Best Practices in Reject Inferencing

Dennis Ash
Experian

Steve Meester
The CIT Group

Conference on Credit Risk Modeling and Decisioning

Wharton FIC, University of Pennsylvania **May 29-30, 2002**

Objectives

Introduction

- What is Reject Inference
- Why we Need Reject Inference
- Literature Review

Reject Inference Techniques

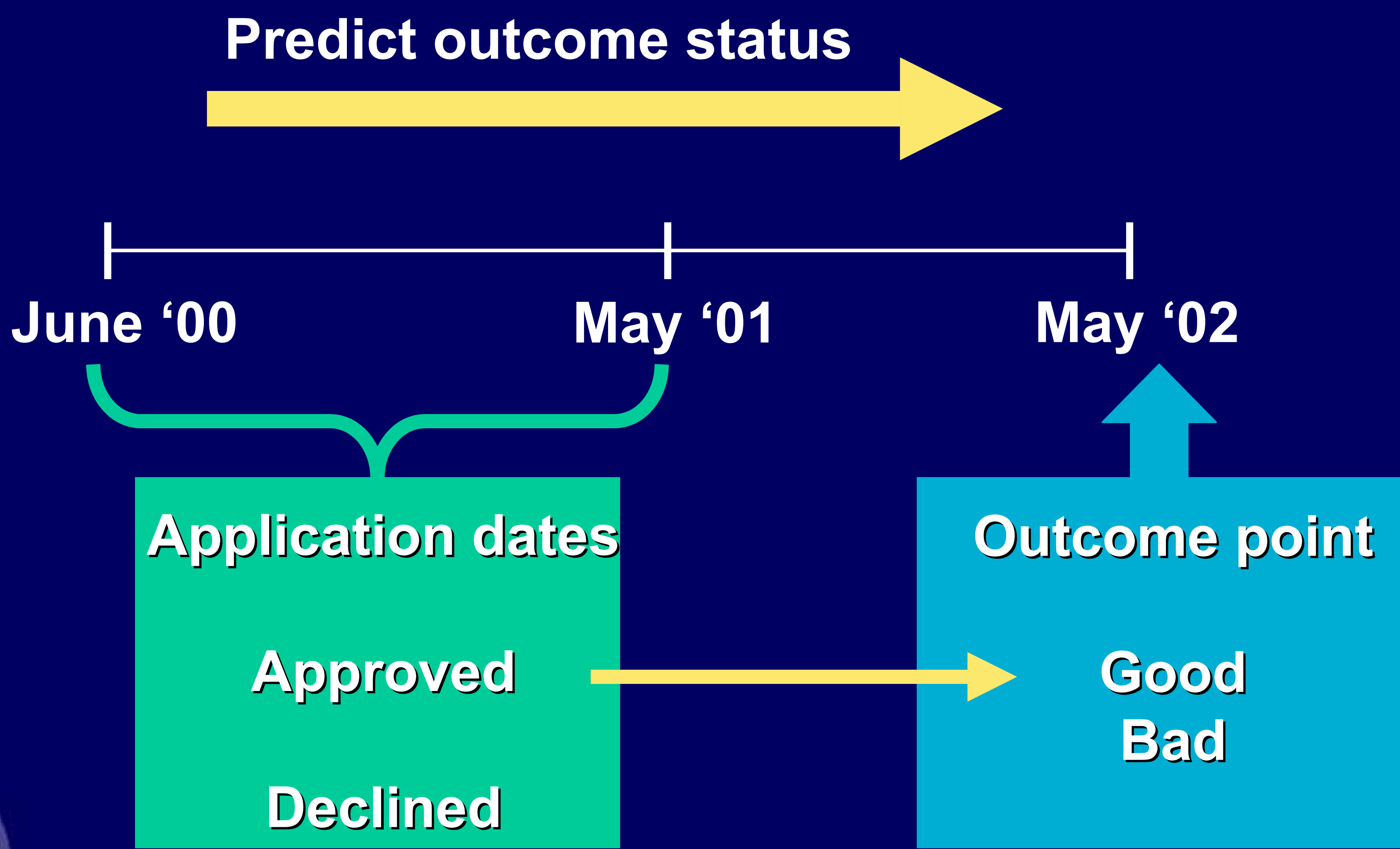
- Description
- Assumptions
- Outcomes

Reject inference: What is it?

Assignment of an inferred status (G/B) to applicants declined for credit

Equivalent to saying “if these applicants had been accepted, this is how they would have performed”

New applicant scoring



Is the missing outcome performance for rejects a problem?

- Sample bias
- Need statistically sound representative scorecard development sample
- Need scorecard to be effective for applications with reject profile
- Depends on past decision making

Why we need it

- If prior screening process used by the lending institution to separate applicants into accepts and rejects was applied in a (stratified) random manor
- Then all applicants would be represented in the accepted population

Why we need it

- A good (stratified) random sample of accepts could then represent the applicant pool
- It would contain some occurrences of bad credit followed by bad performance for all regions of the applicant pool

Why we need it

- Then we would not need reject inferencing.
- This is not often done. It is too expensive because the losses are too high.

Literature Review

Overview of scoring with discussion of reject inference: Hsia, 1978; Alan, Cho, Wagner, 1983; Hand and Henley, 1997.

Literature Review

Theoretical papers on reject inference: Copas and Li, 1997; Hand and Henley, 1993; Hand and Henley, 1994.

Literature Review

Heckman's correction: Heckman, 1979;Heckman, 1990; Greene, 1981.

Bivariate probit: Poirier (1980); Meng and Schmidt (1985); Boyes, Hoffman and Low (1989)

Why do we need reject inferencing?

- Development sample bias
- Forecast bias

Reject inference techniques

Techniques

- No reject inference
- Re-classification
- Re-weighting
- Parceling
- Heckman's bias correction
- Supplemental Bureau Data

1	2	3	4	5	6

No reject inference

- Build model on known bad / good flag
- Ignore rejects in model development
- Incorporate rejects in forecast

1 2 3 4 5 6

Reclassification

- The worst cases of rejects are selected and reclassified as accepts
- A “bad” status is then assigned

1 2 3 4 5 6

Reclassification – How's

The rejects are selected by

- **Reject / accept model**
- **Serious derogatory information**

1 **2** 3 4 5 6

Reclassification – How's

Reject / accept model

- Used to identify the worst rejected applicants
- Apply reject / accept model to approved and rejected accounts
- The lowest scoring rejects are reclassified

1 **2** 3 4 5 6

Reclassification – How's

Serious derogatory information

- Used to identify the worst rejected applicants
- Rejects who have more than a significant number of trades with seriously derogatory information
- Analyze RA and BG cross-tabs

1 **2** 3 4 5 6

Re-weighting

- Based on accept extrapolation
  - Accepted accounts are similar to declines
  - How declines would have performed if approved
- Accepts are weighted up to represent the rejects

Re-weighting – How's

Reject / accept model

- Used to identify similar applicants
- Apply reject / accept model to approved and rejected accounts
- The accounts (rejected and approved) are grouped by similar score
- The behavior of the approved accounts in a score interval can be used to infer what the likely behavior of the corresponding rejects would be, had they been approved

1 2 3 4 5 6

Re-weighting – Example

Score Interval	Rejects	Accepts=	Bads+	Goods
.	.	.	.	.
.	.	.	.	.
601-700	20	100	10	90
.	.	.	.	.
.	.	.	.	.

1 2 3 4 5 6

Re-weighting – Example

- 90% of approved accounts were good, while 10% were bad
- Can infer that
  - 10% of rejects in that interval (0.10*20=2) might have gone bad, had they been approved
  - 90% of reject (0.90*20=18) would be good
- By weighting the approved accounts by 1.2 (120/100) the sample would contain
  - 12 bads and 108 goods
- Therefore, the approved accounts were used as proxies for the rejects

Parceling

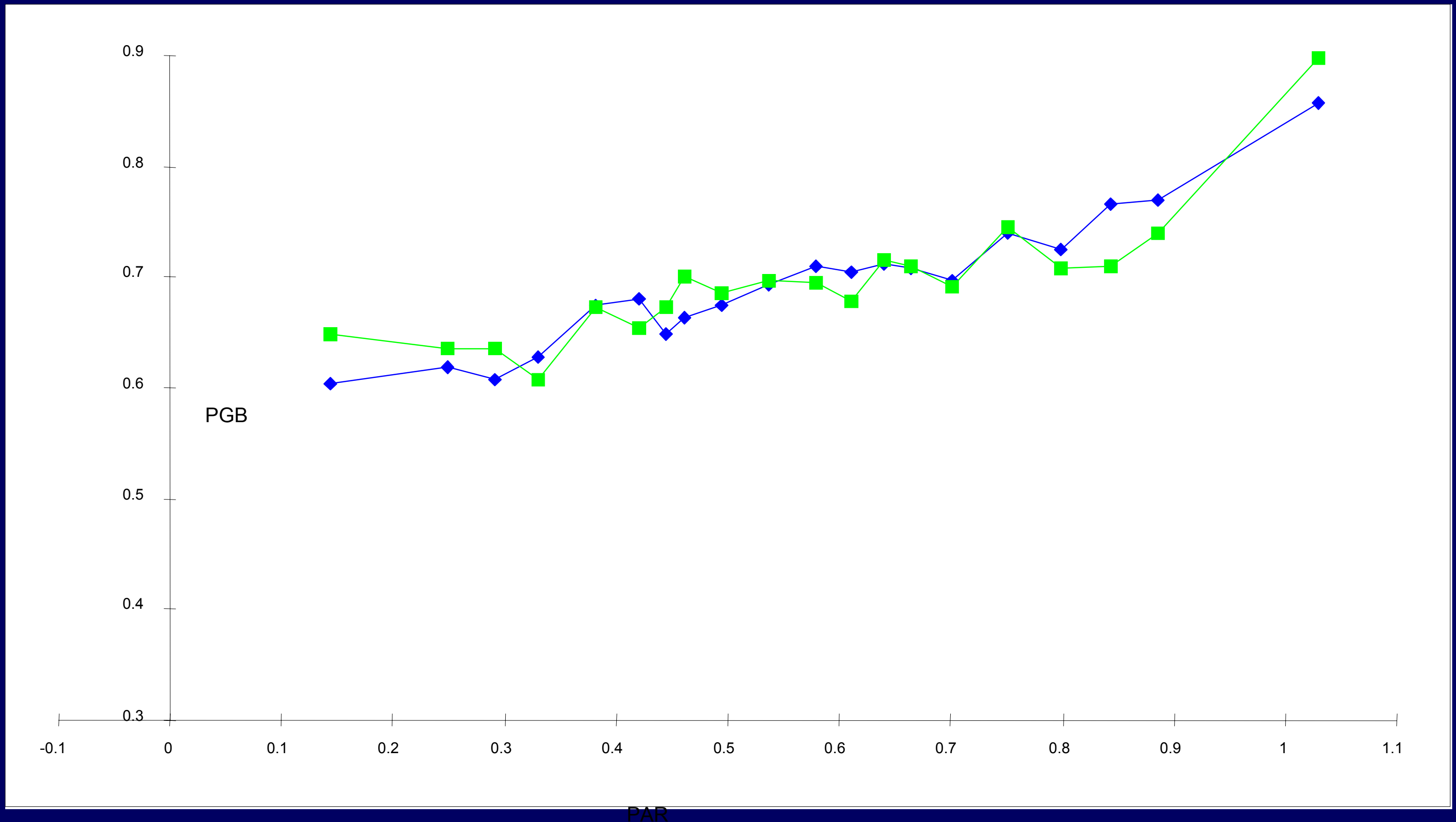
- Rejects are assigned into both bad and good categories, or probability of good
- Based on logical and statistical evidence of the proportion that would have gone bad

Parceling illustration

- Build reject / accept model
- Build known good / bad model
- Plot known good / bad model versus reject / accept model
  - Accepts
  - Rejects
- Adjust performance for rejects to reflect trend

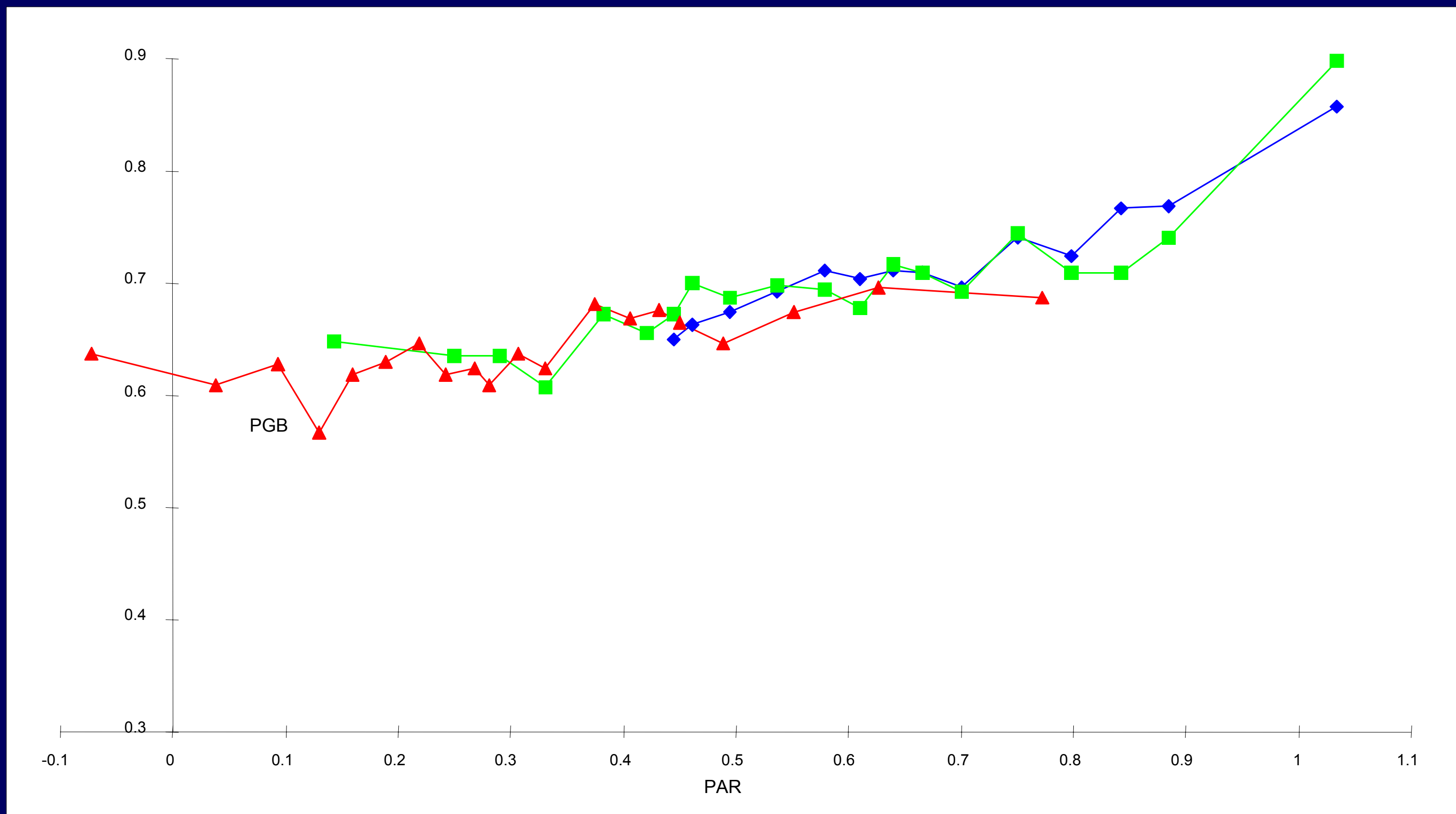
1 2 3 **4** 5 6

GB model based on known goods and bads

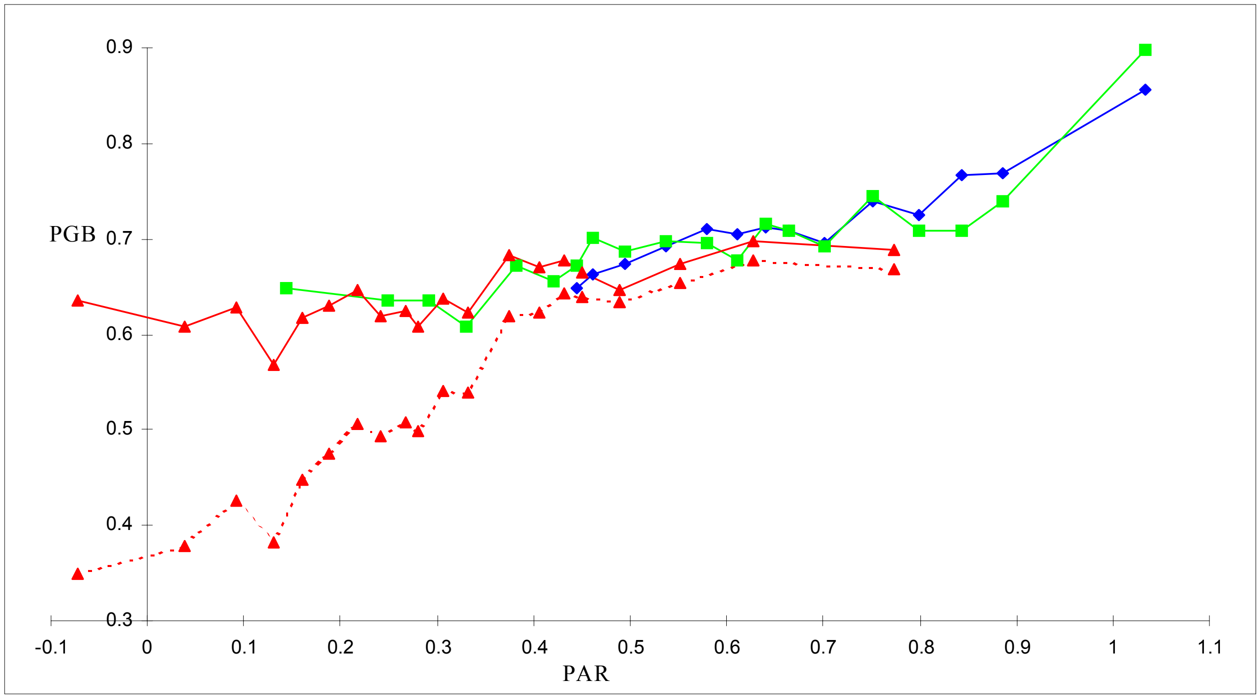


1 2 3 **4** 5 6

GB model applied to rejects



Adjusted performance on rejects



Heckman's Correction - Introduction

Hand & Henley (1993)

- **Lack of theoretical foundation that could justify any claim of bias correction**
- **Additional assumptions could validate RI methods, only if they are reasonable and consistent with statistical theory**

Heckman's correction

Heckman (1979)

- Discussed bias from using nonrandom selected samples to estimate behavioral relationships as a specification error
- He suggests a two stage estimation method to correct the bias
- The correction is easy to implement and has a firm basis in statistical theory

Heckman's correction

- Normality assumption
- Provides a test for sample selection bias
- Formula for bias corrected model

Shortcomings/Assumptions

No Reject Inference

- Does not adjust for sample bias.
- Portfolio quality estimates will be optimistic over the rejects.

Reclassification

- Ad-hoc.
- Implies $P(\text{bad} \mid X) = 1$ over a segment of the covariate space. We know this is not true.
- May bias the scoring model over the accepts.

Re-weighting

- Assumes

$$P(\text{bad} \mid X, \text{rejected}) = P(\text{bad} \mid X, \text{accepted}).$$

This is a very strong and generally unrealistic assumption.

Implies accept/reject procedure provides no discrimination given the bureau data X.

- Must have accepts with the same bureau profile as the rejects.

Heckman/Bivariate Probit

- Accept/reject procedure must be stochastic.
- All factors used in the accept/reject decision must be observable, i.e. no additional factors may be considered by credit managers.

How Well Do These Work?

- Several studies have shown that gains from using correction for sample selection based on observation data are less than expected
- Reliable model based reject inference is impossible - model assumptions are important and are violated
- But the information loss due to selection bias is substantial
- Need real information on rejects

Supplemental Bureau Data

- Obtain bureau data on accepts and rejects at the end of the observation period.
- Use the performance with other creditors over the observation period to infer how the rejects would have performed had they been accepted.

Supplemental Bureau Data Methodology

- Let Z denote the downstream bureau data.
- Fit a model for P(bad | Z) over the accepts.
- Impute P(bad | Z) for the unobserved Y for a reject.
- This is parceling BUT we use payment performance with other creditors over the time frame of interest to determine the parceling for a prior decline.
- The parceling is no longer subjective. It is driven by supplemental performance data.

Assumptions

Key assumption:

$$P(\text{bad} \mid X, Z, \text{rejected}) = P(\text{bad} \mid X, Z).$$

That is, the bureau data at time of application and the downstream bureau data contain all the relevant information about P(bad).

This is a much weaker assumption then required for re-weighting.

Shortcomings

Requires a good bureau match rate.

Supplemental Bureau Data: Cautions

- Models for which the likelihood score is linear in Y, just impute P(bad |X, Z) for Y.
  - e.g. logistic regression model.
- Models for which the likelihood score is non-linear in Y, impute $E[S(\theta) |X, Z)$ for $S(\theta)$, Meester (2002).
  - e.g. linear model.
- Naive standard error estimates are not correct. Bootstrap!

Example

- 9259 leases from a business which approves approx. 98% of applications.
- Create "declines" if any prior liens or judgments.

	"Accepts"	"Declines"	Total
# apps.	8,127	1,132	9,259
Bad rate	6.3%	13%	7.1%

Example

- Fit logistic regression model to full sample with observed response to get the "Gold Standard" model.
- Fit model with no reject inference.
- Fit logistic regression model using the reject inference procedure.

Example: Parameter Estimates

	Gold Stnd.	Excl. Declines.	Reject Inf.
Intercept	-2.250	-2.2161	-2.1472
log(liens+1)	0.7378	NA	0.7824
1(judgments>0)	0.7143	NA	1.0063
X3	0.4668	0.4497	0.4299
X4	-0.2303	-0.1911	-0.1919
X5	-0.4052	-0.4744	-0.4100
X6	0.7744	0.9400	0.7634
X7_1	0.8031	0.7371	0.5883
X7_2	0.5916	0.4677	0.3423
X7_3	0.9795	1.1561	0.9778
X7_4	1.1023	1.0388	0.9592
log(suits+1)	NA	0.4953	NA

Estimated Standard Errors

Parameter	Naive Estimate	Bootstrapped
Intercept	0.3726	0.3638
log(liens+1)	0.0963	0.1688
1(judgments>0)	0.1318	0.4030
X3	0.1910	0.1922
X4	0.0275	0.04164
X5	0.0445	0.0444
X6	0.1302	0.1217
X7_1	0.3667	0.3734
X7_2	0.3738	0.3875
X7_3	0.3871	0.3955
X7_4	0. 3496	0. 3639

Portfolio Quality: Percent Bad

Approval Rate	Full Data Model	Imputed Model		No Reject Inference	
	Actual	Actual	Estimate	Actual	Estimate
90%	5.9	6.0	6.1	6.2	5.3
80%	5.0	5.1	5.2	5.3	4.5
70%	4.1	4.3	4.3	4.5	3.7
60%	3.6	3.6	3.6	3.8	3.2
50%	3.2	3.2	3.2	3.4	2.9
40%	2.7	2.6	2.7	3.0	2.4
30%	2.4	2.4	2.4	2.7	2.2
20%	2.3	2.1	2.1	2.8	2.2
10%	2.6	2.3	2.3	2.3	1.9

Conclusions

- Other reject inference methods
  - require very restrictive assumptions: Heckman/Bivariate Probit, Re-weighting;

or

  - employ adhoc intervention which may lead one astray: Re-classification, parceling.

Conclusions

- Parceling with downstream bureau data uses additional, *data driven* information, for the reject inference.
- Requires fewer assumptions.
- Requires good bureau match.

When to use each method

Method	Use
None	▪ Low reject rate ▪ Random decisioning
Reclassification with derogs	▪ Simulate policy rules
Reclassification with score	▪ Extreme rejects by score
Reweighting	▪ Accept / reject overlap
Parceling	▪ Strict adherence to scorecard
Heckman	▪ To test bias ▪ Easy to use
Bureau	▪ Quality bureau match

When to use each method

Method	Use
None	■ Low reject rate ■ Random decisioning
Reclassification with derogs	■ Simulate policy rules
Reclassification with score	■ Extreme rejects by score
Reweighting	■ Accept / reject overlap
Parceling	■ Strict adherence to scorecard
Heckman	■ To test bias ■ Easy to use
Bureau	■ Quality bureau match

When to use each method

Method	Use
None	▪ Low reject rate ▪ Random decisioning
Reclassification with derogs	▪ Simulate policy rules
Reclassification with score	▪ Extreme rejects by score
Reweighting	▪ Accept/reject overlap
Parceling	▪ Strict adherence to scorecard
Heckman	▪ To test bias ▪ Easy to use
Bureau	▪ Quality bureau match

When to use each method

Method	Use
None	■ Low reject rate ■ Random decisioning
Reclassification with derogs	■ Simulate policy rules
Reclassification with score	■ Extreme rejects by score
Re-weighting	■ Accept / reject overlap
Parceling	■ Strict adherence to scorecard
Heckman	■ To test bias ■ Easy to use
Bureau	■ Quality bureau match

When to use each method

Method	Use
None	■ Low reject rate ■ Random decisioning
Reclassification with derogs	■ Simulate policy rules
Reclassification with score	■ Extreme rejects by score
Re-weighting	■ Accept / reject overlap
Parceling	■ Strict adherence to scorecard
Heckman	■ To test bias ■ Easy to use
Bureau	■ Quality bureau match

When to use each method

Method	Use
None	▪ Low reject rate ▪ Random decisioning
Reclassification with derogs	▪ Simulate policy rules
Reclassification with score	▪ Extreme rejects by score
Re-weighting	▪ Accept / reject overlap
Parceling	▪ Strict adherence to scorecard
Heckman	▪ To test bias ▪ Easy to use
Bureau	▪ Quality bureau match

When to use each method

Method	Use
None	▪ Low reject rate ▪ Random decisioning
Reclassification with derogs	▪ Simulate policy rules
Reclassification with score	▪ Extreme rejects by score
Re-weighting	▪ Accept / reject overlap
Parceling	▪ Strict adherence to scorecard
Heckman	▪ To test bias ▪ Easy to use
Bureau	▪ Quality bureau match

Combinations of approaches!

- Sometimes essential
- Depends on technique used
- Depends on past decisioning
- Depends on sample available

Conclusions

- Need for reject inference influenced by decline rate
- All methods discussed are valid under assumptions
- However, the best method varies case to case and a method may be invalid if assumptions are violated
- Select method according to the portfolio and validity of assumptions
- Use real outcome information on accounts if available
- Frequently require multiple approaches