

Distinguishing between Good and Bad Subprime Auto Loans Borrowers: The Role of Demographic, Region and Loan Characteristics

Dr. *Yaseen Ghulam*^{a,b} (Correspondence author)

a. University of Portsmouth, Portsmouth Business School, Economics and Finance Subject Group,
Richmond Building, Portland Street, PO1 3DE, UK

Tel: +442392844127 E-mail: yaseen.ghulam@port.ac.uk

b. Al Yamamah University 7010 King Fahd Road, Al Qirawan, Riyadh 13541, SAUDI ARABIA

Sophie Hill

Inchcape Fleet Solutions, Haven House, Compass Road, Portsmouth, PO6 4RP, UK

E-mail: sophielh@hotmail.co.uk

Abstract: Research on subprime mortgages has recently been gaining momentum, but subprime auto loans have largely been ignored. By using a unique data set of a very large UK vehicle finance company, this study analyses secured loans extended to the subprime borrowers with impaired or limited credit history. It looks specifically at characteristics in relation to payment history, in order to determine what characteristics make a good or bad borrower. We conclude that married and divorced borrowers as well as borrowers living in low unemployment and relatively prosperous regions such as the South East and London are less likely to default compared to not married, furnished tenants or borrowers living in the North West of the UK who have a high probability of default. Similar to the prime loans, income of borrowers and defaults propensities are negatively associated. Loan and security characteristics with the most impact on default status are price and age of the automobile, effective interest rate measured by APR, loan-to-value (LTV) and term of the loan agreement. The results of this study will help in understanding subprime auto loans and borrowers as well as helping lenders to distinguish between good and bad subprime borrowers.

Keywords: Automobile loans; Defaults; Subprime; Credit risk; UK

JEL Classifications: D11, D81, G23

1. Introduction

Some recent studies have explored the impact of subprime borrowers demographic and loan characteristics on the performance of loans. Studies by Heitfield and Sabarwal (2004), Lax *et al.* (2004), Agarwal *et al.* (2008a), Danis and Pennington-Cross (2008), Amromin and Paulson (2009), Daghish (2009), Demyanyk and Hemert (2011), Capozza and Order (2011), Eriksen *et al.* (2013), and Bellotti and Crook (2013), fall in this category. But these studies are mainly concerned with the subprime mortgage market and ignore subprime consumer finance issues. The studies related to credit card defaults do not address the issues specifically related to secured auto subprime loans. By using a unique data set of a large UK subprime auto finance company supplying car finance agreements, this study intends to establish parameters that can be used by subprime lenders to distinguish a 'good' subprime auto loan applicant from a 'bad' subprime auto loan application. This study is the first UK study in recent times which provides a new insight in understanding the working of this market.

We find that marriage reduces the probability of subprime auto loans default as well as missing payments (arrears). Similarly borrowers who are either divorced and/or live in the South East and East of England and/or are homeowners, have low probability of default. Increase in income reduces the probability of default, going into arrears as well as early pay-off. Higher proportion of income coming from the state benefits relative to total income also increases the chances of default and write-off. Longer duration of the current employment reduces the probability of default and getting into difficulty in paying back loans. Lending for higher value cars increases the probability of defaults and arrears. Longer tenure of the loan reduces default as well as early pay-off occurrences. Surprisingly, higher APR results in fewer loan defaults probably indicating pricing of loans as per their risk. Furnished tenancy, loan origination from a relatively high unemployment areas such as the North West of England, higher LTV and purchase of older car, all increase default probabilities and the chances of missed payments. Similarly, borrowers in full-time employment are less likely to repay loans early while borrowers who live in the North West of England are more likely to repay loans early. We conclude that the behaviour of prime and subprime borrowers in regards to defaults and missed payments does not vary a lot despite the fact that subprime lending carries more risk.

The rest of the paper proceeds as follows. Section 2 contains literature review investigating the subprime loans and the risks of subprime borrowers. The methodology section explains the method chosen and evaluates the data structure. The results section looks at the findings in detail and compares these results with the existing empirical literature. The conclusion section summarises the findings and discusses limitations as well as identifying areas for further research.

2. Literature on Characteristics Associated with Defaults

In reviewing the literature on defaults on loans and its prediction, we individually assess each characteristic that potentially can affect a borrower's propensity of default. We look at personal characteristics, market conditions and macro-economic factors as well as any other factors covered by the literature in this area.

2.1 Individual characteristics

Generally, it is assumed that a borrower with lower income is more likely to default. Borrowers with a larger income and with a bank account are more likely to make payments (Einav *et al.*, 2012). Agarwal, *et al.* (2011) agreed with this and found that, in comparison with individuals with low income, individuals with high income are 17% less likely to default and 22% less likely to declare bankruptcy. Dinh and Kleimeier (2007) also agreed by finding that the individuals who have a lower income had defaulted more. Higher income and number of years in the current employment give a lower annual percentage rate (APR), as higher income indicates more resources available to repay, but also the length of employment shows stability and therefore seen to be less risky (Agarwal *et al.*, 2008b). From another perspective, Kočenda and Vojtek (2009) found that individuals who have high income, but also high expenditure are considered more risky. Similarly, some authors have looked at the borrower's employment status and their propensity to default. A borrower that is self-employed and entrepreneurial is seen as more risky and therefore has a higher chance of defaulting on the loan. This is due to the variable income, which can cause the borrower to be liquidity constrained more frequently (Capozza and Thomson, 2005). Zywicki and Adamson (2008) also looked at self-employed borrowers, and agreed that lenders may charge higher interest rates in such a case where the borrower is self-employed due to their less predictable income. Marshall *et al.* (2010) found that those who do not have a monthly salary are more likely to default.

The rate of default is expected to be lower for a debtor who is a homeowner (Agarwal *et al.*, 2011). Agarwal *et al.* found that a homeowner is 17% less likely to default and 25% less likely to

file for bankruptcy. They also found that an individual who lives in their state of birth are less likely to default by 9%, whereas those who have moved 190 miles from their state of birth are 17% more likely to default. Marshall *et al.* (2010) had similar findings, arguing those who are not homeowners, are more likely to default on their loans. The length of time of a borrower's current address can also affect loan risk. The findings of Marshall *et al.* (2010) suggest that a longer period at the current address reduces the likelihood of defaulting. However, Dinh and Kleimeier's (2007) study showed that default rates increase with an increase of time at current address in Vietnam. They conclude that this could be due to an individual's wealth being high or improving and therefore they may want to move to a better location or a larger home. When looking at residency status, Dinh and Kleimeier (2007) found that those who live with parents are less likely to default.

Agarwal *et al.* (2011) expected to find that debtors who are married to have a lower risk of default. They estimated married individuals to be 24% less likely to default and 32% less likely to become bankrupt. Kočenda and Vojtek (2009) also agreed with this, finding that married individuals are considered as less risky, although they consider that this could be because both incomes of the married individuals are being considered. In contrast to this view, Dinh and Kleimeier (2007) found that those who are married are more likely to default than single borrowers. They conclude that this could be due to an increased number of dependants and therefore more financial pressure. Yap *et al.* (2011) are in agreement with this view, finding that married individuals are more likely to default in comparison to single individuals. They also found that those who have no or one dependant are more likely to default. In a similar study, Marshall *et al.* (2010) found that those who have more children are more likely to default on their loans.

Agarwal *et al.* (2011) found that defaults and bankruptcies rise and fall over an individual's life cycle. Their study looked at age and revealed that the youngest (30 years and under) and the oldest (60 years and older) aged individuals have the lowest bankruptcy risk. The middle aged groups (30-60 years) are more likely to default and have a higher rate of bankruptcy. In other research, Agarwal *et al.* (2008b) had contrasting results, finding that younger and older adults tend to pay higher interest rates and fees, with the least risky age to be early 50's. Yap *et al.* (2011) agree with this, finding that those who are 53 and older have the lowest risk of default. Similarly, Kočenda and Vojtek (2009) conclude that education of the borrower is a strong predictor of default. Their results show that individuals with a higher level of education are less likely to default. Dinh and Kleimeier (2007) expected to find that individuals with a better education have a lower chance of default due to having more stable employment and the potential to earn a higher income. Their results showed that default did not consistently decline to a higher level of education, finding that college graduates have the highest default frequency. Marshall *et al.* (2010) found that students tend to be less risky borrowers.

2.2 Loan characteristics

Adams *et al.* (2009) looked at borrower attitudes to loans, concluding that borrowers are more likely to default on larger loans. This indicates that lenders should be more careful about the amount being lent and that they should cap loan sizes to prevent over borrowing. They also estimate that a larger down payment will decrease the likeliness of default. A larger down payment would decrease the LTV ratio, which is often used when deciding whether to grant credit or not. Bar-Gill (2008) also found that small down payments have been linked to increased delinquency and foreclosure rates. A higher LTV gives a lower percentage of collateral and this implies that a higher APR will be charged (Agarwal *et al.*, 2008b). Demyanyk and Van Hemert (2011) looked at whether the increase in high LTV lending caused the increased default which resulted in the financial crisis, and concluded that lenders were aware of the higher risk these loans involved. These loans with higher LTV also consisted of higher rates by the lenders as they were aware of the extra risk. Mayer *et al.* (2009) found that those with high LTV are more likely to default as they will find it more difficult

to refinance and they have less to lose through default. Rajal *et al.* (2010) looked at low LTV, stating that those who chose a lower LTV are likely to have greater wealth, and therefore are less likely to default. Einav *et al.* (2012) also conclude that loan size is one of the determinants of payment duration and the likeliness of default, with a larger loan, resulting in fewer repayments actually made and a higher chance of default. Marshall *et al.* (2010) conclude that purchasing behaviour is a good indicator of default risk, finding that borrowers who purchase less expensive items and are able to make a larger deposit are more likely to make repayments, therefore decreasing the risk of defaulting. In terms of the length of the loan, borrowers are more likely to default early into the loan, which is in line with the market assumption that default rates are higher in the first 12-24 months of the loan (Malik and Thomas, 2009).

2.3 Macroeconomic conditions

Capozza and Thomson (2005) found that on average, a subprime borrower has lower income and wealth and therefore they expect these borrowers to be less able to cope with stressful trigger events such as unemployment or divorce. Zywicki and Adamson (2008) conclude that a borrower may wish to repay the loan, but is unable to do so, which could be as a result of job loss. They believe that subprime borrowers are more likely to lose their job during recessions, and have fewer savings, making them unable to make the repayments of the loan. The studies by Agarwal *et al.* (2011) and Demyanyk and Hemert (2011) also confirm that unemployment increases the risk of default, looking at credit cards alone; they found that the risk of default and bankruptcy is significantly higher when living in a county with a high unemployment rate. Rajan *et al.* (2010) find that an economic decline, more specifically a fall in house prices, will cause an increase in default rates. Bonfim (2009) also agrees that economic decline increases the rate of default; however, this is due to the loosening of lending standards when the economy is strong. To conclude, she finds that when the economy experiences a downward trend, loans tend to perform worse and therefore lending criteria tightens. Malik and Thomas (2009) used the percentage change in the consumer price index over 12 months to reflect the inflation rate, to determine whether this affected the default probability on loans. They thought that high levels of inflation may cause a rise in the borrower default rates. Zywicki and Adamson (2008) found that some subprime borrowers find the interest rates of subprime loans, including any late payment penalties more attractive than the rates of other personal loans, such as payday lenders or personal finance companies. Capozza and Thomson (2006) also agreed with this view, suggesting that subprime borrowers could skip a payment to meet short-term liquidity needs.

3. Data and Modelling Approach

The data for this study was retrieved from a UK vehicle financing company supplying secured loans to UK borrowers with impaired credit, categorised as subprime. It covers the subprime loan agreements signed for the period October 2012 to October 2013 with more than 7% default rate within thirteen months of the signing of the loan contract. The data consisted of personal information about the borrowers, information about the vehicle which the loan is secured to and their repayment history over a specific time period. The sample covers the whole of the UK borrowers from the company. The company anonymised the data prior to release for our research.

The variables used for the regression contain a large variety of the borrower's information, with a sample size of 10,670. Continuous variables were used for quantitative data such as income, where a log was taken in order to make the interpretation more straightforward. A standard logistic regression function was used for modelling the determinants of defaults. It relates the set of potential predictor variables to a probability of a dichotomous outcome ($Y=0$ for not defaulted or in arrears and $Y=1$ for default or in arrears). This can be written as:

$$\log \left[\frac{P(Y = 1)}{1 - P(Y = 1)} \right] = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon \quad (1)$$

where $P(Y=1)$ is the outcome of interest. This equation is solved to obtain:

$$P(Y = 1) = \frac{1}{1 + e^{-z}}, \text{ where } z = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon \quad (2)$$

Table 1. Construction of default/arrears, early payment and write-offs predictors

Variabl	Description	Expected effect
ms1	marital status = single	+
ms2	marital status = married	-
ms3	marital status = divorced	+
ms4	marital status = separated	+
ms5	marital status = common law	-
ms6	marital status = widowed	+
rs1	residential status = homeowner	-
rs2	residential status = furnished tenant	+
rs3	residential status = unfurnished tenant	+
rs4	residential status = council tenant	-
rs5	residential status = living with parents	-
rs6	residential status = other	+
es1	employment status = full time employed	-
es2	employment status = part time employed	+
es3	employment status = self-employed full time	+
es4	employment status = self-employed part time	+
es5	employment status = temporary employment	+
es6	employment status = unemployed	+
a1	area = area of living - South	-
a2	area = area of living - South East and East	+
a3	area = area of living - Wales and West	-
a4	area = area of living - Midlands	+
a5	area = area of living - North	-
a6	area = area of living - North West	+
a7	area = area of living - North East	-
lmicr	months in current residence (log)	-
lmice	months in current employment (log)	-
ltinc	total income of the borrower (log)	-
lpincb	percentage of total income which comes from government benefits (log)	+
ldeposit	deposit paid by the customer towards the car (log)	-
lltv	loan to value (log)	+
lapr	Annual percentage rate (log and converted into monthly)	-
lterm	length of the agreement in months (log)	+
lcarprice	price of the car which the loan is secured to (log)	+
lcarage	age of the car which the loan is secured to, in months (log)	+
group	categories are numbered 1-6 where: 1 = live deal not in arrears 2 = live deal in arrears 3 = live deal previously in arrears (looking at previous 10 months) 4 = car repossessed or surrendered 5 = loan settled early (by customer, car dealer, insurance company) 6 = bad debt written off	
default	default status of the agreement constructed from group categories default status = 1 for groups 2,4,6 and non-default = 0 for groups 1,3,5	

Equation (2) above is used to determine the conditional probability of a borrower fitting into the category of a defaulter/non-defaulter based on different explanatory variables. The initial model

will fit two outcomes; default or currently in arrears and no-default or no current arrears. This classifies the borrowers into good and bad categories in regards to payment history and current status of the loan. Subsequently, multinomial logistic regression will be used to classify loan performance into further 5 categories (currently in arrears, previously in arrears, defaulted, repaid off, written off), which allows for more specific and useful information. Multinomial logistic regression tends to be used when the response variables have more than two values and there is no ordering of the categories (Chatterjee and Hadi, 2006). More specifically, the multinomial logistic regression will allow the modelling of different categories to specify which characteristics determine the probability that a borrower will end up in which category. Table 1(above) lists the variables used for this analysis. The *variable* column shows the abbreviation used, with a description explaining the variable and how it was constructed. The *expected* column indicates the expected effect each variable will have on default status. The expectations were determined from the empirical literature and also the personal judgement based on one of the authors' placement job with the finance company.

4. Estimation and Discussion

The regression results are split into two sub-sections. The first part discusses the logistic regression model results. The second part is a multinomial logistic regression model which analyses the role of borrower and loan characteristics in determining different loan statuses instead of just default/arrears or otherwise. Table 2 presents 4 different variations of the determinants of the loan default model. Model 1 is our main and complete model that includes all the person and loan specific variables, whereas in models 2-4, we drop some variables to test the stability of the regression coefficients against our main model specification and estimation. The first 5 variables (those with a prefix of 'ms') look at the marital status of the borrower. The marital status equals to common law is the reference category in this regard. Authors such as Agarwal *et al.* (2011) and Kočenda and Vojtec (2009) suggest that married borrowers are less likely to default. Perhaps married status would suggest that borrowers are more financially stable, as well as benefiting from the potential of having two main incomes plus additional state benefits. Referring to the results, the marital status that has a significant effect on default rates are those who are married (ms2) and divorced (ms3). In line with the majority of the literature in this area and expectations, marginal effects estimates indicate that a borrower who is married is 2.5% less likely to default compared to the not formally married subprime borrower. In models 2, 3 and 4, marginal effect estimate is also very close to the original model 1 estimate indicating the stability of these estimates. Similarly, a borrower who is divorced is 3.7% less likely to default. This could be due to a lower number of dependants¹ with less financial pressure (Dinh and Kleimeier, 2007), or a lone parent receiving generous state benefits or money from ex-partner.

The 4 variables with a prefix of 'rs' (rs1-rs4) look at the residential status in relation to default and arrears status. Home ownership is an indicator of financial strength and a signal of creditworthiness. The literature suggests that homeowners are seen to be less risky and therefore have a lower risk of default (Agarwal *et al.*, 2011; Marshall *et al.*, 2010). Hence, *a priori* and consistent with the literature, it is expected that homeowners are less likely to default. There is also suggestion in the literature that those living with parents are unlikely to default as they have minimal expenses (Dinh and Kleimeier, 2007). Looking at the results of the base model, the coefficients of those who are furnished tenants (rs2) are seen to be statistically highly significant and homeowners (rs1) marginally significant (in model 2 only). Referring to model 2, the results show that homeownership reduces the default probability, by 1.9%, showing stability and security

¹ We assume this fact that we do not have information on number of children of these borrowers.

and therefore enabling borrowers to be able to make regular repayments. The results also show that those who are furnished tenants are 1.9% more likely to default. Renting a furnished property is possibly the most expensive type of residency. This could cause high monthly expenditure and therefore rendering less available funds to pay the finance agreement, as rent is likely to be a priority in comparison to an automobile loan. Employment status (es1-es4) did not have a significant effect on default status. A significant number (88% of the 10,670) of borrowers in our sample are in full time employment. Those in full time employment are seen to be less likely to default, as full time employment offers job stability as well as a fixed monthly income to cover the cost of repaying the loan. Nonetheless, whether employed or unemployed, employment status has no significant bearing on default status in our case study.

Table 2. Determinants of subprime auto loans defaults/arrears

Variable ²	(1)		(2)		(3)		(4)	
	Coef.	M.E.	Coef.	M.E.	Coef.	M.E.	Coef.	M.E.
ms1	0.025	0.002	0.028	0.002	0.017	0.001	0.027	0.002
ms2	-0.391***	-0.025**	-0.397***	-0.026***	-0.393***	-0.025**	-0.389***	-0.025***
ms3	-0.577**	-0.037**	-0.578**	-0.037**	-0.584**	-0.038**	-0.578**	-0.037**
ms4	-0.310	-0.020	-0.316	-0.020	-0.332	-0.021	-0.308	-0.020
ms6	0.048	0.003	0.031	0.002	0.001	0.000	0.048	0.003
rs1	-0.268	-0.017	-0.289*	-0.019*	-0.277	-0.018	-0.267	-0.017
rs2	0.296**	0.019**	0.279*	0.018*	0.291**	0.019**	0.294**	0.019**
rs3	0.052	0.003	0.044	0.003	0.046	0.003	0.049	0.003
rs4	0.123	0.008	0.101	0.007	0.125	0.008	0.123	0.008
es1	-0.128	-0.008	-0.115	-0.007	-0.131	-0.008	-0.128	-0.008
es2	-0.428	-0.028	-0.429	-0.028	-0.438	-0.028	-0.428	-0.028
es3	0.201	0.013	0.247	0.016	0.181	0.012	0.197	0.013
es4	0.967	0.062	1.025	0.066	0.974	0.063	0.961	0.062
a2	-0.335***	-0.022**	-0.291**	-0.019**	-0.341***	-0.022***	-0.334***	-0.022***
a4	-0.192	-0.012	-0.189	-0.012	-0.143	-0.009	-0.192	-0.012
a6	0.434***	0.028***	0.528***	0.034***	0.463***	0.030***	0.425***	0.027***
lmicr	0.037	0.002	0.039	0.003	0.035	0.002	0.036	0.002
lmice	-0.119***	-0.008***	-0.118***	-0.008***	-0.122***	-0.008***	-0.120***	-0.008***
ltinc	-0.649***	-0.042***	-0.527***	-0.034***	-0.638***	-0.041***	-0.656***	-0.042***
lpincb	0.056**	0.004**	0.051**	0.003**	0.059***	0.004***	0.056**	0.004**
ldeposit	0.009	0.001	0.023	0.001	-0.036**	-0.002**	0.009	0.001
lltv	2.192***	0.141***	1.857**	0.120**			2.339***	0.151***
lapr	-0.451*	-0.029*	-0.773***	-0.050***	-0.566**	-0.036**		
lterm	-0.829**	-0.053**	-0.553	-0.036	-0.701*	-0.045*	-0.635*	-0.041*
lcarprice	0.728***	0.047***			0.618***	0.040***	0.853***	0.055***
lcarage	1.428***	0.092***	0.994***	0.064***	1.383***	0.089***	1.448***	0.093***
Constant	-17.084***		-8.895**		-6.183*		-20.091***	

Notes: Table above contains coefficients (Coef.) and marginal effects (ME) of a logit regression of a subprime auto loans borrower going into default or arrears. Significance level of each factor is determined by *p*-values. ***, **, and * indicate statistically significant at 1%, 5%, and 10% level of significance, respectively.

The geographical area in which the loan originated is included in our modelling of subprime automobile defaults as geographical location of the borrower could have a significant effect on default status as Agarwal *et al.* (2011) study found that the risk of default is significantly higher when living in a county with high unemployment rates. Looking at variables with a prefix of ‘a’, the regression results show that areas a2 and a6 are statistically significant. There are 7 areas in total in our sample, and these three were chosen for the regression as they had a higher number of defaults

² For the specific description of each variable, please refer to the details in Table 1.

in comparison to other areas. The regression coefficient a_2 which represents South East and East of England has a negative coefficient and therefore borrowers in this area are 2.2% less likely to default. In comparison to this, those living in a_6 , North West of England, are more likely to default by 2.8%. The average unemployment rate in the UK in 2013 was 7.1%. Looking particularly into the areas used in this regression, the unemployment rate for the South East is 5.3% and for East of England is 5.7%. These are both below the UK average unemployment rate and therefore it would be expected that borrowers in this area would be less likely to default, as shown in the results. The unemployment rate for the North West is 7.9% (The Telegraph, 2014), which is higher than the UK average and also explains why the results indicate that borrowers in the North West are more likely to default.

The time period of the borrower's current residence is also an important indicator of risk. The expectation is that a longer period of the current residence will indicate a lower risk of default, which aligns with the study of Marshall *et al.* (2010). A reason for this could be due to stability in personal life and employment status. The regression coefficient attached to the variable $lmicr$ in our regression model however is insignificant and therefore a longer period in the current residence has no statistically significant effect on default status. As we discussed above, employment status indicators were not able to explain the default outcome, we therefore include the length of time in current employment. The longer duration of current employment would indicate that the borrower has good job stability and therefore likely to have a regular income to be able to afford the monthly repayments. The studies looking at unemployment find that subprime borrowers are more likely to lose their jobs more frequently during a period of economic uncertainty and they also have fewer savings to continue to keep up with repayments (Zywicki and Adamson, 2008). Referring to the results in Table 2, $lmice$ is statistically significant and negative indicating that borrowers who have been in their current employment for a longer period of time are less likely to default. A 1% increase in current employment time will result in a 0.8% decrease in default probability.

Looking at the borrower income ($ltinc$) variable, it is expected that a higher income borrower would be less likely to default as they would have sufficient funds available to keep up with repayments. Studies by Morton (1975), Furstenberg and Green (1974), Chinloy (1995), Adams *et al.* (2007), Einav *et al.* (2012), Agarwal *et al.* (2011) and Dinh and Kleimeier (2007), all suggest that higher income borrowers are less risky, and therefore less likely to default on loans. The regression coefficient and marginal effect estimate also show that borrowers with a higher income are less likely to default. A 1% increase in total income would result in the probability of default being decreased by 4.2%. This is as expected, as a higher income could indicate stability because they are likely to be in a secure job.

In addition to looking at income, the contribution of the state benefits to total income, such as working tax credits, child benefits and other allowances could also be an important determinant of default outcome. This is particularly important for subprime borrowers with *ex ante* low income expectations. Income top-up from state benefits not only helps in paying back loans but also smooth current and future consumption. To calculate a percentage of how much income is from state benefits, we use the ratio of total benefits received from the government as a % of total income. In fact, this could be a risk proxy and therefore a higher risk of default if the majority of income is from government benefits rather than employment. The regression coefficient of $lpincb$ shows that a borrower who has more state benefits in relation to income is more likely to default. A 1% increase in benefits/income ratio will increase probability of default by 0.4% (not an alarming contribution to risk but nonetheless statistically significant). Therefore, higher benefits do indicate more risk, as borrowers are likely to be in a less stable economic situation.

LTV could also have an impact on default status. A higher LTV is an indication of larger loan size and less contribution from the borrower. A higher LTV also indicates the borrower will have

more to repay, which could become unaffordable in the time of economic uncertainty and potential loss of employment. Available evidence suggests that a borrower with a larger loan is more likely to default, and also a borrower with a larger deposit, which will decrease the LTV, will have lower probability of default (Adams *et al.*, 2009). Danis and Pennington-Cross (2008), Morton (1975), Furstenberg and Green (1974), Chinloy (1995), Calem and Wachter (1999), Amromin and Paulson (2009), Demyanyk and Hemert (2011) Bar-Gill (2008) also found higher LTV will increase probability of default. The regression coefficient and marginal effect of LTV support above mentioned studies, finding that a borrower with a higher LTV is more likely to default. A 1% increase of LTV will increase probability of default by 14.1%. Hence, lenders should be more careful when lending to subprime borrowers with lower deposits and a high LTV.

Our regression results for the price of loan measured by APR and shown by the variable *lapr* in Table 2 indicate that it has a significant influence on default status. We found that a higher APR will result in the borrower being less likely to default. A 1% increase in the APR will decrease the probability of default by 2.9%. The literature suggests that subprime borrowers would prefer to default on the cheapest loan as this will potentially cost them less (Zywicki and Adamson, 2008). The financing agreement that this study evaluates has a high default charge and also a late payment fee and therefore borrower may default on another loan in order not to incur these charges. These borrowers are also likely to be stable in their residency and employment status to be able to cover these payments. The results may also be affected by the fees charged when taking out the finance. Some borrowers are subject to a large fee at the start of the agreement which is added to the interest, and this could cause the APR to be higher especially for lower principal agreements.

The length of the loan agreement is also an important factor that could have implications for the subprime loan performance. As the majority of literature on subprime loans relates to mortgages, the loan length is likely to be much longer for a mortgage compared to a car financing agreement. The expectation would be that an agreement over a longer period is more likely to result in default. However, increasing the length of repayment can also decrease the payment amount which could help to make the repayments more affordable. Malik and Thomas (2009) found that borrowers are more likely to default early into the agreement. The impact of loan tenure is shown by the variable *lterm* in Table 2. We found that loans over a longer period are less likely to default than those that are repaid over a shorter term. A long term agreement can reduce the monthly payment amounts by spreading the total repayment over more months, which may make it more affordable for the borrower.

The purchased and current price of the collateral asset (car) could have a significant effect on the probability of default. Rajan *et al.* (2010) suggest that in the case of a mortgage, a decrease in the value of home would increase the likeliness of default. Looking specifically at the car price regression coefficient (*lcarprice*) and marginal effect estimate, it emerges that a higher purchased price of the car will increase the probability of default. A 1% increase in car price would increase the probability of default by 4.7%. One could argue that a higher purchased value car is likely to be bought by the borrowers who desire a more prestigious and reliable car and are able to afford it. They could also have paid a higher deposit towards the car. A newer and more expensive car is also likely to have fewer faults which suggest borrowers would have less reason for defaulting on payments. But higher monthly payments could make it harder for these borrowers to keep up with payments schedule over a relatively longer time period.

Looking further at the asset which the loan is secured to, the car age is another variable that could have a significant impact on default status. The car value could be an indicator of the borrower's economic and social status and employment stability, as they are likely to desire a newer car if they have more stable current and future earning and more likely to be able to afford it. A newer car with high value would also require normally higher deposit and hence would attract lesser

intentional defaults (moral hazard). Further, it would be expected that a newer car would have less problems and higher resale value and therefore would generate less defaults. A new car also loses value much quicker in the first few years and this could conversely increase the likeliness of default as the car gets older. The variable *lcarage* shows that car age is a significant indicator of default. A borrower with an older car is more likely to default. A 1% increase in the car age will increase the probability of default by 9.2%. The impact measured by the magnitude of marginal effect is in fact not only substantial, but alongside LTV, the higher, when considering only the continuous variables.

5. Disaggregated Analysis of Arrears, Defaults and Early Pay-off

In the above default model, we combined two outcomes, i.e. currently in arrears and already defaulted and car repossessed to analyse the impact of demographic and loan specific indicators. Rather than combining different outcomes, we model five different outcomes using multinomial logistic regression on the same set of independent variables. By looking at these individual specific outcomes, subsequent analysis shows which variables can predict each outcome? Group 1, live agreements not in arrears, is used for the comparable base category.

Group 2 represents loans currently in arrears but technically still classified as live agreements. Columns 2 and 3 in Table 3 display multinomial regression coefficients and marginal effects. The coefficients for the variables *ms2* (married), *rs2* (furnished tenant), *a6* (North West), total income, purchased car price and time in current employment and residence are statistically significant and follow the same trend as the base model in Table 2. The results show that a borrower who is married is 1.3% less likely to be in arrears, a furnished tenant is 1.6% more likely to be in arrears, and a resident in the North West area is 1.4% more likely to be in arrears. For income, the marginal effect of a 1% increase is -2%, implying higher income borrowers are less likely to be in arrears. For car price, the marginal effect is 3.3% and for months in current employment and residence, the marginal effect is -0.4% and 0.4%. Hence, lending for higher price car increases the likelihood of getting into arrears during the agreement period while longer duration of current employment (residence) reduces (increases) such probability. The finding that a longer period in current residence would increase the likelihood that a borrower would in an arrears status is similar to the study by Dinh and Kleimeier (2007) which discusses how a longer time period in the same residency will increase default as an increase in income for the borrower could cause them to move to a bigger house or a better area.

Group 3 is the category for live agreements that have been previously in arrears. These borrowers have gone into arrears over the previous 10 months but have been able to resolve this and are now up to date with repayments. The results show that borrowers that are unfurnished tenants are less likely to fall into this category by 1.4%. Additionally, council tenants are 1.2% less likely to go into arrears and repay the outstanding balance subsequently. It could be that borrowers who are unfurnished tenants or council tenants are unlikely to have an arrears status and therefore they are also unlikely to be in this situation due to the fact that these types of residential status could be less expensive than a homeowner or a furnished tenant which would require more disposable income to meet expenses as well as to repay the loan. Regarding total income, a higher income indicates that the borrower has a lower probability of being in arrears in the past with the marginal effect being -1.4%. This is the expected result, as a higher income can be used to prove stability and affordability, as stated in the previous results. It would be likely that a borrower with a higher income is unlikely to be in arrears at all and therefore unlikely to be in this situation. Looking at the remainder of the statistically significant variables, the results are consistent with the base model results (presented in Table 2). These variables are car price, LTV, period in current employment and car age. It is expected for these results to align with the base model, as a borrower that is unlikely to

default would also be unlikely to be in the position of arrears needing to catch up on missed repayments.

Table 3. Determinants of Subprime Auto Loans Performance

Variable ³	Group 2 currently in arrears		Group 3 previously in arrears		Group 4 defaulted		Group 5 repaid off		Group 6 written off	
	Coef.	M.E.	Coef.	M.E.	Coef.	M.E.	Coef.	M.E.	Coef.	M.E.
ms1	-0.037	-0.001	-0.106	-0.003	0.131	0.005	-0.046	-0.002	-1.074*	-0.002
ms2	-0.424***	-0.013**	-0.209	-0.005	-0.364**	-0.011**	-0.091	-0.002	-0.832	-0.001
ms3	-0.471	-0.013	-0.521	-0.013	-0.805**	-0.025*	-0.202	-0.005	-0.515	-0.001
ms4	-0.580	-0.017	-0.165	-0.003	-0.164	-0.004	-0.533	-0.018	-0.083	0.000
rs1	-0.307	-0.009	-0.226	-0.006	-0.175	-0.005	-0.034	0.000	-1.645	-0.003
rs2	0.497**	0.016**	-0.146	-0.005	0.185	0.006	0.167	0.005	-0.699	-0.001
rs3	0.225	0.008	-0.493***	-0.014***	-0.082	-0.002	-0.133	-0.004	-1.057*	-0.002
rs4	0.189	0.007	-0.443**	-0.012**	0.118	0.005	-0.221	-0.008	-2.750**	-0.005**
es1	-0.487	-0.016	0.819	0.023	-0.029	-0.002	-0.625**	-0.024*	12.647	0.023
es2	-0.996**	-0.032	0.594	0.018	-0.120	-0.003	-0.475	-0.016	0.399	0.001
es3	-0.158	-0.007	1.289**	0.036	0.254	0.007	-0.552	-0.023	14.149	0.026
a2	-0.269	-0.008	0.149	0.005	-0.346**	-0.011**	0.035	0.002	-0.797	-0.001
a4	0.169	0.007	0.008	0.001	-0.542***	-0.018**	-0.147	-0.005	-1.715	-0.003
a6	0.488***	0.014***	0.266*	0.006	0.457**	0.014***	0.354***	0.011**	-0.037	0.000
lmicr	0.110**	0.004**	-0.054	-0.002	-0.015	0.000	-0.053	-0.002	-0.275	-0.001
lmice	-0.123***	-0.004**	-0.107**	-0.003*	-0.123***	-0.004**	-0.034	-0.001	-0.257	0.000
ltinc	-0.709***	-0.020***	-0.586***	-0.014**	-0.676***	-0.020***	-0.495***	-0.015**	-1.043	-0.002
lpincb	0.011	0.000	0.047	0.001	0.096**	0.003***	0.049	0.002	0.253*	0.000*
ldeposit	0.011	0.000	-0.010	0.000	0.018	0.001	-0.041	-0.002	0.064	0.000
lltv	1.416	0.037	2.789**	0.072**	4.723**	0.156***	-1.175*	-0.054**	-2.494	-0.005
lapr	0.626	0.022	-0.156	-0.003	-0.739**	-0.025**	-0.452	-0.016	-1.950***	-0.004***
lterm	-0.367	-0.009	0.012	0.003	-1.130**	-0.036**	-1.200***	-0.041***	2.622	0.005
lcarprice	1.088***	0.033***	0.975***	0.025***	0.767**	0.023**	0.055	-0.002	-2.059*	-0.004*
lcarage	0.537	0.010	1.916***	0.049***	2.445**	0.078**	0.528	0.012	4.735**	0.008*
Constant	-15.366**		-29.191***		-32.950***		8.686		-9.941	

Notes: Table above contains multinomial logit model coefficient (Coef.) estimates and marginal effects (ME) of factors influencing subprime auto loans performance Reference category in this case is normal loan service. Significance level of each factor is determined by p values. Asterisks *, ** and *** indicate statistically significant level of 10%, 5% ,and 1%, respectively.

Group 4 consists of borrowers who have defaulted or can no longer afford the repayments and have returned the car. This could be by surrendering the car to the company or continuous default which has resulted in the repossession of the car. Looking at marital status, borrowers who are married are less likely to fall into this category by 1.1%. Borrowers who are divorced are also less likely to default by 2.5%. As found in the earlier results, these are characteristics in borrowers that are unlikely to default, and therefore are unlikely to find themselves in this situation where they cannot afford the car. In terms of area, borrowers who live in the South East and East area or the Midlands area are less likely to find themselves in this situation where the car has to be returned or repossessed by 1.1% and 1.8% respectively. However, those who live in the North West area are more likely to be in this category by 1.4%. This proves that there are regional differences in terms of borrower’s attitudes to repaying loans. It could also be due to the high unemployment rates in the North West of the UK, which means that the car is unaffordable and therefore the borrower is likely to surrender the car or default until it is repossessed.

³ For the specific description of each variable, please refer to the details in Table 1.

Some continuous variables have the same effect as explained in the previous regression model. These are total income, state benefits, car price and age, LTV, APR, term of the loan, and time in current employment. This is expected as the previous model looks at default status, and this category where the car is returned, tends to be in a situation where the borrower has defaulted and is having difficulty making repayments. Therefore, it would be expected that these continuous variable effects remain the same. Specifically looking at LTV, a 1% increase in LTV will result in a 15.6% increase of probability of defaulting and repossession of the vehicle. It appears that the LTV has a significant effect on affordability and the likelihood of default. This is likely to be because the loan is too large and therefore unaffordable to the borrower. This is similar to the study by Adams *et al.* (2009) which addressed the issue of over borrowing in the subprime industry. Authors such as Bar-Gill, 2008; Mayer *et al.*, 2009; Rajal *et al.*, 2010 agree that a high LTV is likely to increase the probability of default. Our findings confirm earlier conclusions. Looking at the government benefits included in the total income figure, a borrower who is receiving a higher percentage of benefits in relation to income is more likely to have the car repossessed or need to surrender. This could be due to being less financially stable. A borrower with an older car is more likely to find themselves in this position, as per the previous results. When car age is increased by 1%, the probability of default will increase by 7.8%. These results would suggest that there could be potential problems with the car due to the car age and decide that surrendering the car is easier than paying for repairs.

Group 5 is comprised of borrowers that have repaid off their loan agreement early. These are settlement payments from the borrower, a car dealer or an insurance company. One interesting and somewhat different result compared to other outcomes is that borrowers in full time employment are 2.4% less likely to repay loans. The results show that borrowers who live in the North West area are more likely to be in this category by 1.1%. This could be due to a large inflow of money, which would allow the borrower to make the repayment. Another reason could be inability to repay the loan and therefore a decision to sell the car and repay the loan in full. A 1% increase in monthly income would reduce the probability of early settlement by 1.5%. The results show that larger income earners are less likely to be in this category. A loan with a higher LTV would be less likely to be settled early. Marginal effect estimates show that a 1% increase in LTV would reduce the probability of settling early by 5.4%. If a borrower is unable to make a deposit and therefore have a higher LTV on the loan, they are also unlikely to have free cash in order to repay the loan early. The results show that a longer term agreement will reduce the probability of a borrower being in this category. This aligns with the base model result, as a steady affordable monthly payment which would be lower when split over a longer term would give the borrower no problems with repayment.

Group 6 consists of loans that have been written off as bad debts. These are agreements where the borrower is in default and the car cannot be repossessed, therefore the balance of the loan must be written off. The results show that borrowers that are council tenants are less likely to be in this group by 0.5%. This could be due to lower expenditure which indicates that the borrower is economically stable enough to make repayments. The car price has the opposite effect when measured against write off than in the base model. A 1% increase in the car price would reduce the probability of being written off as a bad debt by 0.4%. This could be because the borrowers who are able to afford a more expensive car are likely to be wealthier and more knowledgeable. They may be more likely to surrender the car if they are unable to afford it in order not to damage their credit history further. Some of the other variables have the same effect on this category of bad debt as the base model. This is expected as the borrowers who are in this category will have a default status as they are likely to have missed many payments at this point. These variables are APR, percentage of income coming from benefits and car age.

6. Conclusion

The main objective of this study is to find what characteristics affect a subprime borrower's ability to repay a loan. By using simple logistic and multinomial regression, the study tries to answer this question by determining which characteristics have a significant effect on default status, falling into arrears and repaying early. Our findings broadly aligned well with other findings in the literature. The results show that married and divorced borrowers living in the South East and East and Midlands with stable and high income employment and longer time to pay outstanding loans are less likely to default compared to very high probability of defaults of the borrowers living in the North West of England with significant portion of income from state benefits, purchasing pricy (old) cars with low deposits. Larger proportion of income coming from the state benefits and lending for old cars could also lead to write-offs. Single marital status, comparatively lower housing rents, risk based higher APR and lending for higher priced cars – all these reduce chances of write-offs. Married borrowers living in tenancy accommodations with stable employments and higher income are less likely to not be able to repay loans and go into arrears. Opposite is true for furnished tenants, self-employed, living in relatively high unemployment regions, and purchasing pricy (or old) cars with low deposits. Similarly, full time and high income employees with longer length loan contracts are less likely to settle loans earlier. Interestingly, borrowers residing in relatively high unemployment localities are likely to do opposite.

References

- [1] Adams, W., Einav, L., Levin, J. (2007). "Liquidity constraints and imperfect information in subprime lending", National Bureau of Economic Research, Working Paper No. w13067.
- [2] Adams, W., Einav, L., Levin, J. (2009). "Liquidity constraints and imperfect information in subprime lending", *American Economic Review*, 99(1): 49-84.
- [3] Agarwal, S., Ambrose, B. W., Chomsisengphet, S. (2008a). "Determinants of automobile loan default and prepayment", *Economic Perspectives*, 32(3): 17-28.
- [4] Agarwal, S., Driscoll, J., Gabaix, X., Laibson, D. (2008b). "The age of reason: financial decisions over the lifecycle", American Law and Economics Association Annual Meetings, Be Press (p. 97).
- [5] Agarwal, S., Chomsisengphet, S., Liu, C. (2011). "Consumer bankruptcy and default: The role of individual social capital", *Journal of Economic Psychology*, 32(4): 632-650.
- [6] Amromin, G., Paulson, A. L. (2009). "Comparing patterns of default among prime and subprime mortgages", *Economic Perspectives*, 33(2), 18-37.
- [7] Bar-Gill, O. (2008). "The law, economics and psychology of subprime mortgage contracts", *American Law and Economics Association Annual Meetings*, 2008.
- [8] Bellotti, T., Crook, J. (2013). "Forecasting and stress testing credit card default using dynamic models", *International Journal of Forecasting*, 29(4): 563-574.
- [9] Bonfim, D. (2009). "Credit risk drivers: Evaluating the contribution of firm level information and of macroeconomic dynamics", *Journal of Banking and Finance*, 33(2): 281-299.
- [10] Calem, P., Wachter, S. (1999). "Community reinvestment and credit risk: Evidence from affordable-home-loan program", *Real Estate Economics*, 27(1): 105-134.
- [11] Capozza, D. R., Thomson, T. A. (2005). "Optimal stopping and losses on subprime mortgages", *The Journal of Real Estate Finance and Economics*, 30(2): 115-131.
- [12] Capozza, D. R., Thomson, T. A. (2006). "Subprime transitions: lingering or malingering in default?", *The Journal of Real Estate Finance and Economics*, 33(3): 241-258.

- [13] Capozza, D. R., Order, R. V. (2011). "The great surge in mortgage defaults 2006–2009: the comparative roles of economic conditions, underwriting and moral hazard", *Journal of Housing Economics*, 20(2): 141-151.
- [14] Chatterjee, S., Hadi, A. S. (2006). *Regression analysis by example* (4 ed.). New Jersey: Wiley.
- [15] Chinloy, P. (1995). "Privatized default risk and real estate recessions: The U.K. mortgage market", *Real Estate Economics*, 23(4): 410-420.
- [16] Daglish, T. (2009). "What motivates a subprime borrower to default?", *Journal of Banking and Finance*, 33(4): 681-693.
- [17] Danis, M. A., Pennington-Cross, A. (2008). "The delinquency of subprime mortgages", *Journal of Economics and Business*, 60(2): 67-90.
- [18] Demyanyk, Y., Van Hemert, O. (2011). "Understanding the subprime mortgage crisis", *Review of Financial Studies*, 24(6): 1848-1880.
- [19] Dinh, T. H., Kleimeier, S. (2007). "A credit scoring model for Vietnam's retail banking market", *International Review of Financial Analysis*, 16(2): 471-495.
- [20] Einav, L., Jenkins, M., Levin, J. (2012). "Contract pricing in consumer credit markets", *Econometrica*, 80(4): 1387-1432.
- [21] Eriksen, M. D., Kau, J. B., Keenan, D. C. (2013). "The impact of second loans on subprime mortgage defaults", *Real Estate Economics*, 41(4): 858-886.
- [22] Furstenberg, G., Furstenberg, V., Green, R. (1974). "Estimation of delinquency risk for home mortgage portfolios", *Real Estate Economics*, 2(1): 5-19.
- [23] Heitfield, E., Sabarwal, T. (2004). "What drives default and prepayment on subprime auto loans?", *The Journal of Real Estate Finance and Economics*, 29(4): 457-477.
- [24] Kočenda, E., Vojtek, M. (2009). "Default predictors and credit scoring models for retail banking". CESifo Working Paper No. 2862.
- [25] Lax, H., Manti, M., Raca, P., Zorn, P. (2004). "Subprime lending: An investigation of economic efficiency", *Housing Policy Debate*, 15(3): 533-571.
- [26] Malik, M., Thomas, L. C. (2009). "Modelling credit risk of portfolio of consumer loans", *Journal of the Operational Research Society*, 61(3): 411-420.
- [27] Marshall, A., Tang, L., Milne, A. (2010). "Variable reduction, sample selection bias and bank retail credit scoring", *Journal of Empirical Finance*, 17(3): 501-512.
- [28] Mayer, C., Pence, K., Sherlund, S. M. (2009). "The rise in mortgage defaults", *Journal of Economic Perspectives*, 23(1): 27-50.
- [29] Morton, T. (1975). "A discriminant function analysis of residential mortgage delinquency and foreclosure", *Real Estate Economics*, 3(1): 73-88.
- [30] Rajan, U., Seru, A., Vig, V. (2010). "The failure of models that predict failure: distance, incentives and defaults", Chicago GSB Research Paper No. 08-19.
- [31] The Telegraph (2014). "Unemployment in the UK: regional breakdown". [Online] Available from The Telegraph: <http://www.telegraph.co.uk/news/politics/10589129/Unemployment-in-the-UK-regional-breakdown.html> (Retrieved on March 25, 2014).
- [32] Yap, B. W., Ong, H., Husain, N. H. (2011). "Using data mining to improve assessment of credit worthiness via credit scoring models", *Expert Systems with Applications*, 38(10): 13274-13283.
- [33] Zywicki, T. J., Adamson, J. D. (2008). "The law and economics of subprime lending", George Mason Law and Economics Research Paper, No. 08-17; *University of Colorado Law Review*, 80(1): 1-86.