Cross-Tab Weighting
for Retail and Small-Business Scorecards in Developing Markets

Dean Caire (DAI Europe)
Mark Schreiner (Microfinance Risk Management L.L.C)

August 24, 2011

DAI Washington | London | Johannesburg | Ramallah

Structure of Presentation

- Why do we use cross-tab weighting?
- How does cross-tab weighting work?
- Results: cross-tab weighting vs. Logit

DAI

Why do we use cross-tab weighting?

- We work in developing markets
- It facilitates the transfer of skills to bankers with limited modelling experience
- It simplifies the scorecard-building process

DAI

How does cross-tab weighting work?

Table 1: Cross Tab: Gender of Borrower

A		Women	Men	TOTAL
B	# Goods	52	40	92
C	# Bads	3	5	8
D	**% Bad**	**5.5%**	**11.1%**	**8%**
E	# Total	55	45	100
F	**% Total**	**55%**	**45%**	**100%**
G	**POINTS**	**5.5**	**11.1**	

DAI

How to build the model?

- Model each potential factor using cross-tab analysis

- Build multi-factor model balancing two goals:
  - To have a comprehensive set of factors – the 5 Cs of Credit
  - To maximize predictive power

DAI

How to understand the model?

- The “score” for a borrower is the sum of points received for each risk factor in the scorecard

- Higher points indicate higher risk

Point scores are the bad rate per risk characteristic

DAI

The full model can also be presented as a cross-tab

Table 2: Cross Tabulation: Multi-Factor Model

A	Risk Group	1	2	3	4	5	6	7	Total
	Points in Range	9.2 - 11.7	11.8 - 14.2	14.3 - 16.8	16.9 - 19.3	19.4 - 21.9	22.0 - 24.4	24.5 – 27.0	
B	# Goods	1,694	11,850	4,937	1,906	1,216	161	10	21,774
C	# Bads	1	35	32	37	96	25	3	229
D	Bad Rate	0.1%	0.3%	0.6%	1.9%	7.3%	13.4%	23.1%	1.1%
E	# Total	1,695	11,885	4,969	1,943	1,312	186	13	22,003
F	%Total	7.7%	54.0%	22.6%	8.8%	6.0%	0.8%	0.1%	100.0%

DAI

Results: cross-tab weighting vs. Logit

- Slightly lower predictive power (AUC)
- Easier to interpret points (coefficients)
- Easier to monitor changes in risk factors by comparing current bad rates to points in model

DAI

Empirical Example 1:
"Group loans" to micro-borrowers in Tajikistan

Data:

31,433 contracts

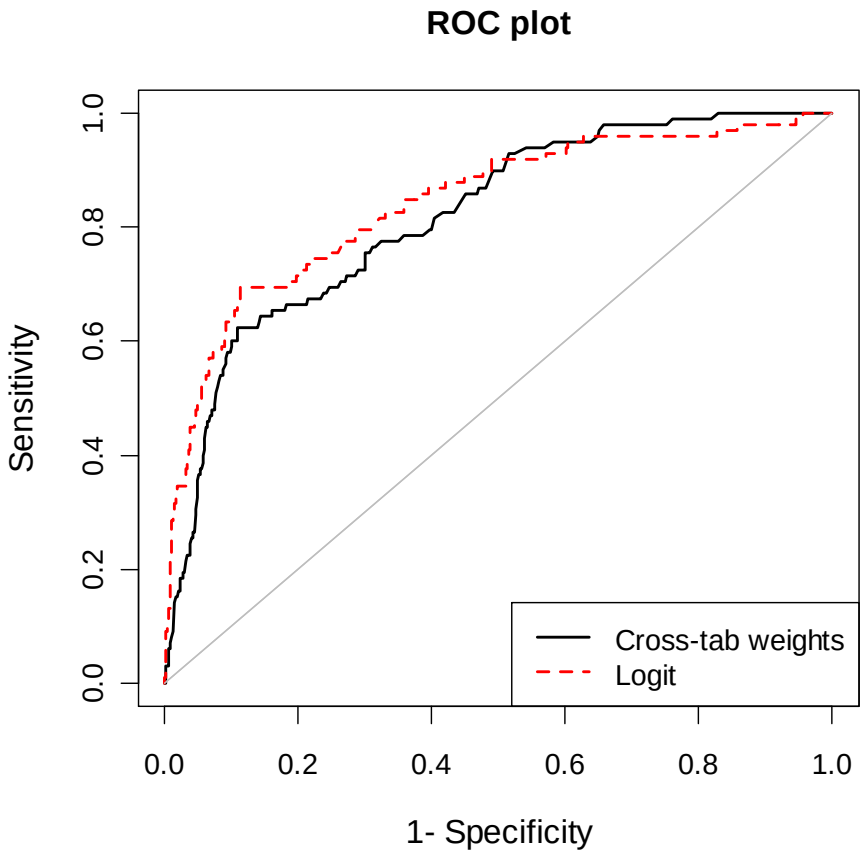
327 "bad" (30 or more days in arrears), bad rate 1%

14 risk factors:

- Loan terms (3 factors)
- Borrower demographics (6 factors)
- Borrower credit history (2 factors)
- Financial ability to repay (3 factors)

DAI

Empirical Example 1:
"Group loans" to micro-borrowers in Tajikistan



Cross –tab AUC: 0.818
Logit AUC: 0.841

DAI

Empirical Example 2:
Overdraft loans to small businesses in Bulgaria

Data:

1,434 contracts

140 “bad” (90 or more days in arrears) , bad rate 10%

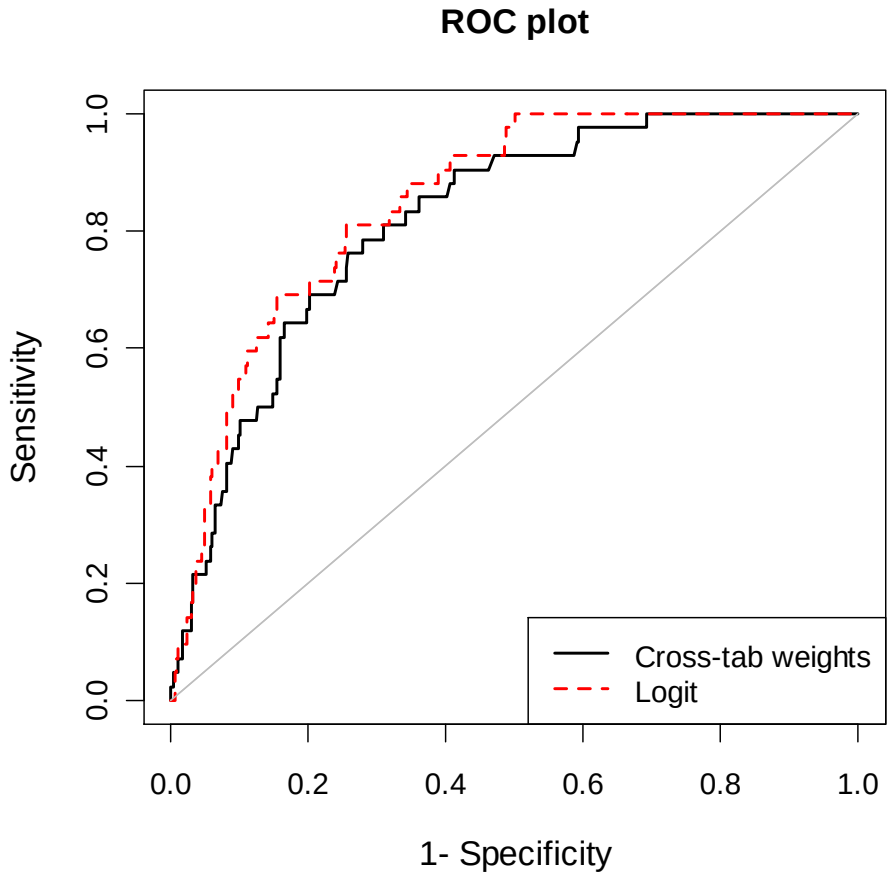
17 risk factors:

- Loan terms (2 factors)
- Non-financial factors (3 factors)
- Borrower credit history (2 factors)
- Financial ability to repay (5 factors)
- Risk mitigation (5 factors)

DAI

Empirical Example 2:

Overdraft loans to small businesses in Bulgaria



Cross –tab AUC: 0.818
Logit AUC: 0.850

DAI

Empirical Example 3:
Microlender in Latin America

Data:

14,039 contracts

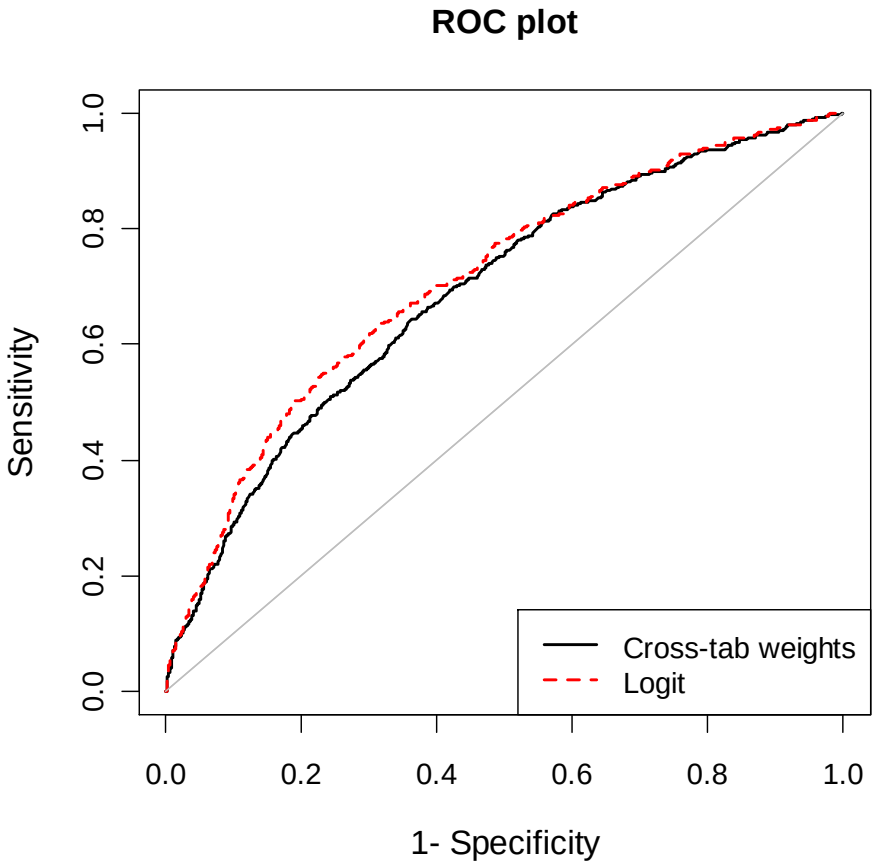
2,187 “bad” (30 or more days in arrears) , or 15% bad

37 risk factors:

- Loan terms (6 factors)
- Borrower demographics (18 factors)
- Borrower credit history (8 factors)
- Financial ability to repay (5 factors)

DAI

Empirical Example 3:
Microlender in Latin America



Cross –tab AUC: 0.691
Logit AUC: 0.710

DAI

Summary of Empirical Results

Table 3: Summary of Empirical Examples

Sample	Model	Tajikistan	Bulgaria	Latin America
Development (70%)	Cross-tab AUC	0.83	0.82	0.73
	Logit AUC	0.87	0.83	0.75
	Difference	5%	1%	3%
Validation (30%)	Cross-tab AUC	0.82	0.82	0.69
	Logit AUC	0.85	0.85	0.71
	Difference	4%	4%	3%

DAI

Summary of Main Points

- Cross-tab weights are much simpler than Logit (or other techniques) and make sense to managers
- Cross-tab weights have comparable predictive power.
- The methodology is very convenient in terms of getting models used and understood n practice

DAI