

Credit Performance: Does Situational Data like the Economy Matter?

Robert B. Avery

Paul S. Calem

Federal Reserve Board

May 29, 2002

Presentation at “Credit Risk Modeling and Decisioning,” a conference jointly sponsored by the Federal Reserve Bank of Philadelphia and the Wharton School

The views expressed are those of the authors and do not necessarily represent those of the Board of Governors of the Federal Reserve System or its staff.

“Our [the industry] credit-scoring models are based on credit histories that reflect only prosperity...loans that were approved during the period of economic expansion may now be affected by changed circumstances.” -

--Fed Governor Olsen, 5/21/02

What are the Implications of this Point?

- Credit Performance is situational and depends on economic environment
- Credit History—and credit scores---since they are cumulations of performance should also depend on past economic environment
- Logic can be extended from national business cycle to regional or local economy
- Could also be extended to individual “exogenous” situational factors such as layoffs, health problems or divorce

Research Question

- Examine the relationship between economic environment and both credit performance and assessment of credit history
- Use detailed “account-level” data for 250,000 nationally representative people selected from national CB in June of 1999

Preview of Conclusions

- Find some evidence that economic conditions matter suggesting that scores of individuals in economically depressed areas may be inappropriately low
- Also find some serious limitations in the nature of the data reported to the CBs that effects ability to use situational data in scores

Background– CB files

- CB data come in four forms: public records, inquiries, trade line, and collection accounts. Each “account” is a separate record
- Focus here on trade lines (e.g. mortgages, credit cards). For each account we have:
 - Date opened, “closed,” and last reported
 - Amount currently owed, past due, credit line, maximum amount owed
 - Current payment status, 48 month history of payments
 - Type of lender and account, various comment codes, and account ownership

Background – CB files

- CB files are like FBI raw files. Purpose is to provide input to analysis not do the analysis
- Judgemental underwriting or a “custom” or application score can be based on other information outside CB. A Bureau or classic “FICO” score based only on CB data.
- CB data appear to be extraordinarily clean with virtually no “validity-type” errors.
- Problems appear to relate more to missing data or accounts which are inconsistently reported

Background – CB files

- Problems seem to arise in following accounts which are transferred, closed, or sold to a collection agency
- Some of these problems are significant and effect the ability of analyst to use situational factors in developing a scoring model
- At the very least problems require assumptions to deal with missing data

Background – CB files

- 35% of trade line files are not currently reported and not reported as closed
 - 13% of these are missing current balance
 - 30% of these are missing current payment status (though 5/6 of these have 0 balance)
 - 2.5% of these show current payment status of a minor delinquency with a positive balance. These represent 57% of all accounts which are “currently” minor delinquent
 - Many of these are closed-end accounts past date due
 - Appears accounts often are not closed when they are transferred, paid off or sent to collection

Background – CB files

- Particularly acute problem with mortgages. 80% of individuals with 2 or more open mortgages showed that one mortgage was opened within 2 months of the last reporting of the other mortgage for approximately the same amount. Often one account is listed as past due. Hard to distinguish between sale of servicing and a new loan.

Background – CB files

- Big problem with major derogatories. Hard to follow accounts when sent to collection department or agency. Cannot tell if one or two accounts. Sporadic reporting of chargeoffs and payoffs.
- Looked at all individuals who took out a new mortgage in first 6 months of 1999. About 5% showed a major derogatory. About ½ of these showed that the account was unpaid. This appears to be an inaccurate representation as new mortgagees are typically required to pay off seriously delinquent accounts

Background – CB files

- Collection agency accounts also a problem.
 - 30 percent of individuals show some collection account.
 - 88% are small (under \$500).
 - Source of creditor not coded. We parsed name of creditor to estimate type. Estimate 52% are medical; 24% are utilities; only 5% are for normal “trade line-type” loans (some of these are double counted).
 - Payoff information sparse and often not linked to the original account. Inconsistency in reporting multiple small charges or single consolidated amount.

Background – CB files

- Credit limit missing in 34 percent of open revolving accounts currently reported
- Account ownership status missing for many non-primary account holders. Cannot tell if authorized user, cosigner, or co-applicant

Implications for Scoring

- Any differential based on the timing of a major derogatory hard to do
- Number of accounts and current minor delinquencies and unpaid collections likely overstated
- Need to estimate credit limit for those missing it
- Hard to differentiate behavior on single and joint accounts (useful for divorce) or accounts opened since a specific date (complicated by transfers which appear to be new)
- Model builders must address these and other problems

Modeling Framework

- Given difficulties in using situational variables it is important to see if there is value in it
- First test is to use economic environment— economic conditions in the borrower's county-- during times when credit history (and credit performance) is measured and see relationship to performance
- Crude tests only approximating analysis that would be done in full development of a score

Modeling Framework

- Traditional scoring model: $y_{it} = F(Y_{it}, \eta_{it})$
 - y_{it} = period t repayment performance of individual i on representative account (perhaps of particular type)
 - Y_{it} = vector of measures of credit usage and repayment performance by individual i prior to period t
 - Y_{it} may include, for example, *number of accounts 30-120 days delinquent (currently and in past two years); number of accounts ever charged off or in collection; number of accounts of different types; number of new credit accounts opened in past year; credit line and revolving account utilization rates*

Modeling Framework

- Our hypothesis: $y_{it} = f(X_{it}, C_i, \mu_{it})$
 - X_{it} = vector of exogenous factors affecting borrower i's ability to pay (*local economic conditions, personal trigger events, etc.*)
 - C_i = measures of borrower's willingness to make timely payments (*measures of reliability, character, financial responsibility, etc.*)
 - Obstacles to empirical estimation: *unobserved borrower attributes; limited information on timing and context of delinquency episodes; serial correlation of economic conditions.*

Modeling Framework

- We estimate: $y_{it} = F(y_{i,t-1}, x_{it}, x_{it-1}, Z_i, \phi_{it})$
 - Period t (“performance period”) is 7/97 – 6/99 and period t-1 (“credit evaluation period”) is 7/95-6/97.
 - y_{it} tracks performance on new accounts opened on or after 7/97 but no later than 3/99
 - $y_{it} = 1$ if account becomes at least 60-days delinquent during the performance period, and 0 otherwise.

Modeling Framework

- We estimate: $y_{it} = F(y_{i,t-1}, x_{it}, x_{it-1}, Z_i, \phi_{it})$
 - y_{it-1} measures maximum delinquency during the evaluation period on any account open as of 7/95
 - $y_{it-1} = 0$ if no delinquencies; =1 if 30 days; =2 if 60 days; =3 if 90-150 days; =4 if collection (extended version)
 - $x_{i,t-n}$ = measures of local economic conditions:
 - Z_i = control variables: Age of the borrower; seasoning of the loan; Census division; Census tract relative median income; Census tract minority percentage

Modeling Framework

- $x_{i,t}$: 1998 county unemployment rate; 1997-1998 MSA or state house price appreciation rate; 1997-1998 percent change in county per-capita income
- If “situation” matters, period t economic environment will be reflected in payment performance

Modeling Framework

- $x_{i,t-1}$: 1994 to 1997 change in unemployment rate
- If “situation” matters, past performance may overstate or understate current credit risk depending on past context.

Modeling Framework

- Clearly, we face the data limitations already noted for scoring models
- Also limited to data from relatively homogeneous economic environment (broad-based expansion)
- Still, we believe we can test whether “situation” matters and whether there may be value added in attempting to control for it

Empirical Procedure

- OLS estimation (General Linear Model)
- Close to 200,000 observations
- Weighting for single vs. joint accounts

Results

Means of variables

% with no prior delinquency:	75.2
% with prior 30-day:	16.1
% with Prior 60-day	4.5
% with prior 90-120 day	4.5
% in low/moderate income ZIP	13.7
% in (95%) non-minority ZIP	52.8
1998 county % unemployed	4.20
1998 house price % change	5.57
1998 % change per capita income	5.57
Change in unemployment 1994-97	-1.22

Results

$$y_{it} = F(y_{i,t-1}, x_{it}, x_{it-1}, Z_i, \phi_{it})$$

<i>Mean of dependent variable:</i>	<i>0.021</i>
Intercept	0.070 (18.68)
No prior delinquency (base 90 day)	-0.062 (38.4)
Prior 30-day (base 90 day)	-0.047 (26.8)
Prior 60-day (base 90 day)	-0.028 (12.7)
Low/moderate income ZIP	0.014 (12.5)
Predominantly non-minority	-0.023 (12.8)
1998 county % unemployed	0.0008 (4.23)
1998 house price % change	-0.0007 (3.16)
Change in unemployment 1994-97	-0.0013 (3.19)

Extensions

- Incorporate past charge-offs, collections
- Borrower-level analysis
- Sample segmented by age-group (over- and under-50)
- Sample segmented by revolving vs. installment credit

Conclusions

- “Situation” matters
 - o Value added to considering current economic environment
 - o Value added to considering past context

Conclusions

- There are limitations to what can be accomplished with available data
 - Ideally, credit reporting systems would collect more data on timing of delinquency, collection processes, situational factors