Low Response Rate from Merchants? Sample and Ask Consumers!

Heng Chen, Joy Wu

Bank of Canada

Overview

We apply indirect sampling outlined by Deville and Lavallée (2006) and Lavallée (2007) to estimate merchant payment acceptance through a consumer payment diary.

- **Key Contribution 1**: Consumers are both the sampling and reporting units
- **Key Contribution 2**: account for three-day diary structure through statistical adjustment

Motivation

	Merchant Surveys	Consumer Surveys
Survey frame	Custom-built	Readily available
Survey methodology	Computer Assisted Telephone Interview	Online
Response Rate	2.5% (2015), 2.8% (2018) and 4.8% (2021/22)	7.4 % (2017)

	Direct Sampling	Indirect Sampling
Sampling Unit	Merchants	Consumers
Response Unit		
S_M	Drawn from sample frame	Mapped from consumer-merchant transactional data
y_m	Reported by merchants	Reported by consumers
w_m	Known (design-based sampling)	Estimated using Generalized Weight Share Method (GWSM)

Indirect Sampling Estimator

$$\hat{u}^{3,cal} = \frac{\sum_{m\in\hat{S}_M^3} \hat{w}_m^{cal}\hat{y}_m}{\sum_{m\in\hat{S}_M^3} \hat{w}_m{}^{cal}}$$

- $\hat{S}_M^3 \equiv \cup_{c\in S_C} \Omega_c^3$
  - where Ω_c^3 the set of merchants visited by consumer c over three-days
- $\hat{w}_m^{cal} = \hat{w}_m F(\hat{\boldsymbol{\lambda}}^{\mathrm{T}}\boldsymbol{x}_m)$
  - where $\hat{w}_m$ is the GWSM weight, $\boldsymbol{x}_m$ is the vector of auxiliary variables (business size, industry, locality, region) and $F(\hat{\boldsymbol{\lambda}}^{\mathrm{T}}\boldsymbol{x}_m)$ is the calibration objective function
- $\hat{y}_m = f(usage, perceived\ acceptance) = f(u_m, p_m)$

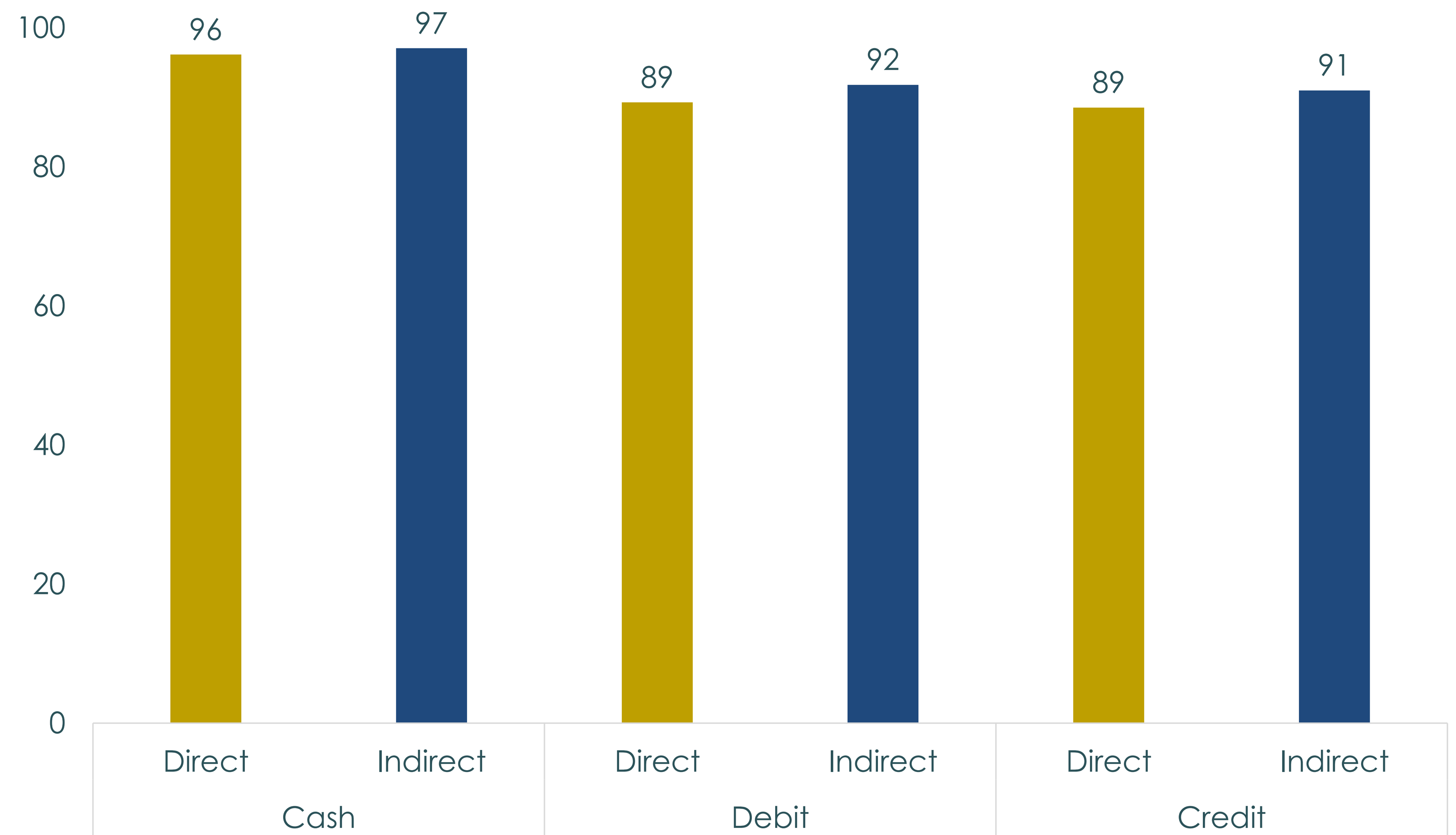
Assumptions

- Good coverage of merchant population
- High quality of consumer responses
- Few non-recorded merchants

References

- Deville, J., & Lavallée, P. (2006). Indirect sampling: The foundations of the generalized weight share method. Survey Methodology, 32(2), 165.
- Lavallée, P. (2007). Indirect sampling. New York: Springer.
- Haziza, D., & Lesage, É. (2016). A discussion of weighting procedures for unit nonresponse. Journal of Official Statistics, 32(1), 129-145.

Indirect and direct estimates are comparable



Direct sampling (2022 MAS):
Indirect sampling Rule 3 (2022 Consumer MOP):

Merchants report MOP acceptance
Consumers report **Merchant's** MOP acceptance

Step 1: Constructing merchant sample $\hat{S}_M$

"What was the name of the business where you made this purchase"

String-matching is performed on all reported merchant names to identify the set of unique merchants, $\hat{S}_M$. Evidence of coverage of merchant population: coverage of consumer sample $\boldsymbol{S_C}$ (Table 1) and coverage of merchant sample $\hat{S}_M^3$ (Table 2)

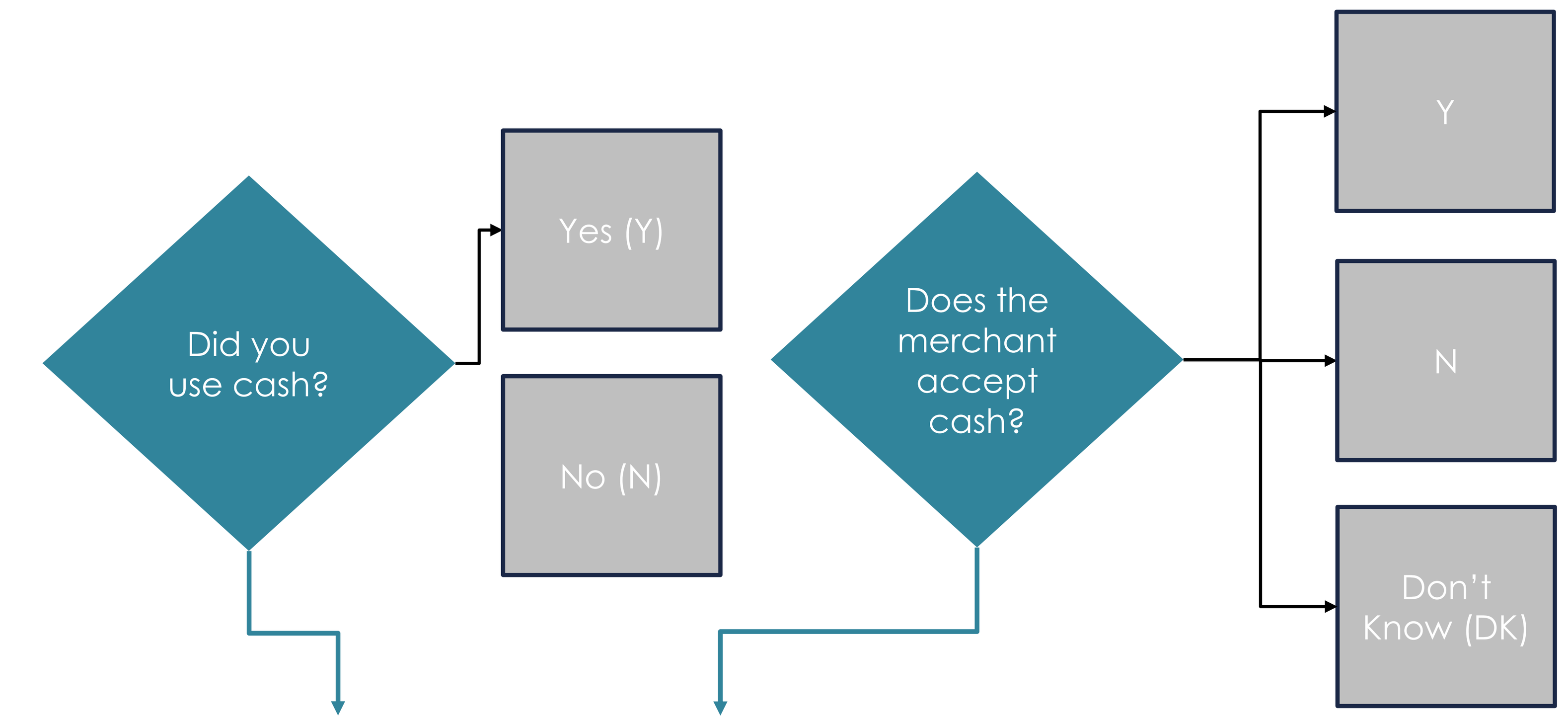
Table 1: Sample Composition of $\boldsymbol{S_C}$

Age	**18-34**	28.10
	35-54	32.00
	55+	39.90
Gender	**Male**	49.41
	Female	50.59
Income	**<$40K**	19.04
	$40K-$80K	28.38
	>$100K	52.58

Table 2: Sample Composition of $\hat{S}_M^3$

Size	**Small (0-5 employees)**	49.83
	Medium (6-49 employees)	50.17
Industry	**Retail trade (NAICS 44/45)**	52.33
	Food services and drinking places (NAICS 722)	40.83
	Other services (NAICS 81)	6.83
Locality	**Rural**	16.17
	Urban	83.83
Region	**British Columbia**	18.83
	Prairies	17.17
	Ontario	38.17
	Quebec	15.00
	Atlantic	10.83

Step 2: Constructing merchant MOP acceptance $\hat{y}_m$



$$f(usage, perceived\ acceptance) = f(u_m, p_m) = \hat{y}_m$$

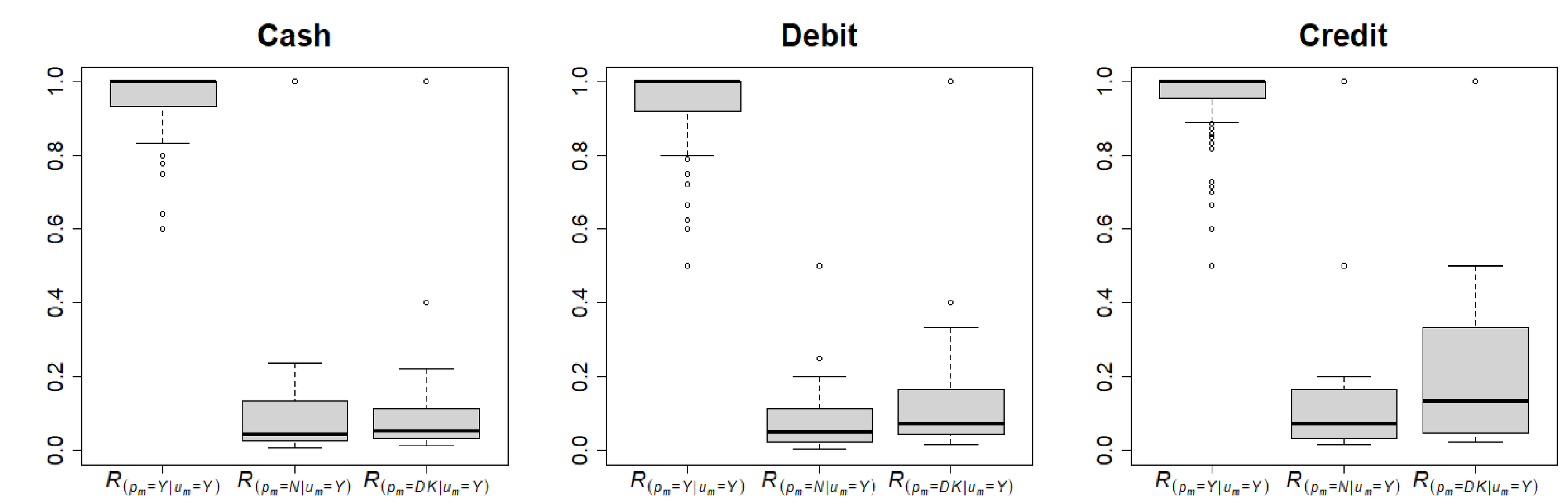
Rule 1: $\hat{y}_m$ determined by the most frequently occurring value across u_m and p_m.
Rule 2: $\hat{y}_m$ determined as the weighted average of all values across u_m and p_m
Rule 3: $\hat{y}_m$ is mapped to Y if usage occurs at least once, otherwise $\hat{y}_m = f_1(u_m, p_m)$

Incidence and Intensity of conflicts are low

Table 3: Incidence of conflict in $\hat{S}_M^3$

	Cash		Debit		Credit	
	1 visit	**> 1 visit**	**1 visit**	**> 1 visit**	**1 visit**	**> 1 visit**
No Conflict	80.73	16.64	80.73	16.26	80.73	17.01
Conflict						
Between		0.09				0.28
Within		0.28		0.56		0.09
Both		2.26		2.44		1.88

Chart 1: Intensity of Conflict



$R_{p_m=Y|u_m=Y} = \frac{\sum_{v=1}^{v+m} 1_{p_{m,v}=Y}}{\sum_{v=1}^{v+m} 1_{u_{m,v}=N}} = \frac{v_{+m}^{p=Y}}{v_{+m}-v_{+m}^{u=Y}}$, $R_{p_m=N|u_m=Y} = \frac{\sum_{v=1}^{v+m} 1_{p_{m,v}=N}}{\sum_{v=1}^{v+m} 1_{u_{m,v}=N}} = \frac{v_{+m}^{p=N}}{v_{+m}-v_{+m}^{u=Y}}$, $R_{p_m=DK|u_m=Y} = \frac{\sum_{v=1}^{v+m} 1_{p_{m,v}=DK}}{\sum_{v=1}^{v+m} 1_{u_{m,v}=N}} = \frac{v_{+m}^{p=DK}}{v_{+m}-v_{+m}^{u=Y}}$. This chart only includes merchants where $v_{+m} > 1$ and $v_{+m}^{u=Y} > 1$.

Step 3: Constructing merchant weights $\hat{w}_m^{cal}$

Table 4: Number of merchants visited

Days Complete	# Consumers	Min # Merchants	Max # Merchants	Avg # Merchants
1	53	1	4	1.47
2	163	1	8	2.09
3	872	1	11	2.79

Longer diary → fewer missing merchants. We treat missing merchants as unit-nonrespondents. Since in practice we are unable to observe these merchants $\hat{S}_M \backslash \hat{S}_M^3$ (our diary only lasts for the maximum three days), we employ nonresponse calibration outlined in Haziza and Lesage (2016), obtaining $\hat{w}_m^{cal}$