



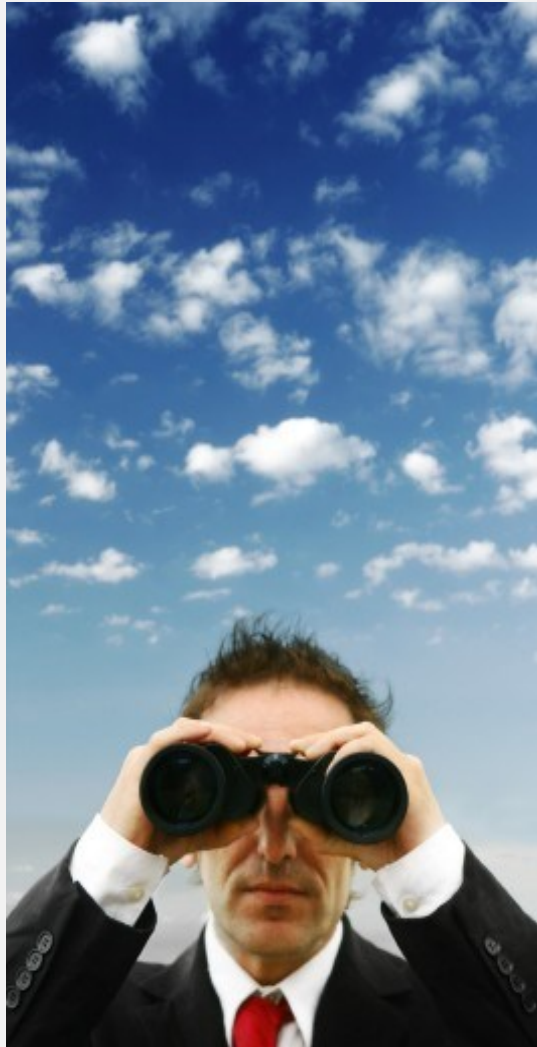
An Empirical Survey of Reject Inference Techniques Credit Scoring and Credit Control XIII

Radostina Etimova, Lee Gregory and
Milena Marcheva

August 2013



Summary



- I. What is Reject Inference (RI)?
- II. Why use it?
- III. Does it matter what technique we use?
- IV. Could Bureau Referencing add any value?



Key Findings



- I. The common 'extrapolation' technique is shown to improve discrimination on both accepts and rejects.
- II. Discrimination on masked accepts using outcome point bureau data to reference performance varies by portfolios' reject rates.
- III. Bureau referencing has shown potential to improve the resultant swap-in sets.



Agenda



- Objectives
- Portfolios
- Assessment methodologies
- Reject Inference techniques
- Results
- Conclusions



Objectives of the Survey



- I. Assess the discrimination on rejects and accepts of performance inferred using a variety of Reject Inference techniques
- II. Use two different actual performance outcome metrics for the assessment:
 - ▶ Outcome point bureau data for rejects
 - ▶ Actual performance on masked marginal accepts
- III. Evaluate the potential of using outcome bureau data as means of inferring performance for rejects
- IV. Examine the results on two different credit portfolios to highlight any differences



Portfolios

Unsecured Personal Loans

- ▶ **Application dates:**
 - Jan 2009 – Dec 2009
- ▶ **Outcome window:**
 - 12 months
- ▶ **Sample size:**
 - ~ 30,000
- ▶ **Reject rate:**
 - 35%
- ▶ **Known bad rate:**
 - 6.7%

Credit Cards

- ▶ **Application dates:**
 - Jan 2009 – Dec 2009
- ▶ **Outcome window:**
 - 12 months
- ▶ **Sample size:**
 - ~ 100,000
- ▶ **Reject rate:**
 - 55%
- ▶ **Known bad rate:**
 - 4.8%



Agenda



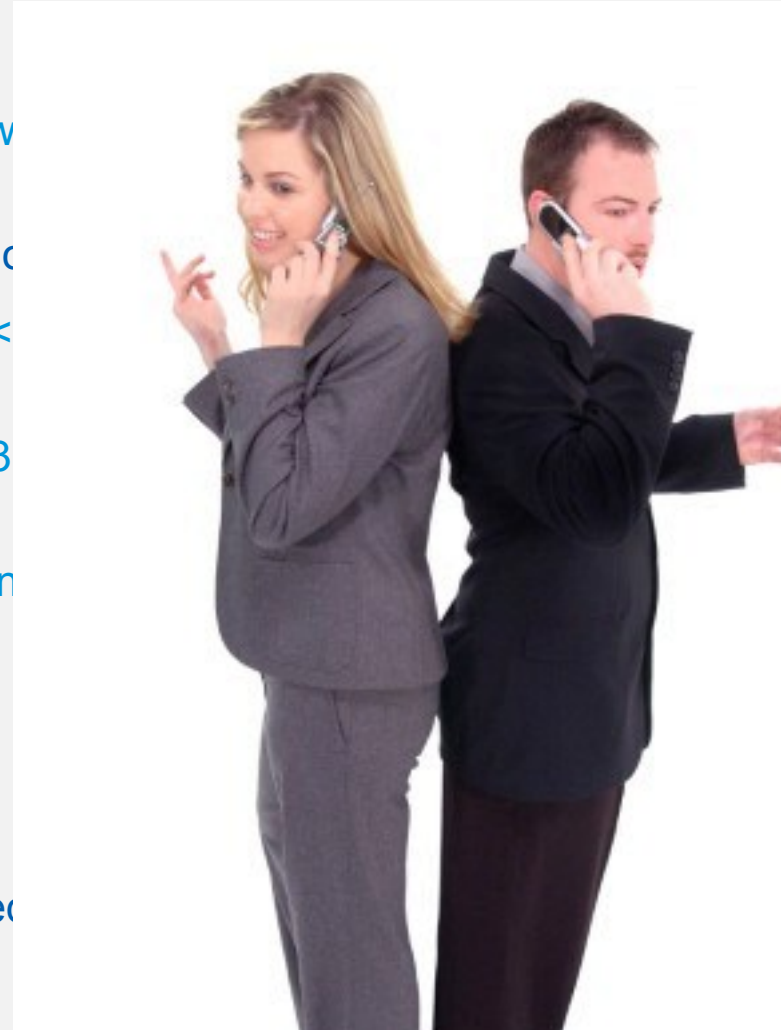
- Objectives
- Portfolios
- Assessment methodologies
- Reject Inference techniques
- Results
- Conclusions



Assessment Methodology

Bureau Referencing

- Identify from outcome point bureau data:
 - ▶ Rejected applicants with similar searches or an account opened with another lender in a window around the application date
- Build a reject performance flag based on this information
 - ▶ Applicants with opened accounts, delinquency < 30 days in next 12 months = **Good** (~ 10-15%)
 - ▶ Applicants with opened accounts, delinquency 30-90 days in next 12 months = **Bad** (~ 1-2%)
 - ▶ Applicants with searches and no opened accounts = **Bad** (~ 33-44%)
 - ▶ Applicants with no searches and no opened accounts = **Indeterminate** (~45-50%)
- Use this reject performance flag to assess the inferred probabilities from various RI approaches





Assessment Methodology

Performance Masking



- Identify accepted applications which were only marginally above the cut-off score
- Mask these marginal accepts' performance and treat them as rejects
 - ▶ There need to be enough masked accepts to make the assessment valid, but not enough to severely impact the bad rate of the remaining accepts
 - ▶ Approximately 5-10% of accepts were masked
- Apply reject inference techniques to infer performance on all rejects, then reveal the performance on masked accepts and use to assess



Agenda



- Objectives
- Portfolios
- Assessment methodologies
- Reject Inference techniques
- Results
- Conclusions



Reject Inference Techniques



Reject Inference techniques:

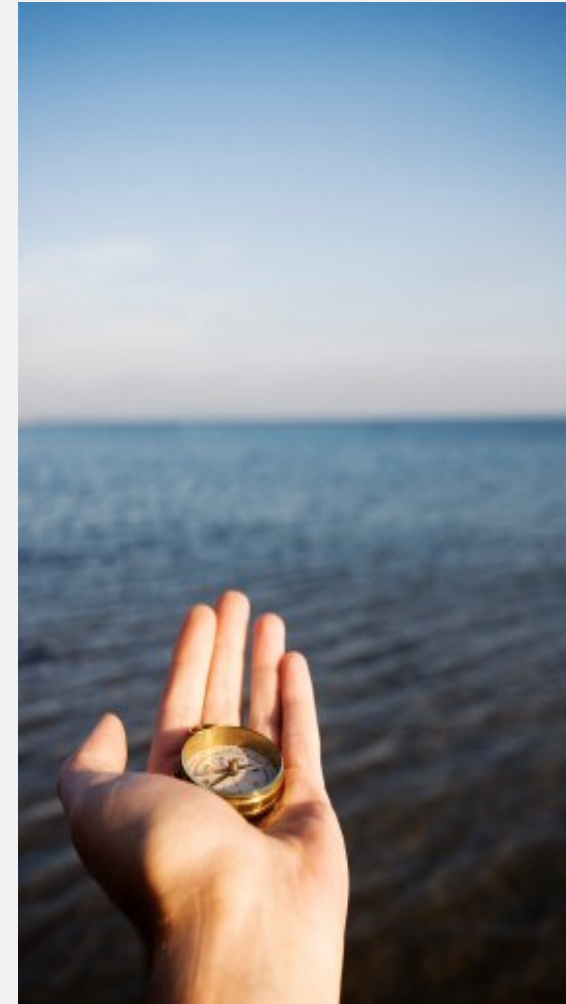
- ▶ [Known Performance](#)
- ▶ [Re-weighting](#)
- ▶ [Re-classification](#)
- ▶ [Outcome Scoring](#)
- ▶ [Extrapolation](#)
- ▶ [Bureau Referencing](#)



Reject Inference Techniques

Known Performance

- **Known Performance**
 - ▶ Build a model on the written population based on known performance only
 - ▶ Make no corrections to reduce the bias from ignoring rejects and not taken ups
- **Advantages:**
 - ▶ Easy
 - ▶ Quick
- **Disadvantages:**
 - ▶ Biased by previous strategies and decisions
 - ▶ Does not consider the “through the door” population





Reject Inference Techniques

Re-Weighting



- **Re-Weighting**
 - ▶ Build a model on the written population using known performance only, but apply a weight to scale up to the total population
 - ▶ The weight is based on a predictive score and ensures the written population has the same score distribution as the through-the-door population
- **Advantages:**
 - ▶ May reduce previous strategy bias
 - ▶ Relatively quick
- **Disadvantages:**
 - ▶ May not work well on small samples with high reject rates
 - ▶ Assumes that booked and declined applications have the same performance by score



Reject Inference Techniques

Re-Classification

■ Re-Classification

- ▶ Build a scorecard on the written population and the worst part of the rejected population
- ▶ Use a score to select the worst rejects to define as bad
- ▶ Re-weight the modelling population so that Known bads / Reclassified bads = 70/30

■ Advantages:

- ▶ May reduce previous strategy bias
- ▶ Relatively quick

■ Disadvantages:

- ▶ Would result in a small swap set because previous decisions are replicated
- ▶ Assumes that the score works well on declined accounts
- ▶ May not work well with high reject rates





Reject Inference Techniques

Extrapolation



- **Extrapolation:**

- ▶ As for Known Performance build a model on the written population based on known performance only
- ▶ Adjust the predicted probability of being good for rejects via manual or functional adjustments
- ▶ Known odds/Inferred need to be between odds = 3/6
- ▶ Inferred trends need to follow known trends

- **Advantages:**

- ▶ Reduces previous strategy bias
- ▶ Allows business logic to be applied and probabilities to be finely adjusted to common expectations

- **Disadvantages:**

- ▶ Time consuming
- ▶ Depends on the score used for extrapolation
- ▶ Subjective



Reject Inference Techniques

Outcome Scoring

▪ Outcome Scoring Inference:

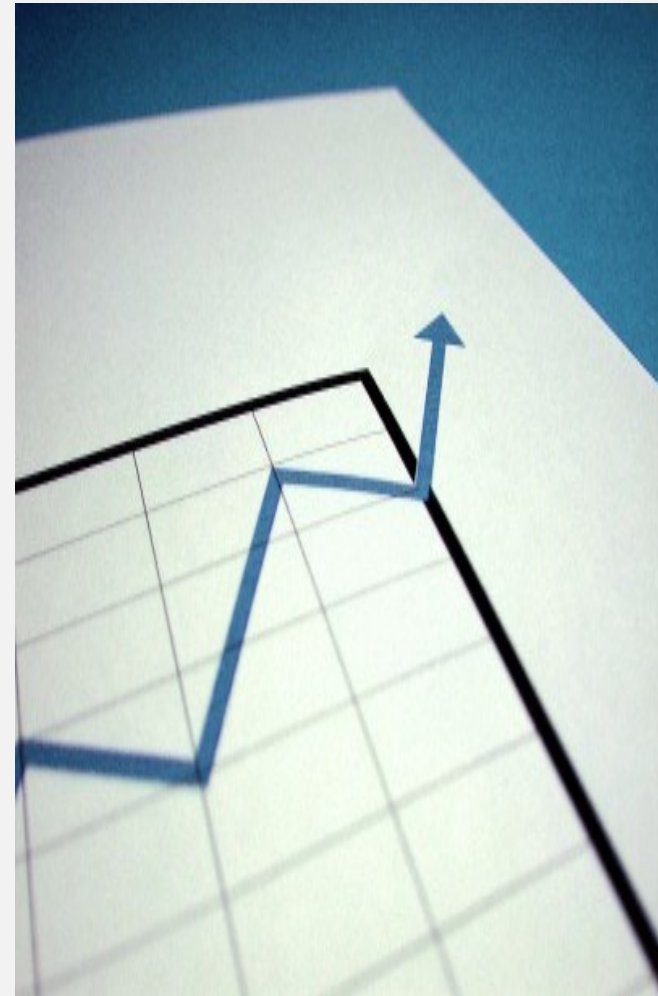
- ▶ Use a predictive generic application score at outcome to infer performance on the rejects
- ▶ The relationship between score and known performance on the written accounts is calculated by score band (a continuous function could be used)
- ▶ The outcome score bands on rejects are then transformed to give inferred performance

▪ Advantages:

- ▶ Very quick
- ▶ Easy and straightforward

▪ Disadvantages:

- ▶ Assumes that written and declined applications have the same performance by outcome score
- ▶ Results in a small score range on the rejects





Reject Inference Techniques

Bureau Referencing



- **Bureau Referencing:**
 - ▶ Build a model on the total population:
 - On known performance for accepts
 - On referenced performance for rejects
- **Advantages:**
 - ▶ Uses real historical data
 - ▶ Gives a more realistic perspective of the market situation and other clients' behaviour
- **Disadvantages:**
 - ▶ More effort for data preparation and extraction
 - ▶ There could be accounts that could not be linked.



Agenda

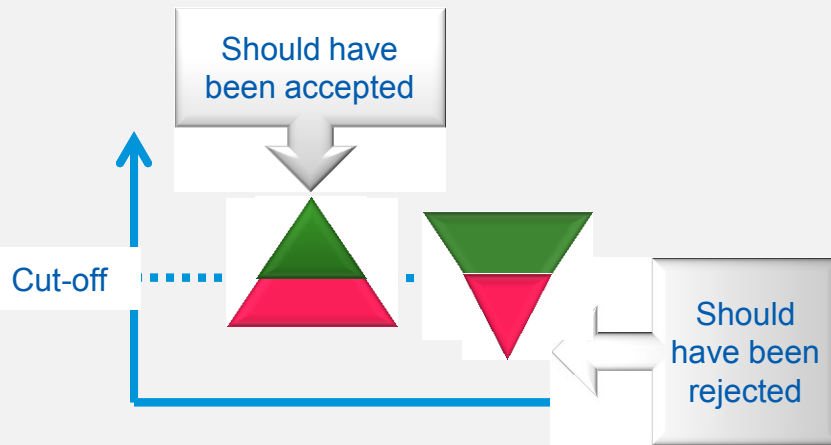


- Objectives
- Portfolios
- Assessment methodologies
- Reject Inference techniques
- Results
- Conclusions



Results

- Discrimination:
 - ▶ **Inferred probabilities on Rejects** – Did we infer performance close to the actual and referenced ones?
 - ▶ **Final Models** – How does discrimination change?
 - ▶ **Bureau Referencing** – Could it improve discrimination any further?



- Swap sets:
 - ▶ How many bads would we accept if we were to accept the same number of goods?



Results

Personal Loans

Gini Coefficients of Inferred Probabilities

	Known Performance	Re-weighting	Re-classification	Outcome Scoring	Extrapolation	Bureau Referencing
Masked Accepts	24	21	22	24	25	34
Rejects*	42	37	42	41	43	47

- Re-weighting, Re-classification and Outcome Scoring perform no better than Known Performance on the masked accepts and rejects.
- Extrapolation shows improvement on both masked accepts and Rejects.
- Bureau referencing performs very well on masked accepts (and, as is to be expected, on rejects):
 - ▶ This portfolio has the higher bad rate and lower reject rate.
 - ▶ The masked accepts in this portfolio have a profile similar to rejects.

* *Reject performance using bureau referencing*



Results

Personal Loans

Gini Coefficients of Final Models

	Known Performance	Re-weighting	Re-classification	Outcome Scoring	Extrapolation	Bureau Referencing
Total	73	71	72	73	73	74
Total Accepts	64	62	62	62	67	63
Masked Accepts	23	22	22	22	27	33
Rejects*	43	41	43	44	42	46

- Extrapolation shows improvement on total and masked accepts, but a slight decrease on Rejects.
- Bureau Referencing still has the highest discrimination on Masked accepts (and, as is to be expected, on Rejects), but it performs worse on total accepts:
 - ▶ Rejects performance inferred by this technique suggests accepts and rejects have quite different profiles.
 - ▶ A scorecard built to discriminate better on Rejects could be expected to discriminate worse on accepts.

**Reject performance using bureau referencing*



Results

Personal Loans

Constant Number of Goods Strategy

PL	Known Performance	Re-weighting	Re-classification	Outcome Scoring	Extrapolation	Bureau Referencing
Known+Rejected Goods	11,228	11,222	11,225	11,225	11,223	11,226
Known+Rejected Bads	2,009	2,004	1,877	1,979	1,929	1,823
Known Bads	493	503	489	490	469	473
Rejected Bads	1,516	1,501	1,388	1,489	1,460	1,350

- Re-classification captures fewer bads for the same number of goods than the extrapolation method, but extrapolation seems to capture around 4% fewer known bads than re-classification. The decrease in discrimination on Rejects, could be the reason it captures relatively more rejected bads.
- Bureau Referencing decreases the overall number of bads captured by ~ 4%.



Results

Credit Cards

Gini Coefficients of Inferred Probabilities

	Known Performance	Re-weighting	Re-classification	Outcome Scoring	Extrapolation	Bureau Referencing
Masked Accepts	34	24	31	22	34	26
Rejects*	39	39	37	40	40	50

- As for Personal Loans, Re-weighting, Re-classification and Outcome Scoring perform worse than Known Performance on the masked accepts, but there is more variation.
- Extrapolation has the best overall performance.
- Bureau Referencing performs better on rejects (as expected), but poorly on masked accepts:
 - ▶ This portfolio has a relatively high reject rate, biasing all models to perform slightly better on rejects.
 - ▶ The masked accepts seem to have a profile closer to the accepts, which may have the same strategic cause as the higher reject rate.

* *Reject performance using bureau referencing*



Results

Credit Cards

Gini Coefficients of Final Models

	Known Performance	Re-weighting	Re-classification	Outcome Scoring	Extrapolation	Bureau Referencing
Total	74	74	74	74	76	80
Total Accepts	68	63	68	63	69	61
Masked Accepts	33	25	31	18	34	26
Rejects*	39	41	39	37	41	50

- Extrapolation has the best overall performance – discriminates better on total and masked accepts and no worse on Rejects.
- Bureau Referencing has the highest overall discrimination, due to the significant increase of discrimination on Rejects, but it performs worse on total and masked accepts.

**Reject performance using bureau referencing*



Results

Credit Cards

Constant Number of Goods Strategy

CC	Known Performance	Re-weighting	Re-classification	Outcome Scoring	Extrapolation	Bureau Referencing
Known+Rejected Goods	40,099	40,109	40,119	40,019	40,157	40,144
Known+Rejected Bads	5,101	5,040	5,126	5,040	4,615	3,782
Known Bads	815	917	823	935	869	1,029
Rejected Bads	4,286	4,123	4,303	4,105	3,746	2,753

- Extrapolation captures around 10% fewer bads for the same number of goods than the rest of the methods.
- Bureau referencing decreases the number of bads by 24% compared to extrapolation.
- It captures more known bads, which could be due to the lost discrimination on accepts, but it compensates by fewer rejected bads.



Conclusions



- For both portfolios extrapolation infers probabilities with highest discrimination on both masked accepts and rejects.
- On both portfolios the extrapolation final models improve discrimination on total and masked accepts. Discrimination on rejects improves for the higher reject rate portfolio, but slightly decreases for the lower reject rate one.
- Bureau referencing has the potential to improve discrimination further on rejects, but the benefit on marginal cases depends on their profile.
- Despite the lost discrimination on accepts for the portfolio with higher reject rate, bureau referencing captures the fewest bads for the same number of goods for both portfolios.



Considerations

- **There are many choices to be made on the criteria for assigning the bureau referenced outcome**
 - ▶ Is counting searches with no opened accounts a bad too harsh?
 - ▶ Are there any other effective methods for referencing outcomes to reduced the number of referenced indeterminates?
 - ▶ What products are appropriate to reference against?
 - ▶ How wide a time window is appropriate?
- **Building final models is a complicated process incorporating business knowledge**
 - ▶ Are quick models representative?
- **Models built using bureau referencing have an accept/reject discrimination trade-off**
 - ▶ Middle risk niche modelling may alleviate this





Time for Questions

THANK YOU FOR YOUR ATTENTION!

Credit Scoring and Credit Control XIII 2013, Edinburgh

