CONFERENCE PAPER – EDINBURGH CONFERENCE 2009

USING EXTERNAL 'PAYMENT INTENT' DATA TO IMPROVE DISCRIMINATION IN CREDIT DECISIONS

OUTLINE

Until recently, the main source of external data used to improve the quality of credit and collections decision-making was credit bureau data. This is a well-established segmentation tool in business to consumer ('B 2 C') field, relying on lenders sharing data on bad debt outcomes. Some lenders also share data on the status of well-conducted accounts. The data are often used in conjunction with credit scoring, and sometimes with other geo-demographic data collected on a survey basis.

These approaches have served the credit industry reasonably well for nearly 50 years. However, recently there has been a paradigm shift to a higher level of consumer debt, undertaken often at historically low rates of interest. Waiting for bad debts to reveal themselves has not always been sufficient to control runaway consumer debt in these circumstances. The situation has been exacerbated by a relaxation in bankruptcy rules, further limiting the consequences of default to the debtor. The introduction of new ways of limiting debtors' exposure, such as Individual Voluntary Arrangements ('IVA') and Debt Relief Orders ('DRO') have made the situation still riskier for the lender.

These trends have led to a search for new and better methods of limiting lenders' credit risk exposure.

A particularly promising approach is to use 'payment intent' data. These data are based on transactional payment information, and so reflect an up-to-date, behavioural picture of who will, and won't, pay their bills. Intent to pay seems to be a psychological trait – people who pay their bills tend to do so, even when there are no adverse consequences from failure to pay. They may not have the ability to pay right now, but they will pay in the end.

This paper explores a method of assessing payment intent, using shared retail payment data, and shows a worked example of the results obtained when a mail order company used the data to predict future credit performance.

CONTENTS

1. SUMMARY

This paper shows that, by using Fraudscreen payment intent coding to supplement an already effective credit scorecard, a large mail order company was able to demonstrate a 2% increase in contribution net of bad debt, at the cost of a 0.5% increase in overall bad rate. For this company the net benefit exceeded £4.5m a year in net profit contribution.

The paper outlines the main features of Fraudscreen payment intent coding, including the source of, and philosophy behind, the data.

It then outlines the test programme, and shares the overall results. The test results are then translated into business results.

Finally, the paper concludes by setting out the main lessons to be drawn from this test, together with some caveats. It suggests a way forward for those who wish to see whether this approach might help their businesses.

2. WHAT IS FRAUDSCREEN PAYMENT INTENT CODING?

Fraudscreen payment intent coding is a coding system based on a large dataset of payment transaction data. The data are drawn mainly from on-approval mail order companies. It is a characteristic of such companies that the transaction amounts are, for the most part, too low to justify a significant level of collections or enforcement action. The companies rely on the honesty of their customers to repay their commitments.

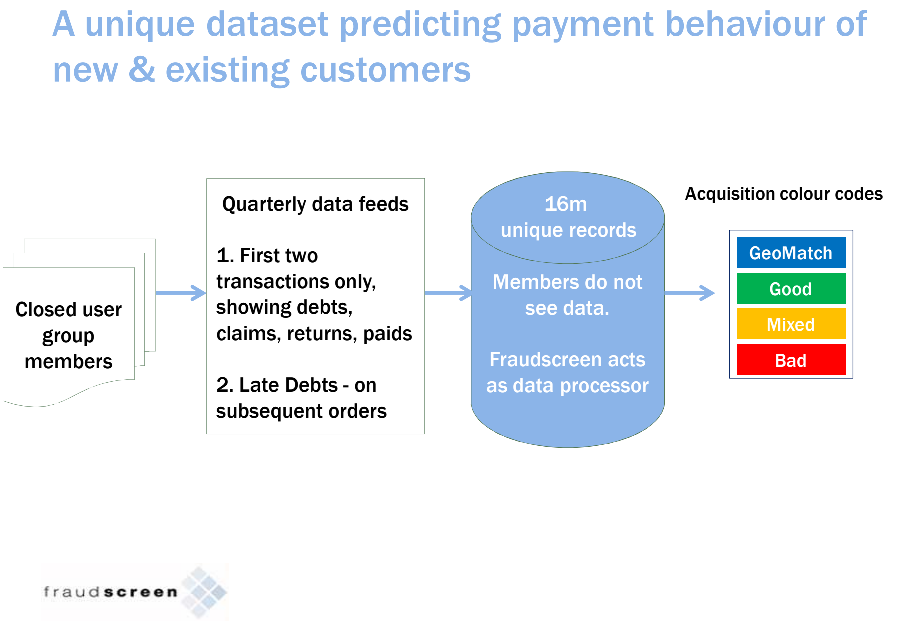
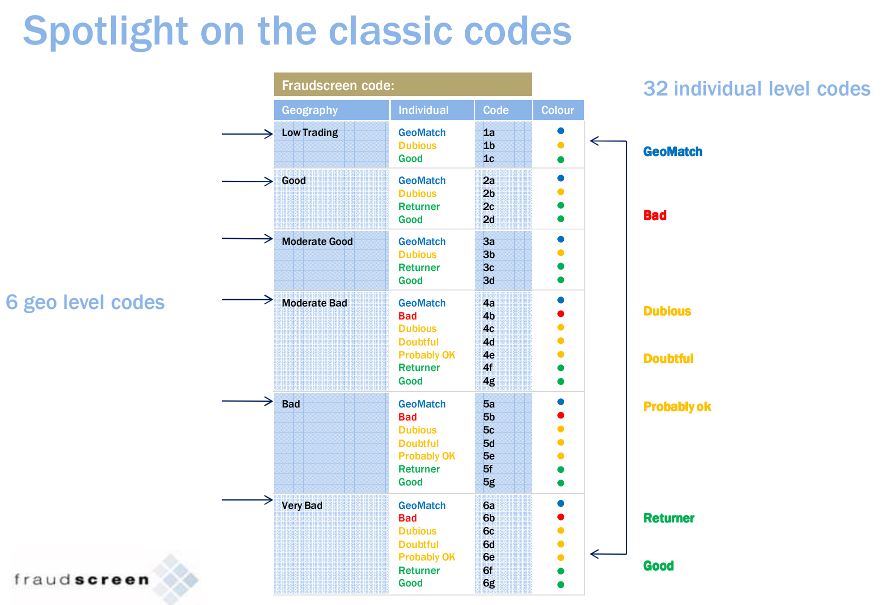


Table 1

By carefully analysing the patterns underlying the payment behaviour of these mail order customers, and by applying these measurements over a very large base of customers and transactions, Fraudscreen has been able to develop a coding system which summarises the behaviour patterns, and classifies them into groups. Each group is associated with a payment intent code.



Geography	Individual	Code	Colour
Low Trading	GeoMatch Dubious Good	1a 1b 1c	
Good	GeoMatch Dubious Returner Good	2a 2b 2c 2d	
Moderate Good	GeoMatch Dubious Returner Good	3a 3b 3c 3d	
Moderate Bad	GeoMatch Bad Dubious Doubtful Probably OK Returner Good	4a 4b 4c 4d 4e 4f 4g	
Bad	GeoMatch Bad Dubious Doubtful Probably OK Returner Good	5a 5b 5c 5d 5e 5f 5g	
Very Bad	GeoMatch Bad Dubious Doubtful Probably OK Returner Good	6a 6b 6c 6d 6e 6f 6g	

Table 2

The coding system is split into two parts, a number, and a letter. The number is a geographic code, based on post-code, splitting the database into groups of about 240 households. Where an individual does not appear on the database, because the individuals have not been users of mail order in the past, the geographic code helps to provide an indicator of likely payment performance. For most credit applicants, however, it is possible to allocate a code on an individual basis, classifying their behaviour into up to six groupings.

The classification 'Returner' probably deserves a word of explanation. When the coding system was conceived, people who returned mail order goods were seen as bad news. The retailer incurred a lot of cost for no return, if the good s were shipped straight back. However, from a payment intent point of view, this is very honest behaviour. The customer has gone to some length to fulfil their part of the bargain. They have not simply held on top the goods and done nothing.

The final part of the coding story is the colour code.

All credit portfolios are different. Some lenders have a substantial risk appetite, whist some are risk averse. Some are dealing in large tranches of lending, others in very small individual transactions. Some are making a turn only on the credit agreement, whilst others can count on a profit from the goods shipped as well. Different channels behave differently again. All of these factors make a substantial difference to whether a particular behaviour is seen as good or bad. The colour code provides a prism, called a 'signature', through which the behaviour can be interpreted for the needs of a particular lender and loan portfolio.

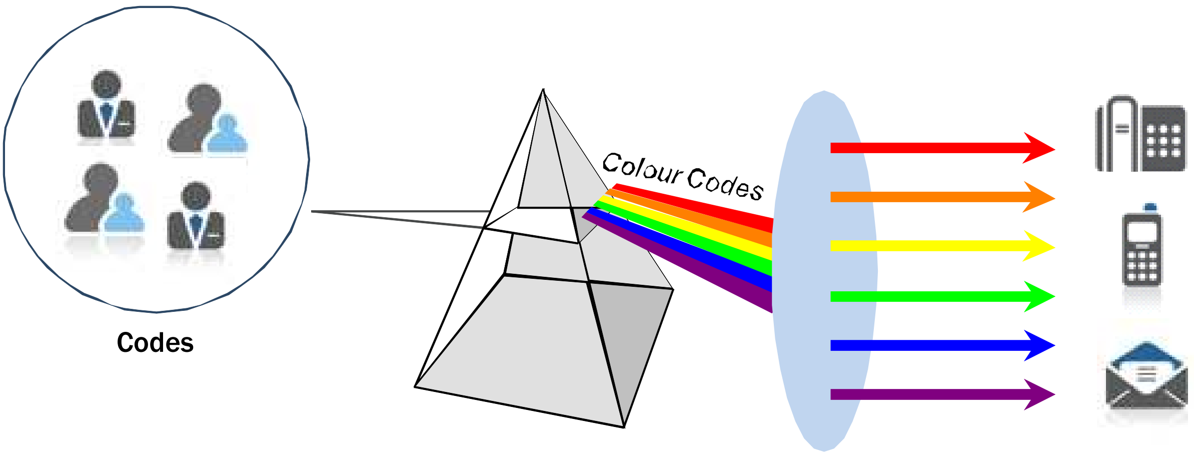


Table 3

Colours range from red, indicating the likelihood of poor payment performance, through amber, to green, indicating likely good performance. The blue colours indicate that the match is based on geography only. So blue colours can indicate either good or poor performance.

3. The Test – Outline

3.1 *Purpose of the test*

The test aimed to verify the assertion that using Fraudscreen payment intent codes produced statistically significant improvements in discrimination when used in conjunction with an existing well-performing credit scorecard.

3.2 *The Trader*

The trader supplying the data is engaged in on-account mail order. This business consists of opening a revolving credit account facility for customers, with a credit limit. In many ways the business is similar to a 'Private Label' credit card, except that the goods sold are exclusively supplied by the trader. Customers are recruited through a variety of channels: direct mail; internet and catalogue.

A crucial part of the operation is deciding who to can safely be allowed to open an account. The trader long ago recognised the benefits of credit scoring and has invested heavily in this technology.

3.3 *The test portfolio*

The portfolio data has been anonymised to protect the trader's identity.

A large number of accounts with known good / bad outcomes were coded using Fraudscreen payment intent codes. The trader retained a hold-out sample, to check results, and passed the rest to Fraudscreen for colour-coding and analysis.

A colour code signature was agreed with the trader, reflecting the trader's risk appetite for this portfolio.

Fraudscreen then analysed the data, comparing the results actually experienced using the scorecard, to those that would have happened, and Fraudscreen payment intent codes been used to supplement the credit score.

Further tests are planned to establish performance against current underwriting and trading conditions.

3.4 *Methodology*

The portfolio was evaluated by comparing the outcomes using the scorecard alone, and then using Payment intent code alone. In addition, outcomes using a combination of score and payment intent code were measured, and a strategy derived to assess value to the trader.

Finally, a simple logistic regression model was derived to calculate whether the payment intent coding made a significant contribution to the score alone.

The score and the payment intent code were the only variables tested. To develop a scorecard, ideally you would need to unpack the component data underlying the score, but this has not been done so far.

4 Test Results

The portfolio characteristics are shown in the two-dimensional matrix in Table 4 below.

TOTAL STARTERS								
	FS Colour Code							
Score	Deep Red	Red	Lower blue	Amber	Mid Blue	Green	Top Green	Grand Total
<430	129	510	1,178	444	920	719	651	4,550
431-490	1,439	4,748	7,173	3,071	3,209	1,668	584	21,890
491-550	1,740	5,438	12,162	4,503	6,729	3,000	1,164	34,736
551-610	1,157	4,224	11,360	4,235	7,185	4,157	1,827	34,143
611-670	827	3,348	10,302	4,202	7,731	5,303	2,717	34,428
671-730	537	2,705	10,454	4,424	8,787	7,010	4,362	38,277
731-790	407	2,370	10,952	4,143	10,587	8,945	6,272	43,674
791-850	291	2,148	11,826	4,572	12,902	11,792	8,841	52,371
851-910	266	2,001	12,855	5,165	15,360	15,251	12,824	63,720
911-970	191	1,817	12,119	5,451	16,251	16,841	16,901	69,569
971-1030	164	1,443	8,582	4,923	12,912	14,936	17,297	60,255
1031-1090	65	761	4,680	3,290	7,496	9,875	13,125	39,290
1091-1150	41	381	1,968	1,932	3,467	5,079	7,742	20,609
1151-1210	14	120	705	905	1,211	2,096	3,672	8,721
>1211	5	51	318	393	522	923	1,731	3,942
TOTAL	7,268	32,063	116,631	51,650	115,266	107,589	99,707	530,172

Table 4

Coarse-classed credit scores are shown on the left of the table, with Fraudscreen payment intent codes, classed into colours, across the top. The most seriously risky accounts are shown in the top left of the table, becoming less risky as you move to the right, and down the table. The cells of the table contain the volume of accounts appearing in each cross-tabulated category. This format will be used in a similar way in the rest of this paper. Results in respect of the in-house scorecard are described as 'A-score', with the results associated with Fraudscreen payment intent codes described as ''FS Colour Codes'.

Tables 5 and 6 below show the outcome statistics – good and bad outcomes, broken down into the same cross-tab format.

TOTAL GOODS								
	FS Colour Code							
Score	Deep Red	Red	Lower blue	Amber	Mid Blue	Green	Top Gree	Grand Total
<430	113	456	1,068	410	893	708	650	4,296
431-490	881	3,005	4,649	2,073	2,244	1,239	444	14,534
491-550	1,130	3,791	8,577	3,255	4,976	2,387	972	25,086
551-610	777	3,191	8,834	3,428	5,855	3,630	1,652	27,365
611-670	650	2,811	8,834	3,657	6,881	4,902	2,576	30,309
671-730	471	2,447	9,686	4,104	8,307	6,743	4,236	35,993
731-790	384	2,258	10,547	4,005	10,310	8,756	6,198	42,456
791-850	284	2,096	11,624	4,493	12,725	11,664	8,784	51,668
851-910	260	1,979	12,735	5,117	15,266	15,192	12,779	63,326
911-970	188	1,805	12,062	5,426	16,185	16,803	16,851	69,318
971-1030	158	1,433	8,544	4,905	12,888	14,912	17,274	60,113
1031-1090	65	759	4,670	3,282	7,487	9,857	13,116	39,234
1091-1150	41	381	1,964	1,929	3,461	5,075	7,736	20,585
1151-1210	14	120	704	903	1,209	2,094	3,671	8,714
>1211	5	51	318	393	522	921	1,730	3,939
TOTAL	5,415	26,579	104,811	47,378	109,205	104,880	98,666	496,932

Table 5

TOTAL BADS								
Score	FS Colour Code							
	Deep Red	Red	Lower blue	Amber	Mid Blue	Green	Top Gree	Grand Total
<430	17	54	110	35	27	11	2	254
431-490	558	1,743	2,525	998	965	429	140	7,356
491-550	611	1,647	3,585	1,248	1,754	614	192	9,650
551-610	380	1,034	2,526	807	1,331	527	176	6,779
611-670	177	537	1,469	545	851	401	141	4,119
671-730	66	258	768	320	480	267	126	2,285
731-790	23	113	405	138	278	189	74	1,218
791-850	8	53	203	80	177	128	57	704
851-910	6	23	120	48	95	59	45	395
911-970	3	12	57	26	66	38	50	251
971-1030	6	11	38	18	24	24	23	143
1031-1090	-	2	11	8	9	18	9	56
1091-1150	-	-	5	3	6	5	6	24
1151-1210	-	-	2	2	2	2	2	8
>1211	-	-	-	-	-	2	2	3
TOTAL	1,853	5,484	11,820	4,272	6,062	2,709	1,041	33,240

Table 6

The marginal bad rate is shown in Table 7.

These statistics show a scorecard which is working well, and providing a strong degree of discrimination against the good / bad outcome.

BAD RATE								
Score	FS Colour Code							
	Deep Red	Red	Lower blue	Amber	Mid Blue	Green	Top Gree	Grand Total
<430	12.8%	10.6%	9.3%	7.8%	2.9%	1.5%	0.2%	5.6%
431-490	38.8%	36.7%	35.2%	32.5%	30.1%	25.7%	23.9%	33.6%
491-550	35.1%	30.3%	29.5%	27.7%	26.1%	20.5%	16.5%	27.8%
551-610	32.8%	24.5%	22.2%	19.1%	18.5%	12.7%	9.6%	19.9%
611-670	21.4%	16.0%	14.3%	13.0%	11.0%	7.6%	5.2%	12.0%
671-730	12.3%	9.5%	7.3%	7.2%	5.5%	3.8%	2.9%	6.0%
731-790	5.5%	4.7%	3.7%	3.3%	2.6%	2.1%	1.2%	2.8%
791-850	2.6%	2.4%	1.7%	1.7%	1.4%	1.1%	0.6%	1.3%
851-910	2.3%	1.1%	0.9%	0.9%	0.6%	0.4%	0.4%	0.6%
911-970	1.6%	0.7%	0.5%	0.5%	0.4%	0.2%	0.3%	0.4%
971-1030	3.7%	0.7%	0.4%	0.4%	0.2%	0.2%	0.1%	0.2%
1031-1090	0.0%	0.2%	0.2%	0.2%	0.1%	0.2%	0.1%	0.1%
1091-1150	0.0%	0.0%	0.2%	0.2%	0.2%	0.1%	0.1%	0.1%
1151-1210	0.0%	0.0%	0.2%	0.2%	0.1%	0.1%	0.0%	0.1%
>1211	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.1%	0.1%
TOTAL	25.5%	17.1%	10.1%	8.3%	5.3%	2.5%	1.0%	6.3%

Table 7

The GINI coefficient for the scorecard and the FS colour codes were calculated, and are shown at the graph at Table 8.

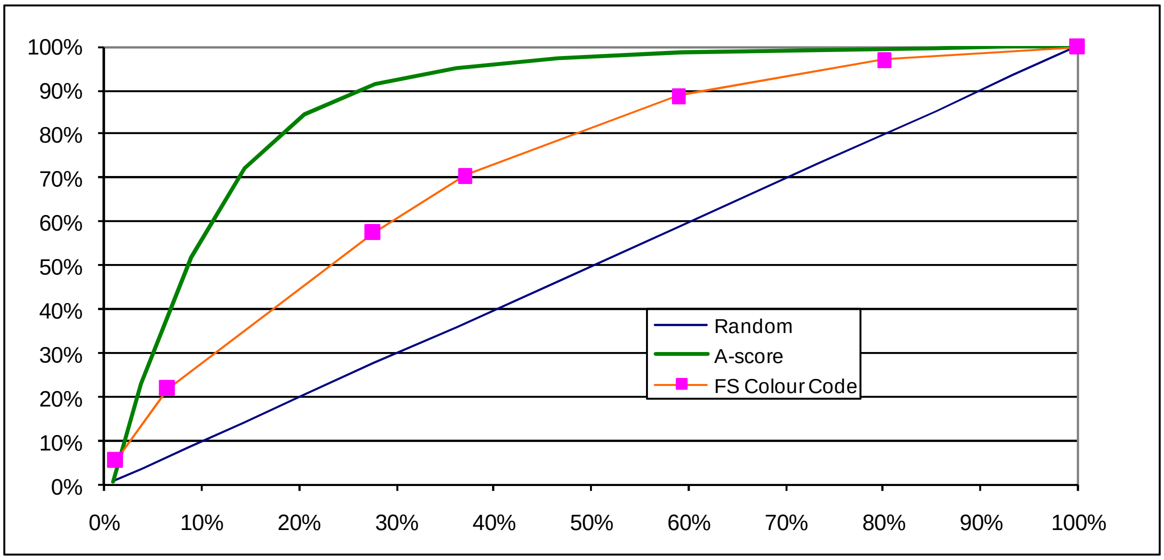


Table 8

The GINI coefficient, 77.6, for the scorecard is higher than that commonly experienced with, say, fixed instalment credit accounts or credit cards. A summary of the results of this initial test appears below:

	Info Value	Gini	KS
A-score	2.843	75.4%	64.3%
FS Colour Code	0.723	42.4%	33.4%
Cross-Char	2.937	77.6%	65.6%

Table 9

The FS Colour code GINI as a stand-alone characteristic has a coefficient of 42.4. The question is, what value, if any, would it add to the scorecard?

To find out, the analyst constructed a Logistic Regression model to establish the GINI coefficient of the combination. These were the results:

		Marginal IV	Info Value	Gini	KS
A-score		0.000	2.843	75.4%	64.3%
FS Colour Code		0.002	0.723	42.4%	33.4%
Cross-Char		0.013	2.937	77.6%	65.6%
Model				78.0%	65.6%

Table 10

The use of the FS Colour code does indeed add value to the overall scorecard.

5 BUSINESS IMPLICATIONS OF THE TEST

Using the additional power of the Fraudscreen payment intent code would require a whole article to itself.

This section contains an example, for demonstration purposes of one way of assessing the effect of the additional code. This method involves constructing a two-dimensional matrix, similar to the ones we have already seen, and varying the cut-off score in the light of the code. Other options include a full re-development of the scorecard to incorporate the additional data, although obviously that is a longer term (and more expensive) project, and using a delta score to 'correct' the existing scorecard in the light of the new data.

The two strategies we are considering are:

5.1 Single cut-off score. This is the traditional method of using a scorecard. Analysts choose a score below which the account is rejected.

5.2 Variable cut-off score. Here the cut-off score varies, depending on another variable in the matrix, in this case the FS colour code.

By varying the cut-off score (as long as the new variable adds value to the decision), you will obtain a better result than by using the single scorecard cut-off alone.

5.1 SINGLE CUT-OFF SCORE

The single cut-off score chosen for this part of the test was a score of 610. Below this it was assumed that accounts would be rejected, and above this score, accounts would be accepted.

Table 7, 'Bad Rate', shows that the marginal bad rate for this cut-off score is 19.9%. By using Table 6, and adding the 'Total Bads' in the first four categories, then expressing this as a percentage of the 'Total Starters' from Table 4, you can calculate the average bad rate for the portfolio as a whole resulting from using this strategy. The overall bad rate is 4.53%, and the number of bad accounts accepted is 24,039.

A similar process, using the 'Total Goods', Table 5, shows a total of 43,918 good accounts would be rejected by this strategy, and 453,014 good accounts accepted.

5.2 VARIABLE CUT-OFF SCORE

Knowing that the marginal bad rate of the single cut-off score strategy is 19.9% will enable us to structure a cut-off strategy for the variable cut-off score strategy which will improve on the results. We have chosen to take the improvement in the form of additional low risk sales volume, rather than improved bad rate. So we will need to examine each cell in 'Bad Rate' Table 7 to see if the cell falls below the marginal bad rate of 19.9%. If it does we should include it by granting an account to the applicant falling in that cell category, even if the previous strategy rejected it. If not, we should exclude the cell, by declining the applications falling into the cell definition.

BAD RATE

	FS Colour Code							
Score	Deep Red	Red	Lower blu	Amber	Mid Blue	Green	Top Gree	Grand To
<430	12.8%	10.6%	9.3%	7.8%	2.9%	1.5%	0.2%	5.6%
431-490	38.8%	36.7%	35.2%	32.5%	30.1%	25.7%	23.9%	33.6%
491-550	35.1%	30.3%	29.5%	27.7%	26.1%	20.5%	16.5%	27.8%
551-610	32.8%	24.5%	22.2%	19.1%	18.5%	12.7%	9.6%	19.9%
611-670	21.4%	16.0%	14.3%	13.0%	11.0%	7.6%	5.2%	12.0%
671-730	12.3%	9.5%	7.3%	7.2%	5.5%	3.8%	2.9%	6.0%

Strategy	
Single	Variable

Table 11

Table 11 shows a comparison of the two strategies.

By including the additional cells shown under the red line, but above the black line, and excluding the cell above the red line but below the black line, as before, you can use Tables 5 and 6 to calculate the volume of good accounts recruited and declined by the new strategy, and the volume of bad accounts accepted.

The strategy will accept 467,901 good accounts; decline 29,031 potentially good accounts, and accept 26,895 bad accounts. The overall bad rate has increased to 5.07%

5.3 COMPARISON

It is now possible to assess the possible financial performance of the two strategies, by putting a value on the accounts. We have assumed that a good account has a value of £500 within the outcome period, and that a bad account will generate on average £1,000 in costs and balances written off.

	Single cut-off		Variable Cut-off		Contribution from Variable Cut-off	Contribution %
Account Category	**Number**	**Amount £'000s**	**Number**	**Amount £'000s**	**£'000s**	
Good a/c at £500	453,014	226,507	467,901	233,950.5	7,443.5	3.68%
Bad a/c at £1,000	(24,039)	(24,039)	(26,895)	(26,895)	(2,856)	(1.41%)
Net Total	428,975	202,468	441,006	207,055.5	4,587.5	2.27%

Table 12

The 'Contribution %' column in the table is the percentage change from the Net Total line under the Single Cut-off Amount, i.e. £202m.

Table 12 shows that, by adding payment intent data to an already effective scorecard, and using the results to select additional accounts that would otherwise have been rejected, this trader was able to generate an additional contribution of £4.6m (2.8%), net of a 0.5% increase in overall bad rate.

6 CONCLUSIONS CAVEATS AND POSSIBLE ACTION PROGRAMME

This paper demonstrates that, by using payment intent data in the form of Fraudscreen codes, some credit portfolios' performance can be significantly improved.

Of course, not all credit portfolios are the same. The impact of using payment intent data varies from one portfolio to another, depending on such characteristics as:

- Sales channel;
- Loan value;
- Loan purpose;
- Demographic make-up of the customer base;
- Risk appetite, and
- Lender's profitability assumptions.

Because of this, it is not possible to take the results of this study and simply apply them to another portfolio. Results may vary from one loan portfolio to another, even for the same lender.

Nevertheless this study gives grounds for asserting that using payment intent data can add significantly to the bottom line results for loan portfolios. Other studies have shown that results can be replicated for a variety of lending situations, including a number of different lending products, and other operational areas such as collections and recoveries.

Clearly the approach taken in this paper would not be the best implementation approach you could take. Creating a two-dimensional table, as we have done here, does help illustrate the swap-sets that account for the forecast improvement in profitability. But setting up a strategy to use such a two-dimensional table, and monitoring the results effectively would be unnecessarily complex. A better approach would be to use a delta-score calculation to work out a correction to each of the combinations of score and colour code. In this way, you can test the modified score against the original score, using a champion / challenger approach, without major changes to the portfolio strategic approach. You can make the credit score summarise all the credit-relevant information you have about the portfolio, so that it becomes a sufficient statistic on its own.

To learn more about Fraudscreen Payment Intent codes, or to explore the possibilities for your own credit portfolio, you should contact Fraudscreen on www.fraudscreen.co.uk.

To learn more about this brief case study, you are welcome to contact the author with comments and suggestions. Please email him in the first instance on PeterWTaylor@talktalk.net enclosing your name and job title, and telephone and postal address contact details.