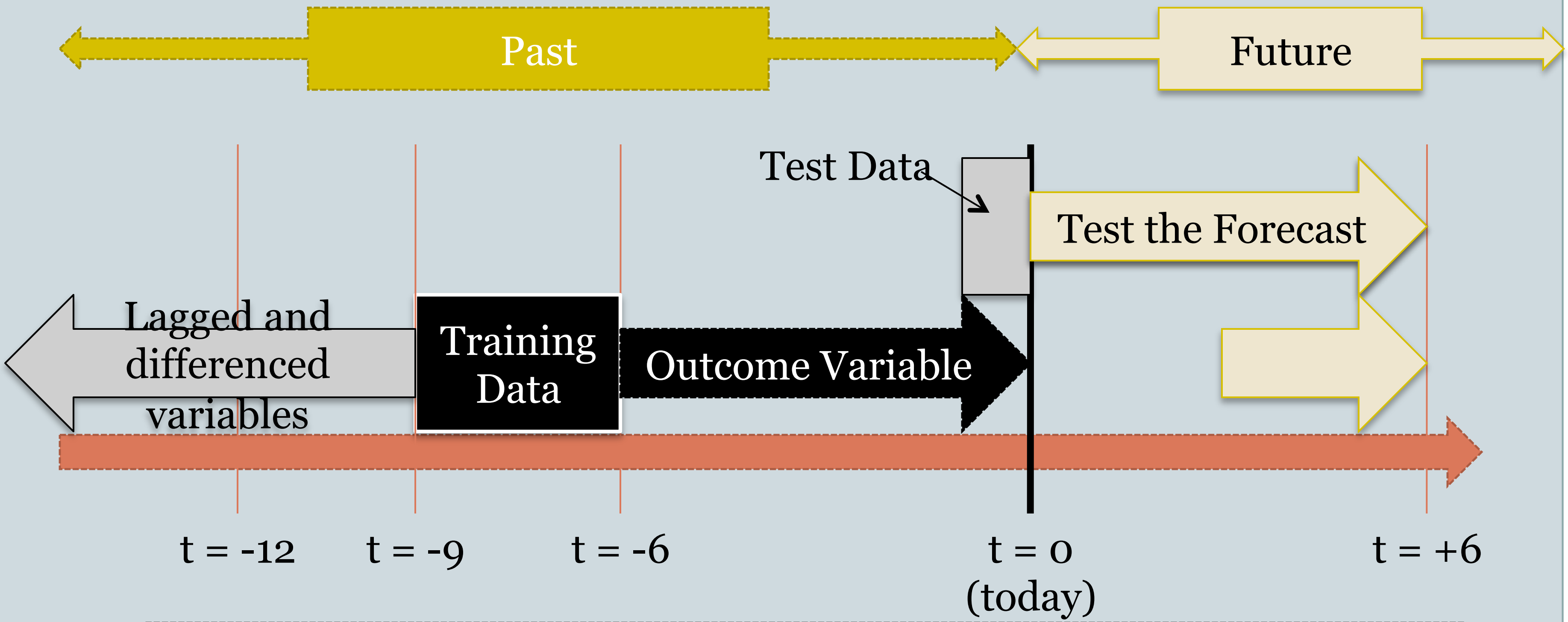
Relative Delinquency Rates



Loss Forecasting- Empirical Designs

- All models trained on historical data so all forecasts are out of sample and out of time.
- Loan –level default models:
  - Goal is to classify individual accounts and study the risk management practices of individual banks- for example:
    - Credit line increases/decreases
- Default is defined as 90 days past due (90DPD)
- Compare performance of our models using a value added analysis (Khandani, Kim, & Lo (2011))
- Analyze the attributes (e.g., decision tree output)

Timing of the Models



All models are tested out of sample!

Modeling techniques

Three approaches

- Decision Trees
  - C4.5 algorithm [J48 classifier in Weka]
- Random Forests
  - Bootstrapping a decision tree with replacement
- Regularized Logistic Regression with a quadratic penalty function

Evaluating the Models: F-Measure Kappa statistic

- Precision = the proportion of positives identified by a technique that are truly positive.
- Recall = the proportion of positives that is correctly identified.
- F-Measure = the harmonic mean of precision and recall, and is meant to describe the balance between precision and recall.
- kappa statistic = performance relative to random classification.

Performance Metrics: More Precisely

		Model Prediction	
		Good	Bad
Actual Outcome	Good	True Positive (TP)	False Negative (FN)
	Bad	False Positive (FP)	True Negative (TN)

Precision = TN/(TN+FN)

Recall = TN/(TN+FP)

True Positive Rate = TP/(TP+FN)

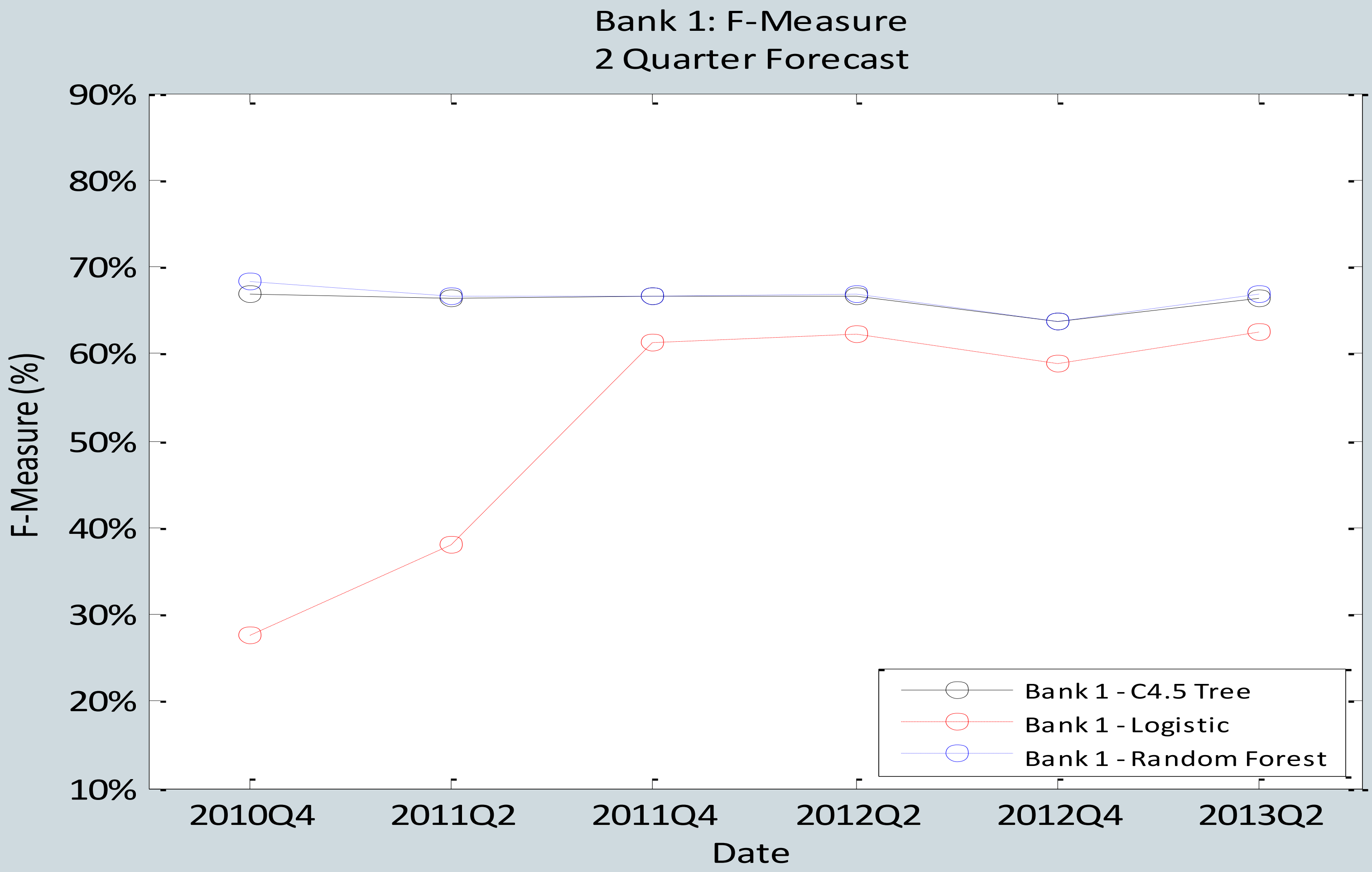
False Positive Rate = FP/(FP+TN)

F-Measure = (2*Recall*Precision)/(Recall+Precision)

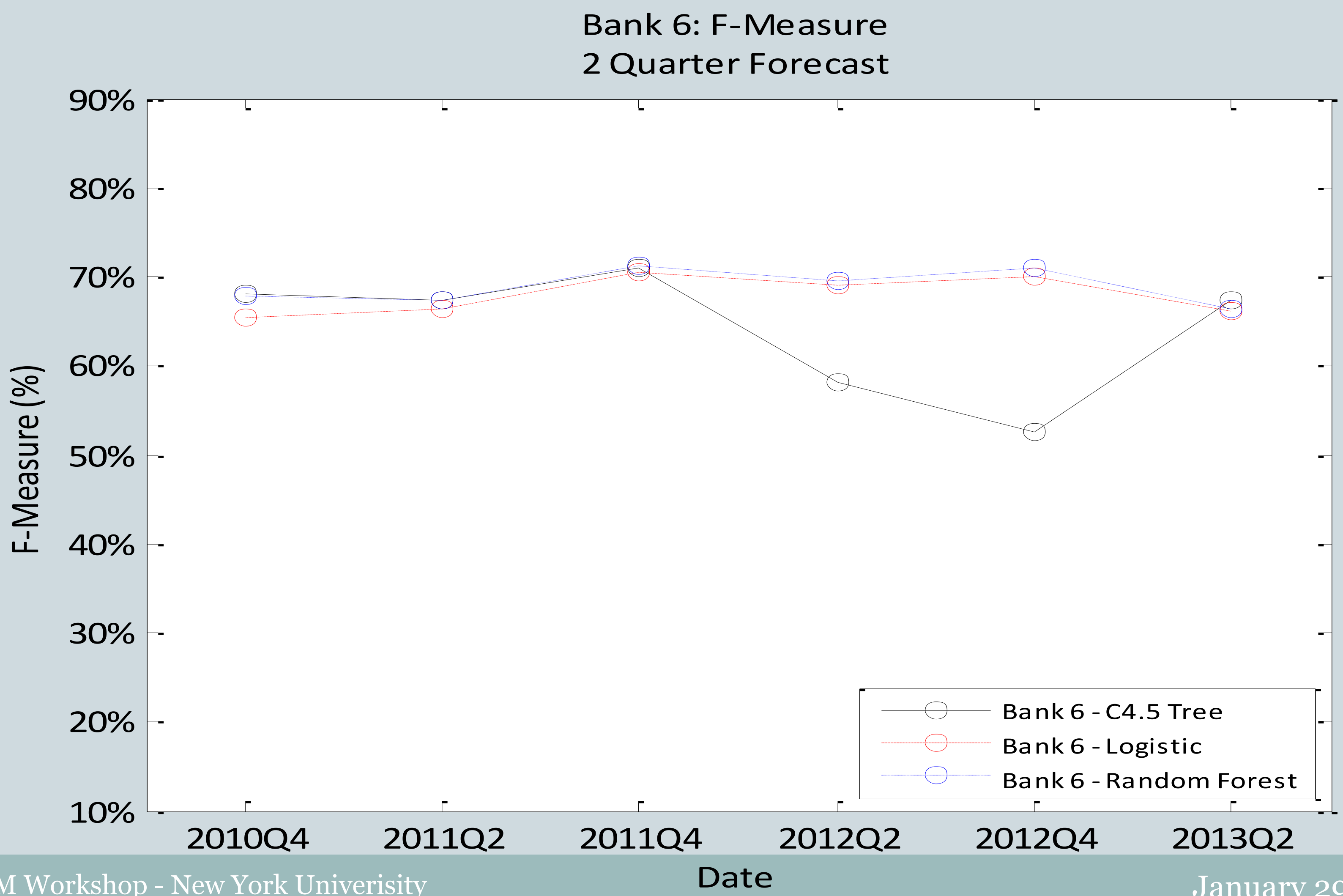
Kappa Statistic = $(P_a - P_e)/(1\text{-}P_e)$,

where $P_a = (TP+TN)/N$ and $P_e = [(TP+FN)/N]*[(TP+FN)/N]$

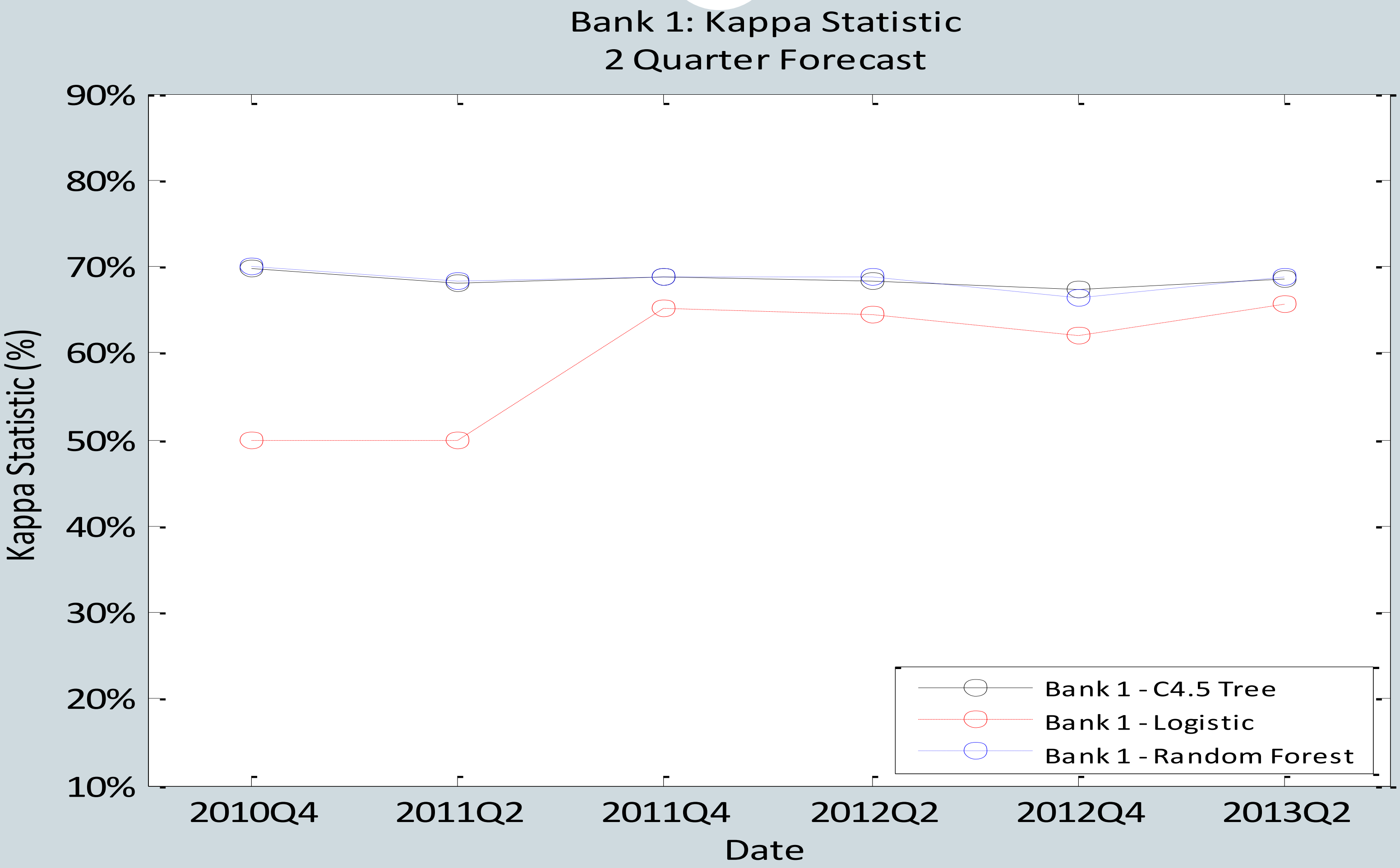
Bank 1


Bank 1: F-Measure
2 Quarter Forecast
90%
80%
70%
60%
50%
40%
30%
20%
10%
F-Measure (%)
2010Q4
2011Q2
2011Q4
2012Q2
2012Q4
2013Q2
Date
Bank 1 - C4.5 Tree
Bank 1 - Logistic
Bank 1 - Random Forest

Bank 6

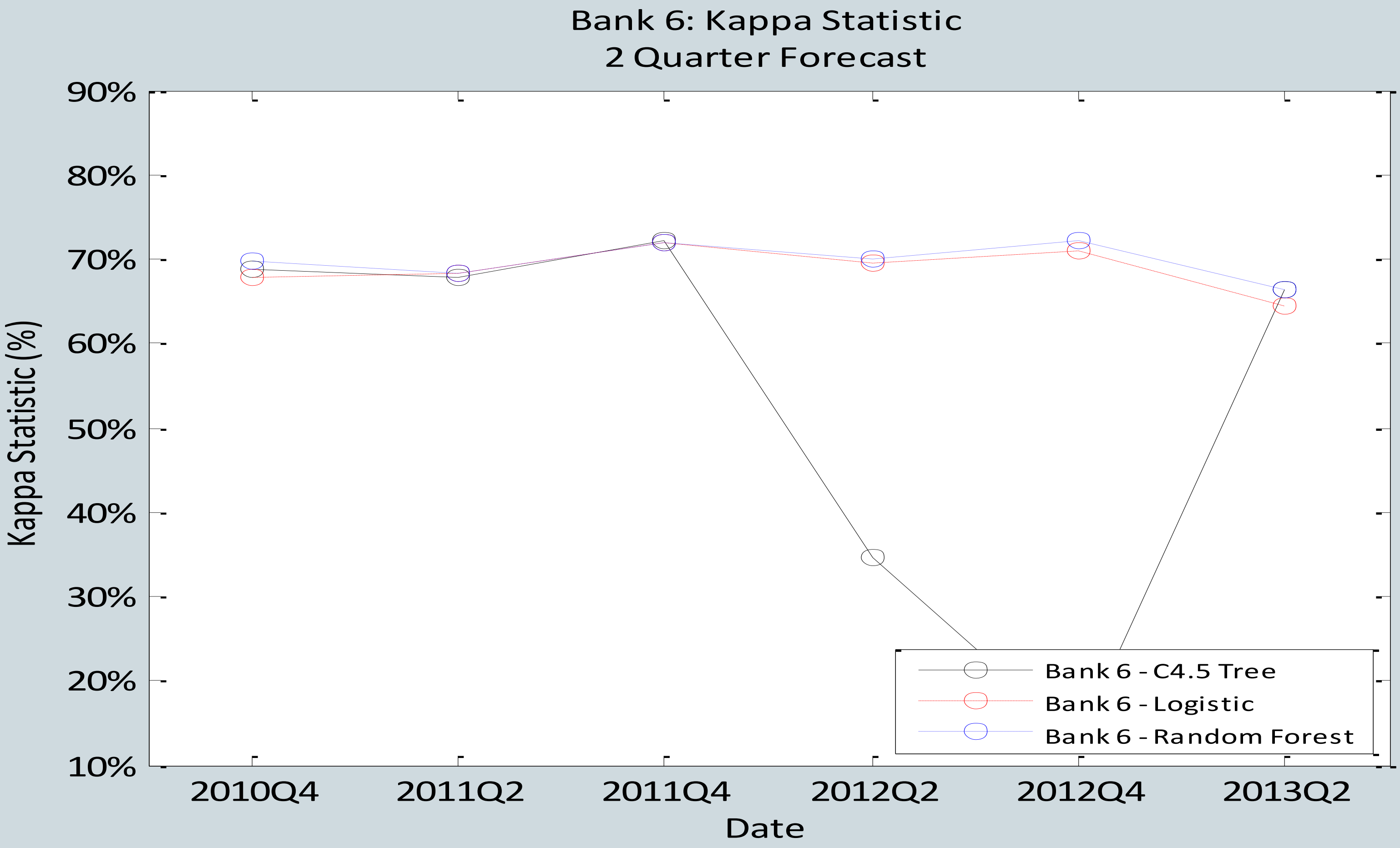

Bank 6: F-Measure
2 Quarter Forecast
90%
80%
70%
60%
50%
40%
30%
20%
10%
F-Measure (%)
2010Q4
2011Q2
2011Q4
2012Q2
2012Q4
2013Q2
Date
Bank 6 - C4.5 Tree
Bank 6 - Logistic
Bank 6 - Random Forest

Bank 1



Bank 1: Kappa Statistic
2 Quarter Forecast
90%
80%
70%
60%
50%
40%
30%
20%
10%
Kappa Statistic (%)
2010Q4
2011Q2
2011Q4
2012Q2
2012Q4
2013Q2
Date
Bank 1 - C4.5 Tree
Bank 1 - Logistic
Bank 1 - Random Forest

Bank 6


Bank 6: Kappa Statistic
2 Quarter Forecast
90%
80%
70%
60%
50%
40%
30%
20%
10%
Kappa Statistic (%)
2010Q4
2011Q2
2011Q4
2012Q2
2012Q4
2013Q2
Date
Bank 6 - C4.5 Tree
Bank 6 - Logistic
Bank 6 - Random Forest

Evaluating the Models: Value-Added Analysis (Loan-level)

- Khandani, Kim, & Lo (2011) develop a method to estimate the potential cost-savings of applying a classification model
  - Assumptions:
    - Without a forecast, the bank will not manage the account (e.g., will not cut credit limits, freeze accounts, etc.)
      - We relax this assumption
    - All customers have an observable current balance B_C at t=0
      - True in our data
    - Customers that default will increase their balance to B_D between t=0 and t=default (call this the "run-up")
      - We can calculate this for each account
    - Cost savings depends on:
      1. **Classification accuracy (i.e., quality of the model)**
      2. "Run-up" ratio
      3. Expected future cash flows (opportunity cost of misclassification)
         1. Refer to this as the "Profitability Margin"

Value-Added Analysis

		Model Prediction	
		Good	Bad
Actual Outcome	Good	TP	FN
	Bad	FP	TN

$$\Pi_{\text{no forecast}} = (TP + FN)B_C P_M - (FP + TN)B_D \qquad [1]$$

$$\Pi_{\text{forecast}} = (TP)B_C P_m - (FP)B_D - (TN)B_C \qquad [2]$$

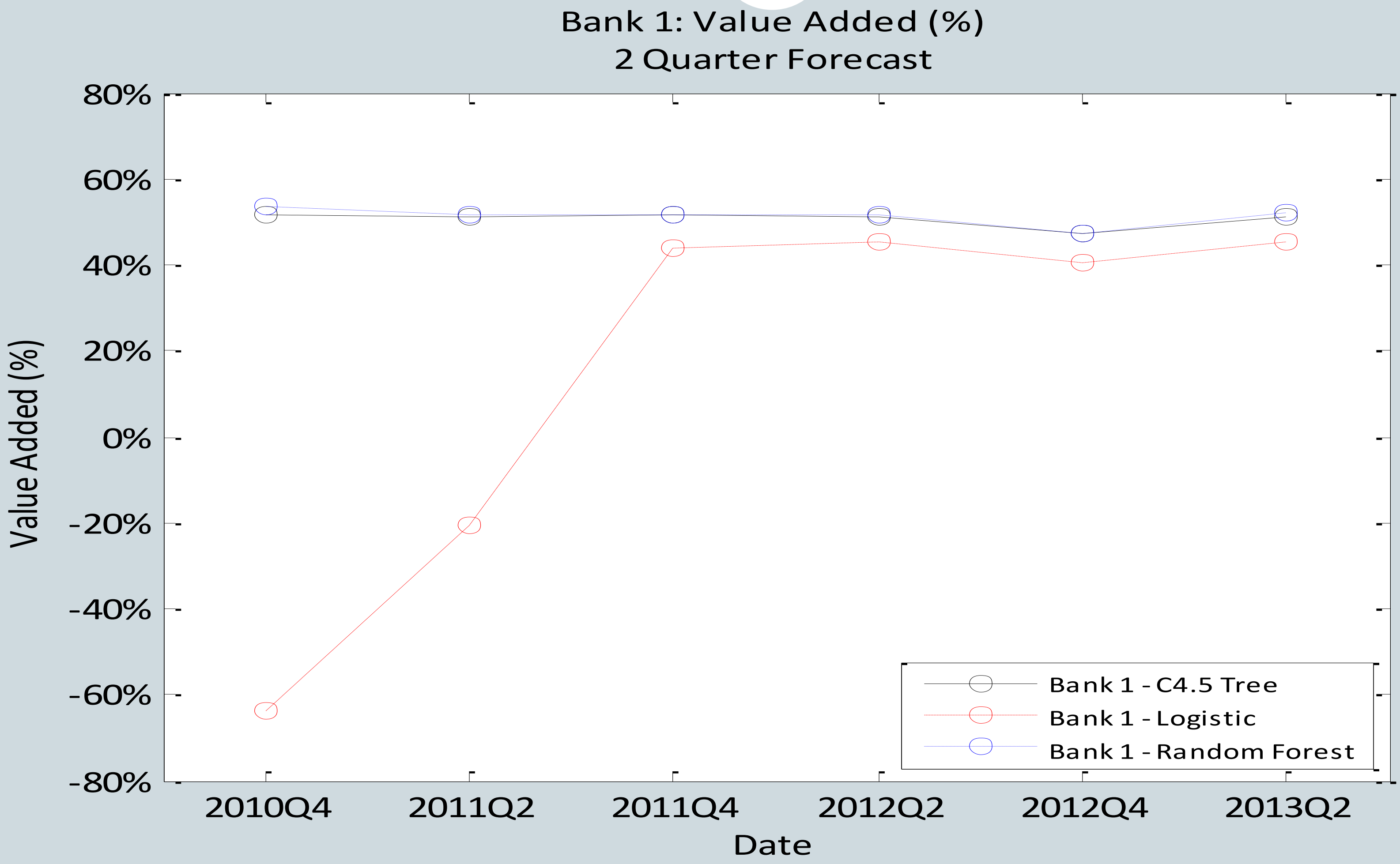
$$\Delta\Pi = TN(B_D - B_C) - (FN)B_C P_m \qquad [3]$$

$$\text{Value Added}\left(\frac{\mathrm{B_D}}{\mathrm{B_C}}, r, N\right) = \frac{TN - FN\left[1-(1+r)^{-N}\right]\left[\frac{B_D}{B_C}-1\right]^{-1}}{TN + FP}$$

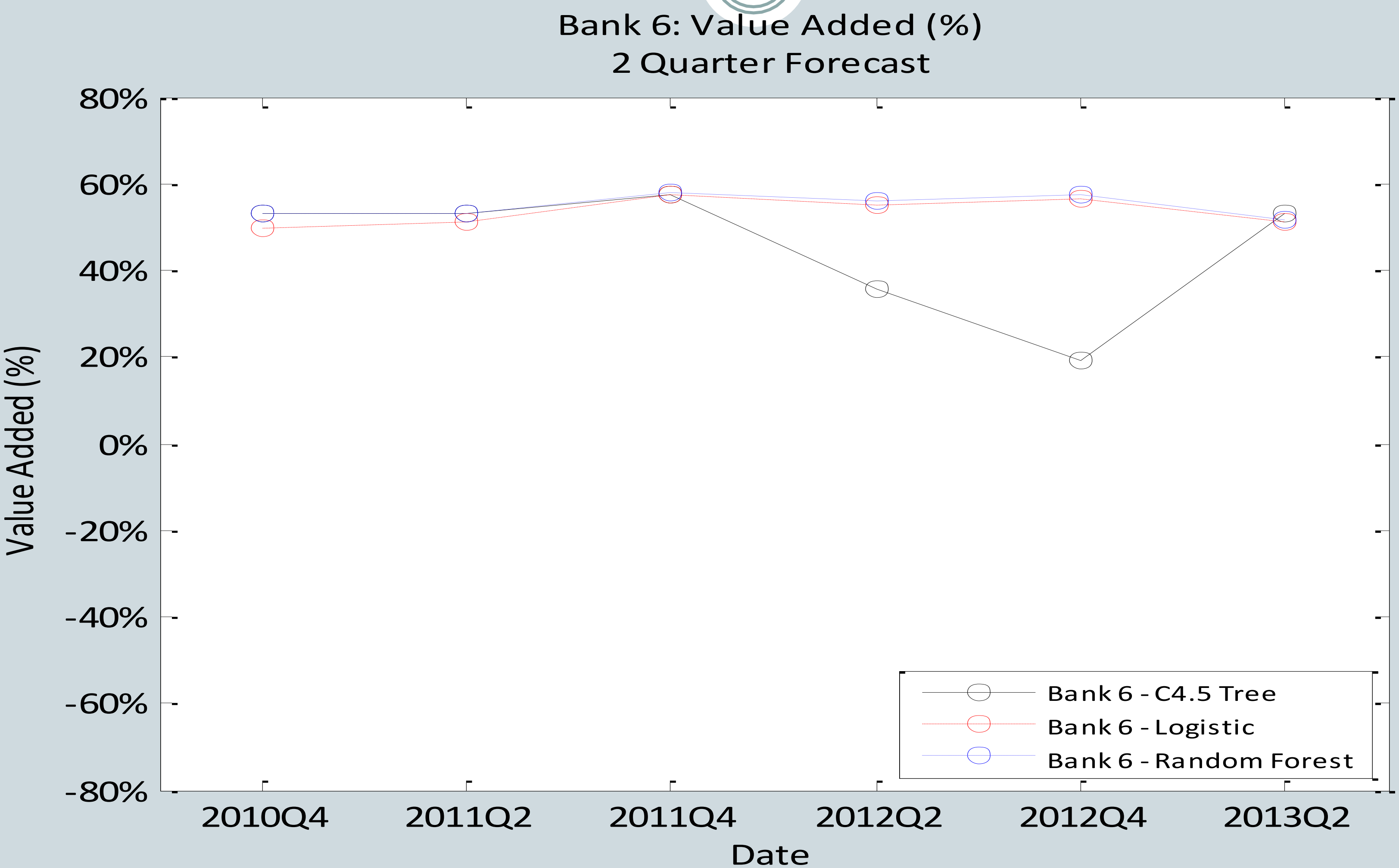
Value-Added Analysis

- We make comparison to passive risk management
- Our results can be compared relative to what the bank is actually doing (regardless of the model)

Value added : Bank 1, 2 Quarter



Value added : Bank 6, 2 Quarter



Risk management via line cuts

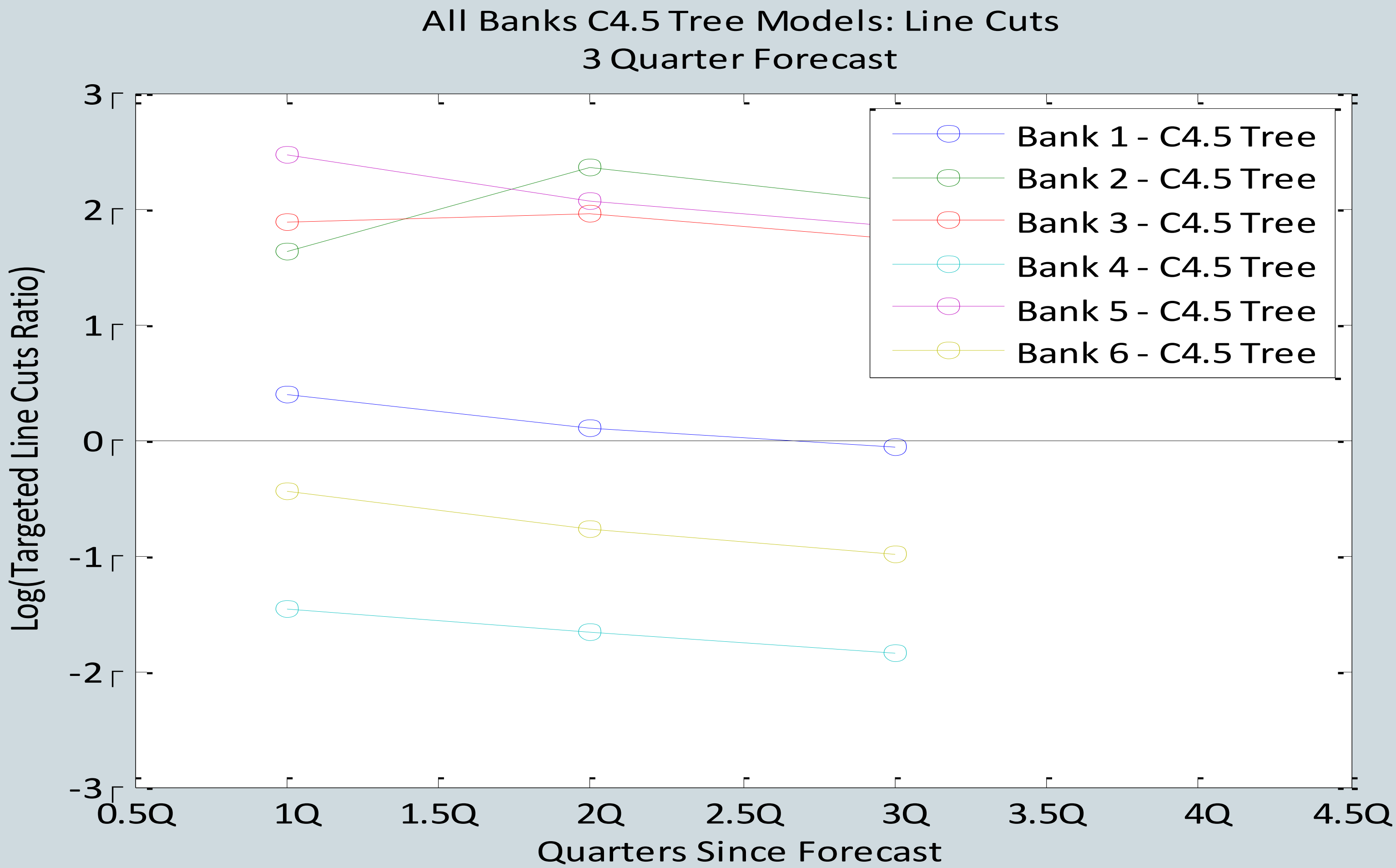
Steps

- Take the accounts which were predicted to default over a given horizon for a given bank,
- Analyze whether the bank cut its credit line or not.

Metric

- Ratio of the percent of lines cut for defaulted accounts to the percent of lines cut on all accounts.

Comparing banks efficacy in line cuts



Attribute Analysis

Understanding the relative importance of attributes.

- *Log of the number of instances classified*
- *The minimum leaf number*:
- *Indicator variable equal to 1 if the attribute appears in the tree and 0 otherwise*:

Attribute	All Models	2Q Horizon	3Q Horizon	4Q Horizon	Bank 1	Bank 2	Bank 3	Bank 4	Bank 5	Bank 6
Days past due	1.4	1.2	1.7	1.5	1.0	1.7	1.0	1.7	2.3	1.0
Behavioral score	3.7	3.2	4.3	3.7	8.3	1.3	2.3	3.0	1.7	5.7
Refreshed credit score	6.3	7.8	6.0	5.0	5.0	8.0	7.0	9.0	4.0	4.7
Actual payment / minimum payment	6.7	5.2	6.3	8.5	9.7	11.3	3.7	5.7	5.0	4.7
1 mo. chg. in monthly utilization	7.8	5.5	7.8	10.0	4.0	3.3	15.7	11.3	2.0	10.3
Payment equal minimum payment in past 3 mo.s (0,1)	8.6	7.8	8.5	9.5	6.3	8.7	6.7	10.3	6.3	13.3
Cycle end balance	9.8	10.8	10.2	8.3	11.0	6.3	12.7	7.7	17.0	4.0
Cycle end balance	11.9	8.8	16.0	11.0	3.0	13.0	13.0	14.0	12.0	16.7
Cycle utilization	12.1	19.3	8.7	8.3	8.7	21.3	4.7	22.3	9.0	6.7
Number of accounts 30+ days past due	12.6	12.8	12.7	12.2	18.7	5.0	10.3	7.3	13.0	21.0
Total fees	15.9	16.2	12.8	18.8	15.0	21.3	8.3	14.3	9.3	27.3
Workout program flag	16.8	23.5	14.2	12.7	6.7	19.3	10.3	4.0	24.0	36.3
Total number of bank card accounts	17.8	18.5	17.5	17.3	22.0	21.7	19.0	14.3	17.7	12.0
Current credit limit	17.9	18.8	18.2	16.7	21.0	7.7	30.7	16.7	10.0	21.3
Line frozen flag (current mo.)	17.9	17.5	15.7	20.5	9.7	16.3	48.7	1.3	9.0	22.3
Monthly utilization	19.9	21.5	15.3	23.0	16.7	30.0	42.3	12.0	13.7	5.0
Number of accounts 60+ days past due	23.2	22.3	27.2	20.0	21.0	19.0	20.7	18.7	19.3	40.3
3 mo. chg. in credit score	24.4	21.8	24.2	27.2	8.7	27.3	28.3	21.7	32.3	28.0
Number of accounts in charge off status	26.3	26.0	27.7	25.2	27.3	17.0	24.0	18.3	39.0	32.0
1 mo. chg. in cycle utilization	27.0	29.3	26.7	25.0	17.7	38.3	10.7	30.3	28.3	36.7
6 mo. chg. in credit score	27.1	28.8	28.3	24.2	12.7	42.3	25.0	41.3	20.3	21.0
Total number of accounts 60+ days past due	27.9	21.5	32.3	30.0	31.7	24.3	18.0	11.3	41.3	41.0
Total balance on all 60+ days past due accounts	30.2	36.5	30.5	23.7	36.3	28.0	19.7	17.7	32.3	47.3
Total number of accounts verified	30.3	32.3	28.0	30.7	46.7	18.7	42.7	31.0	24.7	18.3
Flag if greater than 0 accounts 60 days past due	30.5	36.2	27.2	28.2	39.3	42.3	16.0	36.0	34.3	15.0
Line frozen flag (1 mo. lag)	30.9	15.5	34.5	42.7	16.3	8.0	33.3	29.0	47.3	51.3
3 mo. chg. in monthly utilization	33.4	30.2	34.8	35.2	19.0	22.7	31.7	42.7	40.0	44.3
Number of accounts 90+ days past due	33.7	43.5	29.8	27.8	34.3	25.0	33.3	31.7	36.0	42.0
6 mo. chg. in behavioral score	34.6	34.5	37.2	32.2	36.0	55.7	22.0	45.3	21.7	27.0
Account exceeded the limit in past 3 mo.s (0,1)	35.3	28.5	46.0	31.3	31.0	23.0	64.7	28.3	34.0	30.7

Attribute	All Models	2Q Horizon	3Q Horizon	4Q Horizon	Bank 1	Bank 2	Bank 3	Bank 4	Bank 5	Bank 6
3 mo. chg. in cycle utilization	35.4	28.8	33.5	44.0	29.7	48.0	29.0	38.7	18.7	48.7
Flag if the card is securitized	36.2	35.5	36.7	36.3	24.0	13.7	30.3	28.7	71.7	48.7
Total number of accounts opened in the past year	36.4	41.7	36.0	31.5	41.0	24.0	38.7	45.0	28.3	41.3
Total number of bank card accounts 60+ days past due	37.4	38.5	32.8	41.0	47.3	25.0	23.7	25.3	40.7	62.7
Total balance of all revolving accounts / total balance	39.3	41.0	34.5	42.5	30.0	40.3	43.0	43.3	33.3	46.0
Total number of accounts	41.3	34.2	48.7	41.0	40.7	26.3	35.3	32.3	64.0	49.0
Product type	41.4	38.5	41.7	44.0	20.3	61.0	73.0	71.7	11.3	11.0
Unemployment rate	41.6	41.8	37.2	45.7	42.3	36.7	48.3	54.7	29.3	38.0
Flag if greater than 0 accounts 30 days past due	41.6	47.7	39.7	37.5	55.3	37.3	35.7	44.7	22.0	54.7
Purchase volume / credit limit	43.4	43.5	38.2	48.5	30.3	58.3	32.3	70.3	36.0	33.0
Utilizatiion of all bank card accounts	45.2	53.5	39.5	42.7	39.0	54.0	63.3	52.7	28.3	34.0
Flag if greater than 0 accounts opened in the past year	45.8	49.7	44.0	43.8	64.0	25.7	56.7	58.0	38.7	32.0
Flag if greater than 0 accounts 90 days past due	46.2	47.7	44.8	46.0	42.7	38.3	28.3	54.7	60.0	53.0
Avg. weekly hours worked (private) (12 mo. chg.)	46.2	44.8	49.2	44.5	61.0	37.0	55.7	42.7	52.3	28.3
Avg. hourly wage (private) (3 mo. chg.)	47.7	49.5	43.2	50.3	53.7	56.3	60.0	45.3	36.3	34.3
Avg. weekly hours worked (leisure) (12 mo. chg.)	47.9	49.7	43.0	51.0	53.3	40.0	57.0	60.7	54.3	22.0
Number of total nonfarm (NSA)	48.2	53.2	48.2	43.3	40.7	54.3	52.0	48.7	49.7	44.0
Avg. weekly hours worked (trade and transportation) (1	48.6	46.7	51.0	48.2	49.3	49.0	34.3	52.0	51.0	56.0
Avg. weekly hours worked (private) (3 mo. chg.)	49.8	48.2	44.2	57.0	48.7	46.7	53.0	42.3	50.3	57.7
Number of total nonfarm (NSA) (12 mo. chg.)	50.2	49.7	45.2	55.7	45.3	58.0	50.3	44.3	49.0	54.0
Avg. weekly hours worked (trade and transportation) (3	50.3	50.8	50.3	49.7	52.7	44.0	55.0	61.0	44.7	44.3
Avg. hourly wage (trade and transportation) (3 mo. chg	50.3	48.8	50.0	52.2	55.3	38.0	61.3	38.0	54.3	55.0
Total non-mortgage balance / total limit	50.6	55.0	46.3	50.3	51.7	64.7	55.7	38.7	46.0	46.7
Avg. hourly wage (private) (12 mo. chg.)	51.8	50.3	53.5	51.5	56.0	45.7	59.0	54.0	47.3	48.7
Avg. hourly wage (trade and transportation) (12 mo. ch	51.8	57.2	48.8	49.3	52.0	55.0	60.0	47.3	37.3	59.0
Avg. weekly hours worked (leisure) (3 mo. chg.)	51.9	52.5	50.5	52.7	51.3	43.3	39.7	59.7	64.7	52.7
6 mo. chg. in cycle utilization	52.1	46.7	54.7	54.8	33.0	70.3	48.0	64.7	38.3	58.0
Avg. hourly wage (leisure) (12 mo. chg.)	53.2	49.0	53.5	57.2	47.0	48.3	53.3	46.3	62.0	62.3
Avg. hourly wage (leisure) (3 mo. chg.)	53.6	52.7	52.7	55.5	58.7	60.7	62.3	37.3	66.3	36.3
Total credit limit to number of open bank cards	54.0	52.0	52.3	57.7	68.0	56.0	45.3	41.7	49.0	64.0
Number of total nonfarm (NSA) (3 mo. chg.)	54.2	51.3	55.2	56.0	62.3	45.0	54.0	57.3	40.0	66.3
Flag if total limit on all bank cards greater than zero	54.8	50.0	60.2	54.3	59.3	72.0	43.7	33.3	67.3	53.3
Unemployment rate (3 mo. chg.)	55.0	58.3	53.5	53.2	52.3	56.0	68.3	55.3	52.7	45.3
Number of total nonfarm (NSA) (3 mo. chg.)	55.9	59.2	64.2	44.3	58.0	61.7	50.0	55.0	62.0	48.7
Total private (NSA) (12 mo. chg.)	56.0	57.5	53.5	57.0	53.3	47.7	54.3	64.3	56.3	60.0
Percent chg. in credit limit (lagged 1 mo.)	56.5	57.0	52.5	60.0	66.7	74.3	10.7	68.7	58.7	60.0