



GORDON INSTITUTE
OF BUSINESS SCIENCE

University of Pretoria

Construct of borrower identity and credit scoring: Affordable housing segment

Charles De Winnaar

11356431

A research project submitted to the Gordon Institute of Business Science,
University of Pretoria, in partial fulfilment of the requirements for the degree of
Masters of Business Administration

07 November 2012

ABSTRACT

Credit scoring and risk analytic models have to an extent become purely analytical, faceless and completely inhumane.

Borrowers are classified as a mere number or perhaps better said, a borrower's worth has degenerated to that of derived risk scores.

For banks credit application scoring provides a mechanism for managing risk and hence profits.

From the quantitative study, it was established that imposed risk variables, had a negative selection outcome and provided grounds for arguing that there is a requirement for the further development of risk assessment scorecards that should be more informative and less exclusionary in terms of understanding of the affordable borrower market. These findings were further enforced by the qualitative interviews that were held with risk practitioners.

As a consequence of imposed risk assumptions, it is argued that there are groups of borrowers that are intentionally and unintentionally excluded through credit scoring.

Collectively these findings support the view that borrower identity is systematically fabricated on an on-going basis through the direct application of processes, systems and technologies as imposed by enterprise and credit scoring practitioners through risk assessment models.

It was also argued that the future of application scorecard building would see more advanced or alternative variables being used which will give rise to a self-regulating borrower segment, less overtly disciplinary banks, and more pervasive credit bureau information and data constructs.

KEY WORDS

Debt to income (DTI)

Loan to value (LTV)

National credit regulator (NCR)

Repayment to income (RTI)

Thin payment profile

DECLARATION

I declare that this research project is my own work. It is submitted in partial fulfilment of the requirements for the degree of Masters of Business Administration at the Gordon Institute of Business Science, University of Pretoria. It has not been submitted before for any degree or examination at any other University. I further declare that I have obtained the necessary authorisation and consent to carry out this research.

Charles de Winnaar

07 November 2012

ACKNOWLEDGEMENTS

I wish to express my sincere thanks and appreciation to:

- Kerry Chip, my supervisor, for her assistance and guidance in the preparation of this report and most importantly for her continuous advice;
- the community of interest at the participating bank, which gave generously and unselfishly of their time to this study;
- my wife, Jacquie, for her love, understanding and unwavering support over two very long and eventful years.

Contents

CHAPTER 1:	1
1. Introduction to the Research Problem:.....	1
1.1. Purpose of study	1
1.2. Problem statement	1
1.3. Sub problem statements	3
1.4. Background to the problem	4
1.5. The borrower market.....	5
1.6. Access to credit	6
1.7. Functions of credit risk analytics	9
1.8. Significance of the study	10
CHAPTER 2:	11
2. Literature Review	11
2.1. Introduction.....	11
2.2. Credit Analytics	12
2.2.1. What is credit?.....	12
2.2.2. The origins of credit risk analytics	13
2.2.3. Evolution of credit risk scoring models	14
2.2.4. How credit risk analytics work	16
2.2.5. What credit risk analytics do.....	18
2.2.6. Construct of credit risk scoring models	20
2.2.7. Limitations of credit risk scoring models	22
2.2.8. Limitations of credit risk analytics.....	22
2.2.9. Exclusionary biases of risk analytic models	24
2.2.10. Disciplines of credit risk analytics.....	25
2.2.11. Alternative approaches for credit risk analytics.....	27
2.3. Sociology and Borrower Identity	29
2.3.1. Identity	29
2.3.2. Borrower identity	30
2.3.3. Social identity	31
2.3.4. Issues with the construct of identity	35
2.4. Affordable Housing.....	36

2.4.1.	Why consider the affordable segment?.....	36
2.4.2.	What makes the affordable market different to that of traditional markets?	37
2.4.3.	Why the “underclass”	37
2.4.4.	Societal importance of providing affordable housing	38
2.4.5.	Affordable housing; welfare state or open market.....	39
2.5.	Overview of Literature Review	40
CHAPTER 3:		41
3.	Propositions:.....	41
3.1.	The imperative to research risk assessment parameters and borrower constructs ...	41
3.2.	Research Proposition 1	41
3.3.	Research Proposition 2	42
3.4.	Research Proposition 3	42
CHAPTER 4:		44
4.	Methodology and Design.....	44
4.1.	Research Design	44
4.2.	Choice of Methodology.....	45
4.3.	Scope	46
4.4.	Universe / Population	46
4.4.1.	Participants and size of the industry	46
4.5.	Unit of Analysis.....	47
4.6.	Sample Method	48
4.7.	Data Gathering and Analysis	50
4.8.	Research Limitations.....	53
CHAPTER 5:		55
5.	Results.....	55
5.1.	Introduction.....	55
5.2.	Source of extracting quantitative data	55
5.3.	Statistical data analysis	56
5.4.	Discussion of data obtained pertaining to the first research proposition	58
5.4.1.	Thin payment profile risk variable	59
5.4.1.1.	Response from risk practitioner interviews	62
5.4.2.	Repayment to income (RTI) risk variable.....	63
5.4.2.1.	Response from risk practitioner interviews	66

5.5.	Discussion of data obtained pertaining to the second research proposition.....	67
5.5.1.	Slow payer status	67
5.5.1.1.	Response from risk practitioner interviews	68
5.6.	Discussion of data obtained pertaining to the third research proposition	70
5.6.1.	Default rates	71
5.6.2.	Debt to income	72
5.6.3.	Vintage defaults.....	72
5.7.	Conclusion.....	73
CHAPTER 6:		75
6.	Discussion of Results	75
6.1.	Introduction.....	75
6.2.	Research findings in response to the research propositions and literature	75
6.3.	Proposition 1: Credit risk assessment variables and exclusion factor(s).....	76
6.4.	Proposition 2: Traditional borrower identity and risk assessment model constructs ..	82
6.5.	Proposition 3: Alternative scorecard	85
CHAPTER 7:		89
7.	Conclusion.....	89
7.1.	Introduction.....	89
7.2.	Summary of Main Findings.....	89
7.3.	Future Recommendations	93
7.4.	Implications	94
8.	References	95
9.	Appendix A	103
	Standardised – Open-Ended Interview Guide	103
10.	Appendix B.....	105
	Guide for Interview purpose and process:	105
11.	Appendix C.....	106
	Schedule for Secured and Unsecured Risk Analyst Interviews:.....	106
12.	Appendix D.....	107
	Interview schedule for Borrower Identity Constructs and Credit Risk Scoring:.....	107
13.	Appendix E.....	109
	Interview schedule:.....	109

Tables and Figures:

Table 1: Mortgages Granted - size of agreements (National Credit Regulator, 2011)	8
Table 2: Number of agreements by size of loan (National Credit Regulator, 2011)	8
Table 3: Population size of data	56
Table 4: Population Data applied to Secured Lending Risk Scorecard	57
Table 5: Disbursed Unsecured Lending Loans Default Levels	57
Table 6: Population performance outcome	58
Table 7: Financial Services Charter (FSC) Income Brackets	59
Table 8: Unsecured Lending Population qualifying for AFH - 'thin payment' profile.....	60
Table 9: 'Thin Payment Profile' Loan distribution performance.....	60
Table 10: Thin payment profile - Applied risk assessment decision	62
Table 11: Disbursement - Secured Lending Risk and Affordability Assessment.....	62
Table 12: Disbursed Unsecured Loans (AFH) Segment.....	65
Table 13: Risk and RTI credit analytics outcome.....	66
Table 14: Loan Size scenario and opportunity lost	66
Table 15: Secured lending risk analysis for 'slow payer' borrowers	68
Table 16: 'Slow payer' secured lending risk assessment outcome.....	69
Figure 1: Affordable Market Representation	2
Figure 2: Retail Credit Distribution	6
Figure 3: The General form of a globally connected feed-forward network.....	28
Figure 4: A general theory of underclass formation and persistence	38
Figure 5: Distribution of loan to value of risk and affordability assessment	65
Figure 6: Contractual debt vs. Net Income (Disbursed Loans).....	71
Figure 7: Borrower Categories: Debt to Net Income %.....	72
Figure 8: Borrower Categories: Vintage Analysis	73
Figure 9: Risk Practitioner vs. Borrower: Managing the paradigm	90

CHAPTER 1:

1. Introduction to the Research Problem:

1.1. Purpose of study

The primary purpose of this study is to explore how credit and credit risk analytics are modelled, incorporated and adopted for usage in secured risk assessment models for the affordable borrower market. Emphasis is placed on the construct and functioning of credit risk models, the unintentional effect associated with the usage of third party borrower/applicant risk data and the impact that exclusionary credit risk assessment models infer on borrower identity constructs. The impacts on social and societal exclusions are argued in favour of developing credit risk analytics that are representative of the uniqueness of the affordable borrower retail credit market.

The research project focuses on risk assessment models as used in the retail secured lending operations of a large retail financial institution that provides finance to the affordable housing in South Africa. The study incorporates principles adopted from lending practices of that of secured and unsecured credit scoring analytics.

The discipline of borrower identity constructs contributes to the understanding of the uniqueness of individual and group assimilation behavioural attributes that are incorporated into credit risk decision making models.

Consumer behavioural patterns are becoming more relevant in the discipline of borrower identity constructs as consumer behavioural patterns consistently change and adapt to changing markets. The aforementioned discipline pertains to a multiple of social, societal, behavioural, psychological and psychographic constructs with consumer behavioural patterns.

As such, so is the need for financial institutions to adopt in understanding and meeting the borrowers' needs.

1.2. Problem statement

It is argued that the manners in which credit and risk assessment models are constructed, certain members or sub groups of population cohorts are intentionally and

unintentionally excluded through risk assessment practices, and as such are unable to obtain credit.

It is further argued that these exclusionary factors are unduly biased towards non-traditional and non-representative population groups. These exclusionary factors are believed to be inherent in the basic construct of credit risk assessment model(s).

Owning a house is a national imperative that is aiming to curb inequalities and undue wealth distribution. It is an important determinant for establishing an equitable framework whereby quality of life and wealth creation is promoted within a community (Department: Human Settlements Republic of South Africa, 2009, pp. 21-26).

Statistics South Africa concluded from research undertaken in 2005/2006 that the Affordable and GAP housing market (Level 3b, Figure 1) consisted of approximately 3,8 million households and if combined with the sub market segment known as the *subsidised housing market* (Level 4b, Figure 1), the total number of households in the affordable segment increases to 10 million, which is slightly less than 87 per cent of the total household population in South Africa as depicted in Figure 1 below.

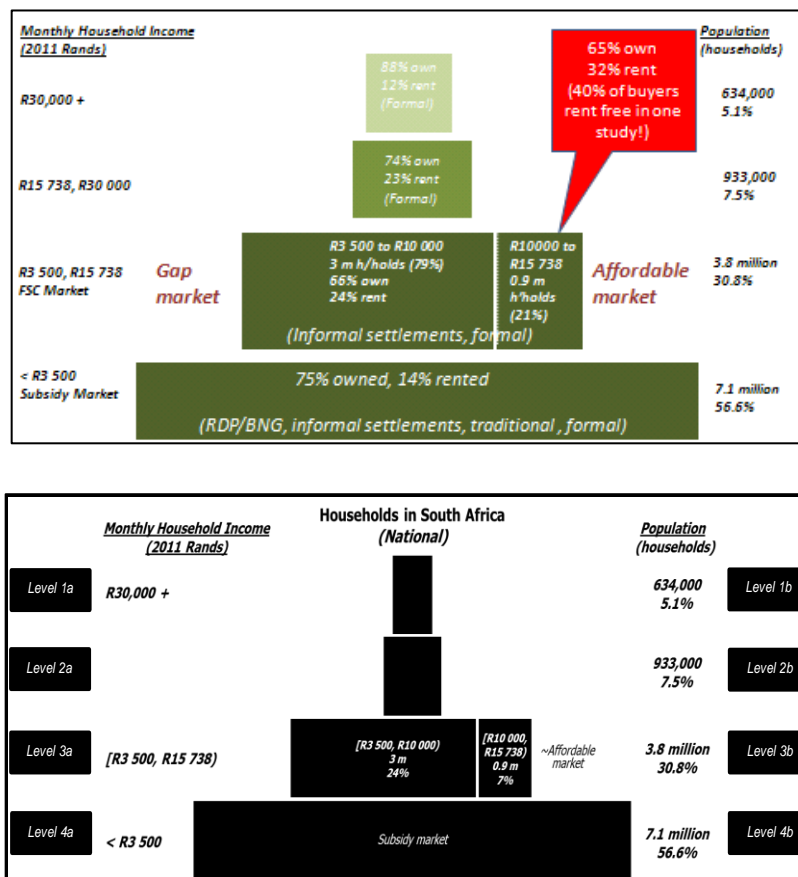


Figure 1: Affordable Market Representation

Source: IES 2005/6 (Incomes have been inflated to 2011 Rands)

The empirical evidence is more than sufficient to highlight the magnitude of the social requirement for credit providers to provide access affordable and responsible finance solutions.

Economic factors do contribute to credit extensions, however, credit participation continues to be a contentious debate at social political levels which constitute to be an imposed factor that influence lending and risk assessment outcomes. This collaborate the fact that assessment models have been developed over time with the primary objective being that of maximising profits attributable to the lender, whilst minimising risk and undue losses as the same time (Anderson, 2007).

As such the argument is formulated off the premises that risk analytics and imposed credit risk models are developed by participants from the middle or upper social class who represent the very group or sub population group who can afford to borrow credit, whilst credit risk models are aimed at the affordable market segment, a segment that has its own unique characteristics which is seldom similar to that of the middle or upper social segments which was identified in the work of Hand (2001).

The consequence is unintentional exclusion from risk assessment practices for non-representative groups since the initial design of stated credit risk assessment models were built on the mental models and believe systems of that of the middle and upper social classes.

The recommendation is to review and develop alternative and more representative models that provide for more inclusive risk variables, and to effectively minimise the effect of unintentional group exclusions.

1.3. Sub problem statements

- (i) Identify credit risk variables that are used in the construction of secured lending credit risk application scorecards, if different to that of the affordable borrower segment.
- (ii) Identify the assumptions made about borrowers through credit scoring design (*in the affordable client segment*).
- (iii) Explore what alternative characteristics to consider for consideration in credit risk assessment and application scorecards (*in the affordable client segment*).
- (iv) Identify which credit applications are included/excluded from the credit application scoring process.

1.4. Background to the problem

With a growing backlog in the supply of housing for South Africa's affordable market currently estimated to be in excess of one million units for households earning less than R15 000 per month (Lightstone, 2011), an emergent need for alternative solutions are required to address such shortcoming and provide for more sustainable and practical solutions to deal with growing societal inequalities.

One of the fundamental challenges is for affordable consumers to secure loan finance. These events are by no means mutually exclusive as the latter requirement also facilitates the necessity of enabling the establishment of secondary markets.

Verstraeten & Van den Poel, (2005), identified the perpetuating exclusionary impact of having a non-representative sample that is biased towards that of the population that consistently participate in the risk assessment process whereby credit facilities are secured whilst non-representative population sub-groups continue to be excluded.

With the continuation of these shortcomings and societal challenges not being met, the disparity between South Africa's poor, middle and upper classes continue to expand and wealth distribution become even more disparate (Bertelsmann Stifting's Transformation Index (BTI), 2012). This is evident in the rather high recorded level of South Africa's measured Gini coefficient, currently recorded at 0.70 and amongst the highest in the world for countries of similar GDP characteristics within economies of similar size (Bertelsmann Stifting's Transformation Index (BTI), 2012). If this noted disparity continues to widen, the reality of a mass revolt can becomes more prevalent.

In meeting the objective to improve overall living standards for the majority of South African's and address the historical legacy of shortcomings, government is committed to the agenda of supporting the Affordable Residential Housing Market (Department: Human Settlements Republic of South Africa, pp. 20-22). The objective of this governmental code is to provide for meaningful housing solutions and access to finance, that will not only benefit the broader societal construct, but also instil property ownership as the basis of creating wealth and security for the majority of households.

Yet, these objectives are far from being met and continue to be a challenged ideology. Access to credit is limited and if obtained such lending facilities will be issued at a price premium well above that of general market levels. Lenders who are willing to take on risks within the affordable markets are predominately price givers in view of pricing for risk (Anderson, 2007; Siddiqi, 2006). The higher the loan required without a deposit or

even with a lower deposit commitment, the more the borrower has to end up taking on of that of a price premium.

1.5. The borrower market

In general applicants looking for credit within the affordable housing segment will fall within the Living Standards Measurement demographic of 4-7 depending on household income and life stage (AMPS, 2012). The affordable market is defined by the Financial Services Charter as households who earn less than R15, 738 (2011) and R16 486 (2012) per month (Financial Sector Charter Council, 2012). It is a rather infrequent occurrence that these borrowers will have the means to provide for a deposit at loan request stage which constitute to be a self-eliminating factor when entering the market.

Government subsidies do assist, however the majority of subsidies are for house hold income sectors of less than three thousand five hundred Rand per month as per Figure 1, and the extension terms for granting subsidies does not constitute to be limitless. As such, the need for alternative approaches and participation from the financial institutions become more prevalent.

In a recently published Mail & Guardian article, Marius Marais, the chief executive of First National Bank housing finance, was quoted in saying that the typical charge rate for the affordable segment is prime (*the rate at which Banks lend above what the Reserve Bank will borrow to Bank's*) plus one to three per cent above and in addition to prime (Steyn, 2012). According to such inflated price charge levels, affordability becomes a challenge as lenders are pricing for perceived risk, which further eliminate borrowers from participating and securing housing finance.

In terms of the national debate and importance of meeting and delivering on housing needs in the affordable market, the Department of Human Settlements policy introduced the *National Housing Programme* with the aim to provide funding to social housing institutions and to provide for:

- (i) *Loss Limit Insurance* which is still under investigation and review; *Fixed Interest Rates* which is directly linked to the financial sector to curb interest rate volatility and changes;
- (ii) *New Product Development* which will review alternative options in terms of assisting individuals without formal income streams; and perhaps one of the most direct measures,

- (iii) *The Monitoring* of lending trends for those of retail financial institutions through the Office of Disclosure which was established through the promulgation of the Home Loan and Mortgage Disclosure Act, 2000 (Act No. 63 of 2000) (Department of Human Settlements: National Housing Code 2009, pp. 20-22).

1.6. Access to credit

To enable the housing market to function and generate the required stock churn, it is up to the financial institutions through which finance extensions are granted to end users to assist and provide the means for consumers to have access to credit. The consumer credit market is still dominated by the Banking sector as of end December 2011 (National Credit Regulator, Fourth Quarter, December 2011, p. 1).

According to the NCR, the banking sector data as depicted in Figure 2 reflects a gradual decrease in the distribution of credit granted for secured lending purposes (yellow line) leading to an increase in the rejection rate of secured lending applications and increases with a counterintuitive increase in credit lending activities in sectors such as unsecured lending and short term lending. It can be argued that potential borrowers in the affordable market segment are excluded from securing credit approvals due to the current nature of applied risk assessment models which does not take into consideration unique borrower identity construct and consumer behavioural patterns (Verstraeten & Van den Poel, 2005).

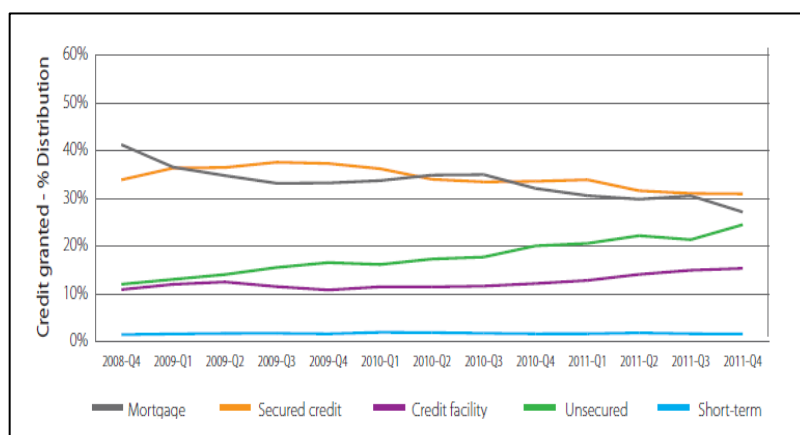


Figure 2: Retail Credit Distribution

(National Credit Regulator, Fourth Quarter, December 2011)

As at the end of 2011 the recorded rejection rate was 45.75 per cent for the quarter and a 44.65 per cent year-on-year change (National Credit Regulator, Fourth Quarter, December 2011, p. 4). As can be seen in Figure 2 (purple line), growth recorded in the volume of approvals of new credit extensions only grew by 27.54 per cent year-on-year as at December 2011 and constitute to be a substantial increase over a relatively short recorded period

In contrast, mortgages accounted for 61.03 per cent of all extended consumer credit by value (National Credit Regulator, Fourth Quarter, December 2011, p. 5), however in terms of volume, mortgages account for only 4.77 per cent of the total number of accounts booked whilst unsecured credit account for 19.60 per cent (National Credit Regulator, Fourth Quarter, December 2011, p. 6), suggesting that credit scoring criteria for mortgages are too strict reducing the number of loans granted. Lenders must continue to develop new methods and criteria for assessing credit requests stemming from the affordable borrower segment.

FinMark Trust commissioned research into the Township Residential Property Market and found that the property market boom of the mid 2000's as experienced in South Africa was not taking place for one fifth of households who lived in former black townships (Rust, 2012).

It was further concluded based on stated research that because of dysfunctional secondary markets (markets where re-sale did not occur or profits were not realised from re-sales) being prevalent in these areas, households that were living in these former black townships were unable to realise the value of their property investment, which at the time (2004) had an estimated value of R68.3 billion (Rust, 2012).

Not being able to access or utilise available equity and funds is largely due to the fact that an inadequate response from the financial sector has failed to address and meet an increase in demand for access to financial services, and in particular to long term lending requirements (Financial Sector Charter Council, 2012). Only the basic needs are being met and as a result no tangible and long term wealth creation disciplines can be entrenched.

The quarterly report as released by the NCR (Fourth Quarter, December 2011) indicates that there was an overall decline in the Rand value of mortgage agreements granted, and what is even of greater concern and topical to the research problem (as can be seen in the last column of Table 1) is that the recorded year-on-year change for mortgage loans granted across segments less than three hundred and fifty thousand

Rand declined on a year-on-year basis. Refer to Table 1 for related segment comparisons.

Table 1: Mortgages Granted - size of agreements (National Credit Regulator, 2011)

Agreements	2010-Q4 R000	2011-Q1 R000	2011-Q2 R000	2011-Q3 R000	2011-Q4 R000	2011-Q4 % Distribution	% Change (Q4/Q3)	% Change (Y/Y)
R0-R50K	65,440	60,362	52,186	57,960	45,935	0.16%	-20.75%	-29.81%
R51K-R100K	245,691	240,188	207,741	232,248	194,269	0.66%	-16.35%	-20.93%
R101K-R150K	303,724	273,714	249,958	280,198	268,629	0.92%	-4.13%	-11.55%
R151K-R350K	2,554,039	2,227,881	2,094,388	2,522,200	2,404,369	8.20%	-4.67%	-5.86%
R351K-R700K	6,817,467	6,095,519	6,215,484	7,667,365	6,994,636	23.86%	-8.77%	2.60%
>=R700K	16,881,609	15,862,251	16,628,758	19,518,416	19,405,986	66.20%	-0.58%	14.95%
Total	26,867,971	24,759,915	25,448,516	30,278,386	29,313,825	100.00%	-3.19%	9.10%

Credit approvals for loans in excess of seven hundred thousand Rand showed the strongest year-on-year growth and do not assist in establishing wealth distribution across the affordable housing segment or for government to meet its housing objectives (National Credit Regulator, Fourth Quarter, December 2011).

Similar year-on-year declining loan approval trends are noticeable as per Table 2 for the volume statistics for all segments under three hundred and fifty thousand Rand.

Table 2: Number of agreements by size of loan (National Credit Regulator, 2011)

Agreements	2010-Q4	2011-Q1	2011-Q2	2011-Q3	2011-Q4	2011-Q4 % Distribution	% Change (Q4/Q3)	% Change (Y/Y)
R0-R50K	2,038	1,758	1,642	1,905	1,411	3.34%	-25.93%	-30.77%
R51K-R100K	2,965	2,901	2,524	2,822	2,374	5.61%	-15.88%	-19.93%
R101K-R150K	2,304	2,076	1,891	2,126	2,025	4.79%	-4.75%	-12.11%
R151K-R350K	9,785	8,607	8,023	9,605	9,031	21.36%	-5.98%	-7.71%
R351K-R700K	13,353	11,921	12,165	14,961	13,743	32.50%	-8.14%	2.92%
>=R700K	12,205	11,281	12,079	14,347	13,699	32.40%	-4.52%	12.24%
Total	42,650	38,544	38,324	45,766	42,283	100.00%	-7.61%	-0.86%

The number of successful credit agreements concluded for households earning in excess of fifteen thousand Rand per month as recorded was in excess of eighty per cent of the recorded volume and ninety two per cent of the recorded value as per lending statistics released by the NCR for the quarter ending December 2011. The statistics contextualises the reality of what borrowers are faced with.

1.7. Functions of credit risk analytics

In order to borrow money from retail financial institutions, borrowers are placed through a pre-defined credit worthiness assessment matrix in order to be evaluated and establish the likelihood that said creditor will be able to meet requested debt repayment obligations over the term of the loan (Anderson, 2007, p. 7).

Finance decisions are made on the premise of extracting maximum economic profits (Thomas 2000), and do not factor social or political requirements into the assessment process, thus exacerbating the perpetuating problem of enforcing the deepening rift of inequality and wealth disparity.

Automated scorecard matrixes are designed to offer consistency in risk assessment modelling by using the same set of rules and applying the same weighted variables to the entire portfolios, permit for regular reviews of the entire account portfolio that can quickly and effectively identify accounts that require intervention, thus isolating customers who warrants human intervention (Oracle Financial Services, 2008, p. 2).

It is however counterintuitive should the data be of poor quality or should a large subset of clients require human interventions or manual assisted overrides (Anderson, 2007, p. 82).

Automated credit risk analytics does however operate with implied generalisation assumptions based on pre-categorised risk groups as defined by the lending organisation. Such inferred generalisation assumptions can further lead to discrimination in terms of scoring decisions. (Anderson, 2007, pp. 621-625).

Anderson concludes that the credit industry should be governed implicitly by specific or general legislation to protect the interest of all borrowers such as rules governing;

- *“(prohibition of) unfair discrimination – To prejudice, based upon knowledge of a group membership, defined by colour, culture, gender, religion, nationality, or some other personal characteristic.*
- *(promotion of) Equal opportunity – to make available, or offer to all, without discrimination.*
- *(protection of) Human rights – to ensure basic freedoms, facilities, and protection (Anderson, 2007, p. 621).*

The study will contribute towards the theory of risk assessment disciplines associated with inclusion and exclusion risk assessment variables which are contradictory to the

implied governance covenants as mentioned by Anderson. Implied exclusionary risk variables will be reviewed and analysed to develop and formulate recommendations for developing future borrower identity constructs and credit risk analytics for affordable housing borrowers.

1.8. Significance of the study

There is evidence of both theoretical and practical relevance for conducting this study:

From an intellectual perspective the contribution to research includes:

- To contribute towards academic research in the Affordable Housing credit risk scoring and risk analyses segment.
- In terms of contributing to broader social benefits, the research contribution will be towards eliciting public policy debates, depicting the characteristics of credit scoring and borrower identity constructs. The objective is to stimulate discussions for the development of alternative credit scoring methods applied to the affordable segment.

The contribution is to stimulate public understanding of what is at stake for individuals in borrower identity construction and what constitute to be viable alternatives (Clement, 2002). Further to assist retail credit lenders in South Africa to develop more informed assessment methods, to better contextualise borrower identity and to assist borrowers in establishing equitable borrower identities to secure housing finance.

CHAPTER 2:

2. Literature Review

2.1. Introduction

To follow is the primary assessment of what is deemed to be the most appropriate literature composition for substantiating this research study which is formulated off the premises of diverse philosophies.

A background discussion pertaining to that of more traditional methods applied to the construct and discipline of credit scoring and risk analytics will be elaborated on, following which the effect and outcome of aforementioned disciplines on borrower segments will be reviewed this contextualising the relation between borrower identity and the construct of borrower identity in credit risk models.

Access to credit is critical to the establishment of a well-functioning modern economy since the majority of consumers rely on credit and the access to credit to alleviate poverty and finance the acquisition of basic assets and opportunities to which every human being is entitled such as an education, or home (Kozolchyk, 2007, p. 729). In most circumstances, the terms and conditions set out by the lender are based primarily or solely on the review and assessment of an individual's payment history, often summarised by denoting such assessment to that of a numerical indicator that constitute to be the borrower's credit score (Avery, Brevoort, & Canner, 2009).

The research report will place particular emphasis on the implied assumptions and variables that are made use of to construct borrower identities in the applied discipline of credit scoring and function as used in credit risk analytics. The report will further discuss the relevance of understanding the uniqueness of certain population subgroups of borrowers by subjecting these subgroups to credit risk assessment criteria for secured risk scorecards based on unsecured credit borrower portfolios. It is to be noted that, the manner in which these risk disciplines are applied to general consumer segments will be assessed, and that the focus will be directed to the affordable consumer segment.

It is argued that in credit markets, it is the lenders, who are at risk of *adverse risk selection* because their customers or potential borrowers have better knowledge of

their own personal circumstances than the lender, and that the “*morale hazard*” is that borrowers become less financially responsible with funds in hand (Anderson, 2007).

As a result of the aforementioned, institutions responsible for extending lending facilities to creditors continuously seek out new ways in which to assess client risk and improve on minimising the risk of selecting disproportionate adverse risk clients.

Financial institutions purpose is to facilitate and manage the allocation with deployment of capital resources in an uncertain environment. (Merton, 1992, p. 12). This function encompasses the payment system and medium of exchange in order for the transfer of capital resources from savers to borrowers through reduction of risk insurance and risk diversification (Merton, 1992, p. 12). The system is proposed to function based on the astuteness of the financial lender or Bank, that acts as an intermediary between the borrower and the depositor by managing the risk for protecting the interest of the depositor who is ultimately providing the means whereby the financial institution can further onward extend credit to borrowers.

Mortgage modelling involves both theoretical and empirical analysis and relates to the fact that borrower behaviour cannot only be determined by theoretical considerations; nevertheless risk model constructs are primarily constructed based on demographic and empirical analysis of a borrower’s risk and repayment profile (Levin & Davidson, 2008, p. 40).

2.2. Credit Analytics

2.2.1. What is credit?

In its current context, “credit” simply means “buy now, pay later”, whether the purchase is for short term consumption, durable goods or other assets that will provide users with services, or productive enterprises (Anderson, 2007). The availability of credit and liquidity is a pre-requisite for economic growth and societal progression, since the centrality of finance in an economy is vital to economic growth, which elevates the importance of financial innovation (Levine, 1997). It further alleviates poverty and a well governed lending environment that is competitive provides for a risk reduction that can lend itself to passing on the savings to the consumer (Kozolchyk, 2007, p. 728).

Credit is viewed as a commodity and human construct, and both are heavily intertwined to the extent that lenders activities extend far beyond lending money, to include managing the transaction process/value chain in order to obtain information to manage risk. Credit is intangible, has no physical form other than documentary evidence and so

too is the information on which it is based (Anderson, 2007). As such the concepts of *information goods* and *information economies*, with the uses in economics and law became relevant (Anderson, 2007).

The application of economics and law in credit extensions can be argued to have traditionally contained discriminatory exclusion factors, and that such factors have led to the exclusion of certain population constructs and market segments from accessing credit.

The questions is then, if finance is virtually the enabler of all production activity and improvements within the financial market sector, how can it be adequately and equitably distributed and utilised to have a positive outcome throughout an economy (Levine, 1997)?

2.2.2. The origins of credit risk analytics

Credit scoring or credit risk analytics is in essence the way in which a lender can recognise the different groups in a population when one cannot see the characteristics that separates the groups, apart from related ones (Thomas L. C., 2000, p. 151). Fisher introduced the concept of credit risk analysis in statistics as early as 1936.

Credit scoring was exclusively associated with “accept/reject” decisions generated by the new business application process known as “application scoring” (Anderson, 2007). Financial institutions also relied exclusively on subjective analysis or so called banker “expert” systems to determine credit risk (Altman & Saunders, 1998, p. 1722).

From the literature review, it was ascertained that the progress in credit scoring was due to greater and increased levels of competition in the financial industry, and with greater advances in computer technology, the potential of large databases were introduced (Koh, Tan, & Goh, 2006, p. 97). This progress has resulted in a pendulum shift seeing relationship based lending activities to be replaced by transactional lending assessments.

This pendulum shift to transactional lending assessments in credit scoring incorporates the use of statistical models to extend and manage credit and includes the measurement of risk, response, revenue, and retention also known as the 4 R’s, whether for marketing, new business processing, account management, collections or recoveries in the credit risk management cycle (CRMC) (Anderson, 2007).

Since credit scoring is there to assist credit providers with the means of deriving an outcome of either accepting or rejecting a loan request relatively quickly compared to assessing each request manually (Akhavain, Frame, & White, 2005), the automation of risk analytics was favoured since a quick automated assessment would benefit the lender in the process where more credit and higher volumes of borrowers require an assessment to determine whether credit should be extended (Morrison, 2004).

Relationships between risk, variable scores and bad rates were identified as integral decision making components during the borrower risk assessment process, required to make realistic risk assessment decisions, very specific calculations, and to provide for set cut-offs (Siddiqi, 2006).

Since the automation of risk analytics was introduced, 36 per cent of respondents in a 2003 Credit Research Foundation survey reported to have actively engaged in using credit scoring techniques and that 72 per cent of respondents that at the time were not using credit scoring were adopting a similar approach within two years (Cundiff, 2004). Virtually all credit card applications and 70 per cent of home loan reviews were conducted by utilising credit scoring techniques (Wendel & Harvey, 2003).

What can be ascertained from the results of the aforementioned survey is that the relationship that automated credit risk analytics draws between risk, variable scores and bad rates is a key driver behind why credit practitioners and providers automated systems to fulfil the majority of credit risk tasks.

2.2.3. Evolution of credit risk scoring models

Originally credit assessment practices were purely judgemental, as a credit analyst read the application and confirmed either “yes” or “no” as a decision outcome. Decisions were based on what was called the 3Cs, then the 4Cs and now also the 5Cs (Thomas L. C., 2000, p. 152) which according to Thomas (2000) consist of;

- (i) The person’s characteristics. Is there an established relationship?
- (ii) How much capital is being asked for?
- (iii) What own contribution in terms of collateral is the applicant willing to put up?
- (iv) What are the borrowers repaying ability, how much disposable income is available?
- (v) What are the conditions in the market?

Today's credit risk analytics is a highly technical and automated process. One example of this is the Basel II accord where advanced risk approaches have been adopted for lenders of credit to measure credit risk capital. The contribution made by the modern day version of credit risk analytics is that each lender or bank now internally rate/score borrowers either individually as a unique borrower or as a collective population in a pool (Oracle Financial Services, 2008, p. 3). According to Oracle Financial Services, banks have to implement and own the necessary analytics to measure risk parameters like probability of defaults (PDs), and loss given defaults (LGD) resulting in enhanced value beyond regulatory compliance and provide a competitive advantage to the lender (Oracle Financial Services, 2008, p. 3).

On the one hand credit scoring benefits the lender, however according to Anderson (2007), it comes with its own obligations and borrowers must be cognisant of set commitments that ultimately bond the borrower through commercial term agreements that relate to;

- (i) creating the impression of trust, with obtaining credit comes responsibility and personal expense and spending control;
- (ii) repaying according to the agreed terms of the loan agreement which is a standard fix repayment obligation; and
- (iii) paying a risk premium for the possibility that they might not repay and default in future.

With further advances in computing, the 1980's saw the introduction of logistic regression and linear programming to credit risk scoring and these disciplines are still being used today which does raise the concern in terms of relevance to current markets and uniqueness of today's borrower requirements (Thomas L. C., 2000, p. 152).

According to Thomas (2000, p. 152), logistic regression measures the relationship between a categorical dependant variable (*variable that can be a limited number of categories*), and a continuous independent variable (*can be several*) by means of converting the dependant variable to probability scores, and is the mathematical method for determining the want to achieve the most favourable outcome.

Today though, the emphasis is more on establishing how to maximise the profit derived from a client, rather than just trying to minimise the probability that a client will default, how to get a return on investment to mitigate the imposed risk and delinquency assumptions (Thomas L. C., 2000, p. 152). The change in methodology can be attributed to the average tenure of a loan actually staying on book with the lender. As a

result of the aforementioned objectives, economic profits and opportunity costs also influence strategic and tactical lending decisions, contributing to the overall robustness of the automation and technical attributes of credit scoring.

2.2.4. How credit risk analytics work

The credit risk scoring model investigates relationships based on historical information, and predicting future credit performance for which this relation can be described as follows (Lahsasna, Aion, & Wah, 2010, p. 116).

$$f(x_1, x_2, \dots, x_m) = y_i$$

Each customer contains;

- (i) m_i attributes, x_1, x_2, \dots, x_{m_i} ; and
- (ii) y denotes the type of risk classification for the borrower such as good or bad; whilst
- (iii) f denotes the function of the credit risk scoring model and maps the borrowers features (inputs) and creditworthiness (output), with the primary task of the credit risk scoring model (function f) being that of predicting the value of y_i , and is the **creditworthiness** of borrowers i by knowing the variable(s) x_1, x_2, \dots, x_{m_i} , and denoted customer features such as age and income (Jagric & Jagric, 2011, p. 76).

This analysis and formulated equation is one of the core foundations that credit risk analytics and decision making models are based on. It makes use of past information and implied assumptions in order to predict possible future outcomes. (Lahsasna, Aion, & Wah, 2010).

Credit worthiness is often mistaken for personal attributes like height, weight, eye colour and as something that can be directly measured (Thomas, Edelman, & Crook, 2002). Application, behavioural and bureau risk scores are sometimes mistakenly believed to be creditworthiness scores or measures however in most instances such assessment mechanisms ignore the loss severity and omit profitability elements (Thomas, Edelman, & Crook, 2002).

Finance theory though typically review four major balance sheet risks;

- (i) liquidity,
- (ii) interest rate,
- (iii) capital, and
- (iv) credit risk.

Of these noted measurements, credit risk is in general identified as the key risk influencing bank performance and bank failure (Boffey & Robson, 1995). Banks can only absorb certain levels of debt losses and through credit scoring principles such losses and the probability of default is managed. Losses are either absorbed through other profitable loans, or bank capital (Boffey & Robson, 1995).

Credit risk analytics can then be described as the method and tools used by lenders to minimise the impact of possible future defaults due to lending activities. Where lending decisions were once based on *relationship lending*, the *transition to transactional lending* has taken place, and although the former is more appropriate in communities where lender and borrower have personal knowledge of each other, it has become ineffective in an era characterised by high customer mobility and high complexity in terms of predicting consumer behaviour (Anderson, 2007).

It is a statistical technology that quantifies the credit risk posed by a prospective or current borrower and seeks to rank order individuals by their credit risk score so that borrowers with a poorer score and expected to perform worse on their credit obligations be ranked in accordance with those who will perform favourably, whilst promoting competition and overall improved market efficiencies (Avery, Brevoort, & Canner, 2009).

In its simplest form, credit risk scoring is defined as the use of statistical models to convert data into numerical measures through data variables in order to guide credit decisions. It is also known to be the industrialisation of trust; a logical development of the subjective credit rating variables as provided by nineteenth century credit bureaus, and driven by the lenders need to deliver fast, objective, and consistent credit decisions which are made possible by advancements in technology and computation (Anderson, 2007).

Credit risk analytics introduced benefits in terms of cost savings, increased access to credit and promoted competition between credit providers to ultimately improve market efficiency (Avery, Brevoort, & Canner, 2009), however leniency towards accepting

poorer risk borrowers ultimately defeats the purpose of risk scoring and fails as a gatekeeper to minimise adverse risk implications.

2.2.5. What credit risk analytics do

Credit risk analytics is the systematic method employed for evaluating credit risk and provides for a consistent variable analysis of the factors or variables that have been identified to cause and, or affect the level of risk a borrower can obtain in terms of creditworthiness and risk behaviour (Fensterstock, 2005). Credit risk scoring is also made use of to assist credit providers with the quantification and management of the financial risks involved in providing credit, to make quick and objective lending decisions, and to make use of credit risk scores as a predictor of risk or loss (Johnson-Speck, 2005). Miller (2003), further argues that actuarial studies concluded that credit scores are one of the most powerful predictors of risk, provided the risk variables made use of are the correct variables and does not lend it to adverse risk selections within risk assessment model outcomes.

Credit risk is a primary area for using scoring and is used for predicting delinquencies and traditional methods of credit risk measurement focused more on predicting the probability that a borrower would default on their repayment commitment (Allen, DeLong, & Saunders, 2004).

Credit scoring has many benefits to lenders and that of borrowers alike, since credit scores assist in the reduction of discriminatory factors. Credit scoring models also provide for an objective analysis of a customer's creditworthiness, which enables credit providers to focus on information that relates to credit risk and to avoid personal subjectivity associated with that of the traditional manual interventions and presumptions of a human credit analyst (Fensterstock, 2005).

The contrary though is that since credit risk analytics also provide for exclusionary practices, the objectivity is questionable, risk models operate and assess borrower risk purely based on what information is provided and what variables are to be made use of. Throughout the literature the benefit of automation has been praised however in dealing with different credit segments different solutions are to be adopted and perhaps that of automation does not completely solve the requirement.

Making use of predictive variables ranging from; application forms, external data suppliers and own company records, statistical models also called credit risk scorecards are used to determine the probability of a borrower defaulting on loan

obligations extended by the lender and is purely subjective in terms of the variables that are being used for accept and reject borrower risk analysis (Verstraeten & Van den Poel, 2005, p. 981).

The argument for making use of credit risk scoring models is that it assists with the speed and consistency of loan application processes and allows for further automation of lending processes (Rimmer, 2005). The aforementioned automation of processes is there to supposedly reduce the need for human intervention in credit risk assessment processes which is associated with that of an additional cost benefit for the lender (Wendel & Harvey, 2003; Diana, 2005).

Financial institutions can quantify the risk associated with granting credit in a quicker time, and Leonard (1995), found that the average credit processing time per application reduced from nine days to three following the implementation of credit risk scoring. It is stated that credit risk scoring can improve and assist with the allocation of resources towards the best equilibrium fits however the challenge remains in establishing the best equilibrium (Jacobson & Roszback, 2003).

Credit risk scoring models have also been credited with the enablement and development of sub-prime lending industries and allowed for consumers with poor credit records to obtain credit (Quittner, 2003). Traditional requirements to exclude financing because of credit impairments, missing data, or poor credit histories with difficulty in validating borrower incomes were suddenly included and accepted as approved risk variables in accepted loans (Quittner, 2003). Ironically but to no surprise, this high risk propensity ultimately contributed to the global financial crises and meltdown of retail consumer credit in the United States and was brought about due to “subprime” lending practices.

Ever since the early 2000’s there has been a significant increase and uptake in sub-prime lending and was brought about the automation factor of credit risk decision making and analytics. Advances made in technology allowed for more intelligent credit scoring models to be developed and enabled lenders to better develop and implement collection strategies and optimise the effective utilisation of resources (Cundiff, 2004). These intelligent risk scoring models are of significance as it allowed for a higher risk client to secure debt which was secured by a secured loan facility and backed by the valuation variable that was believed to never decrease.

There is however strong contradictory literature written on the perils of extending credit to unsatisfactory borrowers such as those in the United States prior to 2008 that received credit purely to stimulate the market. Rapid spending on freely available credit

did have a global recessionary impact and was largely due to the acceptance of below standard risk objectives. The question should then be asked, does automation rather not perpetuate and exacerbate the problem should the risk model pose an error or accept the incorrect risk group? The same can be said for the affordable borrower segment in view of exclusionary factors.

The implementation of credit risk scoring models can lead to improved credit scoring results, and from a management perspective the following benefits can be associated with credit risk scoring (Koh, Tan, & Goh, 2006, p. 114):

- (i) Reduce subjectivity and increase objectivity in credit risk models and assessment practices.
- (ii) Increased speed of risk assessment requests with a higher consistency.
- (iii) Automation of the risk assessment process.
- (iv) Improved or optimal allocation of credit risk resources.
- (v) Improved determination of interest rates, loan duration or term and loan amounts to be offered.
- (vi) Better risk management practices.
- (vii) Assist with implementing improved collection/recovery strategies.
- (viii) Further development of non-traditional borrower markets.

The concern through lies with the mass automation of risk models and the effect of such models should the variables and assumptions employed by such models are incorrect or manipulated.

2.2.6. Construct of credit risk scoring models

Koh, Tan, & Goh, (2006) states that the general methodology for constructing credit risk scoring models should incorporate the following basic model data disciplines and risk assessment constructs:

- (i) Sample of previous borrowers or customers is selected and classified as “good” or “bad” which is based on the repayment performance over a specified period. However it has been identified that declined applications are not included in the performance tracking for portfolio performance measurements and as such leads to reject inferences being applied to risk models. It is not just the credit decision and scoring model that should be accurate, but also the supplementary analysis in order to obtain a better understanding of causal relations (Wu, 2011, p. 297).

- (ii) Data are compiled from loan applications, personal and/or business credit records, and various other sources such as credit bureau reports (Koh, Tan, & Goh, 2006).
- (iii) The final step is the statistical or quantitative analysis that is to be performed on the derived data to formulate a credit scoring model. The model comprises weightings that are applied to the different variables and attributes in the data whilst cut-off points are dependent on the weight being applied to each variable the risk score will be changed with such as for example RTI (Repayment to income) variables that has certain cut-off levels in terms of risk assessment scores in certain scorecards.

The derived sum that is applied to the weights as applied to the variables for that of an individual consumer or borrower constitutes the credit score, and the derived cut-off position will determine whether this customer should be classified as a “good” or “bad” risk for credit extension purposes (Koh, Tan, & Goh, 2006, p. 102).

The complexity though is that for multiple credit scoring models, new explanatory variables are required, and especially for those variables that require averages or proxies for non-measurable data which can be implied based on population demographics (Vogelgesang, 2003).

According to Thomas (2000), credit risk scoring can take on different forms which entail that any one of the following three forms can be adopted:

- (i) Judgemental approach, which is the most common approach still in use by credit lenders and practitioners, which primarily assesses the risk, based on the opinion and experience of the credit risk officer although most risk scores are first obtained from a bureau score which is provided by an external data source to that of the lender.
- (ii) Statistical approaches which is based on historical data and make use of discriminant analytics (characterises or separates two or more classes of objects or events and attempts to express one dependant variable as a linear combination of other features or measurements) and logistic regressions.
- (iii) Non-statistical, non-judgemental methodologies that include operational research methods, genetic algorithms and neural networks.

However, the assessment outcome for those of different credit scoring approaches can be conflicting and contradictory more often, which does not adequately address the performance or implied risk predictability as intended (Baesens, et al., 2003). This is

why a forth risk analytic model is worth further investigating. Unlike previously stated risk models, the Neural Network based system sees the available data as patterns categorised to estimate the future risk propensity for default which is based on the data that was initially used in the training of the system (Fensterstock, 2001, p. 58)

2.2.7. Limitations of credit risk scoring models

In a 1996 journal Desai, et al. already stated that neural network analysis can perform better than linear discriminant analysis but in a more recent journal Yobas, et al. (2000), found that linear discriminant analysis performs better than neural network risk models. Linear discriminant analysis have problems with handling the categorical nature of the credit data and heteroscedacity (Jagric & Jagric, 2011, p. 76). The Yobas, et al. (2000) and Jagric & Jagric (2011) findings further contradicts the validity and consistency argument for implimenting automated credit risk analytics.

Thomas (2000), concludes that credit risk scoring should rather employ a pragmatic approach in that when it works, use it, and in particular Thomas (2000) favours the notion that certain combinations of different approaches can be used.

Maximum likelihood estimators are employed for logistic regressions, and require computationally more demanding procedures than that of discriminatory analysis and linear regression analysis whilst, logostic regression models does not have to be constrained by the assumption of normality and normally distributed data since the model predicts a probability variable (Sharma, 1996).

As a credit risk scoring model, logistic regression (*defined by Sharma in 1996 as a model that deliver an output that is interpreted as a percentage term, directly interpretable and usable for performing operational actions such as establishing and setting cut-off values in risk scores*), has problems with strong model assumptions, such as that of variation homogeneity with limited application for modeling credit scoring constraints (Jagric & Jagric, 2011, p. 76).

2.2.8. Limitations of credit risk analytics

Credit risk scoring does however have its limitations due to data dependencies and the backward-looking nature of scoring models. Whilst credit scores dominate in automated high-volume low-value environments, credit ratings still imply some degree

of subjective input, especially applicable to larger loans issued to companies or governments (Anderson, 2007).

The first problem contributing to credit risk analytics limitations is that with the construct of credit risk scoring models, risk models may be developed using a biased sample of consumer/borrower data that have been granted credit thus eliminating the population that did not participate in the outcome of receiving credit such as that of the perceived high risk affordable borrower market (Hand, 2001).

This limitation is presented since applicants or potential borrowers that are rejected will not be included in the data for constructing and developing the model and the sample will be biased due to an over representation of favourable or good risk borrowers (Koh, Tan, & Goh, 2006, p. 112).

As a result of this over reliance on favourable risk clients, the credit risk scoring model may not perform well and in accordance with the entire population it intends to assess, because the data that will be made use of to build the credit risk model will be different from the data used for the borrower segment that the model will ultimately be applied to for credit risk assessment and analytics (Koh, Tan, & Goh, 2006, p. 113).

The second contributing problem is that when constructing credit risk scoring models, there might be a change of patterns over a period of time, which impacts one of the key assumptions for any implied predictive risk model and that is that the past can predict the future (Berry & Linoff, 2000).

For credit risk scoring and analytics, this entails that the characteristics of past applicants who was subsequently classified as “good” or “bad” creditors can be used as predictors of the current credit status for new applicants, although the tendency for the distribution of the characteristics to change over time is too fast and will require for a constant “refreshing” of the credit risk scoring model in order to remain relevant (Koh, Tan, & Goh, 2006, p. 113).

A third limitation risk is that due to credit risk scoring, end-user risk practitioners become so reliant on the technology and what the technology tells them, that they reduce their need for prudent judgement and the need to exercise their self-obtained knowledge pertaining to special cases. For certain cases, end-users may unintentionally apply more resources than necessary to work the entire risk portfolio manually (Lucas, 2000), or fail to even consider the impact of understaffing borrower portfolios such as the affordable borrower market perhaps

The fourth limitation identified is created by a new emergent industry practices whereby borrowers are assisted in rearranging their finances to improve on their credit risk scores. The latter has the potential to distort the patterns of credit defaults since the true performance indicators remain hidden from the public domain whilst these loans are either under debt review or restructure agreements in times of economic downturns (Timmons, 2002). This is a particular concern in the affordable borrower market with the high usage of informal credit options that is largely unregulated.

There are still omissions or untapped data and resources in the mortgage market that is not made use of in credit risk analytics (Feshbach, 2009, p. 42). Loan level data or debt to income (DTI), remain an important source for risk assessment purposes, but can only go so far in terms of determining the default propensity of a borrower. In the affordable borrower market this variable can be completely misinterpreted due to the omission in applying information for risk analytics that has been available for many years such as borrower tax information and additional complementary variables that can contribute to the establishment of neural credit networks (Fensterstock, 2005, p. 42).

The fifth and arguably main limitation is that risk assessment data does not capture changes in borrower and deposit characteristics. Such changes occur more frequently post origination of the debt, and these dynamic changes will have an impact on the perceived risk levels of the overall borrower pool (Feshbach, 2009, p. 42).

2.2.9. Exclusionary biases of risk analytic models

An important exclusionary bias of credit risk analytics worth mentioning in this section is the inherent discriminatory elements of credit scoring techniques and sample bias. The challenge is in estimating the default probabilities for all future credit applications using a model that is already functioning and making assessments off such premises of a skewed sample construct which only contain records for previously accepted applications (and not rejected applications) (Verstraeten & Van den Poel, 2005, p. 981).

For historically rejected applications, it is not possible to observe the risk decision outcome, being whether or not the applicant was able to repay the debt incurred. Hand & Henley (1997), refer to a set of procedures that were developed to decrease this bias phenomenon that arise from building scorecard models based on accepted applicants only which is known as, “reject inference”.

The reject region differs in a non trivial way from the accepted region, and historically sample bias has been accused of introducing at least one of two major shortcomings in risk analysis model constructs namely, population drainage and biased estimates (Hand & Henley, 1997).

Risk assessment models that are only applied to borrowers that were accepted in the past , whilst those borrowers rejected from obtaining credit will remain rejected, will inevitably lead to a decrease in the volume of the borrower base. The term “population drainage” is what this occurrence is called (Verstraeten & Van den Poel, 2005, p. 982).

The conclusions is that risk models require sufficient records in order to measure outcome validity of applied risk decisions. It is clear that as volumes reduce, the volatility within the model becomes greater thus less of a predictive outcome can be obtained. For risk assessment validity that is then representative of the total assessment population, the risk assessment model has to be re-configured or changed.

2.2.10. Disciplines of credit risk analytics

Multiple methods for conducting credit risk scoring and analytics have been investigated in order to develop accurate credit scoring models, for which statistical methods are still favoured (Lahsasna, Aion, & Wah, 2010, p. 115).

The benefits of collaborative data and analytics should not be underestimated. More sophisticated ways of capturing, integrating and computing data in a disciplined way will be critical to establishing predictability since it will further instil and restore confidence in the mortgage backed and secured lending practices (Feshbach, 2009, p. 41).

The traditional approach of credit scoring required that an underwriter make judgemental assessments of prospective borrowers which encapsulate the 5 Cs:

- (i) *character* (of the applicant);
- (ii) *capacity* (to borrow);
- (iii) *capital* (as backup);
- (iv) *collateral* (as security) and;
- (v) *conditions* (external factors) reference.

The 5 Cs are primarily based upon the underwriters own experience even though the majority of risk decisions are made through an automated risk model as identified by

Lucas (2000), whilst incorporating what was learnt from mentors, taking into consideration not only historical information, but also taking a forward looking view of the borrowers prospects (Anderson, 2007). The latter requirement is once again at the forefront of risk analysis in view of the global liquidity meltdown and new more innovative ways for conducting credit risk analytics are sought.

The misjudgement of the “Black Swan” event which took place when the global economy went into recession from 2006/7 onwards evaporated hundreds of billions of dollars in housing value, toppled once mighty institutions and questioned the stability of the global economy through such difficult recessionary periods (Feshbach, 2009, p. 41). Was this perhaps due to the over reliance on automated models which contained inherent risk acceptance variables without any subjectivity in the decision outcome?

Because of the “Black Swan” phenomena, experience has now been replaced by data and less of reliance is placed on human judgement playing a role in the assessment process. The 5Cs still apply; however borrower credit scores can now capture predominately most of the 5C’s by extracting maximum value out of available information. Scoring does however lack the forward-looking element (Anderson, 2007).

According to Feshbach (2009), no data model could have anticipated the recent credit and liquidity crisis, and suggest that a more responsible and disciplined approach to the utilisation of available data and analytics within firms and the industry should have been followed which could have signalled the warning signs. This might have mitigated certain elements of the onslaught and devastation in 2006 when the first signs of credit deterioration became apparent (Feshbach, 2009, p. 41).

The credit scoring process for applied models, use the observable characteristic variable of the credit applicant for which it calculates a score to present the credit risk which classify the borrowers or applicants into different risk levels. The idea behind developing a credit scoring model is to select variables and determine their respective weights (Jagric & Jagric, 2011, p. 75).

By making use of credit scoring, the focus has shifted from being *relationship lending* to *transactional lending*. Although the former is more appropriate in communities where lender and borrower have personal knowledge of each other, it has become ineffective in an era characterised by high customer mobility and extended branch networks (Anderson, 2007).

2.2.11. Alternative approaches for credit risk analytics

Improvements or gains made in credit risk models such as an improvement in prediction accuracy or an improvement in a fraction of a percentage may translate into huge savings (West, 2000).

Artificial Neural Networks (ANN) is some of the most commonly used soft computing methods used in credit scoring modelling (Lahsasna, Aion, & Wah, 2010, p. 116), however of the credit scoring methods currently in general use, neural networks is perhaps the least known of, and perhaps may be the most powerful credit risk technology available (Fensterstock, 2001, p. 58).

Perhaps the answer is contained in the combination of multiple risk models. The critique and benefit is that combined individual credit scoring models, although more complex in terms of workings and applicability for extracting true value and meaning, may lead to models that perform better than individual models, such models are based on the accuracy rate and additional assessment criteria that can be distorted when combined with another model, and falls outside of the scope of this research report.

One of the identified limitations is that there may be substantial difficulties inherent in how interpretations and rules are generated by combined risk models and that the outcomes produced by such a model may not make sense to the risk assessor since the model built and construct is based on individual model constructs that can generate different rules rather than function in cohesion (Koh, Tan, & Goh, 2006, p. 113).

2.2.11.1. ANN – (Artificial neural networks)

Artificial Neural Networks (ANN's) are mathematical representations inspired by the human brain functioning, and is composed by a number of simple processors known as neurons working in parallel without central control (Lahsasna, Aion, & Wah, 2010, p. 116). A system of weighted connections determines the information flow through the network for which ANN's model complex real world problems (Lahsasna, Aion, & Wah, 2010, p. 116).

A neural network that is patterned after the structure of a human brain, can perform certain human-type tasks, which constitute to be a valuable tool when applied to the problem of credit risk analytics, which essentially is a synthesis of the available information that is extrapolated to produce an estimated risk variable pertaining to the inherent future performance (Fensterstock, 2001, p. 58).

A neural network as depicted by Figure 3 can be represented as a finite of directed graphs and consists of input and output node (Tsakonas & Dounias, 2007, p. 682).

Conventional computing principles are vested on logic and mathematics. The architecture that is used by most computers assume that of a structured sequential process flow of a central processing unit connected to the memory that contains a storage programme and is managed by a central processor. For neural computers an alternative approach is followed in that these systems directly follow and model the biological structure of the brain and the way in which information is processed (Fensterstock, 2005; Fensterstock, 2001).

This requires new ways of thinking to be developed that consist of larger numbers of interconnected processing elements that can operate in parallel as per the illustration in Figure 3 (Fensterstock, 2001, p. 58).

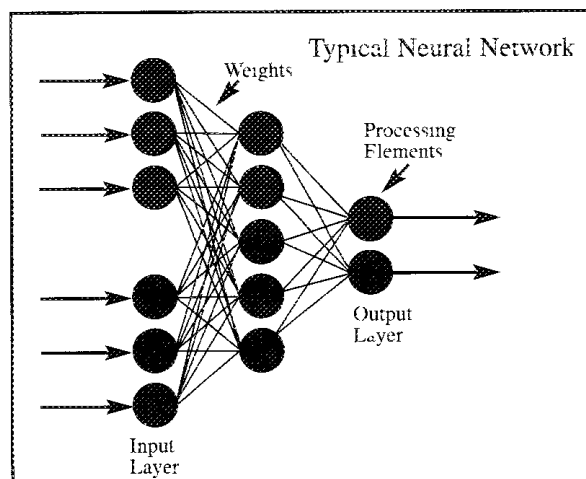


Figure 3: The General form of a globally connected feed-forward network

(Multi-Layer Perception) Neural Logic Network (Fensterstock, 2001)

There is however arguments against the effectiveness of what neural networks have to offer. The rate classification for the models are generally high, however there is criticism in terms of the black box nature of certain models (Baesens, Setiono, Mues, & Vanthienen, 2003, p. 313). The omission of an explanation as to why certain applicants are categorised as either good or bad risk clients (Lahsasna, Aion, & Wah, 2010, p. 117)

Neural networks should be used in parallel as a complimentary technique that functions in conjunction with other operating and computing techniques for conducting risk analytics. A neural network can be used to initially analyse incoming data and to

classify the data into various categories for which each category or variable may require different processing techniques (Fensterstock, 2001, p. 58).

Perhaps one of the most important attributes is that neural networks can be taught and that neural networks can perform a complex task that does not require programming in accordance to that of more conventional systems. Neural networks can learn from experience, generalise from examples, and can extract essentials from noisy data (Fensterstock, 2001, p. 58; Fensterstock, 2005).

The slow market adoption rate for not implementing neural network methods in credit risk analytics is perhaps due to the lack of explanatory skills, funds or capabilities of these techniques enhance the transparency of neural networks which is one of the critical factors for successful deployment (Baesens, et al., 2003). Explanatory capabilities are critical in risk assessments as the evaluator has to justify the reason as to why certain applications are approved and other rejected (Baesens, et al., 2003, p. 312)

As an alternative to credit risk analytics, neural networks as a powerful risk analytic decision tool, is less expensive once developed and perhaps one of the most efficient ways to deal with multiple interlinked and complex decisions. The decision making process will identify the best solution where a trained individual is responsible for the current process and assessment outcome, rather than using multiple unrelated models (Fensterstock, 2001, p. 58).

2.3. Sociology and Borrower Identity

2.3.1. Identity

“Identity”, as classified by Stryker & Burke (2000), “...is the language of identity, is ubiquitous in contemporary social science and cuts across psychoanalysis, psychology, political science, sociology and history.”

There are three relative distinct implied reference paradigms assigned to identity (Stryker & Burke, 2000);

- (i) Identity is used to refer to the culture of a people; and there is no distinction drawn between identity and ethnicity, thus the theoretical purpose of the introduction of identity is obscured. Individual and sub-groups continue to be classified and segmented based on background and especially past performances.

- (ii) The second use of the term identity is to refer to the common identification with a collective or social category as in social identity theory studies (Tajfel, 1982), and contained in contemporary works on social movements and the creation of a common culture among participants with an inherent importance vested in the construct of risk analytics and risk models.
- (iii) The third use of the term identity is to reference to the parts of a self, composed of the meaning that persons attach themselves to the multiple roles they immerse themselves in and in that of highly differentiated contemporary societies.

Identity research is no new phenomenon and has vested roots in the early works of Stryker, (1968) and Stryker & Serpe (1982), whilst *social identity theory* research has its origin from the work of Tajfel, (1959).

The former primarily seeks to explain that individuals are in relation attached to the multiple roles they play in contemporary societies (Stryker & Burke, 2000). As such these contemporary societies become the very statistic that will be incorporated in the predictability element of credit risk analytics for the purpose of modelling risk outcomes.

2.3.2. Borrower identity

The history of and information pertaining to that of the past borrowings of the borrower is used for establishing predictability of a future borrowing outcome, and with the movement towards greater grouping and social cohesion, exclusionary risk assessment models can perpetuate even greater future exclusions.

Tajfel and Turner (1986) also identified three classes of variables that might influence intergroup differentiation and that with imposed pressures for distinctiveness should increase with comparability;

- (i) subjectively identifying people with their in-group;
- (ii) allow for evaluative intergroup comparisons; and
- (iii) the out-group must be comparable in terms of similarity or proximal.

Groups however differ in their endorsement of social diversity, social inequality, and their orientation in approach toward collective action for identifying with the ethnic group (Deaux, Reid, Martin, & Bikmen, 2006, p. 123). Although subjectivity is to be applied in terms of group identification, the concern lies with how to manage the group and especially when the group is classified as an out-group which does not comply or confirm with the unique attributes and demographics of the in-group.

The out-group remains a contentious discussion in terms of credit risk analytics. Risk models have a level of “reject inference” associated with rejected loan applications (such as the affordable borrower segment). As such, it is not possible to observe the risk decision outcome since the borrower was excluded and no measurement of performance took place post such exclusion.

In effect, there is no evidence as to whether or not the applicant was able to repay the debt repayments. This inherent credit risk model attribute with applied risk modelling constructs prohibits the affordable segment from being considered as part of intergroup comparisons or becoming part of an in-group. Further, this phenomenon eliminates the subjectivity of identifying members of the affordable group which does not allow for the out-group (affordable borrower segment) to be compared to that of the assessed or “in group” in terms of similarity and proximal.

2.3.3. Social identity

The perceived entitativity of the social group has important implications in terms of social attribution, stereotyping and group impression (Svirydzenka, Sani, & Bennett, 2010, p. 611). Stryker & Burke (2000), define identity salience as; “...the probability that an identity will be invoked across a variety of situations, or alternatively across persons in a given situation.”

In the case of competing for capital resources and allocations, there will always be a situation where one of the groups will be deprived of participating. Is there an element of invincible competition and believe that the perceived in-group (wealthy cohort) is the only participant of such groups entitled to participate in sharing the proceeds awarded to the market through means of credit modelling and risk analysis? Social identity theory ultimately implies that, indeed a subtle form of competition does exist, and it will continue to exist because only one of the groups could be the better one (Correll & Park, 2005, p. 343).

Contextualising the importance of social attribution and stereotyping, credit risk models continue to be developed with a bias towards judgemental and discriminant statistical models, whilst neural network models continue to lag and has slow or limited adoption rates due to complexities and lack of business understanding and development from the industry.

The effect though is that unrepresentative and underrepresented population cohorts are not factored into the design and modelling of the new credit risk scorecards and

modelling. With this in mind, the question on how to address credit distribution requirements becomes even more topical.

“In-groups” become “out-groups” and can no longer participate or even be considered for inclusion in the in-group unless either the out-group becomes an in-group or individual circumstances change to the effect where such individual is no longer considered to be part of the initial out-group.

Identification with a particular social group may contribute to a sense of self-worth, and in particular when the in-group is positively distinct in comparison with that of the out-group (Vanhoomissen & Van Onverwalle, 2010, p. 84). Projecting features of the self in defining the in-group leads to the derogation of the out-group, which is caused by self-competition, and is pronounced under conditions of self-priming that render characteristics of the self being more accessible. As a result it demonstrate (i) that in-group favouritism is due to self-anchoring or the projection of self-features in the in-group, (ii) that out group derogation is a result a direct consequence of self-competition which is a self-contrasting reflection of the self (Vanhoomissen & Van Onverwalle, 2010, p. 84).

Research has further demonstrated that individuals typically evaluate in-group members as being more favourable than out-group members and that under certain circumstances, unfavourable out-group targets are viewed more negatively than equivalent in-group targets (Lewis & Sherman, 2010, p. 212).

This takes the issue with the approach followed on how credit risk analytic models are constructed even further. Credit risk analytics are highly dependent on predicting the future based on the past through means of grouping cohorts of population constructs and measuring such grouped performances whilst drawing inferences from data that only talks to the past in order to develop future predictability. The individual is lost within the greater population as a sub-group and excluded should the sub-group constitute to be deemed and classified as an out-group.

In certain groups a higher degree of “groupness” are noted than in other groups (Abelson, Dasgupta, Jaihyun, & Banaji, 1998, p. 243), and the groups classified as being more “real” are argued to have greater potential to pose as perceived threat to an individual’s social identity than that of more trivial groups (Lewis & Sherman, 2010, p. 212). According to Mead’s framework in simplified terms, his framework asserted a formula; “Society shapes self, shapes social behaviour” (Stryker & Burke, 2000).

The sense of self is fashioned as to how interpersonal interactions inform the self and is shared by both identity and social identity theory constructs despite their sociological and psychological perspectives pertaining to different levels of analysis.

The importance of the individual within the group referred to as the social group is of fundamental significance. In society we are grouped in accordance with our societal affiliations, however at individual level our own uniqueness is what differentiates us from within the group. Unfortunately, groups can also be stigmatised and assigned undue stereotypes, such as being higher credit risk or traditional poor payers.

A social group is by definition a collection of individuals, however there is a difference as to the extent to which this collection of individuals is perceived as a group (Svirydzenka, Sani, & Bennett, 2010, p. 611). Individual members behave in a different manner to that of the group whilst individuality must not be underestimated in applied risk analytics.

The concern remains that the affordable group is classified based on demographics such as income and employment class, which often fail to recognise the individual attributes of each of its members and impose a pre-defined predictability outcome of what future behaviour will constitute.

Institutionalised status differences between social groups (groups based on characteristics such as gender, race/ethnicity, sexuality and income/social standing) have created and maintained systemic disparities in opportunities and basic human needs such as wealth, health and education (van Zomeren & Iyer, 2009, p. 645). As a result, disadvantaged individuals do not continue to accept the prevailing status quo and history has shown that such groups will challenge social inequality and injustices through collective action (van Zomeren & Iyer, 2009, p. 646).

The work of (Tajfel & Turner, 1979; Wright, Taylor, & Moghaddam, 1990) defines collective action; "...as any action that aims to improve the status, power or influence of an entire group". For societal upliftment as a whole and creation of sustainable wealth this statements is of particular importance. Social inequality is widening in terms of the upper and lower societal segments particularly in South Africa's Gini coefficient, whilst there is a growing notion that this phenomenon can no longer continue and require immediate re-address (Department: Human Settlements Republic of South Africa, 2009).

Social identity theory identifies the repertoire of discrete category memberships that a person will construct in terms of social category affiliation; (nationality, political) in order to belong, or to which one feel one belongs (Hogg, Terry, & White, 1995).

When specific social identity becomes the salient basis for self-regulation in a particular context, it is then that self-perpetuation and conduct become in-group stereotypical and normative reference. Perceptions of out-group members become out-group stereotypical, and intergroup behaviour acquire competitive and discriminatory properties depending on the nature of relations between groups (Hogg, Terry, & White, 1995).

Social identity theory's critical insight is that social groups tend to confer a sense of self-worth to their members (Correll & Park, 2005, p. 343). The argument is thus what constitutes to be a valued group, should the affordable and less affluent group be construed as a valued market or segment and be classified as an in-group for risk analytics?

Relations between different groups are defined against macro contextual factors such as, broader characteristics of society and for which these different groups are embedded, and are often further influenced by micro factors, such as the particular characteristics of the individual in the interaction (Barreto & Ellemers, 2010, p. 433).

Focusing on intergroup phenomena, Brown (2000) identified the following biases;

- (i) Intergroup bias occurs in the absence of objective causes such as conflicts of economic interests, and that group members tend to think that their own group is superior to that of the other groups (Brown, 2000, p. 747). This challenges the premises of how risk analytics and risk models are constructed. Risk models are based on the inclusionary in-group that constitutes the reference group for benchmarking purposes. It is argued that there is an inherent exclusionary component to risk modelling that is perpetuated through the construct of such models by the inclusionary group, which in itself is open to discriminate between the "in" and out-group.
- (ii) Intragroup homogeneity and stereotyping serves as linking observable features to deep inherent characteristics and are most common to all group members (Yzerbyt, Corneille, & Estrada, 2001, p. 144). Representing social division as stemming from the inherent nature of the group, such theories may serve to justify and rationalise social arrangements (Yzerbyt, Corneille, & Estrada, 2001, p. 144) and provide ground for exclusionary practices and

applied variables in risk scoring models. However overgeneralisation results from the observer seeing all group members as being the same, as having the exact same characteristics and properties, and as having the same psychological make-up, thus the complete failure to acknowledge the individuals own and unique attributes (Hamilton, 2007, p. 1078). It is thus argued that the perception of homogeneity is created in the group that combines the actual similarities amongst group members (Hamilton, 2007, p. 749). This is the issue taken with the approach adopted in modern risk analytics, borrowers are merely becoming a number with no individual value.

2.3.4. Issues with the construct of identity

People living their lives in small and specialised networks of social relationships, fulfilling roles that support their participation in those exact networks. The interaction and relationship that are subsequently established is influenced by the larger social construct in which those networks are embedded (Stryker & Burke, 2000).

These social identities are merely not evaluative, they are also descriptive and prescriptive; they are also evaluative and furnishing an evaluation which in general is widely shared within a social category. (Hogg, Terry, & White, 1995). The importance of social identity lends itself to important self-evaluating consequences, that motivate groups and their members to adopt behavioural strategies for achieving or maintaining in-group/out-group comparisons that generally favour the in-group and the self (Hogg, Terry, & White, 1995). The issue with this concept is the phenomenon where in-groups continue to dominate acceptance and usage of resources, thus the debate turns to how in-groups are protective of the benefits they share and enjoy, whilst disallowing the out-group to participate.

The in-group and similarly the out-group create stereotypes that are construed to be mental representations and properties of that group. Perceiving this homogeneity within the group facilitates overgeneralisation, reflecting perceived similarities of in-group and that of out-group members which further encourages stereotyping. The thinking of a stereotype in terms of traits, and abstract properties can encapsulate diverse behaviours, whilst sweeping generalisations are made fostering perceptions of homogeneity in such groups (Hamilton, 2007, p. 1079).

Out-groups are stereotyped as being less variable or homogeneous, and are judged and classified under many different variables (Hamilton, 2007, p. 1080). Most commonly demographical characteristics are used in risk analytics, with applied judgemental and discriminatory exclusionary factors build into models. Such assumptions with applied relevance are being challenged in terms of implied inferences that are purely based on income and demographics.

Does it make a difference what kind of attribute or dimension is used as the primary basis for comparison when it comes to inclusion or inclusion decisions, and if yes, should the variable not rather be tailored in accordance with the uniqueness of the population it is to represent?

As stereotypes are construed to be mental representations of a group and its properties, it came to be viewed as an abstract concept expressed in very general terms and traits. Perceiving homogeneity in a group facilitates overgeneralisation, reflecting perceived similarities of group members which further encourage stereotyping (Hamilton, 2007, p. 1079).

2.4. Affordable Housing

2.4.1. Why consider the affordable segment?

According to more contemporary versions of globalisation, it is assumed that market deregulation is the precursor to domestic development and the growth of liveable cities (Spencer, 2010, p. 431).

The right to housing is a basic need whilst good quality and affordable housing is a reflection of a high quality of life, although it remains a challenge, especially for the bottom billions (Salfarina, Nor Malina, & Azrina, 2010, p. 977).

Relatively narrow understandings of poverty that manifest in the development life-cycle may obscure long term inequality dynamics that will become permanent in rapid developing societies through the creation of new social environments (Spencer, 2010, p. 432).

These socio-spatial inequalities can become cemented in the physical distribution of roads, buildings, water supplies, access to private financing responsible for developing such infrastructure, and other social physical assets, whilst the distribution of these assets are of particular importance when broad economic growth stagnates and are on the decline (Spencer, 2010, p. 432).

2.4.2. What makes the affordable market different to that of traditional markets?

By making the connection to social constructionism pertaining to affordable housing, it draws attention to the frequently contested and socially created nature of social problems by focusing on the role of power and influence that consistently shapes the emergence or disappearance of problems (Heiner, 2006).

It is argued that literature on the representation of poverty in policy and media discourse often obscure the real issues effecting the poor and is presented in an individualistic, un-contextualised way that has its own unique consequences (Bullock, Wychne, & Williams, 2001, pp. 229-230).

In effect, it can be said that the working class and poor have dropped out of popular and policy discourse on housing (Martin, 2011, p. 3). Poverty is most often portrayed as the fault of the poor and unrelated to structural causes which rarely provide for the contextual understanding of the causes of poverty (Bullock, Wychne, & Williams, 2001, pp. 230-231).

Noticeable changes in the demographic conditions has taken place with certain population segments such as that of Asia and Africa currently having much younger populations compared to that of Europe (Ronald & Chiu, 2010, p. 226).

Providing access to housing remains a critical imperative in view of the younger population phenomenon especially in countries such as South Africa where there is a drive to improve on social inequality and improve the overall wealth of the nation. Access to affordable housing is one such imperative and the means of accessing responsible credit for borrowers must be a primary objective.

2.4.3. Why the “underclass”

An “underclass” or affordable market segment is formed through the continued interaction of instability in the labour market, access to assets and in particular home-ownership, and the reinforcement of mentioned instabilities over time through the negative socialisation of spatial and proximal effects shown in Figure 4 (Spencer, 2010, p. 453).

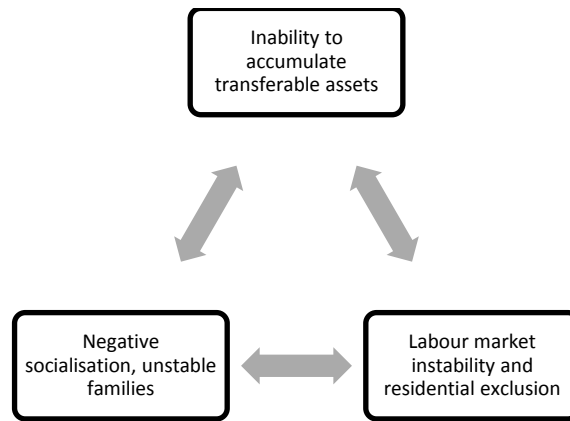


Figure 4: A general theory of underclass formation and persistence

(Spencer, 2010, p. 435)

Through combined practices whereby social and economic disadvantages are vested within an urban space, the re-segregation of cities and poor neighbourhoods are exacerbated as high-achieving members of the residentially segregated poor are able to achieve the goal of residential mobility, and as this segment achieve social mobility they gain the rights and abilities of special mobility (Spencer, 2010, p. 435). The objective is to provide access to the required resources such as credit and secondary housing markets to enable the migration of the poor to that of the middle class.

Massey and Denton (1993) argues that these socio-spatial dynamics associated with the urban business and residential market place were historically enabled through public policies, and most importantly that the institutionalisation of “redlining” practices of poor and minority urban neighbourhoods restricted the urban poor from accessing credit and systematically excluded minority neighbourhoods from affordable homeownership.

The challenge is now how to bridge the gap and provide for a more inclusive society by assessing and responsibly addressing those needs through improved risk scoring and analytics.

2.4.4. Societal importance of providing affordable housing

Apart from the issue of urbanisation, satisfaction in terms of housing and neighbourhood conditions is a further factor that constitutes to be an important indicator to consider in credit risk scoring. These indicators together with quality of life are

important for the process of scoring mortgage applications with an objective to improve and increase the quality of housing and neighbourhoods (Salfarina, Nor Malina, & Azrina, 2010, p. 977).

Where previous accessibility to housing was the predominant factor driving housing delivery, the emphasis has shifted to be more orientated and concerned with affordability and quality of housing that is being provided (Svirydzenka, Sani, & Bennett, 2010, p. 977). Quality not only relates to physical design, size and facilities, but also to the quality of neighbourhoods, and the affordability and location of the housing (Svirydzenka, Sani, & Bennett, 2010, p. 977).

It is thus argued that exclusionary or biased credit risk models tailored to the in-group, can lead to serious inequality problems and the distribution of basic infrastructure (Spencer, 2007). As a result long term socio-economic consequences will be more difficult to mitigate, and in particular social and spatial residential segregation will develop (Spencer, 2010, p. 437).

2.4.5. Affordable housing; welfare state or open market

According to the social constructionist view of affordable housing, it is acknowledged that the issue of biased credit risk scoring is not necessarily to be viewed as problematic in society, however even if affordable housing is to be seen as problematic by the public, policy makers and stakeholders, the nature of the problem with its causes, consequences and proposed solutions are to be contested and continue to be socially constructed on the in-group (Martin, 2011, p. 5).

A fundamental imperative is the continued creation and build-up of housing wealth for poorer affordable communities in order to offset the need to develop onerous and costly welfare states that cannot be sustained over the long term (Ronald, 2007; Ronald & Doling, 2007). The answer is sustainable wealth creation and not continued state participation.

The concept of private sector participation developing in terms of construction delivery and finance is seen as the natural response to the growing need for affordable housing and eliminating some of the discriminating factors of credit risk scoring models. Such support and growth acts as a natural catalyst and provide for much needed macroeconomic stimulus even though the public sector remains committed to leading affordable housing reform initiatives by assuming the role of real estate developer (Svirydzenka, Sani, & Bennett, 2010, p. 2009).

Thus, even though market forces do shape the housing market, for the affordable segment, a network of interlinked, interdependencies linking public and private housing sectors with that of financial institutions must be developed in order to manage the effectiveness of providing sustainable solutions in the need to meet the challenges faced by this market segment (Svirydzenka, Sani, & Bennett, 2010, p. 2010). Financial institutions are the key to providing credit and risk analytics who should ultimately drive the imperative to support to affordable borrower market.

2.5. Overview of Literature Review

The connection between credit risk analytics, sociology and affordable housing provides for well vested argumentative debate and research objectives, since such disciplines continue to emphasise the importance as to why it's important for society as a collective to prosper, and emphasis that only once real social upliftment has occurred within a population or subgroup, that real benefit will be created for all of society.

Literature on collective action and that of social identity theory is heterogeneous and provides grounds for substantiated debate to be initiated as to why collective social upliftment must be, and remain a primary imperative concerned and entrenched in societal policy and decision making process.

CHAPTER 3:

3. Propositions:

3.1. The imperative to research risk assessment parameters and borrower constructs

It is evident that the literature addresses the structure and function of the socially constructed self (identity or social identity) as a dynamic construct that negotiates the relationship between social structure or society and individual behaviour. The interconnected links between society and self are acknowledged whilst behaviour is considered to be an organised construct of meaningful units that are incorporated by specific self-definitions. Identity theory discusses the organisation of behaviour applicable to roles and social identity theory links norms and stereotypes (Hogg, Terry, & White, 1995).

Deep rooted historical social and racial/ethnic inequalities has led to the growth of persistent urban underclasses, recent settlement patterns have helped to reinforce societal inequalities and concentrate them amongst subgroups even as cities have prospered in the long term (Spencer, 2010, p. 434).

Credit analytics is the method used by lenders to minimise the impact of possible future defaults, and where lending decisions were once based on *relationship lending*, the transition to *transactional lending* has taken place. Although the former is more appropriate in communities where lender and borrower have personal knowledge of each other, it has become ineffective in an era characterised by high customer mobility and high complexity in terms of predicting consumer behaviour (Anderson, 2007).

3.2. Research Proposition 1

There is credit risk assessment exclusion factor(s) contained in the credit scoring methodologies that excludes certain members of the population from obtaining credit:

- (i) The understanding of the uniqueness and unique behaviour in terms of the affordable borrower segment is not adequately incorporated into borrower risk assessment scorecards and risk analytics.

- (ii) Historical, stereotyping assumptions made against borrower group;
- (iii) Applied risk assessment variables that are applied to credit risk analytics and based on borrower behaviour and affordability parameters has an exclusionary impact due to being benchmarked against the more affluent traditional middle market. Repayment to income ratios (RTI) are questioned in terms of relevance for predicting probability of default and actual loan performance.

3.3. Research Proposition 2

The over-reliance on automated credit risk scoring models unintentionally eliminates certain population groups (*Affordable Housing borrowers*) from gaining access to credit. Traditional risk assessment model constructs does not provide the analytics for assessing the unique nature of borrower performance in terms of less favourable loan repayment performances (*missed payments or slow payments*), and merely excludes borrowers from participating in the credit decision

- (iv) Imposing system analytics above societal, and individual psychographic attributes for making assessment decisions;
- (v) Historical, negative perceptions associated with borrower segment.

3.4. Research Proposition 3

That alternative scorecard characteristic are to be incorporated in automated credit scoring methodologies, specifically designed to cater for the uniqueness of the affordable borrower segment:

- (vi) The uniqueness of the borrower is absorbed in the review of traditional markets and thus has no individual developmental characteristics pertaining to that of the affordable borrower identity construct in scorecard development.

(vii) Vintage loan performance analysis for unsecured loan extensions is a predictive variable that can be applied to secured credit risk analytics and risk assessment models.

CHAPTER 4:

4. Methodology and Design

4.1. Research Design

The preferred research design opted for was firstly quantitative in nature in order to explain the relationship between variables (Saunders & Lewis, 2012, p. 113), and was conducted off the premises of reviewing credit application data that was collected from the retail banks un-secured borrower lending portfolio (quantitative), following which said collected data was applied to the credit risk scorecard of the secured retail lending division for further analysis of the obtained risk assessment outcome.

In addition to collecting quantitative data supportive of establishing the primary research discussion base, expert interviews were held with the secured lending divisions risk practitioners (qualitative).

The research approach thus followed a two stage approach by being descriptive and quantitative (analysis of secondary data of retail credit unsecured lending borrowers), and exploratory through qualitative analysis (expert interviews held with credit risk practitioners).

The research emphasis was in particular directed towards population and sample segments that were classified under the income definition of “affordable markets” as classified in the South African borrower context. The quantitative approach further allowed for data to be collected on pre-determined instruments that could yield statistical data for analysis and outcome discussions pertaining to a multiple of applied variables (Saunders & Lewis, 2012, p. 85).

Secondary data was used as it constituted loan application data from one of the financial institutions which was suitable for explaining the outcome of risk models that were based on pre-stated and implemented assessment criteria and further assist in the development of alternative recommendations pertaining to this research project (Saunders & Lewis, 2012, p. 84). As previously stated, analysis was applied to secured and unsecured financial instruments and specifically to client portfolios within the affordable income segment.

4.2. Choice of Methodology

In adopting the approach towards undertaking this quantitative and qualitative research study, the implied process for firstly conducting quantitative research was that (i) the study was undertaken within a pre-defined framework of a set of philosophies, that (ii) uses methods and techniques that has been critically validated for reliability and (iii) to be proven to be objective and unbiased (e.g. Dawson 2002; Kothari 1985, 2004; Kumar 2005).

Associated with the development of the qualitative research study, the diffusion of information at organisations and societal level was the challenge posed to the research questions in examining the relationship of how information was applied, and the purpose for which such information was applied. As a result, considerable consideration should be given to ways in which such research examination and analysis are conducted (Howcroft & Trauth, 2004, p. 198).

Triangulation was adopted in an attempt to map out and explain the richness and complexity associated with the discipline of credit risk analytics through adopting two distinct yet different methods to cast light upon the topic (Olsen, 2004).

Research requirements consist of (i) examination and assessment of the ways in which information systems are used and affect people; (ii) to examine and assess the research approach adopted for conducting such assessments (Howcroft & Trauth, 2004, p. 200).

An explanatory and qualitative approach was followed in order to assess the following:

- (i) the effect and impact of identity constructs on risk assessment models pertaining to that of the secured lending practices of a retail bank;
- (ii) the manner in which identity constructs were developed, what fundamental principles were incorporated in the design attributes of borrower identity constructs (if any), and how these implemented risk models functioned as applied and integrated risk assessment models' and
- (iii) to further enhance the richness and contextualisation of interviewee viewpoints in support of the quantitative data review.

A descriptive and quantitative approach was followed in order to assess the following:

- (iv) to further assess the outcome of credit risk approval/rejection decisions by applying the actual unsecured borrower data set to the secured lending credit scoring model and extrapolating the accept / approved and reject

risk review outcome, whilst establishing what exclusionary variables influenced credit risk decisions and borrower exclusions at a secured lending risk portfolio level.

4.3. Scope

Applied risk model assumptions as used by the lender's credit lab was analysed as part of understanding borrower identity construct with applied risk assessment usage.

4.4. Universe / Population

South African retail financial and banking institutions that had diverse lending operations and that could further successfully service the majority of retail consumer credit segments at a national level. The institution was further deemed to be an active participant in the retail consumer credit market, and that the institution processed sufficient volumes of secured and unsecured loan applications on a continual basis. The financial institution was considered as a strategic enabler for providing South African households with consumer credit solutions.

Emphasis was placed on data from 2006 onwards as it had to coincide with new legislation that came into effect as of 2006/2007 that governed credit lending practise with emphasis being placed on the abolishment of reckless credit lending practices which ultimately placed the onus and vesting responsibility of acting as a responsible custodian for the granting of credit with financial institutions.

4.4.1. Participants and size of the industry

In South Africa the consumer credit market is worth an estimated R1.30 Trillion as at the end of the quarter for December 2011. The credit market is segmented into corporate, business and retail banking of which the affordable housing segment falls within retail banking. As at the end of December 2011, mortgage lending comprised of R791, 1 Billion (61.03per cent) and unsecured credit at R112, 98 Billion or 8.72 per cent of total credit (National Credit Regulator, 2011).

The South African banking industry consists of FNB (First National Bank), ABSA (a subsidiary of Barclays Bank UK), Standard Bank (Subsidiary of the Chinese Commercial Bank) and Nedbank.

The total household population for South Africa was estimated to be 12, 6 Million households of which 7, 1 Million is within the subsidised housing market and 31 per cent or three comma eight Million households fall within the GAP and affordable market (Statistics South Africa, 2005/2006).

The primary universe was a retail banking institution that had diverse retail lending operations, whilst the secondary universe was;

- (i) the secured lending division and,
- (ii) the unsecured lending division of the identified retail bank.

The unsecured lending division provided the data for sample composition purposes.

The design, testing, execution and review of credit application scoring decisions primarily vest with the secondary universe of the secured lending divisions, constitute to be highly technical and specialised in nature. Associated credit practitioners operating from within the secured business units constructs, were to be engaged for research gathering purposes in order to conduct exploratory research analysis gathered from interview questionnaires.

Engagements were also not limited to credit practitioners operating within the affordable market segment since this sector of risk analytics was still underdeveloped in terms of assigned resources. As such credit risk practitioners from more traditional market segments were interviewed and all interviewees identities were not to be disclosed for which the research report will merely make use of pseudonyms.

4.5. Unit of Analysis

The following units of analysis were included in the research

- (i) Secondary universe, credit practitioners responsible for developing, testing, implementing and maintaining credit scorecards.
- (ii) Retail credit borrowers as classified under the definition; “Affordable Market” requesting unsecured lending facilities from the retail bank.

Applied assumptions in terms of borrower identity construct and credit risk modelling formed the basis for conducting this research, and was a requirement to engage directly with the practitioners that were responsible for the design and implementation of such assessment models.

4.6. Sample Method

4.6.1. Sampling technique – unit of analysis for qualitative data gathering from retail credit risk practitioners:

The secondary universe pertaining to the secured lending business units of the identified retail banking institution (Primary universe) was used in compiling this sample. From the noted sample, credit practitioners responsible for developing, testing, implementing and maintaining credit risk assessment scorecards were interviewed (Appendix D).

Solicitation in terms of securing participation was based on convenience. Due to confidentiality reasons, identities of delegates that participated in direct interviews were kept anonymous.

The specialised technical nature of the value analytics credit scoring business units for secondary universe constructs necessitated the need for the credit / risk practitioner sample to have been derived from a non-probability sample. Limited resources employed within this segment and participant availability for conducting direct interviews further necessitated the need for convenience sampling techniques to be used.

4.6.2. Sampling technique – unit of analysis for quantitative data gathering purposes for retail borrowers requesting unsecured loan facilities:

Retail unsecured credit borrowers were identified in accordance with pre-defined unsecured loan application borrower data fields. The identified data fields was extracted in accordance to reporting and analysing those required fields to be completed at loan submission stage as required for risk assessment and risk model analytics.

Data obtained was for the period 01 July 2007 to 31 July 2012.

The total population of unsecured lending clients applying for a unsecured loan facility over stated period was made use of for purposes of conducting research data analytic (Total Population) and included the following application / borrower risk assessment data;

- (i) Reject/Declined application status (Affordability, Poor Risk Record, Missed Payment or Slow Payer Indicator)
- (ii) Approved application status
- (iii) Default and or vintage loan status
- (iv) “Thin” payment and credit risk profile indicator
- (v) Loan disbursed indicator

4.6.3. Sample size:

The data obtained was for a period of no less than four years (2006/7 – 2012), and include all the transacting records for borrowers that requested an unsecured loan facility from the retail bank over stated period (Total population).

Actual data of total population of borrower data records obtained for the period 01 July 2007 to 31 July 2012 for unsecured lending borrowers.

Total population of unsecured lending clients applying for an unsecured loan facility was;

N 4,895,417 unique records.

Further sub sets of data samples were derived from the total population and was then grouped based on unique data variables which included the following unique loan identifiers.

- (i) Affordable housing income segment based on income data,
- (ii) “Thin payment profile” and derived based on low or failed bureau scores with unique borrower risk scores.
- (iii) Repayment to income ratios (RTI)
- (iv) Loan payment commitment indicator (Slow payer or late payer)
- (v) Default ratios
- (vi) Debt to income ratios
- (vii) Vintage default ratios

4.6.4. Research Instrument/Measurement:

The research instrument consists of the following component(s):

(i) Credit Practitioners

- a. One-on-one interviews conducted with current practitioners responsible for designing and maintaining risk assessment lending scorecards within the secured lending divisions of the identified institution.

The questionnaire consisted of;

- i. Unstructured; open questions that allows for the practitioner to rather provide input and feedback based on own interpretations and believes rather than to guide the questions based on pre-defined questionnaire questions.

(ii) Retail Credit Borrowers – Unsecured Lending

- a. Obtained and analysed the total population of borrowers that applied for an unsecured loan facility between the period 01 July 2007 to 31 July 2012.

(iii) Secured Lending - risk assessment scorecard decisions:

- a. Applied the data obtained from the unsecured lending division to the risk analytics scorecard of that of the secured lending division in order to assess the accept/reject outcomes of the total population and sub populations based on the unique risk assessment variables.

4.7. Data Gathering and Analysis

- 4.7.1. Unsecured Lending data applied to secured lending risk assessment scorecard:

- (i) Gathered data from unsecured lending divisions for stated period and as per pre-identified data requirements.
- (ii) Identify common field denominators such as, risk score, risk assessment decision, loan amount approved/rejected, for the purpose of conducting a detailed data overlay and data matching analysis by applying secured lending risk assessment criteria to that of the unsecured borrower data construct.
- (iii) Analyse the outcome of the unsecured borrower data simulations applied to the secured risk lending model.
- (iv) Identify the unique number and value of unsecured loans approved/rejected for stated period against the outcome of applying unsecured data to that of secured risk assessment requirements in order to analyse the accept/reject trends.
- (v) Identify the unique risk exclusionary variables that eliminate borrowers from securing loan facilities from that of the secured lending division in support of the proposed research propositions

A structured process to data analysis with interpretation was followed (Saunders & Lewis, 2012) and comprised of the following steps.

4.7.2. Interpretation

Whilst the quantitative data were being collected, relatively low interpretations and analysis were made in terms of analysing and interpreting the quantitative data.

Once the data collection process was completed, more systematic, second level interpretations were made from the qualitative interview questionnaires in support of the discussion pertaining to the research propositions.

Interpreting qualitative data:

A constant comparison method was followed as the preferred qualitative analysis tool for analysing the questionnaires. It entailed the grouping of answers to common questions as contained in the interview questionnaire, and analysed different perspectives on central issues. The qualitative synthesis is a way to build theory through induction and interpretation and to explore lessons learned (Patton, 1990, p. 425).

Inductive analysis meant that patterns, themes and categories for analysis were able to emerge out of the data, rather than being imposed prior to data collection and analysis (Patton, 1990, p. 390).

The logic was for the development and formulation of the questions to precede the design of the method for answering these questions (Gillham, 2005, p. 159).

Mathematical data have no name or identity, and the interpretation only signifies the implied influence or outcome such data possesses and signifies and is no different to a qualitative researcher that assigned different statements to an interpretive category where there was an apparent justification in doing so (Gillham, 2005, p. 9)

The data collection process and analysis for the questionnaire and interview processes were conducted in accordance as proposed by Gillham 2005 and entailed the following:

- (i) Conducted the initial first interview.
- (ii) Made interview notes and transcribed the interview discussion.
- (iii) Captured the researcher notes, and developed themes using the words as used by respondent(s).
- (iv) Obtained an overall sense of the data, and started to extract interpretations and meanings from collected data.
- (v) Synthesized the data and transposed findings into propositions.

Interpreting quantitative data:

The obtained quantitative data was interpreted in terms of the total population and then further in terms of the identified sub-population that contained the following unique risk or assessment variables:

- (i) Borrowers classified as affordable segment borrowers.
- (ii) Borrower, debt to income ratios.
- (iii) Borrower payment status.
- (iv) Loan default status.
- (v) Repayment to income indicator.
- (vi) Thin or no risk bureau payment profiles for borrowers.

4.7.3. Cleaning and Organising the data for analysis

This step involved the checking and logging of the data for accuracy followed by transforming the data to be able to develop and document a database structure integrated the various measures.

4.7.4. Describing the Data

Quantitative data:

Descriptive statistics was used to describe the basic features of the data in this study by providing summaries about the population and the sub population with applied measures that formulated the basis for the quantitative analysis.

The use of descriptive statistics was also be made use of to present quantitative descriptions in a manageable form and was used to simplify large amounts of data in a sensible and more usable format.

4.7.5. Testing the research propositions

Statistics to be provided consisted of:

- (i) Borrower statistics in terms of borrower identity groupings that applied for un-secured lending facilities and being rejected through the secured lending risk assessment scorecard whilst being able to secure either one or multiple unsecured loan facilities over same time period.
- (ii) Exclusionary statistics that are based on traditional demographic risk model constructs. Analysis of unsecured acceptance statistics in relation to that of secured risk assessment models and noted differences in accepted/rejection trends.

4.8. Research Limitations

The following limitations have been identified in terms of the intended research design:

- (i) The inability to develop and answer the research question.

- (ii) The usability and correctness of the data for quantitative analysis purposes.
- (iii) The data is only applicable for one Retail Credit Lender and may not be universally accepted or relevant. Legislative changes implemented in 2006/2007 may have further implications due to changed risk assumptions and risk assessment parameters that falls outside of this time frame.
- (iv) Response and non-response bias pertaining to borrower interview questionnaires.
- (v) Focus will be on conducting interviews with credit scorecard practitioners who primarily operate in highly technical and statistically rich environments. As an unintended consequence, limited insight in terms of consumer behaviour and preferences may be extracted from these interviews.
- (vi) Research is cross sectional in nature and will be limited to what has happened in the past. Forward implementation is based on historic and time-stamped data inputs. Can only build based on predictive and applied future assumptions.
- (vii) To gain access to the right industry participants and at the right decision making levels.

CHAPTER 5:

5. Results

5.1. Introduction

Chapter five presents the sequential process adopted for obtained data which was through qualitative interviews held with credit risk practitioners and the subsequent quantitative data analysis of the identified research population.

Exploratory questionnaires were made use of once the general research propositions were identified through unstructured interviews which were conducted at a business level, and in support of the adopted triangulation research method to obtain supplementary qualitative data on how credit risk practitioners respond to the design and implementation of credit risk models.

The qualitative component of the data gathering process is supplementary to that of the quantitative analysis that forms the basis of the data gathering process and was limited to four interviews that were held with subject matter experts employed by the financial institution from which the quantitative data was obtained from and who are also or were working in the secured and unsecured lending divisions of the retail lender.

Results of the empirical quantitative study are presented through the use of statistical analysis, followed by the discussion of the findings.

An overview is offered of the statistical techniques used in terms of gathering the quantitative data. The total population of data pertaining to the unsecured lending division was made use of and as such constituting statistical relevance in terms of the size of the population.

The data obtained from the unsecured lending division was further applied to the risk scorecard as used by the secured lending division in explaining the research propositions as per chapter 3.

5.2. Source of extracting quantitative data

The primary data was obtained from the unsecured lending division in the Bank and consisted of the following attributes:

- (i) Client demographic data was extracted from the front-end submission system where all new requests for loan funding facilities are administered and captured prior to risk assessment.
- (ii) Supplementary data was obtained from the Bank’s mainframe which was for the purposes of existing client data for those clients that already had a banking relationship with the lender and for which the data constructs further supplemented client behavioural patterns.
- (iii) Credit risk bureau data was obtained for the total population of submitted unsecured loan facility requests as per item (i) above.
- (iv) Unsecured lending practices and analytical risk scores make exclusive use of only one bureau data source in South Africa.

5.3. Statistical data analysis

For the preliminary data analysis unsecured lending client data was extract and compiled as follows:

- (i) Data was obtained for the period 01 July 2007 to 31 July 2012.
- (ii) Total population of unsecured lending clients applying for a unsecured loan facility was;

N 4,895,417 unique records;

and depicted in Table 3 is the breakdown as to how the borrower population is segmented in terms of disbursed loan facilities (32 per cent) and not disbursed loan facilities (67 per cent) of total N.

Loan applications that are indexed as “not disbursed” does not constitute a declined loan facility and although the loan facility was not taken up by the borrower, the facility might still have been approved.

Disbursed Flag	N = 4,895,417	%
Disbursed	1,628,052	32%
Not Disbursed	3,267,365	67%

Table 3: Population size of data

Scorecard Decisions	N = 4,895,417	%
Scorecard Accept	4,005,500	82%
Scorecard Decline	889,917	18%

Table 4: Population Data applied to Secured Lending Risk Scorecard

Disbursed Flag	N = 1,628,052	%
Indeterminate	50,299	3%
Default	157,628	10%
Good	1,420,125	87%

Table 5: Disbursed Unsecured Lending Loans Default Levels

- (iii) The “not disbursed” population segment of 3,267,365 records or 67 per cent of total as per Table 3 were further analysed in terms of the propensity to have obtained a secured loan facility once applied to the initial secured lending risk assessment scorecard without applying affordability assessments. The results of applying the total population to the Secured Risk Assessment scorecard is depicted in Table 4 and based on the initial risk assessment 82 per cent of the total population of unsecured loan applicants would have been accepted from a behavioural risk perspective.
- (iv) The distribution of unsecured loan performance is as per Table 5 and confirms that 87 per cent of the population of disbursed loans are currently performing in accordance with agreed payment and loan obligations.
- (v) The population performance outcome in terms of credit disbursements for unsecured loan facilities is segmented as per Table 6 and the *indeterminate* data segment constitute client performance data that has been disbursed in terms of loan facilities and where the term on book is not sufficient to determine the performance outcome of the risk which is only 1 per cent of disbursements and for data or borrower records that were less than twelve months on book.
- (v) The disbursed data was then further analysed in terms of records classified as being in default with missed payment in excess of >3 consecutive months and was 3 per cent of the total population and 10 per cent of the disbursed population only, whilst borrower records that reflected as “good” payment behaviour and has no

default status is 29 per cent of the total population and 87 per cent of the disbursed population.

Note: A facility is classified as a default when the arrears status on the facility is equal to or in excess of three months of missed instalments.

Disbursed Flag	Default flag	N	%	%
		4,895,417	Total	Disbursed
Disbursed	(Indeterminate)	50,299	1%	3%
	Default	157,628	3%	10%
	Good	1,420,125	29%	87%
Not Disbursed	TOTAL	1,628,052		
		3,267,365	67%	

Table 6: Population performance outcome

5.4. Discussion of data obtained pertaining to the first research proposition

The first research proposition is to establish and confirm through empirical data analysis of the retail credit lenders unsecured borrower data base, that risk assessment scorecards in the secured lending environment contains attributed in terms of applied variables that has an exclusionary effect and eliminates certain segments of the population from accessing credit due to;

- (i) The understanding of the uniqueness and unique performance behaviour in terms of the affordable borrower segment is not adequately incorporated into borrower risk assessment scorecards and risk analytics (*Thin Payment Profile and RTI – Repayment to Income characteristics*);
- (ii) Historical, stereotyping assumptions made against borrower group;
- (iii) Applied risk assessment variables that are applied to credit risk analytics and based on borrower behaviour and affordability parameters has an exclusionary impact due to being benchmarked against the more affluent traditional middle market.

5.4.1. Thin payment profile risk variable

The first concept introduced is that of the term known as “thin payment profile” and entails that a borrower applying for credit does not have sufficient contractual data on debt and retail borrowings contained at the bureau that provides for the data to be used in the risk assessment process as the process and discipline of credit risk analytics and credit scoring is highly dependent upon data in the decision making process (Anderson, 2007).

This occurrence can be attributable to multiple factors such as new entrants to the credit market, students that has just graduated and started with their first formal employment and enquiring about credit borrowings for their lifestyle start-up, or borrowers that has made use of less formal means to access credit such as communal stockvels, or family borrowing that constitutes to be less formal means of obtaining access to credit although much more risky and expensive to the borrower since these loans are used more to settle short term living debt with no real asset accumulation of wealth creation taking place.

The assessment of the result of the data presented in support of that of the first research proposition was obtained from extracting data based on the FSC (financial services sector) income target definition which is for borrower earning less than R16 486 per month in gross income per applicant. Table 7 refers to the income bracket changes for this segment.

Year	Maximum Gross Household Income
2006	12 210
2007	12 770
2008	13 600
2009	15 142
2010	15 498
2011	15 738
2012	16 486

Table 7: Financial Services Charter (FSC) Income Brackets

Table 9 and Table 10: Thin payment profile - Applied risk assessment decision depicts the implied impact of the exclusionary variable pertaining to the term known as “thin payment” profiles or also classified in the theory as borrowers who does not have adequate or sufficient credit borrowing(s) data which can constitute that this borrower segment may be excluded from obtaining credit as the secured lending scorecards has

exclusionary variables that eliminate or rejects borrower records that is categorised as having a “thin-payment” profile or no prior consumer credit.

Unsecured Data Population	N = 4,895,417
AFH (<i>Based on income criteria</i>)	3,978,899
AFH % of Total Population	81,3%
Thin Payment Profile	311,091
Thin Payment Profile % of AFH	7,8%

Table 8: Unsecured Lending Population qualifying for AFH – “thin payment” profile

From the total population 81,3 per cent of borrowers were classified as borrowers that fall within the affordable housing income segment. Of the population sub-group that is deemed to qualify for applying under the affordable housing segment borrower market, 7,8 per cent or 311,091 potential borrowers had a thin payment profile.

This segment classified as “thin payment profile” risk clients were further analysed in terms of loan performance. Of the initial 311,091 loan facilities that were identified as FSC clients with a thin payment profile, 34 per cent or 106,145 loans were disbursed and 78 per cent of disbursed loans had a favourable loan performance indicator (*no defaults or arrear payments*) as depicted per Table 9.

	Default	Good	No Outcome	Total	%
Disbursed	19,208	83,181	3,756	106,145	34%
Not Disbursed			204,946	204,946	66%
				311,091	

	Default	Good	No Outcome	Total	%
Disbursed	19,208	83,181	3,756	106,145	34%
	18%	78%	4%		

Table 9: “Thin Payment Profile” Loan distribution performance

Table 10 highlights the risk assessment outcome for borrower data with “thin payment profiles” that were applied to the secured lending risk scorecard in order to ascertain

the initial reject/accept ratio which is based on risk categorisation principles as used in the secured lending risk modelling analytics.

Risk categorisation principles is dependent on the obtained or derived risk score that is derived from the third party bureau data, internal to lender data and the customer to obtained a derived risk score, which when applied will calculate and make an appropriate risk accepted offer or reject to the borrowers request (Anderson, 2007).

Risk mitigation factors are pre-constructed decision models which entail that although should a borrower apply for a loan to value of 100 per cent, due to the borrowers implied risk assessment and risk score as obtained from the risk assessment model outcome a lesser offer can be extended to that borrower. Offers below 50 per cent of the initial loan to value request unless specifically requested is an automatic reject decision outcome and is automated for applied risk assessment scorecard/model functioning.

The data as per Table 10: Thin payment profile - Applied risk assessment decision reflects an initial risk acceptance rate of 98,52 per cent for borrowers that is classified with a “thin payment profile” and is based on risk assessment drivers that was then overlaid with the applied income data as obtained from the unsecured borrower dataset.

The issue with the identified variable is that on 5,32 per cent of the population classified as having a “thin payment” profile will obtain a possible risk assessment outcome of 100 per cent loan to value. The bigger concern is that borrowers with a “thin payment” profile indicator is excluded from passing though the risk assessment scorecard and is pre-excluded from the start of the borrower application process.

Applicant decline decisions are to be appealed through the risk acceptance process for manual review from a risk assessment practitioner which in the majority of cases is not done and as such the opportunity to do business is lost. The potential still remains that even at a 100 per cent loan to value requirement in excess of 16 thousand potential borrowers are excluded from the credit risk assessment process.

Risk models are constructed to reject records where there is insufficient borrower behavioural data to construct a risk variable and generally defaulted the risk indicator to that of a poor behavioural score thus eliminating the borrower from the credit application process.

The minimum loan scenario applied through the risk scorecard analytics was modelled at extending a minimum loan to value of 50 per cent (*Purchase price R100K * LTV 50% = R50K*) and a maximum of a 100 per cent loan to value of the potential asset to be purchased. The initial modelled repayment obligation was for an average loan of R250,000, prime lending rate of 9 per cent, and a loan term of 240 months with a monthly repayment of R2,249,32 +R50 administration fee. The secured lending risk analytics applied LTV intervals of 5 per cent increments for modelling purposes. The cut-off of the minimum loan extension was at a 50 per cent ratio to purchase price.

Table 11: Disbursement - Secured Lending Risk and Affordability Assessment depicts the breakdown of loan affordability and credit extensions modelled on the outcome of the current secured risk assessment scorecard and confirms that 72 874 borrowers or 69 per cent of disbursed unsecured loan facilities would qualify and be approved for secured lending facilities in excess of 80 per cent of the purchase price. The 100 per cent loan to value approval volume would be 9 056 which was previously excluded and dependant on loan amount requirements this is still a conservative review based on the median loan amount requested.

Loan To Value (LTV)	65%	65%	70%	75%	80%	85%	90%	100%	Total
Approval Volume - Risk		128	117,366	5	9,188	111,577	51,668	16,549	306,481
Approval % - Risk		0.04%	37.73%	0.00%	2.95%	35.87%	16.61%	5.32%	98.52%
Decline Volume - Risk	4,610								4,610
Decline % - Risk	100%								

Table 10: Thin payment profile - Applied risk assessment decision

LTV	50%	65%	70%	75%	80%	85%	90%	100%	Total
	1 051	39	32 179	2	2 608	42 867	18 343	9 056	106 145
	0.99%	0.04%	30.32%	0.00%	2.46%	40.39%	17.28%	8.53%	

Table 11: Disbursement - Secured Lending Risk and Affordability Assessment

5.4.1.1. Response from risk practitioner interviews

In conducting the qualitative one on one interviews, one of the interviewees referred to by the pseudonym name of “Alex” a senior risk practitioner and the head of department stated in response to the question on the adopted risk propensity of South Africa’s retail lending banks to be;

“...I think that lending financial institutions are different from institution to institution in terms of their risk, level of risk they want to take on but I also think that they are a little bit too risk averse taking into account a 100 year event that took place in terms of the world economy. And so based on that, they then make the assumptions going forward that they shouldn't, they should have a lot more collateral deposits and things like that”

In response to the same question however directed at the affordable borrower segment, the senior risk assessment practitioner referred to by pseudonym as “Sheila” responsible for constructing and modelling credit risk analytics stated that;

“So...affordable housing is higher risk than the rest of the population. We've seen that clearly over time. So we...so maybe we are being too harsh on them but it is because of past experience.... But I do believe there are pockets of business in the affordable housing space those maybe we can open up on and give more loans too. So I think there is better risk affordable housing out there and we need to do something to...”

This finding is supportive of research proposition one in that implied risk assessment variables does seem to have an exclusionary effect on borrower's credit application outcomes as thin payment profile bureau data and risk scored does impact on the acceptance of loan referrals and is part of an automated risk assessment process.

From the interviews it was established that for this particular credit borrower, there is only one risk model with the purpose of servicing both that of traditional markets and the affordable borrower segment. The need for a new and unique risk model has been identified however adoption rates are slow in terms of the associated economics and lack of data for building such risk models. The irony though is that due to imposed exclusionary risk assessment variables there will not be sufficient data coming through the risk models. This was evident from the large fall-offs experienced in the unsecured data once applied to the secured risk scorecard.

5.4.2. Repayment to income (RTI) risk variable

The second exclusionary factor contained in credit risk scoring is that of repayment to income (RTI) affordability assessment and capped ratio requirement (Anderson, 2007, p. 279).

Repayment to income criteria is inherently contained in most credit analytics and is a set ratio in terms of the maximum capped ratio of income that is allowed or allocated as the portion to be re-paid on a specific debt instrument.

For affordable housing borrowing and credit risk analytics, the industry specific ratio for RTI ratios are set at proposed cut-off levels of 30 per cent of income. Such cut-off ratios are lender specific and constitute to be the generally accepted risk cut-off rules for assessing secured lending borrower risk.

As a result of the RTI benchmark set at a pre-determined level, certain borrowers are excluded from securing credit facilities. The benchmark is an automatic reject decision and if breached an automatic loan reject will be issued in response to the borrower's loan request.

RTI ratios are only indicative of the level of debt a borrower has to pay in relation to that of income and that these ratios do not accurately reflect a consumer's financial position and ability to repay debt. Furthermore, by not allowing for credit facilities to be extended above certain thresholds borrowers are unfairly penalised and especially those who are genuinely able to repay, which again confirms the statement that what does matter is a proper assessment of affordability at an individual level.

- Loan to value (LTV); ratios are still a relatively consistent predictor of default however it must be measured in terms of the market it is applied to.
- Loan to income (LTI) ratios are not a strong or consistent predictor of defaults and is dependent on multiple factors such as borrower's life stage etc.

Table 12 depicts the analysis of the data extraction obtained from the main population and was modelled on the disbursed unsecured loan facilities. Of the total population 1,359,051 loans were disbursed from the unsecured loan facility population.

The analysis thus used the disbursed loan facilities as the primary analytics premises from which the exclusionary factor of an applied RTI ratio is illustrated and Figure 5 depicts the initial risk acceptance of borrowers and the fall-off once the applied RTI ratio of 30 per cent is applied to income which constitute to be a fixed risk assessment and credit scoring variable in risk assessment models.

The RTI ratio is thus biased towards loan reject decisions as per Table 13 from which it is further evident that since most of the risk is accepted through the initial secured lending risk score which will include 87,6 per cent of unsecured disbursed loan facilities that will pass the secured behavioural risk assessment, once the RTI rule is applied only approximately 31 per cent of borrowers will be able to access credit from the initial

87,6 per cent that passed the secured lending behavioural risk scorecard due to the exclusionary factors imposed by RTI credit risk assessments.

Table 12 depicts the breakdown of the total unsecured lending portfolio, segmented down to the affordable housing segment (AFH), and the further down to the level of disbursed unsecured loans which equate to 1,359,051 unique records. The data was then further analysed to assess the impact of the applied RTI cut off ratio at 30 per cent of borrower income.

Unsecured Data Population	N = 4,895,417
AFH (Based on income criteria)	3,978,899
AFH % of Total Population	81,3%
Disbursed Unsecured Loans (AFH)	1,359,051
Thin Payment Profile % of AFH	34%

Table 12: Disbursed Unsecured Loans (AFH) Segment

RTI exclusions does not solve the requirements of borrowers in their start-up phase or unique circumstances such as the affordable housing market where the only secured debt the borrower might have or incur, will be that of an asset such as the borrowers primary residence should the borrower be in a position to secure such debt.

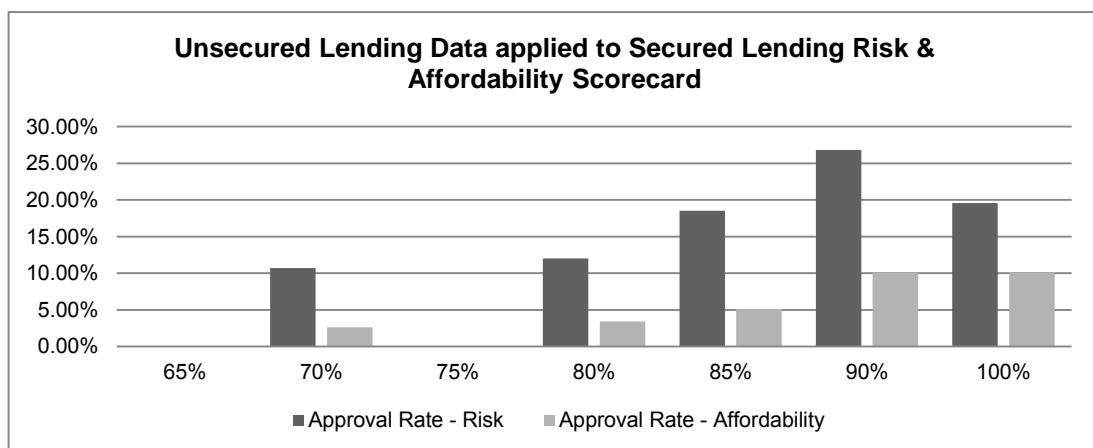


Figure 5: Distribution of loan to value of risk and affordability assessment

LTV	65%	70%	75%	80%	85%	90%	100%	Total
Approval Volume - Risk	181	145,309	154	163,259	251,571	364,360	266,126	1,190,960
Approval % - Risk	0.01%	10.69%	0.01%	12.01%	18.51%	26.81%	19.58%	87.63%
Approval Volume - Affordability (RTI)	47	35,638	56	46,285	69,463	137,194	136,966	425,649
Approval % - Affordability	0.00%	2.62%	0.00%	3.41%	5.11%	10.09%	10.08%	31.32%
Net Loss of opportunity	-134	-109,671	-98	-116,974	-182,108	-227,166	-129,160	-765,311

Table 13: Risk and RTI credit analytics outcome

(Loan R250K, 240 Month Term, Rate = Prime)

The predictive outcome as per Table 14 and the loss of opportunity due to the maximum RTI risk assessment exclusion factor at a proposed lending level of 100 per cent loan to value is depicted as per Table 13, and Table 14 depicts the scenario analysis at different loan sizes and loan term which is modelled on the median affordable housing prices in South Africa. A further exclusionary factor that will be discussed under the third research proposition is that of the term of the loan which has direct impact in terms of the affordability of the loan facility.

Loan Scenario = R250k over 240 term	-129160
Loan Scenario = R250k over 300 term	-115914
Loan Scenario = R310k over 240 term	-175299
Loan Scenario = R310k over 300 term	-129 160

Table 14: Loan Size scenario and opportunity lost

5.4.2.1. Response from risk practitioner interviews

In posing the question; “Does the current applied assumption in terms of risk assessment criteria cater for the unique needs of the affordable borrow segment per se?”, Sheila responded in saying;

“No, I don’t believe so because I think they are a population on their own, if I can call it that way, because they’ve got different challenges to the rest of the market. Um,

obviously they've got lower income so I think the RTI rules on some of their cases must be more strict than for your normal population because obviously cost will be the same for them if they buy groceries or whatever, will be exactly the same as for your other market. The thing is, I don't think a lot of them are very credit active. ...We've seen that...the way they conduct credit is very different to the rest of the population. I think they are... I think we need something specific to them and I think at the moment we are not doing that." I think if you get to a point if you go below 10 000 you must be more strict on your RTI rules so I think they will be able to afford less because groceries will be more expensive to them than any of the other populations and then between 10 and 16 thousand you'll have to do a proper analysis on that to see is the normal 30 per cent RTI acceptable or do you need to do something different for that group."

Repayment to income or RTI ratio variables are powerful risk assessment components and incorrect use in a market segment that is not clearly understood can have severe exclusionary implications for the broader population. In general all interview respondents elaborated on the exclusionary effects of RTI and the need for such measurements to be reviewed.

5.5. Discussion of data obtained pertaining to the second research proposition

For the second research proposition the data is to collaborate the proposition in that credit risk analytics contains exclusionary practices that may occur due to the over utilisation/reliance on automated credit scoring models. Traditional borrower identity and risk assessment model constructs does not provide grounds for assessing the nature of borrower performance that is deemed to be less favourable due to perceived adverse risk performances as a result of the exclusionary risk assessment practices and the over utilisation/reliance on automated credit scoring models

- (i) Imposing system analytics above societal, and individual psychographic attributes for making assessment decisions;
- (ii) Historical, negative perceptions associated with borrower segment

5.5.1. Slow payer status

The data that was obtained for the second research proposition was extracted in order to assess the impact of a borrower classified as a "slow payer" by the bureau.

Client profiles classified as “slow payers” although paying late are still paying their monthly instalments and technically does not go into arrears to fall within the criteria for collection purposes to be initiated.

The data that was obtained (Table 15) in support of the discussion pertaining to the second research proposition reflects records for the sub population of borrowers that had a missed payment of up to a maximum of two consecutive months and was only for borrowers with missed payments which were then rectified to bring the account up to date in the unsecured lending loan segment.

This is representative of 11,47 per cent of the total unsecured lending population that is classified as affordable market home loan applicants and 33,95 per cent of the disbursed unsecured loan population.

Row Labels	N	
Scorecard Accept	420,061	92.02%
Scorecard Decline	36,437	7.98%
Grand Total	456,498	100%

Table 15: Secured lending risk analysis for 'slow payer' borrowers

Slow payer N = 456,498

From the “slow payer” data, the initial risk score as obtained from the secured lending risk score, 92 per cent of the payment missed population of <2 months, would pass through the initial risk scorecard assessment.

The data was then further overlaid to incorporate the RTI component of the risk scorecard at 30 per cent of income and applied the income actuals for the borrowers falling within this segment to ascertain if there is an opportunity for possible future lending opportunities.

5.5.1.1. Response from risk practitioner interviews

From the data contained in Table 16: “Slow payer” secured lending risk assessment outcome, it can be established that depending on which option the borrower exercise, should risk analytics and risk models be able to review the “late payment” variable and apply risk analytics to such variable, there would be an opportunity to extend credit to this segment.

In addressing the question with the senior risk practitioner Alex, on what new risk assessment assumptions should maybe be incorporated or considered to be incorporated into risk models that look at specifically the affordable housing sector or affordable sector?, Alex responded in that;

“I don’t think there’s anything new. I think what ... we got to do is revisit what we’ve got, and try and, adopt a slightly lenient approach in certain things, like for example, in this market you might find that people have a poor record of non-payment on specific things which is not necessarily, gives you a poor result at the end, it just might be a lack of understanding. They might miss 3 months, pay a lump sum and then miss...what they don’t understand is the implication of that in terms of what the NCA is saying. So I think there’s a lot of..., I think there’s a lot of education that needs to take place. You need to somehow build a model that can incorporate that and try and sift out the bad from the good in terms of that.”

This is of particular importance since the affordable market does have a trend for paying late which is circumstantial due to the nature of the labour these borrowers provide, access to banking facilities and the fact that there is still a substantial portion of the population that is not primary banked through the appropriate banking medium.

Depending on the loan amount and term opted for, Table 16 depicts the analysis based on the median affordable loan amounts applied for and loan term options. The opportunity of further lending based on the population of date the analysis is being applied to has an opportunity of between ninety thousand and hundred and seventy thousand additional loan opportunities.

R250k, 240 term, prime rate			R250k, 300 term, prime rate		
Row Labels	N		Row Labels	N	
Scorecard Accept	149,957	32.85%	Scorecard Accept	169,764	37.19%
Scorecard Decline	306,541		Scorecard Decline	286,734	
Grand Total	456,498		Grand Total	456,498	

R310k, 240 term, prime rate			R310k, 300 term, prime rate		
Row Labels	N		Row Labels	N	
Scorecard Accept	90,911	19.91%	Scorecard Accept	110,347	24.17%
Scorecard Decline	365,587		Scorecard Decline	346,151	
Grand Total	456,498		Grand Total	456,498	

Table 16: “Slow payer” secured lending risk assessment outcome

5.6. Discussion of data obtained pertaining to the third research proposition

The third research proposition is specific in that the propositions objective is to review alternative risk assessment variables. Issue is taken with the omission of alternative scorecard characteristics for inclusion and integration in automated credit scoring methodologies, specifically designed to cater for the uniqueness of the affordable borrower segment:

- (i) Reliance for scorecard development is purely based on demographics and discipline is fairly dated;
- (ii) The uniqueness of the borrower is absorbed in the review of traditional markets and thus has no individual developmental characteristics in scorecard development.
- (iii) Vintage loan performance analysis for unsecured loan extensions is a predictive variable that can be applied to secured credit risk analytics and risk assessment models.
- (iv) The uniqueness of the borrower is absorbed in the review of traditional markets and thus has no individual developmental characteristics in scorecard development.

In response to the applied methodology of existing risk scorecards, one of the senior risk practitioners stated that the imperative is to find alternative methods for risk variables and assessment purposes to assist the affordable borrower segment. It is further evident that current risk variables are not sufficient to meet the current requirements of the affordable borrower as stated by Dean (pseudonym) in response to adopted methodologies;

“...the current score card and they can't basically prove income. So it's very difficult. Half of this thing is cash so it's very difficult from a credit point of view. Secondly, of your lower income groups, the, I say average their credit ratio there, they don't have proper credit background, they don't have proper education if I can call it that way, exposure to credit, and um, therefore the first time they get credit it's very difficult, especially for the bigger, um, loan, it's very high risk for the banks.; ... definitely, I do believe there's a lot of people which we as banks we've got a responsibility to help the government to house the people, so therefore yes I believe there's a market out there

and if it's priced right and the risk determined correct there's a big market out there for all the banks to get on board."

A further response from Dirk (pseudonym) further elaborated and stated that; "... *it's difficult to say so, obviously like I said, we would have to try and understand that market better. But the one aspect for instance, perhaps not so much on the score card side but definitely when we do the assessments, would be for instance, to cater for clients that does not necessarily receive a payslip at the end of the month, perhaps are self-employed, they run a small business somewhere in the rural area, work for themselves, generate their own money. So I don't know if that, many of the clients won't have a payslip at the end of the month or the week or the day, so that's the one aspect in terms of the assessment policies that comes to mind immediately. Even then that, things like job stability, income stability, those are all issues which under the sort of normal assessment policies are the issues for concern or that people would see as slightly more negative.*"

5.6.1. Default rates

Default rates has two functions in that there is the possibility of default that is built into the initial risk assessment process and then secondly there is the actual default rate which is referred to as the vintage and is the actual reflection of performance pertaining to the actual loan performance,

It is retrospective and can be used for predictability testing and calibration of risk scores based on population risk behaviour constructs.

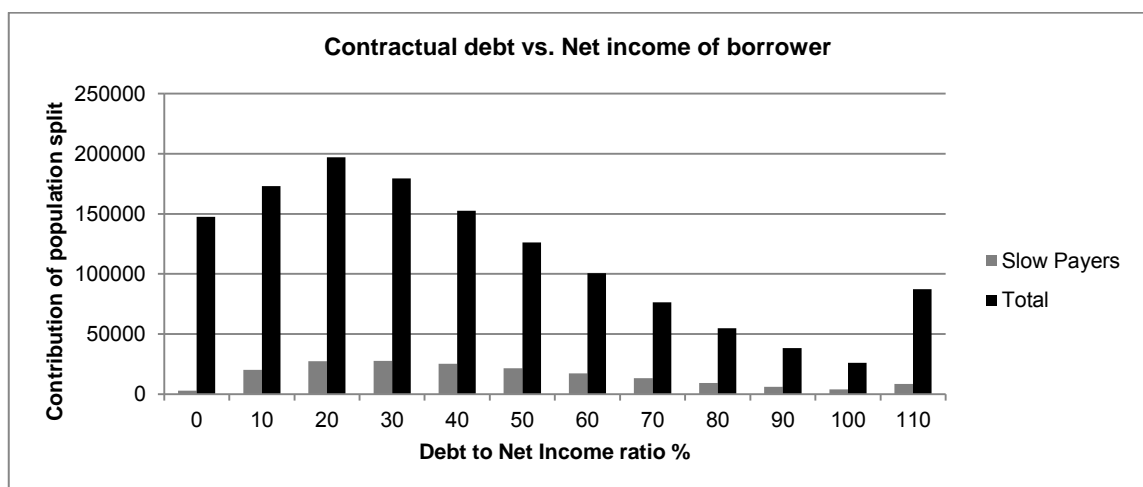


Figure 6: Contractual debt vs. Net Income (Disbursed Loans)

Figure 6 depicts the percentage contribution of total disbursed unsecured loan population and further analyses the contribution of slow payers per debt to income ratio band. The analysis reveals that there is between a 1 per cent and 2 per cent difference per debt bucket difference between slow payers and the total debt bucket population. The proposition thus further set out to disprove the assumption that the higher the borrower debt to income levels the more risky the borrower become in terms of risk analytics.

5.6.2. Debt to income

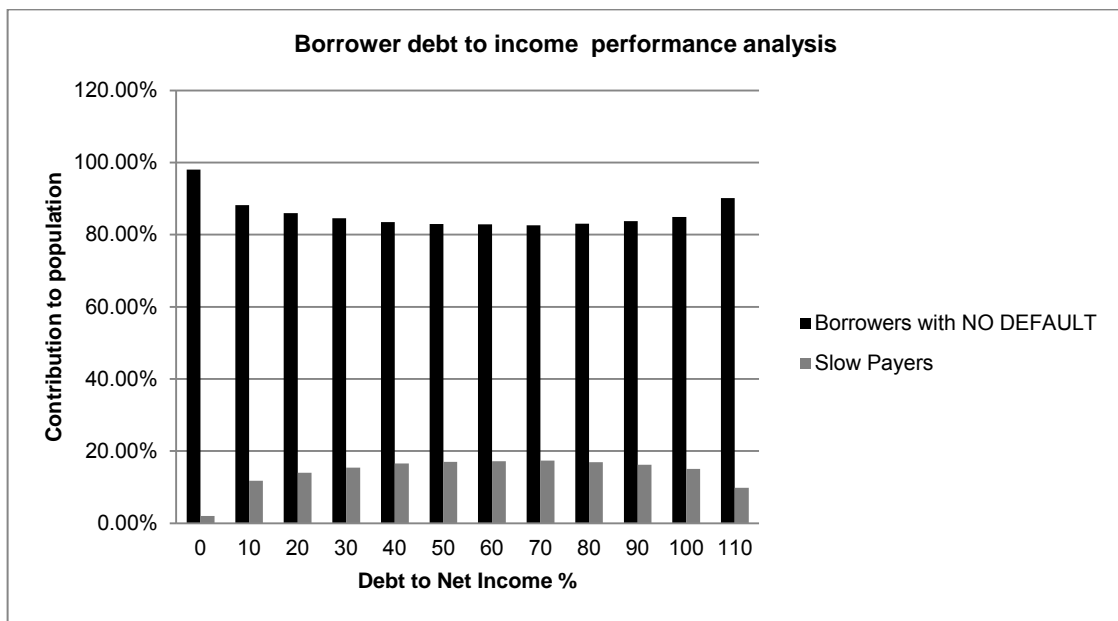


Figure 7: Borrower Categories: Debt to Net Income %

Figure 7 indicates that the contribution of slow payers per debt to income bucket does not increase as the debt levels to income increase. This analysis is supportive of the third research proposition and disproves the implied risk assumption that as borrower debt levels relative to income increases the borrower constitute to become and be more of a higher risk in terms of repayment and debt service capability.

5.6.3. Vintage defaults

In analysing the vintages for disbursed loan facilities based on regular payers and that of slow payers it is evident that the vintages actually cure or perform more favourably for the slow payer borrower segment than that of the regular payer borrower segment.

Figure 8 is representative of the graphical representation of the vintage analysis at the respective debt to income levels.

Once again, the data extracted from the unsecured lending population is supportive of research proposition three in that the implied and applied risk assumptions are not supported by the data and in fact is supportive of allowing for borrowings to take place at this level.

It is thus established that even though certain borrowers are paying later than agreed, payment obligations are still honoured and that the supposable riskier borrower segment actually cures more favourably than that of the traditional non-slow payer borrower segment.

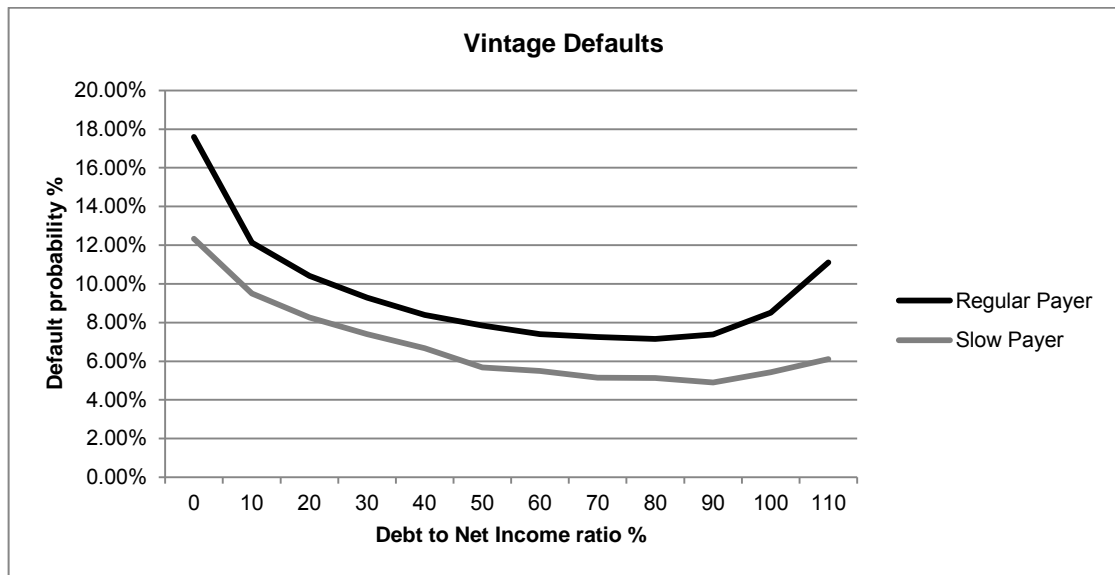


Figure 8: Borrower Categories: Vintage Analysis

5.7. Conclusion

Concluding remarks in terms of what was obtained from both the qualitative and quantitative data it is evident that there are multiple challenges in terms of the adopted and rather dated approach towards credit risk modelling and analytics.

From the qualitative interviews it was established that there is still a lot of work required in terms of actually understanding the affordable consumer and market prior to the development and adoption of a new approach towards assessing credit risk in this segment.

The quantitative data substantiated the exclusionary risk variables that in terms of the questionnaires and quantitative data should rather be developed to be less exclusionary. A detailed conclusion and write-up will be provided for in chapter 7.

CHAPTER 6:

6. Discussion of Results

6.1. Introduction

This chapter sets out to discuss the results in terms of the identified research propositions and the findings obtained from the qualitative and quantitative data in support of emphasising the validity of identified research propositions.

One of the most significant findings was the approach that was adopted towards assessing credit risk in the affordable and secured lending environment. Even though the social and societal imperatives have been identified for the importance of uplifting communities and provide the basics in terms of owning a home to create wealth and general social improvement, the challenges remain in that of managing risk and extracting maximum shareholder value and profitability. As a result, this segment of credit risk analytics has remained relatively neglected in local market context with slow progress being made amid a growing inequality and housing backlog.

Unless the social imperative is elevated at enterprise level, the challenge will remain in that borrowers in the affordable sector will continue to be marginalised and access to credit will continue to be a challenge. The riskiness of borrowers are purely measured, modelled and simulated without proactive engagement to establish and implement new alternative methods for addressing such social constraints (Orlikowski & Iacono, 2001).

Credit scoring and risk analytics is primarily about statistics, probabilities, mathematical algorithms and the data that is used for computation. The objective of credit risk analytics is to predict the loss given default, economic capital and overall improve and deliver organisational and shareholder profit. Credit risk scoring is certainly not about the borrowers or about the uniqueness and unique attributes of borrower identity within the affordable market segment.

6.2. Research findings in response to the research propositions and literature

An analysis of the interview material and observations with that of the collected quantitative data favourably supported and answered all three of the research propositions:

- (i) Identify typical characteristics used for building secured lending credit risk application scorecards (*in the affordable client segment*).
- (ii) Identify the assumptions made about borrowers through credit scoring design (*in the affordable client segment*).
- (iii) Explore what alternative characteristics to consider for consideration in credit risk assessment and application scorecards (*in the affordable client segment*).
- (iv) Identify which credit applications are included/excluded from the credit application scoring process.

The presentation and interpretation of the results is guided and structured around the following sub questions to the research:

- Characteristics typically used in building scorecards;
- The predicted variable;
- Credit scoring and scorecard assumptions;
- Borrower assumptions;
- Groups excluded through credit scoring;
- Consequences of exclusion; and
- Alternative scorecard characteristics.

6.3. Proposition 1: Credit risk assessment variables and exclusion factor(s)

The first research proposition investigated the exclusion factor(s) contained in credit scoring methodologies that excludes certain members of the population group from obtaining credit. The primary objective was to establish how borrowers that are classified as affordable segment borrowers will be assessed through current risk assessment models in the secured lending business by applying unsecured lending data for borrowers that had a loan performance outcome.

The implications of implied variables were assessed in support of understanding loan performance and exclusionary impacts at a secured lending risk scorecard assessment stage.

6.3.1. Three broad sources of credit scoring characteristics, namely bureau, relationship behavioural and application data, were identified.

For credit risk analytic models and credit risk practitioners, the richest source and the most predictive of data were the information obtained from the bureau,

which enabled risk practitioner and risk assessment models to draw literally thousands of potential variable fields.

Examples of commonly used variable fields included:

- The number of credit enquiries which have been made against an applicant;
- the number of credit products applicants have also referred to as active payment profiles;
- the age of credit products (how long they have been open and what is the performance outcome per product); and
- the number of legal judgments or credit infringements which have been made against an applicant (number of adverse decisions or listings).

What was obvious was that little was said about behavioural data since the primary variable was a derived behavioural score and was constructed based on mathematical algorithms that was derived from feedback and credit risk information obtained from the bureau data.

Based on the feedback received from the credit risk analysts and risk assessment practitioners one of the fundamental limitations of this form of data was that it could only be applied to scoring an applicant if the applicant was already a client of the financial institution. Primarily behavioural scores for non-clients were obtained from the bureau and then re-applied and modelled based on the institutions own risk models to derive a risk score for that applicant. Once again there was no uniqueness in terms of the affordable borrower segment and all applicants were treated as being similar with the same level of population “uniqueness” being applied.

Behavioural scores are thus subjective and implied that if the client was deemed to be a primary client of this particular lender, the implied bureau risk score was used to assess the deal which often resulted in borrowers being omitted due to low score outcomes and the lack of having an internal behavioural risk score.

For this reason, separate scorecards are to be built which at the time of conducting this research did not exist nor was any substantial new model development initiatives underway. The requirement and proposal is thus for the

development of a risk assessment model with the first assumption being that no behavioural data exists and the second assumption that such data does exist.

- 6.3.2.** It was argued that there were groups of borrowers who were unintentionally excluded through credit scoring, including new and existing borrowers who are unable to demonstrate regular income streams or who failed to adhere to the exact requirements of meeting monthly debt repayments for other contractual debt instruments and products.

Application information, for most practitioners, offered the opportunity to inflate bureau and behavioural data with information not containing or what was provided by the bureau. It also offered the opportunity to cross-question an applicant and establish what information the applicant is prepared to divulge in terms of the risk model decision outcome and credit risk practitioner trying to establish the applicants' true riskiness.

Examples of commonly used application fields included:

- ID number. from which age is derived;
- Salary;
- number of years with employer;
- education;
- residential status;
- living with parents; and
- home ownership status;
- number of dependants; and
- sector of employment.

From the interviews as well as the review of the literature, general characteristics which would be considered for building scorecards by the community of scorers differed little from that proposed within international literature (Schreiner, 2004, 2002).

What does differ is the abundance of the data that is available for consideration in South Africa. When there is such an abundance of data at our disposal, the question should be what are we doing with all this information and why is there still such a poor response to credit demands and the manner in which we assess risk and credit decisions.

6.3.3. Thin Payment Profile exclusions constitute to be the exclusionary variable factor for not allowing borrowers to gain access to credit. These findings are also in supportive of the statement made under point (6.3.2). Thin payment profiles are the direct result of a borrower not possessing an adequate credit payment profile, thus in essence, the borrower either has no debt or alternatively the borrower has debt through rather informal mediums which are not recorded on the credit bureau listings.

Multiple factors can influence a borrower's risk or bureau profile and include;

- New entrants to the credit market, such as students that have just graduated and started with their first formal employment,
- First credit profile start-ups such as borrowers that have previously made use of less formal means to access credit such as communal stockvels, or family borrowings. These types of borrowings are excluded from the bureau data analysis.

A direct result of “thin payment profile” exclusions is that multiple borrowers are excluded from gaining access to credit.

This is one of the exclusionary variables that pre-screen an applicant and failure to meet the minimum risk score as compiled from the bureau data and available client records which true to the implied name of this variable will yield a reject decision.

The reject decision is derived well in advance of the actual repayment stress test being conducted and as a result there is no reject inference to be applied to further on review and test loan performance of this population subgroup.

Little has been done to understand and review the possible performance of such borrowers as they simply don't form part of the credit approval population. Sample bias was also identified by (Verstraeten & Van den Poel, 2005) and associated risks in excluding certain subgroups from credit risk assessment and loan or borrower performance parameters.

From the analysed data derived from the unsecured lending portfolio, it is evident that different retail credit markets and client segments require different credit risk analysis solutions.

However, Table 11: Disbursement - Secured Lending Risk and Affordability Assessment depicts the breakdown of loan affordability and credit extensions modelled on the outcome of the current secured risk assessment scorecard of the financial services company in question confirms that 72 874 borrowers or 69 per cent of disbursed unsecured loan facilities would be able to qualify, first from a risk perspective (established Unsecured Loan performance history) and secondly be approved for a possible secured loan facility in excess of 80 per cent of the purchase price.

The impact in terms of the affordable market though is that an additional 9,056 borrowers would have been extended credit at a 100 per cent loan to value should the “thin payment profile” not have rejected the loan request upfront.

The exclusionary factors brought about by the implementation of this variable is significant in that it firstly fail to provide for retrospective data analytics that will assist the financial services industry to assess past loan decisions and borrower performance outcomes. Secondly it further leads to the continuous exclusion of able and qualified borrowers that has the means to repay the loan facilities, and that the variable fails to award credit to the actual borrowers who is in need to establishing own wealth and contribute to the start-up of a secondary housing market that as a collective will assist in the improvement of the overall wellbeing and improvement of society as a whole.

- 6.3.4.** Repayment to income threshold levels was argued to have an adverse exclusionary impact on borrowers being allowed access to credit. Imposed cut off levels continue to eliminate certain borrowers from accessing credit even though the performance tracking of borrowers that fall within the unsecured lending ambit disproves the notion that higher repayment to income ratios does affect loan and default performances.

At present the cut-off levels in terms of repayment to income is set at 30 per cent of income.

Figure 8 clearly depicts the curing rate (which refers to the number of borrowers who were able to repay their defaults) of default vintages in the unsecured lending environment and based on the analysis contradicts conventional through in term of higher debt repayment levels that results in poorer cure or repayment performances to be recorded.

Repayment to income ratios are also self-imposed and is based on traditional risk assessment models assessing middle market client with higher levels of income and affordability in terms of disposable income.

For the affordable sector this is however an applied exclusionary ratio for borrowers that fall within this sector and who seldom obtains credit for luxury goods such as motor vehicles. In general the income thresholds are lower (Blue colour workers), and the debt incurred is more in line with short term borrowings. As a result, due to the cost of securing fixed assets such as a primary residence, most of the debt to income ratios is above the imposed threshold of the credit risk assessment model.

Although 87 per cent of the disbursed unsecured lending borrowers will pass the initial risk score of the secured lending scorecard, only 31 per cent will be able to secure a loan facility as a result of the RTI ratio variable as per Table 13. Higher debt repayment levels are synonymous with the affordable borrower segment and as a result borrowers are not incurring additional luxury asset long term debt, it is thus my view that consideration should be given to higher debt to income ratios and cut-off levels for borrowers that qualify as part of the affordable borrower segment.

- (v) Exclusion consequences of being marginalised include not being granted credit or attracting higher interest rate charges for the amounts advanced, and further hold implications for both borrowers and credit risk scorers.

Without exception, the implications of scorecard exclusions and the impact it can have on restrictive consequences for applicants who are on the less favourable side of the decision process are recognised, since, not having access to credit fundamentally means not having access to certain goods and services. Expensive credit, particularly where an applicant is over extended, on the other hand, threatens to engulf borrowers in a cycle of ever increasing debt and erode any possibility of real wealth creation.

Borrowers then have to borrow money to pay off the money they have already borrowed and because they can't afford to pay it back quick enough, borrowers suddenly end up in a position where their borrowings are less than what is being owed and as a result borrowers can't live anymore. The danger lies in the

method and product that is used for extending credit and if credit is extended for incorrect purposes, real wealth creation is eroded rather than created.

6.4. Proposition 2: Traditional borrower identity and risk assessment model constructs

Traditional risk assessment methods and risk model constructs does not provide grounds for assessing the nature of borrower performance that is deemed to be less favourable due to the lack or omission of certain variables that are required to assess traditional risk models.

Further, that such perceived adverse risk performance variables are hard coded into risk assessment models and as a result certain borrower groups are being excluded from accessing credit. This is due to implied risk assessment practices and the over utilisation/reliance on automated credit scoring models become more prevalent.

The literature has made for a compelling case as to why there has been such a rapid demand for risk assessment automation, however the counter argument, based on the analysis that was done with the financial services organisation, is that these models are not capable of adapting to the uniqueness of different borrower behavioural constructs and requirements and that the method to manually assess such risk decisions continue to be inconsistent and subjective.

- (i) By exploring the makeup of scorecards and in particular examining the predicted variable(s) it was shown that scorecard characteristics are correlated with risk and that the inclusion or omission of certain variables will have an accept/reject implication on risk assessment decisions. The predicted variable was demonstrated to be clients who have in some form or another defaulted on the payment, or fallen into arrears in paying, the debt obligation, that repayment to income and debt to income ratios have a direct impact in terms of exclusionary practices on risk scoring and is further exacerbated by the fact that these variables are purely based on traditional and more western middle market borrower segments.
- (ii) Four broad assumptions underpinning credit scoring and scorecards were identified namely, that human credit decisions are faulty and unsustainable thus the rapid industry adoption of automated risk assessment methods, that the past predicts the future, that building scorecards is an imperfect science and only predicts an outcome based on the strength and correctness of the data and

the relevance of the applied variables , and that credit scoring's primary responsibilities are to create and protect corporate profit. Now with understanding that the past predicts the future, surely certain scenarios can be modelled such as that the uniqueness of certain borrower groups can be established, verified and adopted in terms of credit risk analytics and scoring.

The literature and subsequent interviews with the members of the credit risk fraternity confirmed the purpose of credit scoring to measure the inherent risk of credit applicants, to rank and categorise this risk and, finally, to make judgments both on whether to accept or reject a loan, as well as at what cost rate to the borrower to extend credit.

The enterprise therefore aims to include certain of the borrower applicants while on the other hand also exclude others. (Exclusion. in this context is used in a broader sense to refer to both those to whom credit is withheld from either through risk scorecards or the intended biased nature of the risk practitioner responsible for developing the risk scorecard as well as to those who attract higher interest rates, for having been judged higher risk).

While the exclusion of bad risk clients may be argued to be an intentional outcome of the process, there are other groups of clients who are excluded and the objective of this research report is to argue that such exclusions are in certain risk model designs unintentionally.

As has been previously posited, since credit scoring and risk analytics only takes applicants that are already within the credit cycle an entire group of people is excluded who have never received credit before.

As spelt out per above analysis it is evident that the practices of risk practitioners had no appreciation or regard for the discriminatory implications of their models. Yes, there is a level of self-reflection in that more needs to be done to assist this sector in gaining access to credit however the empirical evidence at the predictability of default and economic profit level is for now to prevail.

This leads into the next variable that although not unique in behaviour, does show an alternative approach towards adopting credit analytics and the manner in which borrowers in the affordable segment are being able to cure arrears or

defaults in relation to that of the traditional borrower market. The outcome was that even at high debt to disposable income levels, these borrowers still manage to cure arrears and default statuses and further manage to perform even better in terms of default losses.

- (iii) Slow payer status, is the assigned client segment with the coined name allocation for clients that have missed a payment on the due date as per the contractual loan agreement. The payment obligation is however still met and honoured thus the loan does not default into the actual collection status however in terms of risk analytics and secured lending scorecard characteristics, borrower requests with a “slow payer / missed payment” status is excluded from obtaining credit.

The scorecard building process is primarily based on that of being just logistic regression techniques applied and the measurement of logistic regression outcomes. What the risk model or practitioner is predicting is a one or a zero and that the actual model will only be predicting probability.

Through the process of describing how scorecards are built with specific characteristics, the risk practitioners primarily spoke about using predictive regression techniques. These techniques involve taking a multiple of variables, and passing them through a process which effectively stripped out the least predictive, or least correlated, variables in favour of the better.

In order to perform this process, however, a predicted variable had to be identified in advance. The objective of applied linear regression models is that it tries to tell what the output of the decision or variable would be. The outcome is either “GOOD” or “BAD”. Then it will tell what the factors are that that predicted the “bad”, and what it will do is, the model will fit a linear line through these different variables and will try and minimise the error of this variable.

Credit risk practitioners still work of the good versus bad status for that of the borrower, even though the definition varied across products and sectors of lending (Secured and Unsecured), and that this variable was defined by international compliance to the Basel II Capital Accord.

From the different accounts of the good versus bad definition, as obtained from the literature and interviews with the risk practitioners, the understanding that

the risk scoring practitioners are exploring, per product, clients who have in some form or another either defaulted on their payment obligation, or fallen into arrears in paying. Due to the nature of the variable being binary, you either default or you don't. What credit risk analytics predict is the actual negative outcome of the borrower's behaviour.

Based on the literature and stated data, I conclude that;

- Credit scorecards associate variables with a particular outcome and such variables if applied unjust or without a true understanding of the unintentional exclusionary consequences will eliminate borrowers in error. Such variables can also be imposed to knowingly eliminate certain borrowers from accessing credit even through the understanding of implied future performance outcomes for imposing such a variable is or can be incorrectly assumed.
- In order to understand the future performance outcomes, any variable which might reliably and cost effectively associate with the outcome of the credit model decision be used.
- The predicted variable is a rarefied technical definition which is tantamount to whether a credit applicant is a poor payer or not. For risk scorers' correlation *is* causality despite the fact that the linkage between the actual cause and effect is little known or understood.

6.5. Proposition 3: Alternative scorecard

Alternative scorecard characteristics are to be incorporated in automated credit scoring methodologies, specifically designed to cater for the uniqueness of the affordable borrower segment.

- (i) Assumptions which are made about borrowers by scorecard practitioners were identified namely, that borrowers have essences, that their behaviour mirrors their essence, that they can be measured and represented mathematically, that they derive meaning through their belonging to the aggregate, that their essence is unchanging, that they have an essential riskiness, and lastly that good borrowers are good borrowers and are economically active.

- (ii) Consequences of credit exclusion were entertained. It was argued that being excluded from credit may mean not being able to realise one's aspirational and material dreams. For another, it was maintained that exclusion from inexpensive credit threatened to entrap borrowers in cycles of ever increasing debt that is costly and fails to address social inequality by not creating enduring wealth and in general improve society as a whole, especially that of the poorer communities with the greatest income disparity.

- (iii) Alternative scorecard characteristics were proposed and it was concluded that practitioners were not making progress quick enough when it comes to innovation on the front end and to be able to introduce alternatives for enabling more inclusionary risk assessment models.

What is required is that there is an urgent requirement to have better data capturing in the retail banking sector and affordable housing borrower segment. The database structure, the inter-linking of information between data sources is absolutely critical and must stem across multiple products which include transactional, secured and unsecured lending facilities.

If the aforementioned can be established, it would greatly contribute to the quality and the predictive power of the credit risk scorecards and subsequent borrower analytics.

Based on the literature review, the risk analyst and risk practitioners feedback it can be hypothesised that the credit risk scorecards are relatively moderate in terms of assessment efficiency which can be attributed to problematic data, inconsistent data, the omission of certain data fields and the implied risk assumptions inferred on certain risk variables

Risk models and analysts require clean data to work from and merging information across data sources is rather difficult at the moment.

As a result very few accounts are being used for risk modelling and analytics, however it fails to support the greater debate in terms of allowing more borrowers to have access to credit and the process of credit analytics.

At present the industry is limited to only those borrowers that are captured, thus the opportunity for improvement especially from a data point of view is recognised.

Credit risk scorecards and the science thereof are considered to be deficient by its risk practitioners and analysts and especially for borrowers that fall within the affordable segment.

The fraternity of risk analysts and practitioners have little to add in terms of the above findings and alternative suggestions for improvement.

Characteristics which might be considered in building scorecards of the future included:

- Tax information
- Geo-coding
- Intelligence variables (IQ tests)
- Personality variables

These variables share in common the fact that various forms of measure exist for them already yet in different disguises perhaps

Why not then develop alternative variables and why not experimenting with alternative application and risk variables in scorecards already? The answer is that it is simply not practical, it comes with excessive cost and that there is a pre-requisite first for developing an understanding of such variables which is relatively unsupportive by conventional thinking and risk analytics. Alternative variables are expensive to manage in a sustainable way and constitute to be an overly onerous administrative burden.

- (iv)** Debt repayment to income ratios are identified as predictive indicator and should be explored in more detail. Conventional thinking has stayed true to the implied analogy that high debt to income levels predicts higher levels of loan delinquency. Figure 7 provides for an informative analysis of loan performance. It is evident that in the unsecured lending portfolio, debt to income levels are not necessarily the correct measurement for imposing that high debt to income levels are a predictive variable of loan performance, since the slow payers as a

proportion of the total population under review, remains rather static in relation to that of the performing loans even as the debt to income ratios increase.

It was previously stated that for the affordable borrower sector, the uniqueness of the borrower and the circumstances that these borrowers are faced with, means that debt repayment to income levels will always be in excess of the preferred or imposed 30 per cent cut-off threshold and that in order to fulfil and service the needs of this market segment, conventional thinking will have to be adopted to gain a better understanding of the uniqueness and unique circumstances that this segment is faced with.

- (v) Default Rates and Vintage Analysis of the unsecured lending book for disbursed loans reflected that borrowers classified under the affordable borrower segment showed improved cure rates and further that the borrowers who were classified as slow payers had even further improved levels of vintage cure rates.

This in itself is further evidence that the exclusionary risk variable factors are making incorrect implied assumptions and that the behavioural construct of the affordable segment is different to that of the traditional middle market for borrowers. There is a substantial opportunity to extend credit to additional borrowers through risk mitigation practices and relaxation of certain of the exclusionary variables.

The key in terms of unlocking future potential is in how the various risk variables are incorporated and applied to the risk scorecard. It has been proven that to simply disallow a borrower from gaining access to credit is incorrect and that an in-depth analysis of performance variables does provide for substantiating proof for a detailed review.

What is certain is that the affordable borrower is not the same nor do they behave in a similar manner to that of what is deemed the conventional middle market borrower on which the majority of risk scorecards are based.

CHAPTER 7:

7. Conclusion

7.1. Introduction

The final chapter sets out to present the key findings and highlights of the study.

In particular, by leveraging off the findings entertained thus far this chapter will highlight the value extracted from the research and its aim is to summarise and address the main findings of the research question: namely, how the identity of the affordable borrower is formed through credit scoring practices of a retail bank and further, how such risk models are applied to risk assessment decisions in today's terms.

With all propositions been addressed, the chapter identifies future recommendations and managerial implications. It then concludes with suggestions for further research.

7.2. Summary of Main Findings

Perhaps the most significant finding was that of a pure lack of understanding. Understanding of the market, implied variable inferences and what purpose such variables serve, and that imposed credit risk models may actually have an incorrect and exclusionary effect for borrowers in the affordable segment.

Since these credit risk models are designed for one purpose and one purpose only; *for that is to be predictive of a future risk and probability of borrower default*, practitioners at design and concept level fail to understand the consequences of such models and is not capable of encapsulating the uniqueness of borrower risk characteristics associated with the affordable borrower, for which designed and implemented risk models are purely there to serve.

There was a clear understanding in terms of the social dilemma and the fact that further exclusions will only perpetuate the self-fulfilling prophesy of excluding borrowers due to a lack of understanding of their uniqueness. The riskiness of borrowers are purely measured, modelled and simulated, thus there is no immediate urge or imperative to challenge the conventional wisdom imposed on risk practitioners.

They are highly analytical individuals and only look at their outcome perspective in terms of being black or white, with no objective to develop an in-between grey.

The literature has provided for a rich and well documented overview of how the traditional through process has evolved, yet very little has been done in the context of the affordable segment. This is partly due to associated perceived factors such as those all too familiar stereotypes associated with this market of being very risky, poor payers and generally not a profitable segment to engage in. Figure 9 was developed from the research and supportive literature obtained from Lightfoot (2006), to provide for a schematic overview of the perceived approach to risk analytics from that of a risk scoring perspective and a borrowers perspective.

Economics drive strategy and organisational tactics. For too long, the approach in terms of risk model development and associated risk analytics was focused on the traditional middle market for which those exact same models are now being applied to the affordable market as identified in chapter 5.

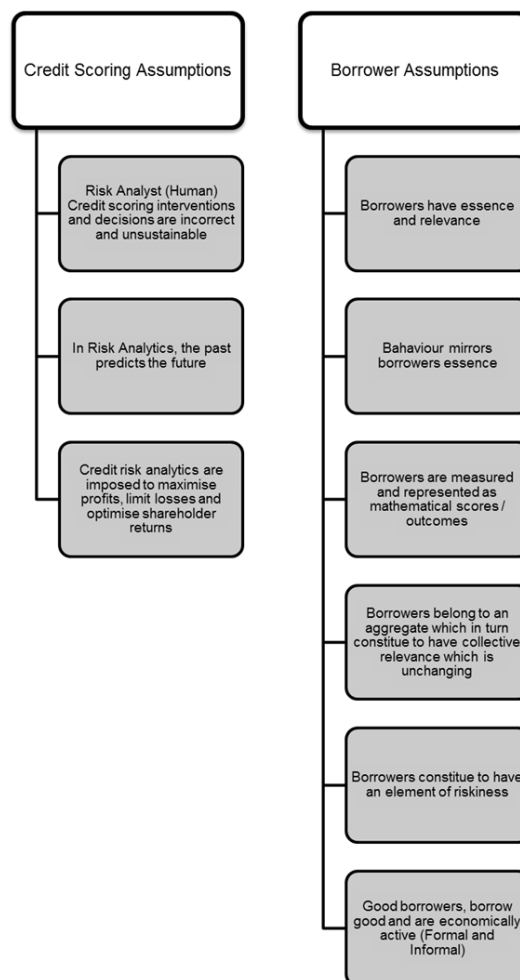


Figure 9: Risk Practitioner vs. Borrower: Managing the paradigm

What was forthcoming from the interviews and substantiated by the quantitative data was the lack of appreciation for which risk practitioners and imposed risk models was withholding or excluding certain borrowers from credit. Since there is no real impact in terms of recourse from a borrower, risk practitioners does not fully appreciate the interconnectivity of their roles and responsibilities in the process of risk analytics and risk model composition.

Although the interviews substantiated the forward thinking notion of certain risk practitioners, unless the impact is made real and there is a direct impact imposed on risk practitioners, they will simply continue to design and develop risk scorecards as per traditional conformist thought process.

Scorecards are simply simple tools, instruments for issuing a response or responses to a request. Therefore, to the question of how the identity of the borrower is formed and incorporated in credit risk scoring practices, the response is rather dull and uneventful. Risk analytics is purely based on computational ability with perceived sense of individuality or uniqueness applied to it. Perhaps it is an unconscious and accidental outcome, keep pushing volume through the system and eventually it will give you an answer you are looking for.

Credit scoring is about statistics, probabilities and mathematical algorithms that are applied to a pre-constructed model for the purpose of data analysis and computation.

Using the past to predict the future is flawed in conventional wisdom, as humans we perpetuate mistakes from our past over and over again however, we seldom learn from it and in credit it is the same. We merely find new alternatives for doing exactly the same with much bigger consequences. When it all goes wrong such as the recent subprime credit lending risk that was cleverly constructed through means of on selling poor risk only to take insurance for such risk through the mere institutions that on sold the risk, which subsequently led to the demise of some of the major and oldest financial institutions in the US and Europe.

In the absence of presenting an equitable alternative risk practitioners continue to look to the past, and in the case of an affordable borrower there is not much of a successful past to commence your work from. Predicting loss given default, economic capital and bank profit is all in a day work for a risk practitioner, however where are the borrower analytics and market factors that should also be incorporated; it is all about the past for future profits. Credit scoring is *not* about borrowers, it is about putting enough through

the system to make the model statistically relevant and then to extract the maximum amount of profits from it through consistent tweaking interventions.

Borrower identity starts in its infancy to be formed through the individual's every interaction with money institutions that are purely for profit, taking of money, borrowing of money, organisations such as retail chain stores for borrowing on credit and retail banks.

Reverting back to Figure 9, the perceived worth and expectation of that of the borrower far exceeds that of the credit risk practitioner and can be said that the risk practitioners outlook is lacking substance in view of what its intended purpose is set out to be.

Credit borrower's transactional experiences are populated for mass computation and centralisation that allows for data to be centralised, sorted, tracked and prioritised based on certain pre-conceived behavioural attributes.

As a result, borrowers are immediately contributing to the data construction warehouse for credit risk analytics through what is called the credit risk bureau. The shortcoming through is that these bureaus can only function provided they are kept abreast with the latest consumer and borrower behavioural data. It is evident from the affordable market that this segment does have prevalence for not making use of traditional risk options. Why blame them, when the exact institutionalised function that is there to protect and purports to responsibly manage the risk positions of such borrowers is slow to adapt to their unique and specific needs and continue to exclude them from accessing credit through their networks.

The emphasis of this research report has been to argue that there are exclusionary variables contained in traditional risk scoring models that are applied to the affordable borrower segment. Imposed risk assumptions underpinning credit scoring and risk scorecards, are flawed in their approach by not understanding and incorporating the uniqueness of the borrower market it is there to assess.

It was established that the individual has *no* meaning in the credit application scoring process and that the individual merely serve the purpose of having relevance for credit application practitioners within groups of other individuals either from the same or different social and risk behavioural constructs which is then further used in terms of excluding and assessing borrower risk behaviours.

Explicit risk behaviour and reinforcement of such behaviours allows for Individuals to be split from themselves and assigned to particular behaviour cohorts, such as regular debt repayment groups, who are recognised in terms of risk participation whilst the non in groups are excluded.

7.3. Future Recommendations

Credit risk scoring is to further develop an understanding of the unique behavioural constructs associated with that of the affordable borrower segment. Throughout the write-up of this report, this narrow focus on behaviour emphasises just how far off from understanding the true market context credit risk practitioners are and that such narrow minded exclusions also comes at the cost of a full appreciation of human complexity.

Borrowers are further circumscribed to risk buckets, and credit applicants are described and discussed as collective cohorts by scoring practitioners. It is time that the humanity be brought back into the risk assessment decision for those autonomous groups which move collectively as a mass.

The National Credit Act and the advent of borrower rights through the promulgation of this bill constitute to be a restrictive burden in terms of what risk practitioners believe to constitute the dis-empowerment of the borrower due to imposed legislative restrictions on credit lending. I believe that research exploring borrower experiences and the imposed impact of the National Credit Act imposed on both borrowers and lenders and what exclusions are brought about this legislative requirement.

The future of credit risk and borrower identity constructs. The advancement of credit application scoring will include using scorecard variables which currently exist beyond the horizon of possibility such as tax information, utility bill information, geo-coding, personality dimensions, and DNA. That the post-sale of the initial residential development be assessed and studied in terms of the outcome on loan delinquencies and contributing factors to better understand cause and effect to enhance and minimise future risk from new developments.

In summary, borrower identity is the systematically fabricated and reinforced process that through an on-going basis through the innocent application of processes, systems and technologies of credit scoring practitioners assess the likelihood of a borrower securing debt. Perhaps the future of application scorecard building will see more

ambitious and more sophisticated variables being used which perhaps will give rise to a self-regulating borrower that is adequately informed, and less overtly disciplinary banks.

7.4. Implications

The implications are severe for both that of the borrower and the risk practitioner. What the research has highlighted is that the current status quo won't be sufficient in addressing the societal needs and widening inequality should current risk assessment models continue to be applied.

Risk practitioners can work with the current variables and even enhance such variables to be more inclusive in terms of risk assessment, however in order to do so, a proper understanding of this risk segment and market should be formulated in order to adapt and develop more efficient ways for assessing and extending credit.

8. References

- Abelson, R. P., Dasgupta, N., Jaihyun, P., & Banaji, M. R. (1998). Perceptions of the collective other. *Personality & Social Psychology Review (Lawrence Erlbaum Associates)*, 2(4), 243-250.
- Akhavain, J., Frame, W. S., & White, L. J. (2005, March). The discussion of financial innovations: An examination of the adoption of small business credit scoring by large banking organisations. *The Journal of Business*, 78(2), 577-596.
- Allen, L., DeLong, G., & Saunders, A. (2004). Issues in the credit risk modelling of retail markets. *Journal of Banking & Finance*, 28, 727-752.
- Altman, E. I., & Saunders, A. (1998). Credit risk measurement: development over the last 20 years. *Journal of Banking & Finance*, 1721-1742.
- AMPS. (2012, June 17). *Eighty20 reports - LSM*. Retrieved from Eighty20: <http://www.eighty20.co.za>
- Anderson, R. (2007). *The credit scoring toolkit*. New York: Oxford University Press Inc.
- Avery, R. B., Brevoort, K. P., & Canner, G. B. (2009). Credit scoring and its effects on the availability and affordability of credit. *The Journal of Consumer Affairs*, 43(3), 516-537.
- Baesens, B., Setiono, R., Mues, C., & Vanthienen, J. (2003, March). Using neural network rule extraction and decision tables for credit risk evaluation. *Management Science*, 49(3), pp. 312-329.
- Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J., & Vanthienen, J. (2003). Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the Operational Research Society*, 54(6), 627-635.
- Banking Association of South Africa. (2005, February 01). *Research into housing supply and functioning markets - Final report*. Retrieved from banking association of South Africa: http://www.banking.org.za/resource_centre/banking_association/publications/publications.aspx
- Barreto, M., & Ellemers, N. (2010). Current Issues in the study of social stigma: Some controversies and unresolved issues. *Journal of Social Issues*, 66(3), 431-445.

- Basel Committee on Banking Supervision. (2000). *Principles for the management of credit risk*. Basel: Risk Management Group of the Basel Committee on Banking Supervision.
- Basel Committee on Banking Supervision. (2006). *Basel II: International convergence of capital measurement and capital standards: A revised framework - comprehensive version*. Basel: Bank of International Settlements.
- Berry, M. J., & Linoff, G. S. (2000). *Mastering Data Mining: The art and science of customer relationship management*. New York: John Wiley & Sons Inc.
- Bertelsmann Stifting's Transformation Index (BTI). (2012). *BTI 2012 - South Africa Country report*. Gutersloh: Bertelsmann Stiftung.
- Boffey, R., & Robson, G. (1995). Bank Credit Management. *Managerial finance*, 21(1), 67-78.
- Bostic, R. W. (2004). *A primer on mortgage lending practices*. Los Angeles, California: USC Lusk Centre for Real Estate, School of Policy, Planning, and Development, Marshall School of Business.
- Brown, R. (2000). Social identity theory: past achievements, current problems and future challenges. *European Journal of Social Psychology*, 30, 745-778.
- Bullock, H. E., Wychne, K. F., & Williams, W. R. (2001). Media images of the poor. *Journal of Social Issues*, 57(2), 229-246.
- Clement, A. (2002). *Digital identity constructions, Unpublished doctoral dissetation*. Toronto, Canada: University of Toronto.
- Correll, J., & Park, B. (2005). A model of the ingroup as a social resource. *Personality and Social Psychology Review*, 9(4), 341-359.
- Cundiff, K. (2004). Closing the loop: How credit scoring drives performance improvements along the financial value chain. *Business Credit*, 106(3), 38-42.
- Dawson, C. (2002). *Practical research methods: A user - friendly guide to mastering research techniques and projects*. Oxford, United Kingdom: How To Books Ltd.
- Deaux, K., Reid, A., Martin, D., & Bikmen, N. (2006). Ideologies of diversity and inequality: Predicting collective action in groups varying in ethnicity and immigrant status. *Political Psychology*, 27(1), 123-146.

- Department: Human Settlements Republic of South Africa. (2009). *Department of human settlements: National housing code 2009*. Retrieved from Department: Human Settlements Republic of South Africa: <http://www.dhs.gov.za/live/>
- Diana, T. (2005). Credit risk analysis and credit scoring - now and in the future. *Business Credit*, 107(3), 12-16.
- Fensterstock, A. (2001, March). The application of neural networks to credit scoring. 103(3), pp. 58-60.
- Fensterstock, A. (2005). Credit scoring and the next step. *Business Credit*, 107(3), 46-49.
- Feshbach, D. (2009, March). The new frontier of mortgage analytics. *Mortgage Banking*, 69(6), 40-46.
- Financial Sector Charter Council. (2012, April 4). *Financial sector charter council*. Retrieved from Financial Sector Charter Council: <http://www.fscharter.co.za>
- Gillham, B. (2005). *Research Interviewing: The range of techniques*. Berkshire, England: Open University Press.
- Hamilton, D. (2007, May). Understanding the complexities of group perception: Broadening the domain. *European Journal of Social Psychology*, 37, 1077-1101.
- Hand, D. J. (2001). Modelling consumer credit risk. *IMA Journal of Management Mathematics*, 12(1), 139-155.
- Hand, D. J., & Henley, W. E. (1997). Statistical classification methods in consumer credit scoring: A Review. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 160(3), 523-541.
- Hawkins, P. (2003). *The Cost, Volume and Allocation of Consumer Credit in South Africa*. Johannesburg: FEASibility.
- Heiner, R. (2006). *Social Problems: An introduction to critical constructionism*. New York: Oxford University Press.
- Hogg, M. A., Terry, D. J., & White, K. M. (1995). A tale of two theories: A critical comparison of identity theory with social identity theory. *Social Psychology Quarterly*, 58(4), 225-269.

- Howcroft, D., & Trauth, E. M. (2004). The choice of critical information systems research. *IFIP International Federation for Information Processing*, 143, 195-211. doi:10.1007/1-4020-8095-6_12
- Iyer, A., Schmader, T., & Lickel, B. (2007, April). Why individuals protest the perceived transgressions of their country: the role of anger, shame and guilt. *Personality & Social Psychology Bulletin*, 33(4), 572-587.
- Jacobson, T., & Roszback, K. (2003). Bank lending policy, credit scoring and value-at-risk. *Journal of Banking and Finance*, 27(4), 615-633.
- Jagric, T., & Jagric, V. (2011, November-December). A comparison of growing cell structures neural networks and linear scoring models in the retail credit environment. *Eastern European Economics*, 49(6), 74-96.
- Johnson-Speck, C. (2005). Abstract of significant cases bearing on the regulation of insurance. *Journal of Insurance Regulation*, 27(4), 81-84.
- Koh, H. C., Tan, W. C., & Goh, C. P. (2006). A two-step method to construct credit scoring models with data mining techniques. *International Journal of Business and Information*, 1(1), 96-118.
- Kothari, C. R. (2004). *Research methodology: Methods and techniques*. Daryaganj, New Delhi: New Age International (P) Limited, Publishers.
- Kozolchyk, B. (2007). Secured lending and its poverty reduction effect. *Texas International Law Journal*, 42(3), 727-749.
- Kumar, R. (2005). *Research methodology: A step-by-step guide for beginners*. California: SAGE Publications Inc.
- Lahsasna, A., Aion, R. N., & Wah, T. Y. (2010, April). Credit scoring models using soft computing methods: A survey. *The International Arab Journal of Information Technology*, 7(2), 115-123.
- Leonard, K. J. (1995). The development of credit scoring quality measures for consumer credit applications. *International Journal of Quality and Reliability Management*, 12(4), 79-85.
- Levin, A., & Davidson, A. (2008). The concept of credit bias in valuation of mbs. *The Journal of Portfolio Management*, 34(3), 39-55.
- Levine, R. (1997). Financial development and economic growth: views and agenda. *Journal of Economic Literature*, 35(2), 688-726.

- Lewis, A. C., & Sherman, S. J. (2010). Perceived entitativity and the black sheep effect: when will we denigrate negative ingroup members. *The Journal of Social Psychology, 150*(2), 211-225.
- Lightstone. (2011). *Lightstone maps South Africa's most affordable housing*. Johannesburg: Lightstone.
- Lucas, P. (2000). Why recoveries are on the rise. *Credit Card Management, October*, 71-72.
- Martin, L. (2011). Good deals for homebuyers, not for the poor: erasing poverty from affordable housing discourse. *Journal of Poverty, 15*, 3-21.
- Merton, R. C. (1992). Finance innovation and economic performance. *Journal of Applied Corporate Finance, 4*(4), 12-22.
- Merton, R. C. (1995). Financial innovation and the management and regulation of financial institutions. *Journal of Banking & Finance, 19*(3-4), 461-81.
- Miller, M. (2003). Research confirms value of credit scoring. *National Underwriter, 107*(42), 30.
- Morrison, J. (2004). Introduction to survival analysis in business. *The Journal of Business Forecasting Methods & Systems, 23*(1), 18-22.
- National Credit Regulator. (2011, July 1). *NCR - Consumer credit report*. Retrieved from National Credit Regulator: <http://www.ncr.org.za>
- National Credit Regulator. (Fourth Quarter, December 2011). *Consumer credit market report*. Johannesburg: National Credit Regulator.
- Olsen, W. K. (2004). Triangulation in social research: qualitative and quantitative methods can really be mixed. In M. Holborn, *Developments in sociology: An annual review*. Ormskirk, UK: Causeway Press.
- Oracle Financial Services. (2008). *Credit risk analytics: A cornerstone for effective risk management*. White Paper, USA.
- Orlikowski, W. J., & Iacono, C. S. (2001). Research commentary: desperately seeking the "IT" in IT research - A call to theorizing the IT artifact. *Information System Research, 12*(2), 121-134.
- Patton, M. Q. (1990). *Qualitative evaluation in research methods* (Vol. 2nd edition). Calif: Sage Publications.

- Quittner, J. (2003). Credit cards: sub-prime's tech dilemma: with delinquencies and charge-offs on the rise, the industry examines the role of automated decisioning. *Bank Technology*, 16(1), 19-23.
- Rimmer, J. (2005). Contemporary changes in credit scoring. *Credit Control*, 26(4), 56-60.
- Ronald, R. (2007). Comparing homeowner societies: Can we construe an East-West model? *Housing Studies*, 22(4), 473-493.
- Ronald, R., & Chiu, R. L. (2010, September). Changing housing policy landscapes in Asia Pacific. *International Journal of Housing Policy*, 10(3), 223-231.
- Ronald, R., & Doling, J. (2007). Shifting East Asian approaches to home ownership and the housing welfare pillar. *International Journal of Housing Policy*, 10(3), 233-254.
- Rust, K. (2012, February 16). *Understanding the ENTIRE residential property market*. Retrieved from Centre for Affordable Housing Finance in Africa: <http://www.housingfinanceafrica.org>
- Salfarina, A. G., Nor Malina, M., & Azrina, H. (2010). Trends, problems and needs of urban housing in Malaysia. *International Journal of Human and Social Sciences*, 5(14), 977-981.
- Saunders, M., & Lewis, P. (2012). *Doing research in business & management, An essential guide to planning your project*. Harlow, England: Pearson Education Limited.
- Schreiner, M. (2004). Scoring arrears at a Microfendr in Bolivia. *Journal of Microfinance*, 6(2), 65-88.
- Sharma, S. (1996). *Applied Multivariate Techniques*. New York: John Wiley Sons.
- Siddiqi, N. (2006). *Credit risk scorecards, developing and implimenting intelligent credit scoring*. Hoboken, New Jersey: John Wiley & Sons, Inc.
- Spencer, J. H. (2007). Health, human security and the peri-urban transition in the Mekong Delta: market reform, governance and new analytic frameworks for research in Southeast Asia. *International Journal of Asia-Pacific Studies*, 3(2), 43-64.
- Spencer, J. H. (2010). an emergent landscape of inequality in southern asia: Cementing socio-spatial inequalities in Vietnam. *Globalizations*, 7(3), 431-443.

- Statistics South Africa. (2005/2006). *Income & expenditure of households*. Pretoria: Statistics South Africa.
- Steyn, L. (2012, June 15). *Business: Mail & Guardian*. Retrieved from Mail & Guardian: <http://www.mg.co.za/article/2012-06-15-banks-focus-on-affordable-home-loans>
- Stryker, S., & Burke, P. J. (2000). The past, present and future of identity theory. *Social Psychology Quarterly*, 63(4), 284-297.
- Stryker, S., & Serpe, R. T. (1982). *Commitment, identity salience and role behaviour: Theory and research example, in personality, roles, and social behaviour*. New York: Springer-Verlang, 199-218: W. Ickes & E.S Knowles (Eds).
- Svirydzienka, N., Sani, F., & Bennett, M. (2010). Group entitativity and its perceptual antecedents in varieties of groups: A developmental perspective. *European Journal of Social Psychology*, 40, 611-624.
- Taifel, H. (1982). Social Psychology of intergroup relations. *Annual Review of Psychology*, 33(1), 1-39.
- Thomas, L. C. (2000). A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers. *International Journal of Forecasting*, 16, 149-172.
- Thomas, L., Edelman, D., & Crook, J. (2002). *Credit scoring and its applications*. Society for Industrial and Applied Mathematics. Philadelphia, PA.: SIAM Publishing.
- Timmons, H. (2002, November). The cracks in credit scoring: As loan defaults rise, worries grow about how creditors pick borrowers. *Business Week*, 25.
- Tsakonas, A., & Dounias, G. (2007, August). Evolving neural-symbolic systems guided by adaptive training schemes: Applications in finance. *Applied Artificial Intelligence*, 21(7), pp. 681-706.
- van Zomeren, M., & Iyer, A. (2009). Introduction to the social and psychological dynamics of collective action. *Journal of Social Issues*, 65(4), 645-660.
- Vanhoomissen, T., & Van Onverwalle, F. (2010). Me or not me as source of ingroup favoritism and outgroup derogation: A connectionist perspective. *Social Cognition*, 28(1), 84-109.
- Verstraeten, G., & Van den Poel, D. (2005, August). The Impact of sample bias on consumer credit scoring performance and profitability. *The Journal of the Operational Research Society*, 56(8), 981-992.

- Vogelgesang, U. (2003). Microfinance in times of crisis: The effects of competition, rising indebtedness, and economic crisis on repayment behaviour. *World Development*, 31(12), 2085-2114.
- Wendel, C., & Harvey, M. (2003). Credit Scoring: Best practices and approaches. *Commercial Lending Review*, 18(3), 4-7.
- West, D. (2000, October). Neural network credit scoring models. 27(11/12), pp. 1131-1152.
- Wu, W.-W. (2011). Improving classification accuracy and causal knowledge for better credit decisions. *International journal of Neural Systems*, 21 (4), 297-309.
- Yobas, M., Crook, J., & Ross, P. (2000). Credit scoring using neural and evolutionary techniques. *IMA Journal of Management Mathematics*, 11(2), 111-125.
- Yzerbyt, V., Corneille, O., & Estrada, C. (2001). The interplay of subjective essentialism and entitativity in the formation of stereotypes. *Personality and Social Psychology Review*, 5(2), 141-155.

9. Appendix A

Standardised – Open-Ended Interview Guide

Introduction and statement to clarify the background and motivation for the study:

Dear

I am *conducting research on the **Construct of Borrower Identity and Credit Scoring: Affordable Housing Segment***, as part of my MBA and I would be extremely grateful if you could answer the following questions:

Thank you in advance for your time and effort.

By way of introduction to the interview, there is a growing urgency to develop sustainable solutions on how to address the growing need in which retail consumer credit is provided to the affordable housing market.

The aim of this conversation is to gain insights into;

- *The understanding if there is unintentional exclusionary factor(s) contained in credit scoring methodologies that inaccurately excludes certain members of the affordable borrower population group from obtaining credit;*
- *The method as to how borrower identity constructs developed and implemented for utilisation/reliance on automated credit scoring models;*
- *Ascertain if there are perhaps alternative scorecard characteristics to be incorporated in automated credit scoring methodologies, specifically designed to cater for the uniqueness of the affordable borrower segment.*

The research is particularly interested in information about the South African market.

Before we are to proceed kindly note the following:

- I hereby confirm that the interview will be recorded for academic purposes.
- As the interviewee kindly provide acceptance and confirmation to be interviewed.
- Your participation is totally voluntary and you can withdraw from the process at any time without penalty.
- There is no cost or benefit attributable to participating in the research process and will purely be for academic data gathering purposes.
- It is confirmed that data collected will be kept confidential and only used for the purpose of the study. (Appendix E – Informed Consent Form to be signed)

Using the information presented in the problem definition and the rationale for the study to be conducted, the background and motivation of the study will be explained.

10. Appendix B

Guide for Interview purpose and process:

Borrower Identity Constructs and Credit Risk Scoring:

1. How did credit risk scoring and come into existence and for what purpose?
2. How did ***borrower identity construct principles*** and ***credit risk-scoring methodologies*** develops or changed since its initial inception?
3. What is the theoretical basis for constructing borrower identities?
4. Is applied risk scoring methodologies based on a specific view of what risk scoring discourse is which is typical of a particular content domain? If so, can you describe it?
5. What is credit risk scoring intended to do in terms of applied business practice?
6. What are the elements that are important in terms of borrower identity constructs and applied credit risk scoring models?
7. What would you say are the strengths and weaknesses of borrower identity construct principles and credit risk-scoring methodologies based on its current format?

11. Appendix C

Schedule for Secured and Unsecured Risk Analyst Interviews:

1. How long have you been a risk analyst in the secured/unsecured risk assessment businesses of your organisation?
2. What do you think is the main purpose of the work you do and the recommendations you make?
3. What elements in risk assessment analytics do you give most prominence to?
4. What elements, if any, do you consider of less importance?
5. In your experience what are the biggest problems in developing borrower identity models that then gets applied to risk assessment models?
6. Are there any special 'tricks' or techniques that you use in your profession?
7. Have you encountered any issues of personal dissatisfaction as to how risk models are being applied? If so, how do you deal with it?
8. Do you think that clients understand the purpose of constructing borrower profiles with the further development of associated risk assessment models?
9. How successful do you think risk models are in its capability to predict correct risk assessment behaviours and to meet client needs/expectations?
10. How does the implied assumption of borrower identity constructs affect risk scoring?
11. If you could improve or change anything about the way in which risk assessments and borrower identity constructs are being developed, what would it be?

12. Appendix D

Interview schedule for Borrower Identity Constructs and Credit Risk Scoring:

In order to get some additional insight into the way as to how secured lending and unsecured lending risk analysis and practitioners view, develop and design borrower identity constructs and risk assessment models, I'm interested in finding out more about what views risk practitioners hold.

CONTEXT, TONE SETTING AND SCREENING QUESTIONS

- Name;
- Organisation segment (Secured/Unsecured Lending);
- Designation;
- Brief background to the work of the Organisation you are responsible for?

INTERVIEW QUESTIONS

Question 1	How much risk is financial institutions willing to take, are they too risk averse?
Question 2	What are the factors encumbering the delivery of housing to the affordable market?
Question 3	Does the current applied assumption in terms of risk assessment criteria cater for the unique needs of the affordable borrower segment?
Question 4	What new risk assessment assumptions should be incorporated into the risk model design (if any)?
Question 5	Should the current risk parameters be reassessed to cater for the affordable segment or even be re-designed?
Question 6	How much relevance is placed on consumer behaviour and identity construct vs. demographics in the risk assessment decision?
Question 7	What are the unique parameters to be assessed in the affordable borrower segment?
Question 8	Does consumer behaviour differ across borrower segments based on societal influencing factors?
Question 9	Any general comments in terms of how to approach risk assessment lending in the affordable borrower segment?

Interview evaluation and closure

1. Have I missed anything out – is there anything else you think I should ask?
2. Are all the questions clear?
3. Does the sequencing of the questions seem logical?
4. Would you like to make any other comments?

THANK YOU VERY MUCH FOR YOUR TIME AND COOPERATION

13. Appendix E

CONSTRUCT OF BORROWER IDENTITY AND CREDIT SCORING: AFFORDABLE HOUSING SEGMENT

Interview schedule:

Interviewee:	Designation
Interviewee 1	Senior Risk Analyst – Secured Lending
Interviewee 2	Senior Risk Analyst – Unsecured Lending
Interviewee 3	Senior Business Analyst – Secured Lending
Interviewee 4	Senior Business Analyst – Unsecured Lending
Interviewee 5	Senior Credit Manager – Secured Lending
Interviewee 6	Senior Credit Manager – Unsecured Lending

