Not If Affordability data adds value but how to add real value by Leveraging Affordability Data:

Enhancing Predictive capability of Credit Scoring Using Affordability Data

By Dhruv Sharma (Independent Scholar), George Overstreet (UVA), and Peter Beling(UVA).

Corresponding author email: ds5j@excite.com

Keywords: improving credit scoring, logistic regression, random forests, credit risk, systemic risk, safe mortgage products

Abstract

The aim of this paper is to focus on the role of affordability of loans on credit risk at 3 different levels. The first being short term practical approach to optimally using affordability data to improve current credit scorecards by showing how Random Forests can be used to improve logistic regression. After this discussion we step back and look at how affordability fits into ensuring systemic risk and use the credit union industry as a base case model for using affordability measures to reduce systemic risk. Finally we turn to the question of whether credit risk is a source of sustainable competitive advantage and what measures would create greater welfare for the financial system as whole. It is the belief of the authors that proprietary black box credit systems and increasing complexity of unpredictable factors make it necessary to build an open credit systems viewpoint. Using this weltanschauung we conclude with a simple example of restructuring mortgage products using affordability and borrower well being as central to sustainable credit risk management.

The article is comprised of survey and quantitative analysis showing how affordability has been used in the credit industry or not used. The article shows that credit decisions

can be improved by leveraging affordability optimally and then also discusses policy implications of using affordability to ensure soundness and well being of consumers instead of simply determining whether the consumer will repay the loan successfully.

Improving Credit Score Models by Leveraging Affordability Calculations

In all credit scoring applications the construct of affordability has a subtle but pervasive impact on the variables used in econometric models which predict borrower default. Econometrically speaking the question is not whether affordability data are important but rather how best to utilize affordability data in credit scoring models. This is a topic which this paper addresses. The history of financial ratio usage in predictive models can be viewed from a general viewpoint as a search for the appropriate calculation or interaction term. We show that viewing affordability data as an interaction term search is fruitful. This makes sense from an econometric sense as affordability could have behavioral interactions and effects on many other credit scoring variables. Traditionally affordability data has been used in underwriting logic viewpoint for cut offs and loan overrides along with its roots as being a basis for sound judgmental underwriting and automated underwriting. The approach described here can improve current credit scoring models and industry usage of affordability in a statistically significant way. The theory underlying the use of cash flow surrogates is sound and has been shown to be effective in numerous corporate bankruptcy prediction studies but has been absent in the consumer credit and mortgage credit scoring space beyond the nominal use of expense ratios and required liquid reserves (Buist, Yang, Megolube, 1998). We believe, that consistent with

Wilcox's model of bankruptcy prediction, free cash flow and assets comprise an econometrically sound bases for underwriting loans and that these variable have meaningful interaction effects with other credit application data used in credit scoring (Wilcox, 1971). The focus of this paper is on credit risk in the retail consumer lending space as opposed to corporate credit risk models which already use cash flow related affordability techniques to evaluate credit risk. Unlike corporate lending, consumer credit scoring literature has reported mixed results on the predictive power of affordability data (Wilkinson and Tingay , 2001). That said the approach outlined here can be used in other settings to improve predictive power for credit risk in general.

In this paper we outline an approach to best leverage affordability via a case study using the German credit data. In examining the German Credit data we show that by properly leveraging the affordability variables one can improve the model. This data set is used as it an open data set available for research. It is also urged that more data should be sanitized and made available for public credit research. The procedures described here have been used in various proprietary credit data sets and found to be effective on large data sets as well.

In contrast to prior research asking whether affordability adds predictive value in consumer credit scoring we show how affordability data can add significant value to traditional credit scoring models(Wilkinson and Tingay, 2001). In the past Gayler has shown that including large amounts of interactions terms can lead to overfitting (Gayler, 1995). Building on this research we find that focusing only on econometrically sound interactions related to affordability to drive increased predictive power is a fairly safe approach which does not lead to overfitting.

Affordability Metrics

From this viewpoint affordability information related to borrowers, whether in the form of income or assets, should maximize the predictive value of existing variables present for consumers. Borrower assets, such as checking or savings, can be seen as a signal of the borrower's income i.e. future cash flow capacity and ability to save.

Much like the well used notion of coarse classification via discretization or binning of continuous variables into bins interaction terms using affordability ratios is critical in squeezing predictive performance out of one's data. It is well known that binning variables helps improve credit scoring. It is especially helpful for consumer expenditure data when the data is noisy, unverified etc. By removing outliers of suspicious expense ratios (monthly expenses/income) modelers can see higher risk concentrated in loans with higher ratios which otherwise may be obscured. By improving data collection and tightening guidelines to not approve >80-100% ratios as a sanity check can be helpful in management of credit as this forces loan officers and underwriters to use correct calculations which can result in more accurate loan approvals. That said there is not a strong sense of standardization of unsecured credit lending, for products like credit cards, as there is for mortgage products.

Interactions of Affordability Data can Offer more predictive Power

An econometrician must carefully think through the sources and uses of affordability for borrowers. This is a task usually performed by judgmental underwriters but should be used in the praxis of econometrics as well. We now discuss some examples of thinking through routine loan application data which can yield some insights that help in predictive model development.

Traditionally expense ratios are used to as a single measure of affordability. Expense ratios in this context are the monthly expenses (mortgage, rental, other debt etc) divided by gross monthly income. By thinking about the same data from the viewpoint of savings one can use expense ratio data to make inferences on borrower savings habits. For example a growth rate of savings can be computed for borrowers using the expense ratio where 1- monthly expense ratio is the rate that is not spent and can be allocated to savings. Taking a portion of this rate and comparing it against existing assets(checking and savings) can provide greater evidence of borrower strength in the form of propensity to save.

In the mortgage industry expense ratios and borrower reserves in terms of how many months an applicant can make payments (if they lose their job), down payment, and full asset information is collected.

From a statistical viewpoint these are examples of meaningful interaction terms of affordability data. From a broader view all financial ratios are interactions. Affordability as a construct can have effects or signals on many aspects of a borrower's behavior or data. In this sense, examining interactions of affordability with other variables is an area not studied much but which can provide greater predictive power. An example of a carefully constructed interaction term is the number of maxed out credit cards which can be calculated by the dividing total unsecured debt outstanding by total unsecured number of tradelines which yields an average balance on each unsecured trade and then the total outstanding unsecured debt balance can be divided by the average revolving balance thus yielding a proxy of how many lines of credit are maxed out by a borrower and is predictive of higher risk (Dunn, 1999)..

This is a simple example of how thinking through calculations and data to construct a meaningful interaction term can yield analytic insight. This is something brute force data mining algorithms have not focused on. Traditional recursive partitioning trees, logistic regression, and other techniques would never be able to ferret out this relationship of interaction.

The problem of determining which variables are most predictive and the importance of affordability data has been clouded over the history of credit scoring literature. We now turn to a comprehensive review of credit scoring to show how conflicting results are due to multi-collinearity present in credit data and logistic regression and discriminant analysis which have been used to build most credit scorecards in production today are inadequate tools for assessment of variable importance. Following this we discuss a relatively recent statistical technique called Random Forests which allow econometricians to get much more accurate assessments of variable relationships. Using this tool we discuss how to make better use of traditional regression tools once statistically significant variables relationships have been confirmed via the use of Random Forests.

Literature Survey of Corporate Financial Distress Prediction and Interactions aka Financial Ratios

In analyzing corporate credit risk literature it is clear that affordability and affordability related interaction terms are used widely to predict credit risk. We now review that literature. Hodges, Clusky and Line found that in analyzing different bankruptcy methods ranging from Altman's z-score to Ohlson's equations that a significant predictor of risk was 'assets growing faster than cash "causes the deterioration of the cash to total assets ratio" (Hodges, 2005). Mossman found across various bankruptcy models for corporations that "only cash flow was a consistent predictor 2 to 3 years out" (2005). Financial ratios found to be predictive in corporate bankruptcies by Altman, Beaver and Deakin were: decreases in net income, z-score, capital expenditures, cash to total assets, and income taxes to total assets (2005). The z-score is comprised of the following variables: working capital to total assets, retained earnings/total assets, EBIT/total assets, equity to liabilities, and sales to assets (Altman, 2003). The most extensive literature review of cash flow usage in bankruptcy prediction was conducted by Gombola, Haskins and Ketz in which the authors conclude that “in general …proxy for cash flow has been found to be useful in bankruptcy classification” and while cash flow from operations was not statistically significant the following variables were significant: income+depreciation as a percent of assets (especially), cash to sales, current debt to total debt, sales/assets, and working capital from operations to assets (Gombola etal, 1987). Gentry, Newbold, and Whitford also found that including cash flow components and ratios improved bankruptcy prediction and that outflows were more important in bankruptcy prediction such as dividends, investments, and receivables (Gentry etal, 1985). Most recently Zhao, Sinha, and Ge in 2009 found that usage of financial ratios, using FDIC data, significantly improved bankruptcy predictive accuracy if the ratios added were based on domain knowledge (Zhao etal, 2009). Sound foundational theory supporting free cash flow as a driver of affordability and risk prediction was established by Wilcox’s class paper which used the gambler’s ruin model to assess risk, in which net liquidation value and average adjusted cash flow was used along with the estimated average cash flow and estimated statistical variance of the adjusted cash flow (Wilcox, 1975). Another important contribution this literature was that of Henebery who showed that cash flow variables improve predictive accuracy of cox proportional hazard for long horizon models and not short term models (Henebery, 1996).

Comprehensive Literature Survey of Credit scoring Variables Found to be predictive

We now turn to a review of consumer credit risk scoring literature. What follows is the most extensive and broad literature review to date to help consolidate the disparate body of work in this field. From a systems point of view it becomes clear that credit applications should be standardized like the mortgage credit application, the 1003 form. That aside, we shall review this literature with the aim of identifying common themes such as the econometric foundations of the variables. The categorization by Overstreet and Kemp of cash flow variables, stability variables, and payment history variables is a useful mapping in understanding the variables (Overstreet and Kemp, 1992). Besides these 3 categories, collateral value is another category which is critical for mortgage and secured loan products. From the finance viewpoint “sources of funds are available to repay a consumer loan” comprise of “net disposable income or 'free cash flow', liquid or liquefiable assets, and additional borrowings” (1992).

The great grandfather of credit scoring research begins with Durand, who in 1941 undertook a comprehensive study which resulted in the following findings: “income was only moderately related to default risk….possession of life insurance, bank account or real estate was a better indicator of credit quality” and that women were better credit risks, older applicants were less likely to default, and that “classification by occupation and industry was important” (Sullivan, 1987). Subsequent to Durand’s research the next major period of credit scoring research occurred in the 1960s with work in applied psychology journals on credit scoring which has been largely ignored.

Hassler, Myers, and Seldin analyzed department store credit and found the following variables to be statistically significant in predicting bad credit risks in terms of high collection expenses (i.e. accounts resulting in loss): # of months payment made, # of months no payment made, # of months payment < amount due,# of months limit exceeded by 100 or more dollars, # of months limit exceeded by $50 or more,# of months since first collection notice sent, # of months second collection notice sent,# of months of collection notices sent, trend up or down in balance, and pattern of balance to prior balance (Hassler, Myers, and Seldin , 1963). Myers conducted a study in the same year in predicting consumer delinquency and found the following variables to be statistically significant in predicting delinquency: # of dependents, marital status, time at present address, have telephone, rent or own, # of accounts in well known stores, occupation, income, bank account, and personal reference given (Myers, 1963). In this study Myers found that people with phone lines were less risky. At the time phone lines would have been a surrogate or proxy for wealth or stability. Myers also conducted one of the first known analyses of published attempt at factor analysis of the credit application variables into factors in 1964. His study had a small sample of approved applications with variables which appeared to have low intercorrelation and resulted in the following factor segments: Factor a: size of transaction; higher amount of purchase higher indebtedness

Factor b: age and stability; time at property address and job
Factor c: prior use of installment credit
Factor d: income, telephone, # of bank accounts
Factor e: loan duration, size of transaction (longer term and lower downpayment associated with higher risk. (Myers, 1964).

As suspected the possession of a phone tended to be part of income factor and showed to be a type of cash flow surrogate. Interestingly Myers in 1967 did one of the first studies in optimal credit scoring cut off decision making based on profit and loss (Myers,1967).

In addition to unsecured credit Myers also performed an early analysis of mobile home using discriminant analysis to classify mobile home loans and found the following items to be predictive in the scoring system: Age, Phone, new parking address, mailing address different then new parking address, time at present job, bank account?, previous trailer, cash, auto-clear, own real estate, previous high credit account, # of unsatisfactory credit references, # of repossessions, bankruptcy, width/length of trailer, form of down payment, unpaid balance, new or used, and term of contract(Myers & Forgy, 1963).

Smith in 1964 found similar findings and echoed the 'surprise' that "traits designed to measure financial ability of the borrower to carry monthly payments were the least effective in discriminating between good and bad accounts" while time on last job and time in last residence were significant predictors (Smith, 1964). Throughout the history of credit scoring the inability of traditional statistical methods such as discriminant analysis and logit to isolate variable importance has been an issue(Eisenbeis, 1977). The recent advent of Random Forest variable importance techniques provide a better lens into true variable importance now clearly show cash flow surrogates and ability to pay to be statistically significant credit predictors.

Lane in 1972 found the following to be predictive of sub-marginal borrowers: months on job, monthly income, months to acquire debt, household assets, auto (car) assets, back rent, collection agency, bank, account, consumer finance trades, department store trades, gas company trade and possession of mortgage (Lane, 1972).

The US Survey Consumer Finances ushered a new wave of consumer credit research in the 1980s to 1990s. In studying consumer who fall behind in payments Sullivan surveyed existing credit literature which stated that per R.C. Peterson's research source of employment was a highly significant indicator of credit quality (government-civilian employment, banking, finance, and real estate professionals were lower sources of risk) and that manufacturing, construction, retail, wholesale industry workers had above average credit risks (Sullivan, 1987). Sullivan cited that C.L. Peterson and R.L Peterson's research on automobile loans found that "young and cyclical industries" had higher than average default rates, and also interestingly that "higher down payment reduced probability of default for young borrower but did not reduce risk for cyclical workers" (1987). Sullivan also cited the importance of Overstreet and Kemp's work on developing one of the first "theoretically derived scoring model (base don economic, credit history, willingness to pay) which had found monthly income to be statistically significant along amount of debt outstanding " was significantly associated with credit risk and also that human "loan officers did not put much weight on " these variables but instead "focused on size of monthly payment" which itself "was not statistically significant" (1987). Overstreet and Kemp built one of the earliest theory based credit scoring systems in 1986 (Overstreet and Kemp, 1986). In their study they found following to be significant: loan type, loan amount, # of monthly installment, monthly installment payment of requested loan, deposits, length of employment, monthly income, monthly fixed expenses, monthly mortgage, net income (income-fixed expenses), total indebtedness, loan history (1986). In 1996 Overstreet and Bradley found the following to be predictive as well “various debt ratio and other cash flow oriented surrogates, employment time, home ownership, major credit card ownership, and representations of past payment history” along with the number of tradelines utilized over 75% (Overstreet & Bradley, 1996).

Sullivan’s own analysis of the consumer finances concluded that:

“risk of payment difficulties is higher for lowest income and low amongst highest income families, the 25-34 age group has the “highest probability of payment difficulties”,”2 of 3 income brackets with before tax income >20K in age 45-54 had above average probability of default, liquid assets was critical as the “incidence of slow pay highest (50%) among families with 0 or no liquid assets”, while the lowest risk with shown with families with >$2000 in liquid assets” (1987). Sullivan also found that families with income >$20K but liquid assets $500-$1999 had higher risk, less educated consumers had higher risk as the “probability of difficulty was 1.5 times higher for high school educated only borrowers” (1987). Interestingly Sullivan found that “young unmarried women less likely to have problems but women in all other groups were more likely to have problems”, “young married couple with children more likely to have problems”, “renters were twice as likely to report debt payment difficulty than home owners”, “ratio of total debt amount to income is inversely related to probability of slow or missed payment”, for male head of household families as “total ratio higher (debt to income) then the probability of missed payment declined”, “consumer debt ratio increased probability of missing payment”, “debtors with mortgage repayment >20% of income were less likely to have trouble", and that consumer with "fewer or no credit cards riskier" (1987). Sullivan also found that "debt repayment problems were associated with source of consumer credit...borrowers who obtained credit from finance or stores more likely to have been late or missed payment than those who borrowed from banks, credit unions or savings/loans" (Canner, 1990).

By 1990 Sullivan's work was criticized by Canner as it was based on univariate analysis and Canner performed analysis using a multivariate analysis of the consumer finance survey data (Canner, 1990). Canner used a multivariate logit analysis and found the following statistically significant results: "the variable which has the greatest significance is one that indicates whether a person has previously been rejected for credit","age of head of household is the next most significant (older people are lower risk)", families with "more children higher risk of missed payment", "liquid assets to debt, receipt of government assistance"," type/source of loan was significant", inverse relationship between consumer ratio and timely payment", and "consumer durable loan types are more risky". Interestingly Canner concluded that "job stability was marginally significant", and that "variables for household income, amount of liquid assets, ratios of debt to income add little to explanation of late payment" and "housing tenure and education not significant" (Canner, 1990). These conclusions again derive based on the traditional use of logistic regression on data with multi-collinearity.

It is important to note that until 1998 monthly expense to income data was not available in the survey of consumer finances. In 2003 Getter used the 1998 survey of consumer finances data to find the important findings that “consumer delinquency problems are mainly the result of unexpected negative events neither the borrower/lender could anticipate”, “size of household payment burden has insignificant effect on delinquency risk/default risk”, and that “household financial assets...used as buffer against negative shocks serve as very important predictor of delinquency risk” (Getter, 2003). These findings were consistent with “Elmer and Selig’s (1999) trigger event theory” (2003). This survey had monthly expense to income ratios data which showed that the probability of being 2 or more months behind rose as the ratio bins rose.

Getter also showed that cash flow surrogates like “assets, home owner status, and college education had negative coefficients with serious delinquency, while “divorce/separation, low income, and family size” had positive coefficients with risk, and that “unusually low income” families and also families with expense ratios between 50-74 had higher risk (2003). Interestingly expense ratios higher than 74% had a negative coefficient and were not statistically significant. Goodwin also used the same data and found higher risk was associated with “younger, non white, had larger households, more positive attitude toward credit, installment debt (mortgage, auto or durable goods debt), had financial support from friends, relative, major real estate transactions” (Goodwin, 1999).

Another important study of credit scoring using actual application data was conducted by Stine, Lewis, and Jones and used probit instead of logit to analyze risk (Stine, Lewis & Jones, 1990). Their probit study found following significant variables: “(savings or checking or none), # of credit references on credit report (trade lines), and TLD (# of minor delinquencies+# of major delinquency on credit report)" (1990). Again curiously they found that salary (income aka cash flow surrogate) had a "marginal p value of .2-.15 but when removed performance of model dropped significantly" (1990).

An interesting study using hurdle box-cox models by Moffatt found an interesting relationship in the credit variables (Moffatt, 2005). Moffat's approach predicted default using 2 hurdles: "one hurdle for customer being a potential defaulter and another hurdle for extent of default" (2005). The key conclusion of this study was that "personal characteristics are important in first hurdle and economic characteristics more important in second" (2005). In particular his study finds that "males are less likely to pass the first hurdle (less likely to be potential defaulters) but conditional on default males tend to have higher losses, borrowers aged 50+ are most likely to be in the never default category, marriage lowers probability of potential default, occupation is important in first hurdle and loan purpose is more important in the second hurdle, and the loan amount relationship is u-shaped" (2005). The statistically significant variables along with their direction towards increasing (+) or decreasing risk (-) is as follows:

- First hurdle: male-, age-, age sq.-, married -, time in occupation-, time at bank-, # of credit searches +, term of loan+, loan amount-, loan amount squared+
- Second hurdle: male+, homeowner-, tenant+, gross income-, # of credit searches+, loan amount+, purpose of loan(+/-)

One of the first studies to compare application scoring to behavior scoring was done in 1992 by Crook, Hamilton, and Thomas (Crook, Hamilton, and Thomas 1992). The authors segmented borrowers into 3 segments: “those who miss at least 1 payment and those who have missed 1 or 2 payments, subsequently miss 3” (1992). This study found that, in descending order the most powerful 6 predictors of borrowers who miss 3 consecutive payments were: “applicant's employment status, spouse's income, years at bank, residential status, years at present employment, and check amount” (1992).
The factors which separated borrowers who would miss 1-2 payments only vs. those who would miss 3 payments were: “years at the bank, spouse's income, applicant's employment status, years at present employment, and deposit account” (1992). Crook etal found in that in 3 predictive models the “proportion of good to bads has a W shape as income increases” (1992). Within this income framework the authors concluded that the chance of moving from 2 month delinquency to a 3 month delinquency was lowest for those earning $15K or more and greatest for those between $7500-$10K (1992). It is important to note the non linearity of the income and the tricky relationship it has. The borrowers most likely to miss 1 payment had: “<3 years bank account, self employed, <1 year employment, large mortgage balance, and >=4 children” (1992). The factors in all models which were critical were: applicant's employment status, years at present employment, years at bank, and residential status. This is consistent with the other studies reviewed thus far. Using proportional hazards for behavior scoring Stepanova and Thomas also found the following variables to be statistically significant: “amount of loan, time at current address, time with current employer, residential status, # of dependents,

net income, performance variables: delinquency, delinquency status, current month end balance, worst status" (Stepanova & Thomas, 2001).

Other variables found to be predictive for credit cards in particular by 3 new financial variables were: 1) ratio of total minimum required payment from all credit cards to household income 2) % of total credit line used by consumer
3) the number of cards on which the consumer has reached the borrowing limit.(Dunn & Kim, 1999). Dunn and Kim's study also found the following other statistically significant variables: credit line, max cards (# of cards maxed out), min pay (total min required payment), balance carried/income, balance carried/total credit line from all CC, age, # of children, and married status (199).

In analyzing the credit losses Monaghan found the following tradeline variables from credit reports to be predictive for credit card, automobile, and home lines of credit: amounts past due (unpaid collections), derogatory public records, collection records, status of tradelines, age of oldest tradeline, non promotional inquiry count (purged every 24 months), leverage ratio on revolving type accounts (sum of revolving debt/revolving limit), sum of credit limits for all revolving trades (Monaghan, 200). This variables seem typical of what are used in most credit scorecards like the Fair Isaac Credit Score. Another study found similar credit report variables to be effective using cluster analysis:

bank card high balance, # of bank trades, # of open bank cards, # of bank always satisfied within 6 months, # of trades always satisfied, # of open trade, # of inquiries in 6 months (Hababou, Cheng & Falk, 2006).

Along with credit report variables geographic risk of credit risk model has been done. In the mortgage space home price volatility drives risk while for unsecured credit unemployment and state laws have impact(Worden & Sullivan,1995). In particular Worden and Sullivan found bankruptcy to be higher in the most debtor friendly states: FL, NC,ND,PA,SC,SD, and TX (1995). IN regards to bankruptcy as a driver for credit risk Weiss and Paquin find that most bankruptcies could be explained by 4 key variables: supply of consumer credit (annual change in # of bank card accounts), consumer capacity to service debt (household debt to income), condition of job market (unemployment insurance claims) and interest rates (using a 2 year lag variable) (Weiss & Paquin, 1998). This study was supported by Grieb who found consumer debt ratio and amount of total revolving debt to be significant predictors (Grieb, 2001).

Overindebtedness as a causal risk factor in credit scoring has been studied recently by Finlay (Finlay 2006). In this important work Finlay shows that "that using only data captured on a typical application form, combined with data from a credit bureau, it is possible to develop good predictive models of expenditure and over-indebtedness" (2006). Finlay bases affordability on a cash flow like construct by using 'disposable income available after regular household expenditure and existing credit commitments have been taken into account' (2006). An important insight from this study was that "most credit scoring models forecast 12-24 months … this may be insufficient to cover the period of over which an over-indebted customer can maintain payments" (2006). It is important to note the in economic bubbles and expansionary periods marginal borrowers can survive as long as they receive additional borrowings for lengthy periods of time allowing them to be appear as goods as banks and credit vendors may be hungry to grow market share in a market for credit lemons. Due to this reason it may appear that the FICO score predicts better than affordability calculations as it factors in availability of credit but in a tightening market the FICO will not adjust as quickly being built on past data as will affordability data. This is an explanation also of why Liu's 2001 study showed property value models to outperform cash flow for the 1995-2001 time period when the economy was booming.

Finlay uses of after tax and after expense pay is consistent with the approach recommended by Langrehr which advocated using residual income based expense to income ratios (Langrehr etal, 1989). In US studies of household insolvency DeVaney studied 1983 Survey of Consumer Finances and found 2 different drivers of household insolvency depending on statistical technique (DeVaney, 1994). DeVaney found that the "liquidity ratio was most important for predicting… insolvency in logistic regression" where the "liquidity ratio=liquid assets/disposable income ." (similar to months reserves used in mortgage underwriting) where 3-6 months of liquid reserves were important in dealing with shocks (1989). Using decision tree recursive partitioning, CART, DeVaney found assets to liability was the most important factor in predicting insolvency along with the debt expenses to income ratio, which also predictive in the logistic regression (1989).

Hanna's study of US overspending found that income level was the most predictive of overspending, using a multivariate logistic regression, and that other statistically significant factors were "size of city >3 million, income, net assets, spending, income, income squared, income cubed, education, mortgage owner status, size of family >4, and age (Hanna etal, 1994). Greninger has performed a study where using a delphi technique with finance experts the following guidelines were determined based on judgmental expertise as criteria for sound financial ratios for families:

- liquid assets should be >=2.5 times monthly expenses
- savings to gross income >=10%
- liquid assets to net worth>=15%
- net investment assets to net worth >=50%
- foreign investment/net worth >=10%
- rent to gross income<=30%
- mortgage exp to gross income<=35%
- non mortgage debt to after tax income <=15% (>=20% dangerous)

(Greninger etal, 1996). The calculations for these variables were as follows:

- Liquid assets = cash and cash equivalents, checking accounts, savings accounts, money market accounts, money market mutual funds, and CDs with maturities of <=6 months.
- Investment assets = all other assets held for investment purposes, not including use assets or equity in a home.
- Monthly expenses = average fixed and variable living expenses including debt/ credit repayment, taxes, and monthly allocations being set aside for irregular

expenses such as auto insurance, vacations, gifts, etc.

- Current debt = all debt/credit obligations, charges, bills and payments due within 1 year.
- Payroll taxes = federal, state, and local income taxes and social security taxes.
- Property taxes = real estate and personal property taxes.
- Renter's expenses = rent, renter's insurance, and utilities.
- Homeowner's expenses = principal, interest, taxes, insurance, homeowner's association fees, utilities, maintenance, and repairs. (Greninger etal, 1996).

Mortgage Credit Scoring

Mortgage credit scoring has long acknowledge the importance of liquid reserves, debt expense to income ratios, and loan to value as key predictors of risk (See Buist,1996). There is a **"**strong correlation between FICO, DTI, disposable income and defaults" (Fabozzi, 2005).

In the late 1960s Buel performed one of first credit scoring studies on mortgage loans and waded through a large number of interaction terms i.e. ratios. In this study Buel predicted delinquency on purchase mortgage loans had a potential of 31 variables and 210 ratios (interaction term calculations). Using judgment Buel tested 79 of the 210 ratios along with application variables. Buel's study found 4 items on the mortgage loan application and 4 ratios differentiated delinquent from non-delinquent accounts effectively. The variables found to statistically significant in predicting delinquency were:wife employed, # of children, # of other accounts, credit rating, monthly mortgage payments, age of husband, total loan/total income, age of husband/# of dependents other than self and age of husband/# of other accounts (Buel, 1968).

Hakim's study in 1999 found loan to value, number of dependents, years on job, loan amount, property age and income-expenses to be predictive of default risk in mortgages.

Freddie Mac's scorecard discussed by Avery indicated that "an index (score) of consumer credit history is roughly as predictive of default as initial housing equity (Avery Freddie Mac Scorecard 1996)." (Buist, 1996). This again was due to multi-collinearity. The recent 2007 and on mortgage crisis shows that equity in the home and home price risk along with ability to pay risk proved to be high indicators of risk and more so than credit scores which did not factor this in.

Fannie Mae's 2002 Annual Report listed the following factors used in mortgage credit decisioning:

> Loan-to-value (LTV) ratio, Product type (term), Property type:(Mortgages on one-unit properties tend to have lower credit risk than mortgages on multiple-unit properties, such as duplexes, all other factors held equal.), Occupancy type: Borrowers may purchase a home as a primary residence, second or vacation home, or investment rental property. Mortgages on properties occupied by the borrower as a principal or second residence tend to have lower credit risk than mortgages on investment properties, all other factors held equal), Credit score, Loan purpose, Geographic concentration: Local economic conditions affect borrowers' ability to repay loans and the value of the collateral underlying a loan, all other factors held equal) ,and Loan age (2002 Annual Report).

In regards to the credit score “borrowers with FICO <620 have 20 times higher probability of foreclosure than borrowers with >660 FICO” (Taff, 2002). In addition to these factors self-employed borrowers are deemed to have higher mortgage risk due to 'uneven cash flows' and stability to pay (Taff,2002)

Findings of review

The review of credit scoring application variables finds that cash flow surrogate variables consistently show up across studies across time. In addition it is difficult due to multi-collinearity to assess the impact of variable importance via the traditional logistic regression or discriminant functions. Despite this most credit scoring research has been done and is practiced using these tools alone. Econometrically credit scoring variables can be segmented into: cash flow variables, stability variables, and payment history variables (Overstreet, 1992). Unfortunately too much of credit scoring has had to work with 'biased estimation in data ...[which] has been shown to predict and extrapolate better when predictor variables are highly correlated...’ as this is common to credit scoring (1992) .

Importance of Random Forests to Credit Risk and Economics in general

To date the majority of credit scorecards used in industry are linear models despite the known issues of the flat maximum and multicollinearity (Wainer, 1978; Overstreet etal 1997; Lovie, 1986). Random Forests are a powerful tool for economic science as they are able to successfully deal with correlated variables with complex interactions (Breiman, 2001).

A simple example of the power of Random Forests was shown by Breiman in the binary prediction case of hepatitis mortality in which Stanford medical school had identified variables 6, 12, 14 and 19 as most predictive of risk using logistic regression. Subsequently using the bootstrap technique Efron showed that none of these variables were significant in the random resampling trials he ran. The Random Forest variable importance measure, created by Breiman, showed variables 7 and 11 to be critical and improved the logit regression results simplifying the model and by reducing error from 17% to 12% (Breiman, 2002).

As Random Forests are non parametric the linear restrictions of the flat maximum do not come into play as such. That said predictive models tend to perform well with regards to pareto optimal trade offs in true positive and false positive rates which look like an asymptote like the flat maximum effect. The complex interactions of economic variables such as macroeconomic forces and affordability are too complex to be studied for simple linear regression anymore. Random Forests serve as good estimate for asymptote of possible predictive power in this regards and help us get past the psychological limit we may believe to exist for predictive power as Roger Banister was able to do with preconceived limit on minimum time for completing the mile run. The way Random Forests work by building large quantities of weak classifiers with random selection of variables grown with out of sample testing is analogous to the way humans make decisions in a market place (See Gigerenzer's work on "Fast and Frugal trees" on human judgment models). Humans each look at the data available to them and make quick inferences and take actions based on these data. Random Forests then take votes from these large quantities of predictors and use decisions of all the predictors to make the final decision. The fact that diverse models built on different variables and samples of data when combined outperform other simple linear models is profound and may help explain why diverse models are essential to a healthy and efficient marketplace.

That said the critical aspects of Random Forests of interest to economic scientists are the features Breiman intended such as :

- Random Forests never overfit the data as they are built with out of sample testing for each submodel
- Variable importance (a measure based on the importance in accuracy each variable provides to the overall model based on permutation tests of removing variables)
- Being able to see the effects of variables on predictions (2002).

Breiman's insight was that although Random Forests are complex they provide insight on complex phenomenon being modeled. Random Forests have been used successfully in complex biomedical and astronomical data sets. Credit data sets abound with correlated variables and Random Forests now allow us to clearly see that traditional regression erroneously considered cash flow variables to not be statistically significant, and that once the full impact (interactions) of affordability data is assessed, this data is predictive and usually to be found in the top 10 most predictive variables in credit score models using Random Forest variable importance tests.

Random Forests help us see the true impact of complex interrelated variables. As Breiman mentioned in his Wald lecture, complex phenomenon cannot be modeled well with goodness of fit models with simplifications. A more scientific approach is to build as complex a model to fit the phenomenon being studied and then to have tools like variable importance to understand the relationship inside the phenomenon(Breiman, 2002). This is an important point as economics is based more and more complex realities. One effort in the credit scoring space recently has been to add macroeconomic variable interactions to credit scoring and efforts in this path have improved credit scoring performance. (Crook,2008) This is yet another example of how models are getting more complex and how interaction terms play an important role in credit model evolution.

A Simple Methodology of Leveraging Interactions

Random Forests are one of the most powerful out of the box statistical techniques available to econometricians. The strength of Random Forests is that they can take a large amount of predictor variables and quickly generate robust variable importance measures which logistic regression alone can miss (Breiman, 2001). Given that we now propose how to integrate Random Forests into current credit scoring practice to improve existing credit scorecards.

A meta algorithm, called the Sharma approach, which we propose for optimally making use of affordability measures in credit scoring using Random Forests is as follows:

- Run Random Forests on all variables to find the most predictive variables.
- Take the predictive variables related to affordability shown to be effective in Random Forests and test out interactions with all variables in logistic credit score model and keep statistically significant terms.

This approach works by adding predictive interactions to existing logistic regression scorecards. This gives us the best of multiple statistical techniques as using this approach we are essentially taking insights from the Random Forests, which linear regression would not provide, and help create a more robust regression model. We show that using this approach logistic regression can be tuned to perform as well as Random Forests or even better than Random Forests, as long as the interaction terms are added carefully with out of sample testing (to avoid overfit).

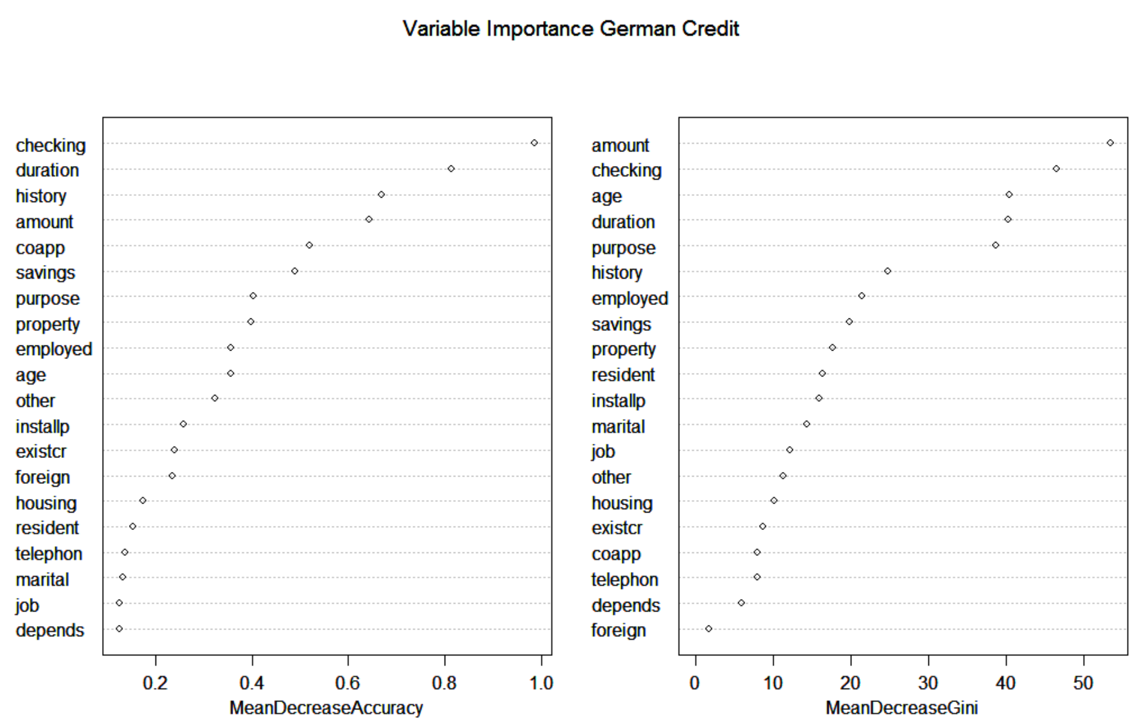
Case Study

Using the German data the following approach from above is used (. *The german credit data has 300 bad loans and 700 good loans and the details of the application data are described in the appendix..*

The potential affordability related data in this data set is: checking account, savings account, install payment ratio to income.

Step 1: Explore Interactions Via Random Forest

Using Random Forest on the German data reveals the following to be predictive in predicting bad loans:

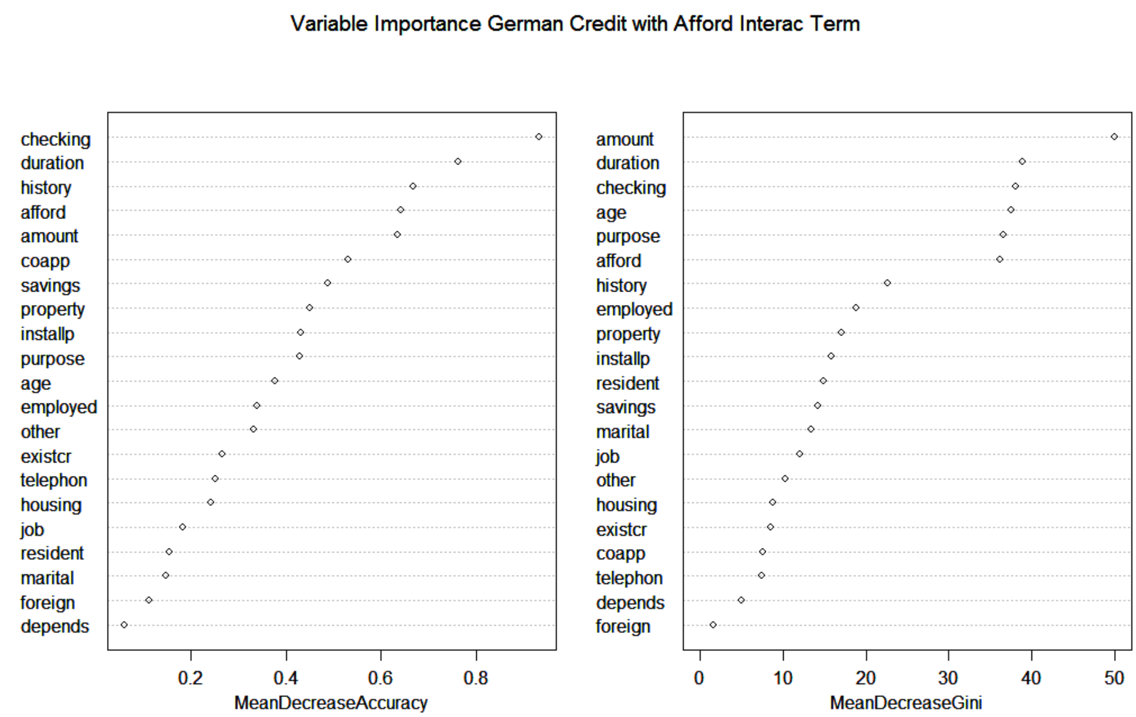


Using Random Forests we can see checking account is the most predictive variable and that savings/ install payment ratio are also very predictive. In consumer datasets, using Random Forests often shows that income is one of the most important predictors of bad loans (note how this is in contrast to the consumer credit scoring literature reviewed earlier).

Step 2: Test out interactions.

In this case we can create an interaction term using all 4 affordability related measures: installment expense ratio, housing, checking, and savings into a new variable called afford.

Step 2a:
We can run Random Forest to now include this new attribute and see the variable importance:



The new afford[ability] variable is in the top 10 most predictive variables. It is the fifth most important predictor in terms of model accuracy. Next the credit score model using a simple logistic regression of all variables is run on the entire data set. This will be our base case credit scorecard without affordability interaction terms.

Base Case Scorecard: The credit score model without the interaction term on the whole data set is as follows:

```
glm(formula = good_bad ~ . - afford, family = binomial, data = c)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.6491  -0.7309   0.3983   0.7063   2.0347

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.812e+00  1.081e+00  -4.452 8.51e-06 ***
checking     5.765e-01  7.261e-02   7.939 2.03e-15 ***
duration    -2.598e-02  8.925e-03  -2.911 0.003603 **
history      3.956e-01  8.889e-02   4.451 8.56e-06 ***
purpose1     1.594e+00  3.585e-01   4.446 8.74e-06 ***
purpose2     7.363e-01  2.510e-01   2.933 0.003354 **
purpose3     8.547e-01  2.392e-01   3.573 0.000353 ***
purpose4     5.414e-01  7.467e-01   0.725 0.468434
purpose5     2.338e-01  5.500e-01   0.425 0.670728
purpose6    -1.361e-01  3.864e-01  -0.352 0.724786
purpose8     1.686e+00  1.154e+00   1.461 0.144148
purpose9     7.167e-01  3.193e-01   2.245 0.024788 *
purposeX     1.242e+00  7.285e-01   1.705 0.088123 .
amount      -1.157e-04  4.185e-05  -2.765 0.005688 **
savings      2.460e-01  6.023e-02   4.085 4.41e-05 ***
employed     1.464e-01  7.385e-02   1.982 0.047455 *
installp    -3.031e-01  8.504e-02  -3.564 0.000365 ***
marital      2.399e-01  1.196e-01   2.005 0.044962 *
coapp        2.997e-01  1.824e-01   1.644 0.100277
resident    -3.518e-02  8.024e-02  -0.438 0.661065
property    -1.741e-01  9.427e-02  -1.847 0.064803 .
age          1.210e-02  8.476e-03   1.427 0.153531
other        3.393e-01  1.134e-01   2.992 0.002767 **
housing      2.983e-01  1.716e-01   1.739 0.082112 .
existcr     -2.060e-01  1.648e-01  -1.250 0.211322
job         -5.301e-02  1.414e-01  -0.375 0.707660
depends     -1.121e-01  2.371e-01  -0.473 0.636390
telephon     3.226e-01  1.936e-01   1.666 0.095618 .
foreign      1.371e+00  6.123e-01   2.239 0.025139 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1221.73  on 999  degrees of freedom
Residual deviance:  919.21  on 971  degrees of freedom
AIC: 977.21

Number of Fisher Scoring iterations: 5
```

Now we build a modified scorecard using logistic regression including the new affordability interaction term.

Adding Affordability Term to Regression

```
glm(formula = good_bad ~ ., family = binomial, data = c)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.6551  -0.7307   0.3982   0.7071   2.0363

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.779e+00  1.162e+00  -4.112 3.92e-05 ***
checking     5.722e-01  9.135e-02   6.264 3.75e-10 ***
duration    -2.603e-02  8.948e-03  -2.909 0.003629 **
history      3.952e-01  8.910e-02   4.435 9.20e-06 ***
purpose1     1.595e+00  3.586e-01   4.447 8.73e-06 ***
purpose2     7.363e-01  2.510e-01   2.934 0.003349 **
purpose3     8.549e-01  2.392e-01   3.574 0.000352 ***
purpose4     5.423e-01  7.472e-01   0.726 0.467943
purpose5     2.327e-01  5.503e-01   0.423 0.672392
purpose6    -1.346e-01  3.870e-01  -0.348 0.728063
purpose8     1.683e+00  1.155e+00   1.457 0.145242
purpose9     7.162e-01  3.194e-01   2.243 0.024917 *
purposeX     1.243e+00  7.283e-01   1.707 0.087864 .
amount      -1.157e-04  4.186e-05  -2.762 0.005736 **
savings      2.406e-01  9.312e-02   2.583 0.009785 **
employed     1.463e-01  7.386e-02   1.981 0.047620 *
installp    -3.060e-01  9.325e-02  -3.282 0.001032 **
marital      2.402e-01  1.197e-01   2.007 0.044792 *
coapp        2.990e-01  1.826e-01   1.638 0.101416
resident    -3.531e-02  8.025e-02  -0.440 0.659982
property    -1.738e-01  9.433e-02  -1.843 0.065359 .
age          1.208e-02  8.478e-03   1.425 0.154200
other        3.393e-01  1.134e-01   2.992 0.002770 **
housing      2.938e-01  1.813e-01   1.620 0.105199
existcr     -2.053e-01  1.650e-01  -1.244 0.213387
job         -5.331e-02  1.414e-01  -0.377 0.706147
depends     -1.140e-01  2.384e-01  -0.478 0.632615
telephon     3.228e-01  1.936e-01   1.667 0.095481 .
foreign      1.370e+00  6.123e-01   2.237 0.025271 *
afford       3.751e-04  4.891e-03   0.077 0.938872
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1221.7  on 999  degrees of freedom
Residual deviance:  919.2  on 970  degrees of freedom
AIC: 979.2
```

Note that just like in prior research, as in “The use of Affordability data-does it add real value? By Wilkinson and Tingay), this new affordability interaction term itself does not appear to not be statistically significant at an initial pass. But we know from the Random Forest that this field has predictive power.

The trick is in finding the interactions which yield this predictive power.

So now we complete this step by testing out interactions with other variables in the credit scorecard and the affordability interaction term. The yields the following regression model.

Testing out Interactions of Affordability and Credit Scorecard Variables

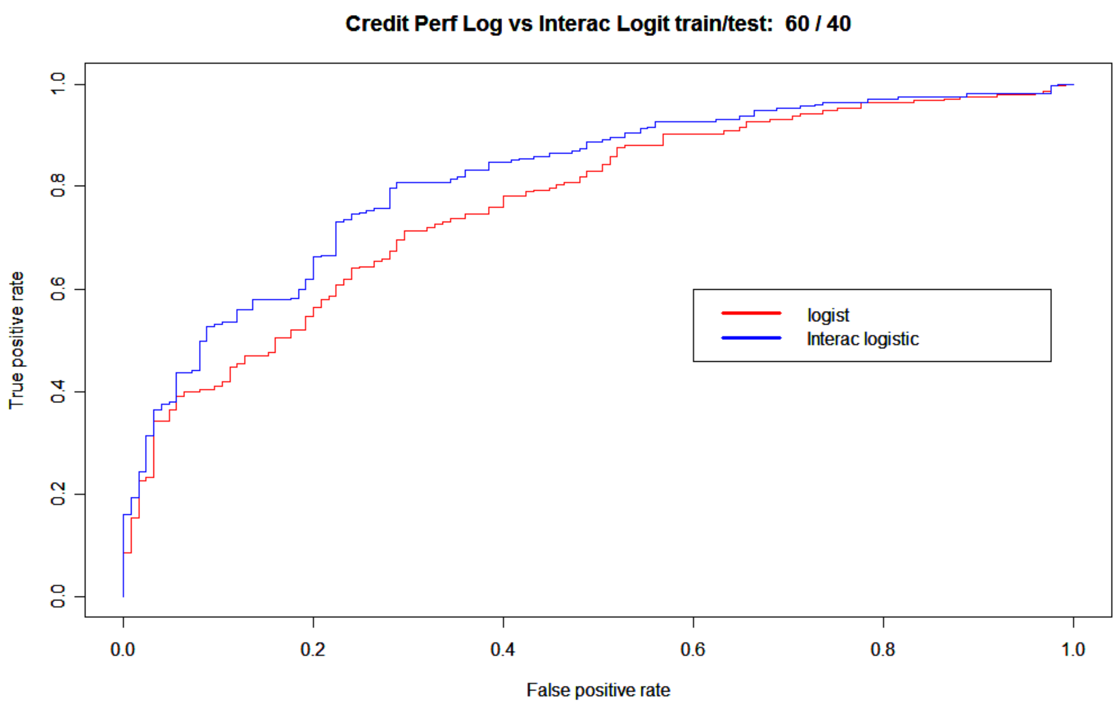
```
> m<-glm(good_bad~.*afford,data=c,family=binomial)
Warning message:
fitted probabilities numerically 0 or 1 occurred in: glm.fit(x = X, y = Y, weights = weights, start = start, etastart =
etastart,
> summary(m)

Call:
glm(formula = good_bad ~ . * inc, family = binomial, data = c)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.5522  -0.6605   0.3526   0.6869   2.2813

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -5.438e-01  1.867e+00  -0.291 0.770821
checking      4.399e-01  1.529e-01   2.878 0.004007 **
duration     -5.718e-02  1.229e-02  -4.651 3.31e-06 ***
history       5.122e-01  1.230e-01   4.165 3.12e-05 ***
purpose1      1.650e+00  4.605e-01   3.583 0.000339 ***
purpose2      3.839e-01  3.546e-01   1.083 0.278951
purpose3      8.344e-01  3.275e-01   2.548 0.010829 *
purpose4     -7.027e-01  1.258e+00  -0.559 0.576344
purpose5      7.922e-01  6.976e-01   1.136 0.256134
purpose6      4.723e-02  5.324e-01   0.089 0.929312
purpose8     -4.148e+01  1.465e+03  -0.028 0.977411
purpose9      1.033e+00  4.516e-01   2.289 0.022108 *
purposeX      9.976e-01  1.059e+00   0.942 0.346207
amount       -6.498e-05  5.260e-05  -1.235 0.216690
savings       6.963e-02  1.645e-01   0.423 0.672098
employed      5.131e-02  1.001e-01   0.512 0.608392
installp     -4.426e-01  1.345e-01  -3.291 0.000998 ***
marital       1.933e-01  1.601e-01   1.207 0.227348
coapp         4.956e-01  2.221e-01   2.232 0.025639 *
resident     -7.290e-02  1.079e-01  -0.676 0.499163
property     -2.979e-01  1.279e-01  -2.329 0.019870 *
age           1.065e-02  1.116e-02   0.954 0.340008
other         1.490e-01  1.584e-01   0.940 0.347064
```

```
housing       3.379e-01  2.489e-01   1.357 0.174643
existcr      -2.169e-01  2.214e-01  -0.980 0.327220
job           1.871e-02  1.925e-01   0.097 0.922575
depends       7.812e-02  3.139e-01   0.249 0.803471
telephon      1.628e-01  2.629e-01   0.619 0.535707
foreign      -1.139e+00  1.262e+00  -0.903 0.366669
afford       -3.918e-01  2.638e-01  -1.485 0.137473
checking:afford -2.292e-03  4.506e-03  -0.509 0.611025
duration:afford  1.306e-03  3.816e-04   3.421 0.000624 ***
history:afford  -3.248e-03  3.677e-03  -0.883 0.377047
purpose1:afford  7.535e-03  1.181e-02   0.638 0.523412
purpose2:afford  2.110e-02  1.403e-02   1.504 0.132549
purpose3:afford  9.581e-04  8.743e-03   0.110 0.912741
purpose4:afford  6.771e-02  6.372e-02   1.063 0.287980
purpose5:afford -1.380e-02  1.400e-02  -0.986 0.324194
purpose6:afford -1.330e-03  1.218e-02  -0.109 0.913044
purpose8:afford  7.067e+00  2.219e+02   0.032 0.974596
purpose9:afford -9.181e-03  1.156e-02  -0.794 0.426919
purposeX:afford  3.609e-03  3.428e-02   0.105 0.916168
amount:afford   -2.103e-06  1.416e-06  -1.485 0.137453
savings:afford  -1.366e-03  3.637e-03  -0.375 0.707331
employed:afford  4.077e-03  2.896e-03   1.408 0.159176
installp:afford  1.072e-03  5.416e-03   0.198 0.843062
marital:afford   1.349e-03  4.993e-03   0.270 0.786977
coapp:afford    -9.129e-03  6.627e-03  -1.378 0.168350
resident:afford  3.990e-03  3.404e-03   1.172 0.241171
property:afford  4.565e-03  3.910e-03   1.167 0.243041
age:afford       1.903e-05  2.998e-04   0.063 0.949379
other:afford     7.445e-03  4.264e-03   1.746 0.080769 .
housing:afford  -1.625e-02  9.396e-03  -1.730 0.083663 .
existcr:afford   7.805e-04  6.485e-03   0.120 0.904197
job:afford      -1.064e-03  5.978e-03  -0.178 0.858760
depends:afford  -8.995e-03  7.532e-03  -1.194 0.232409
telephon:afford  4.783e-03  7.630e-03   0.627 0.530765
foreign:afford   3.900e-01  2.597e-01   1.502 0.133159
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1221.73  on 999  degrees of freedom
Residual deviance:  869.67  on 942  degrees of freedom
AIC: 985.67

Number of Fisher Scoring iterations: 19
```

Using this model we can see that the affordability interactions terms show a significant relationship between duration and the affordability term.

Out of Sample Testing

The score card is rebuilt using the original base case scorecard with the affordability term alone and compared with a model with the interaction term of affordability and duration. These models are then tested out on a 40% hold out sample. This is important as large data sets are needed for logistic regression to ensure the interactions are not overfitting the training data. Also it is good to point out that Random Forests never overfit and thus by using them as a guide it guarantees a smaller and stable search space for variable interactions.



Conclusion

As seen in the results above unlike past consumer credit risk research we show that clearly affordability is an important factor in any credit decision. Given the fact that lenders do not have complete or clean data collection for consumer credit scoring (as it is not customary yet to collect/verify income for consumer debt unlike mortgages) the affordability data is usually messy or noisy. Techniques such as binning or discretizing affordability data used in traditional in credit scoring deal with this issue well but do not deal with multi-collinearity in an effective manner. The fact that affordability is pervasive and has potential interactions with other credit scoring variables has not been studied to date and is based on reasonable a priori expectation. From this viewpoint affordability data like income, assets, and ratios have predictive power which can been seen using Random Forests variable importance metrics. This highlights the fact that the predictive power of affordability data has not been well studied and exploring interactions is a productive approach of finding how to leverage the power of affordability data in credit scoring. We show that affordability does matter in a significant way if it is properly used/analyzed. This is in contrast to the past work stating affordability only provides marginal predictive power (Wilkinson and Tingay,2001).

Note: This approach has been used on proprietary application scoring, behavior scoring, and collection modeling on large data sets and found to be effective in building statistically significantly better models. Some interactions found to be useful are as follows: Free cash flow surrogates based on Monthly Discretionary income* expected life of loan*debt ratio + deposits, Discretionary income as % of debt

Deposits to Debt ratio, Loans in States with higher than average unemployment (macro variables),interaction of cash flow variables and Credit scores such as FICO and bankruptcy scores was especially strong similar to the superscorecard work by Hand of combining scores(Hand,2002).

Future Work on Algorithmics

Extensions of this approach naturally lend to exploring permutations of affordability variables using a genetic algorithm or creating automated approach to mining interactions where math functions (+,-,/,multiply are attempted on a small set of carefully chosen affordability or numeric datasets).

Normative Suggestions for Policy implications

Using the approach described here econometrician's can improve current credit scorecards in a statistically significant manner and obtain greater predictive accuracy from current data. That said it is the view of the authors that from a systemic view of risk simply improving models will not be sufficient in ensuring soundness of credit markets. Instead of having credit risk being managed by proprietary models which predict consumer behavior it is more optimal to build a white box credit system based on economic theory which ensure no customer is given an unaffordable or unsustainable loan. This is imperative given the prisoner's dilemma problem which occurs in the industry resulting in companies competing on credit risk and growing by deteriorating credit standards(Joseph, 2007). The recent work on showing the inevitability of underpricing of mortgage risk shows that risk based pricing is not reliable alone and there is an inherent tendency for systemic risk(Pavlov, 2005). Given this, a mandated open approach to ensuring affordability as a basis for loans alone is a necessity. This would be a simple model and for a great example of what such a standard would look like see Warren and Tyagi's model for consumers based on extensive research on personal bankruptcy (Warren, Tyagi, 2005). Although this book was written as self-help it has comprehensive worksheet for determining ratios for soundness for e.g. borrowers should spend 20% of after tax pay for savings, and have must have expenses below 50%, leaving borrowers with 30% of after tax free cash flows for discretionary consumption spending). These cut offs allow for temporary adjustment due to life events and also provide cushions against income shocks. This is inline with Guttentag's recommendation for underwriting systems to have to min/max cut offs on risk variables along with explicit projections for macroeconomic scenarios as part of underwriting (Guttentag, 1992). One extension to this would be an axiomatic specification of risk calculus to tell lenders how to make pareto optimal trade offs in credit risk. One approach to this would be making risk function monotonic and specify allowable risk trade offs (e.g. self employed borrowers can only borrow $10,000 for credit cards, if the LTV is high then a higher credit score is needed; any trade off in risk must be pareto optimal and layering of risk parameter combinations must be specified in this white box credit rule book based on affordability). This will help ensure credit markets stay on the right track. In terms of pricing risk once a sound credit guideline is adopted universally, then pricing strategies can result in break even credit pricing as should exists in a competitive market (Emms, Haberman, and Savoulli, 2007).

Credit Unions as a Model for Credit Markets built on Affordability

A wonderful example of the success of using affordability as a criterion for loan decision is the credit union industry. Although the industry is not as sophisticated as for profit banks it is thriving in the current marketplace and has delinquency and default rates well below the banking industry(see NCUA data vs. FDIC metrics). Credit Unions use crude rough calculations such as discretionary income and traditional ratios to underwrite loans along with some semi-custom or generic scores (Overstreet, Rubin, 1991; Desai, Crook, Overstreet, 1997). Emmons argued the reason credit unions are risk averse are due to the fact risk taking is not rewarded in credit union objective function and compensation and also that managers have more to lose in the credit union by taking on risk and as such there is pressure to make good loans (Emmons, 1999). In fact Emmons found that as credit unions grew their allowance for loan losses shrank as proportion of assets and members (1999). This is consistent with Smith, Cargill, and Meyer's work showing the objective function of credit unions is balanced by net gain to savers and net gain to borrowers (Smith etal, 1981). Interestingly even Fair Isaac agrees that FICO is not sufficient and must be used within sound underwriting (Quinn, 2000).

Credit Risk is Not a Basis for Competitive Strategy: Why People should not compete on Credit Risk

From a welfare point of it makes sense to disallow competition no credit risk as this does not create benefit to the entire system. Once an open lending criterion for credit policy is adopted then competition would be simply on cheaper service and greater automation and convenience and not on credit standards. This is based on the premise that the net benefit to all participants with open credit standards, open to borrowers and lenders, will create greater benefit than a closed system of proprietary competition. The most optimal approach to vending credit would be an open source transparent one, with the second most optimal solution being shared credit systems used by all credit vendors.

Note that in the best economic boom period a lot of loans which credit unions turned down, for not meeting affordability criteria despite high credit scores, would appear as false positive but in fact were loans in which borrowers or members could not withstand income or macroeconomic shocks.

Case Study of Safe Mortgage Product

To show an example of the type of open credit standards built on affordability we conclude this paper with an example of a 100% LTV loan that is safe (Stein, 1995). A competing issue for mortgage lending has been that downpayments greatly reduce default i.e. credit risk . The down side of this can be that home owners are exposed to housing price risk and volatility and also as housing represents the largest asset in many household portfolios it is unfortunate that housing is correlated to income as this creates even more risk for families (Caplin, 1997). One simple way to create mortgage loans which ensure that borrower well being is not compromised is to take the 20% down payment and have it put in a risk free investment account and have borrowers pledge to only use it in case of unemployment or other family emergencies (medical, death etc).

Under this approach the borrower would get a 100% LTV loan but have the risk profile of a 80% LTV borrower. The lender would benefit with higher loan amount to charge interest on, the mortgage insurer would still get mortgage insurance, and the borrower would have built in liquid reserves in case of income or employment shocks. This ensures welfare for borrowers and also reduces systemic risk. Systemic risk avoided by better risk management at the micro level instead of trying to managing unpredictable macro risks. This approach also addressed the risk pointed out by Ortalo-Magne's finding that "relaxation of down payment constraint triggers a boom-bust cycle" (Ortalo-Magne, 2001). Macroeconomic risks are black swan like to quote Nassim Taleb. Who could have predicted the patterns found in asian crisis of 1990s by Jian Chen would occur in the US as well (Chen, 2004). The pattern of "high foreign borrowing …[along with a] real estate boom while the economy has slowed down considerably … while the currency becomes weaker" seem eerily familiar now in the US (Chen, 2004).

This is a relatively simple example and yet shows the power of open affordability based credit standards and complex interactions of affordability and macroeconomic risks (home prices, unemployment etc).

This paper started with empirical technique to improve proprietary credit models. We then stepped back and questioned whether this is the best approach to tackle the problem of systemic risk and concluded that affordability should be built into the foundation of loan underwriting. By building a robust foundation of micro-motives and individual welfare, economic systems can be engineered to jointly optimize more holistic goals such as lack of financial crisis as well as borrower and lender well being. As recent work on constraint theory has shown adding constraints to utility maximization can result in greater utility being maximized than if utility maximization was treated as a unconstrained optimization problem (Sharma, 2009).

Bibliography

Altman, E. (2003). Z-Score: The Old Reliable Method for Predicting Bankruptcy Still Works. Ioma's Report on Managing Credit, Receivables, and Collections.

Asuncion, A. & Newman, D.J. (2007). UCI Machine Learning Repository [http://www.ics.uci.edu/~mlearn/MLRepository.html]. Irvine, CA: University of California, School of Information and Computer Science.

Breiman,L. (2001). Random Forests. Machine Learning, 45(1): 5–32

Breiman,L. (2002) Wald 2: Looking Inside the Black Box. Retrieved from www.stat.berkeley.edu/users/breiman/wald2002-2.pdf

Buel, W.D. (1968) Evaluating Mortgage Loan Risk. Journal of Applied Psychology.

Buist, H. Yang, T. Megbolugbe (1998)
An Analysis of the Ex Ante Probabilities of Mortgage Prepayment and Default
Real Estate Economics, Vol. 26

Caplin, A. Chan, S. Freeman, C. Tracy, J. (1997) Housing Partnerships. MIT Press.

Crook, J. Bellotti,T. (2008) "Credit scoring with macroeconomic variables using survival analysis".Journal of the Operational Research Society advance online publication 10 December 2008; doi: 10.1057/jors.2008.130

Chen, Jian. (2004) Credit Distortion and Financial Risk. International Review of Financial Analysis 13

DeVaney, S.A. (1994). The usefulness of Financial Ratios as Predictors of Household Insolvency: 2 Perspectives. FINANCIAL COUNSELING AND PLANNING Vol 5, 1994.

Dunn, L. (1999) AN EMPIRICIAL INVESTIGATION OF CREDIT CARD DEFAULT. Retrieved from http://www.econ.ohio-state.edu/pdf/credit.pdf

Dunn, L.F; Kim, T.H. (1999) An Empirical Investigation of Credit Card Default.

Emmons, W.R. (1999) Wages and Risk Taking in Occupational Credit Unions: Theory and Evidence. Federal Reserve Bank of St. Louis. Review.

Emms, P. Habermann, S., Savoulli,L. (2007). Optimal Strategies for Pricing General insurance. Insurance Mathematics and Economics.

Eisenbeis, R.A. (1977) Pitfalls in the Application of Discriminant Analysis in Business, Finance, and Economics. Journal of finance.

Fannie Mae 2002 Annual Report

Finlay, S.M. (2006). Predictive Models of Expenditure and Over-Indebtedness for Assessing the affordability of New Consumer Credit Applications. *Journal of Op. Res. Society* 57.

Gayler, R. (1995) Is the Wholesale Modeling of interactions Worthwhile? (Proceedings of Conference on Credit
Scoring and Credit Control, University of Edinburgh Management School, U.K.)

Gentry, James, Newbold, P., Whitford, D.T., (1985). Predicting Bankruptcy: if Cash Flow's not the bottom line, What is? Fin. Analysts Journal. 87.

Getter, D.E. (2003) Contributing to the Delinquency of Borrowers. Journal of Consumer Affairs

Gombola, M.J. Haskins, M.E., Ketz, J.E., Williams, D.D. (1987). Cash Flow in Bankruptcy Prediction. Financial Management.

Goodwin, D. (1999) Predictors of Household's Debt Repayment Difficulties
Canner, Glenn (1990) Consumer Debt Repayment Woes: Insights from a Household Survey;

Greninger, S.A; Hampton, V.L; Kitt, Karrol; Achacoso, J.A. (1996) Ratios and Benchmarks for Measuring the Financial Well-Being of Families and Individuals. Financial Services Review, 5(l): 57-70

Grieb, Terrance (2001) macroeconomic factors, consumer behavior, and bankcard default rates. Journal of Economics and Finance

Guttentag, J. (1992). When Will Residential Mortgage Underwriting Come of Age?. Housing Policy Debate. Vol 3.

Hababou, M; Cheng, A; Falk, R., (2006) Variable Selection in the Credit Card Industry. Royal bank of scotland.

Hakim, S. Haddad, M. (1999) Borrower Attributes and the risk of Default of Conventional Mortgages. Atlantic Economic Journal, 1999, vol. 27, issue 2, pages 210-220

Hand, D. J. and M. G. Kelly, 2002, “Superscorecards,” IMA Journal of Management Mathematics, 13, 273-281

Hanna, Sherman & Bae, MiKyeong & Lindamood, S. (1997) Patterns of Overspending in U.S. households. Financial Counseling and Planning, 4

Hassler, H.W.; Myers, J.H., Seldin, M. (1963). Payment History as a Predictors of Credit Risk. Vol 47. Journal of Applied Psychology

Henebry, K.L. (1996). Do Cash Flow Variables Improve the Predictive Accuracy of a Cox Proportional Hazards Model for Bank Failures? Quarterly Review of Economics and Finance. 36.3.

Hodges, C. Clusky, G.R, Lin, Bing-Xuan (2005) Analyzing Bankruptcy Predictors Using Time Series Data. Journal of Accounting and Finance Research. Vol 13. No. 1.

Joseph, M. (2007)Attaining sharper Pricing-for-Risk in the UK Sub-Prime Mortgage Market. Retrieved from
http://www.crc.man.ed.ac.uk/conference/archive/2007/abstracts/joseph.pdf.

Lane, S. (1972) Submarginal Credit Risk Classification. Journal of Financial and Quantitative Analysis Vol. 7

Langrehr, F.W. & Langrehr, V.B. (1989). Measuring the Ability to Repay: The Residual Income Ratio. Journal of Consumer Affairs Vol. 23

Liu, Yi-Kang, Jabbour, G.M., Green, R.K. (2001). The Performance of Default Risk Structural Models on Commercial Mortgages: An Empirical Investigation

Lovie, A.D., Lovie, P., 1986. The flat maximum effect and linear scoring models for prediction, Journal
of Forecasting, 5 (3) 159 - 168.

Moffatt, P.G. (2005) Hurdle Models of Loan Default. Journal of Royal Operational Research Society. 56.

Monaghan, J.E. (200) The impact of Personal Credit History on Loss performance in personal lines. http://www.casact.org/pubs/forum/00wforum/00wf079.pdf

Myers, J.H. (1963) Predicting Credit Risk with a Numerical Scoring System. Journal of Applied Psychology

Myers, J.H. (1964). A factor analysis of retail credit application data. Journal of Applied Psychology vol 48 168-170

Myers, J.H., Forgy, E.W. (1963). The Development of Numerical Credit Evaluation System. Journal of the American Statistical Association. Vol. 58.

Myers, J.H. (1967). Decision Models for Credit Scoring System Cutoff Determination. Journal of Applied Psychology

Ortalo-Magne, F., Rady, S. (2001). Housing Market Dynamics on the Contribution of Income shocks and Credit Constraints. Cesfo. Retrieved from http://www.sfbtr15.de/dipa/50.pdf

Overstreet, GA; Kemp, RS; (1986) Managerial control in Credit Scoring Systems. Journal of Retail Banking

Overstreet, G.A.J., Bradley, E.L., Kemp, R.S., 1992. The flat-maximum effect and generic linear scoring
models: a test, IMA Journal of Mathematics Applied in Business & Industry, 4 (1) 97 - 109

Overstreet, GA; Bradly, E. (1996) Applicability of Generic Linear Scoring Models in the U.S credit-union environment. IMA Journal of Math Applied in Business and Industry. 7.

Overstreet, G. and G.M. Rubin, 1991, Blurred Vision: Challenges in credit union research and modelling (The Filene Research Institute, Madison, WI).

Overstreet, G. Desai, V., Crook, J.,(1997). Credit scoring models in the credit union environment using neural networks and genetic algorithms. *IMA Journal of Mathematics Applied in Business and Industry*, 8(4), 324–346.

Paquin, P.; Weiss, M.S. (1998) Personal Bankruptcies: Study Finds Four Key Determinants

Pavlov, Andrey and Susan Wachter. (20050. “The Inevitability of Marketwide Underpricing of Mortgage Default Risk.” *Real Estate Economics*, 34(4), 479-96

Quinn, T., Disk, L. (2000) How will scores perform in a recession? . Retreived from www.camb-gmba.org/How_FICO_Scores_Perform_in_a_Recession.pdf

Sharma, P. (2008) Constraint Theory and Economics. Retrieved from http://idisp.parsuram.com/ConstraintTheory1.htm

Smith, D.J., Cargill, T.F., Meyer, R. (1981) Credit Unions: An Economic Theory of a Credit Union. Journal of Finance XXXVI

Smith, Paul (1964) Measuring Risk on Consumer Installment Credit. Management science.

Stein, J.C.(1995). Prices and Trading Volume in the Housing Market: A Model with Down Payment Effects. Quarterly Journal of Economics.

Stepanova, M.; Thomas, L. (2001) PHAB Scores: Proportional Hazards Analysis Behavioral Scores. The Journal of the Operational Research Society, Vol. 52

Sullivan, A.C.; Fisher, R. (1987) Consumer Credit Delinquency Risk: Characteristics of Consumers Who Fall Behind.

Taff, L.G. (2002) Investing in Mortgage Securities . CRC Press .FL

Taleb, N. (2007) The Black Swan. Random House.

Wainer, H. (1978). On the Sensitivity of Regression and Regressors. Psyc. Bulletin Vol. 85.

Warren, E. Tyagi, A.W. (2005) All Your Worth: The Ultimate Lifetime Money Plan .Free Press

Wilkinsin, G. Tingay, J.(2001) The use of affordability data-does it add value? Proceedings of Credit Scoring and Control VII, Credit research centre, University of Edinburgh. Retrieved from http://books.google.com/books?id=urdhL74nZOAC&pg=PA64&lpg=PA64&dq=does+affordability+add+real+value+to+the+lending+decision&source=web&ots=fXKu2_2HVm&sig=9FDVkdCzcJzIwWouU6wpcTnvOOo&hl=en&sa=X&oi=book_result&resnum=1&ct=result#PPA63,M1

Wilcox, J.W. (1971)] A gambler's ruin prediction of business failure using accounting data. Sloan Management Review

Worden, Debra; Sullivan,A.C. (1995) Credit Cards and the Option to Default. Financial Services research.

Zhao, H. Sinha, A.P., Ge, W. (2009). Effects of Feature Construction on Classification Performance: An Empirical Study in Bank Failure Prediction. Expert Systems with Applications.

Appendix: German Credit Data

http://ocw.mit.edu/NR/rdonlyres/Sloan-School-of-Management/15-062Data-MiningSpring2003/94F99F14-189D-4FBA-91A8-D648D1867149/0/GermanCredit.pdf

Variable Type Code Description

1

OBS#
Observation No.
Categorical

2

CHK_ACCT
Checking account status
Categorical
0 : < 0 DM
1: 0 < ...< 200 DM
2 : => 200 DM
3: unknown

3

DURATION
Duration of credit in months
Numerical

4

HISTORY
Credit history
Categorical
0: no credits taken
1: all credits at this bank paid back duly
2: existing credits paid back duly till now
3: delay in paying off in the past
4: critical account

5

NEW_CAR
Purpose of credit
Binary
car (new) 0: No, 1: Yes

6

USED_CAR
Purpose of credit
Binary
car (used) 0: No, 1: Yes

7

FURNITURE
Purpose of credit
Binary
furniture/equipment 0: No, 1: Yes

8

RADIO/TV
Purpose of credit
Binary
radio/television 0: No, 1: Yes

9

EDUCATION
Purpose of credit
Binary
education 0: No, 1: Yes
10

RETRAINING
Purpose of credit
Binary
retraining 0: No, 1: Yes
11

AMOUNT
Credit amount
Numerical
12

SAV_ACCT
Average balance in savings account
Categorical
0 : < 100 DM
1 : 100<= ... < 500 DM
2 : 500<= ... < 1000 DM
3 : =>1000 DM
4 : unknown
13

EMPLOYMENT Present employment since
Categorical
0 : unemployed
1: < 1 year
2 : 1 <= ... < 4 years
3 : 4 <=... < 7 years
4 : >= 7 years
14

INSTALL_RATE Installment rate as % of disposable
income
Numerical
15

MALE_DIV
Applicant is male and divorced
Binary
0: No, 1: Yes
16

MALE_SINGLE
Applicant is male and single
Binary
0: No, 1: Yes
17

MALE_MAR
Applicant is male and married or widower Binary
0: No, 1: Yes
Page 2
Var. # Variable Name

Description

Variable Type Code Description

18

CO-APPLICANT Application has a co-applicant
Binary
0: No, 1: Yes

19

GUARANTOR
Applicant has a guarantor
Binary
0: No, 1: Yes

20

TIME_RES
Present resident since - years
Categorical
0: <= 1 year
1<…<=2 years
2<…<=3 years
3:>4years

21

REAL_ESTATE
Applicant owns real estate
Binary
0: No, 1: Yes

22

PROP_NONE
Applicant owns no property (or unknown) Binary
0: No, 1: Yes

23

AGE
Age in years
Numerical

24

OTHER_INSTALL Applicant has other installment plan credit Binary
0: No, 1: Yes

25

RENT
Applicant rents
Binary
0: No, 1: Yes

26

OWN_RES
Applicant owns residence
Binary
0: No, 1: Yes

27

NUM_CREDITS Number of existing credits at this bank
Numerical

28

JOB

Nature of job
Categorical
0 : unemployed/ unskilled - non-resident
1 : unskilled - resident
2 : skilled employee / official
3 : management/ self-employed/highly
qualified employee/ officer

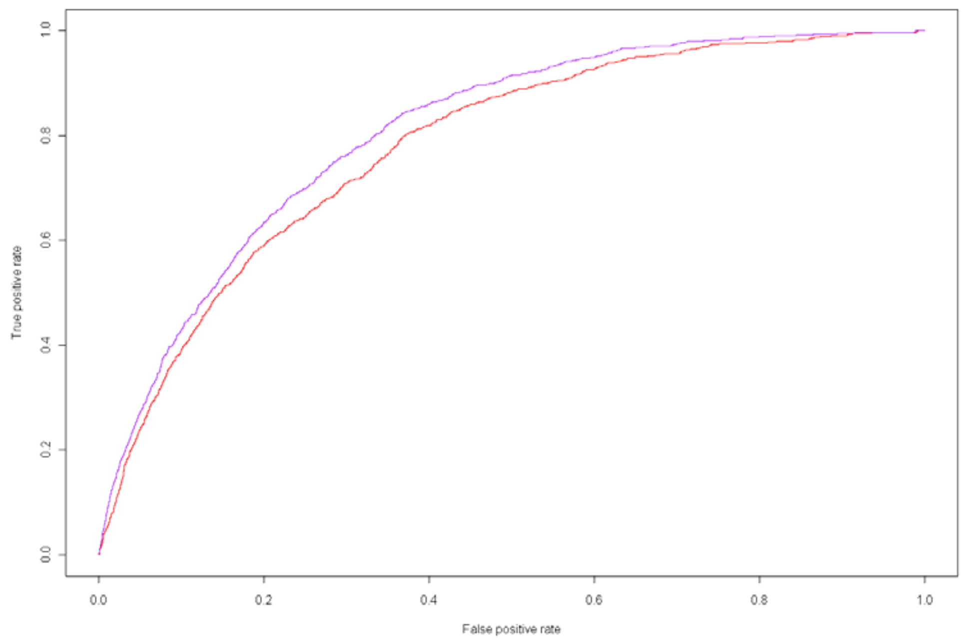
29
NUM_DEPEND Number of dependents
Numerical

30
TELEPHONE
Applicant has phone in his or her name Binary
0: No, 1: Yes

31
FOREIGN
Foreign worker
Binary
0: No, 1: Yes

32
RESPONSE
Fulfilled terms of credit agreement
Binary
0: No, 1: Yes
Binary
0: No, 1: Yes

Appendix of Improvements Using interactions on Large Proprietary Data Set



Appendix Showing That Logit Improved with Insights of Appropriate Interaction terms can even outperform Random Forests (RF based interactions is the logit with interactions, simple is the logit without interactions, and RF is Random Forests).