Recent development and analysis on Household Travel surveys, Quebec

Catherine Morency
Associate Professor
Head of the Mobilité Chair
Departement of civil, geological and mining Engineering
Polytechnique Montréal

Outline

- Context
  - As part of Mobilité research Chair mandate: Evaluation of the typical survey process (before Montreal OD 2013)
- Web-based surveys
  - 9 experiences: respondent behaviours
  - Web vs phone: key findings
- Chronical issue of proxy respondent
- Perspectives

Essential questions ??

Sample size required for each question?

Target population for each question?

Recent challenges vs usefulness of surveys → Mobilité research Chair – formulate recommendations regarding travel survey methods

CONTEXT

General issues (1)

- Declining response rates
- Difficulties vs recruiting interviewers
- Lack of resources (human + financial)
- Importance of survey data = always to be demonstrated (business case)
- Increasing availabilities of other sources (passive stream, technology) – what are the contributions of each source

General issues (2)

- Phone surveys:
  - Harder to reach participants + declining representativeness of typical sampling frame → heterogeneous issue among population segments
    - Cell phones # land line (HH → people)
    - Web-based phone service
    - Answering machines, etc.
- Comparability of surveys over time is compromised

Web as a potential survey tool (1)

- Declining attractivity of classical survey modes among certain segments (paper, phone)
- Increasing availability of internet services :
  - 2010: 79.3% of households have access to internet in Montreal (73% in Québec)
  - Highest penetration rate: **16-24 years old (98.3%)**
  - 2010, 13% of households only use cell phones (vs 8% in 2008). This proportion is **50% among the 18 -34 years old** (vs 34% en 2008)

Statistique Canada: Enquête canadienne sur l'utilisation d'Internet 2010 et Enquête sur le service téléphonique résidentiel

Main questions

- What questions are essential to the conduct of typical activities of the transport authorities (analysis and models)?
- What is the required sample size for each question, why do we ask this question, for which purpose, expected use?
  ... *No answer yet!*
- What is (should be) the target population for each question (should all questions be asked to all participants)? ... *No answer yet!*

Hierarchical analysis of the questionnaire

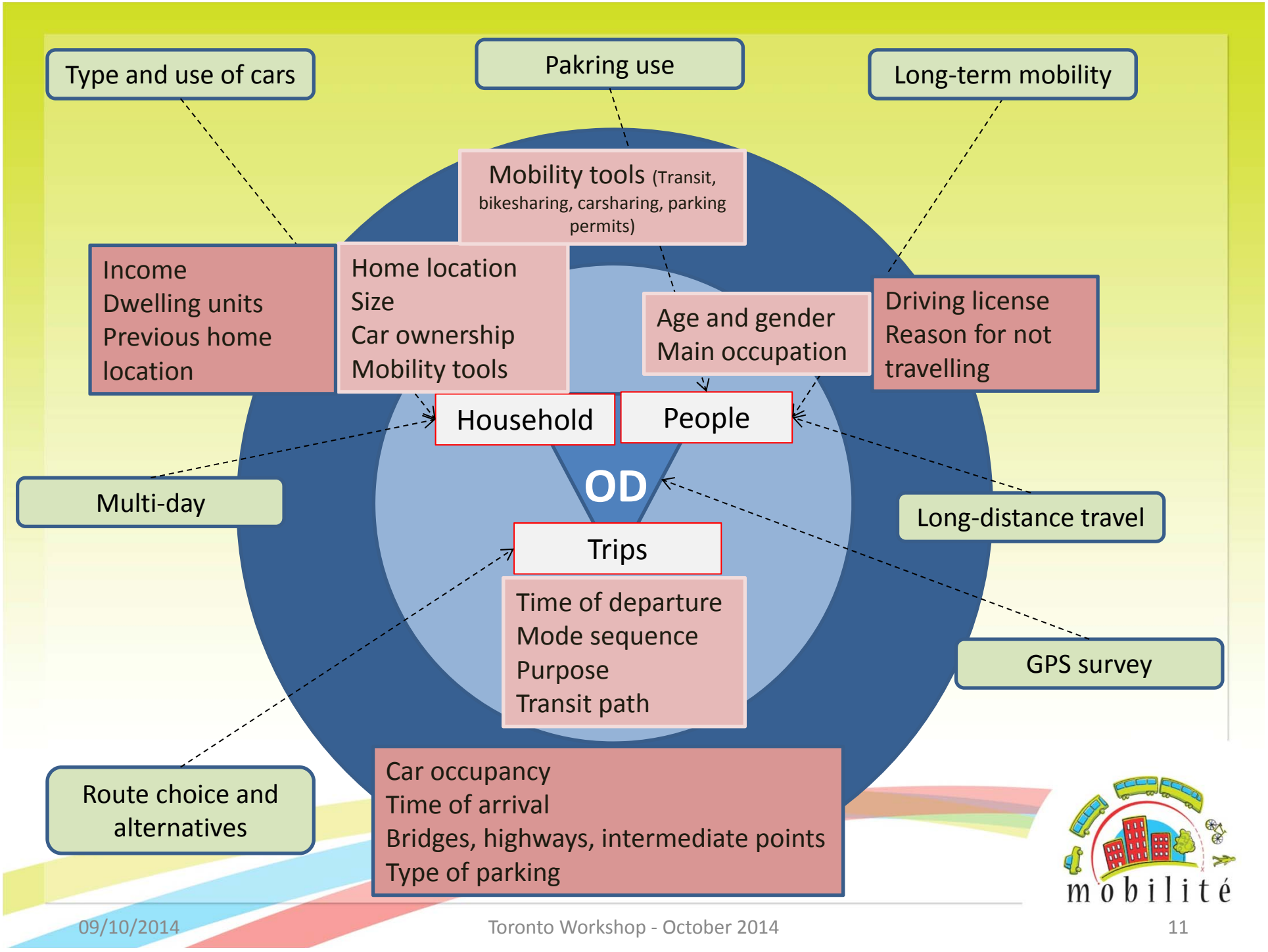
- Essential questions (official uses and publications)
  - Diffusion products
  - Models
- OF COURSE: all questions are relevant for a researcher..

Hierarchical analysis of the questionnaire

HOUSEHOLD	PEOPLE	TRIP
Essential for « key facts » (diffusion products)		
Home location	Age	Purpose
Number of people	Gender	Time of departure
Number of cars	All trips	Origin and destination
		Mode sequence
Essential for key facts but not in questionnaire (currently derived)		
		Occupancy ratio (number of people in the car)
Relevant question and used by partners		
	Main occupation	Transit line and boarding stations
	Driving license	
Questions that are still under examination by partners or others (validation still required, relevance to validate)		
Income	Main activity location	Bridges
	Reason for not conducting any trip	Highways
		Parking type at destination
		Transit pass

Some ideas discussed (1)

- Core-satellite concept
- Vs required sample size: Rotating (50% of households for instance) or optional questions (would you be willing to..)
- Vs proxy respondent bias: Should certain questions be asked solely to self-respondents?
- Questions with spatial filter (use of bikesharing for instance)
- Cross-section + panel (survey some households of the previous survey)


Type and use of cars
Pakring use
Long-term mobility
Mobility tools (Transit, bikesharing, carsharing, parking permits)
Income
Dwelling units
Previous home location
Home location
Size
Car ownership
Mobility tools
Age and gender
Main occupation
Driving license
Reason for not travelling
Household
People
OD
Trips
Time of departure
Mode sequence
Purpose
Transit path
Multi-day
Long-distance travel
GPS survey
Route choice and alternatives
Car occupancy
Time of arrival
Bridges, highways, intermediate points
Type of parking

Some ideas discussed (2)

- Additional attributes of the sample to monitor throughout the survey + include in weighing process: main occupation (workers)
- Insure storing of contextual and reference variables (all explanatory variables that will be required afterward)
- Insure updating of previous surveys vs:
  - Field definition and dictionary
  - Weighing process
- Test a web-based version!

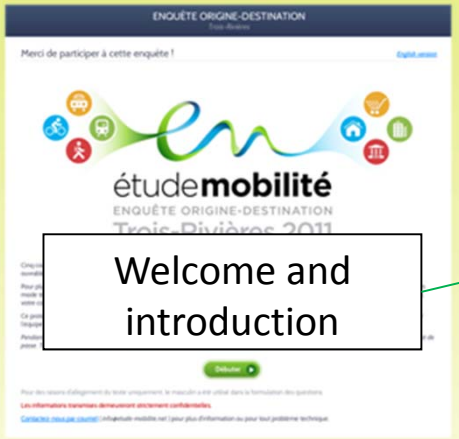
Patterning Respondent Behaviours from 9 Web Travel Surveys (Pierre-Léo Bourbonnais, Catherine Morency, to be presented at ISCTSC 2014 - Australia)

Bourbonnais, P.-L., Morency, C., 2013. Web-Based Travel Survey: A Demo, in: Munizaga, M., Carrasco, J.A., Zmud, J., Lee-Gosselin, M. (Eds.), Transport Survey Methods. 9th International Conference on Transport Survey Methods 2011, Bingley

Recent challenges vs usefulness of surveys

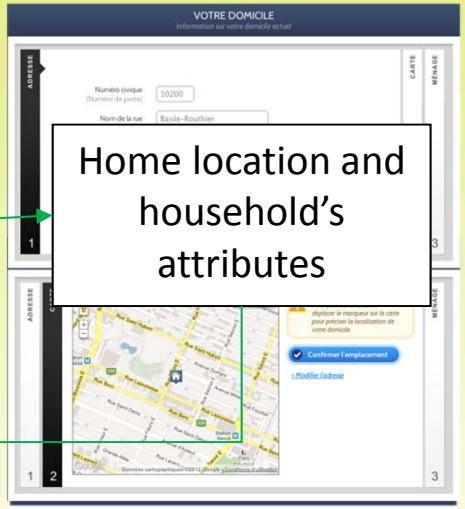
WEB-BASED TOOL

Web-based household / people travel survey tool

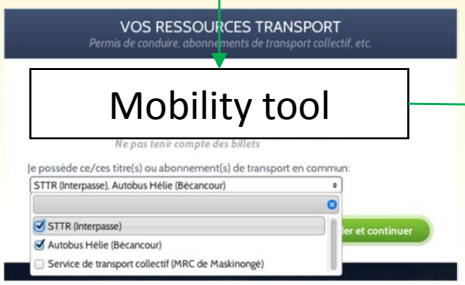


Welcome and introduction

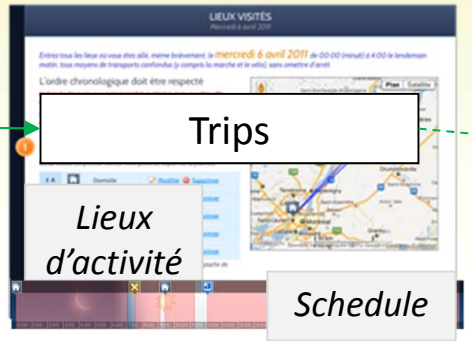
Responden t profile



Home location and household's attributes



Mobility tool



Trips

Lieux d'activité

Schedule

Opinion

Transportation mode

mobilité

10 web-surveys conducted to date (household and people)

- Fall 2010: first web survey among Polytechnique Community – development of a tool inspired by the typical large-scale travel surveys in Quebec
- **Spring 2011: experimentation of a web PERSON survey as part of the Trois-Rivieres regional travel survey**
- Fall 2011: second at Polytechnique and first survey among the University of Montréal community
- **Fall Automne 2011 experimentation of a web HOUSEHOLD survey as part of the Trois-Rivieres regional travel survey**
- Fall 2012: experimentation of a web HOUSEHOLD survey to validate opportunity of adding this mode during the 2013 regional survey in Montreal
- **Fall 2012 + Spring 2013: web survey among university and college students of the Sherbrooke region (as part of regional travel survey)**
- Fall 2013: Montreal (regional + Bixi community + Communauto community)

Interview duration (household questionnaire)

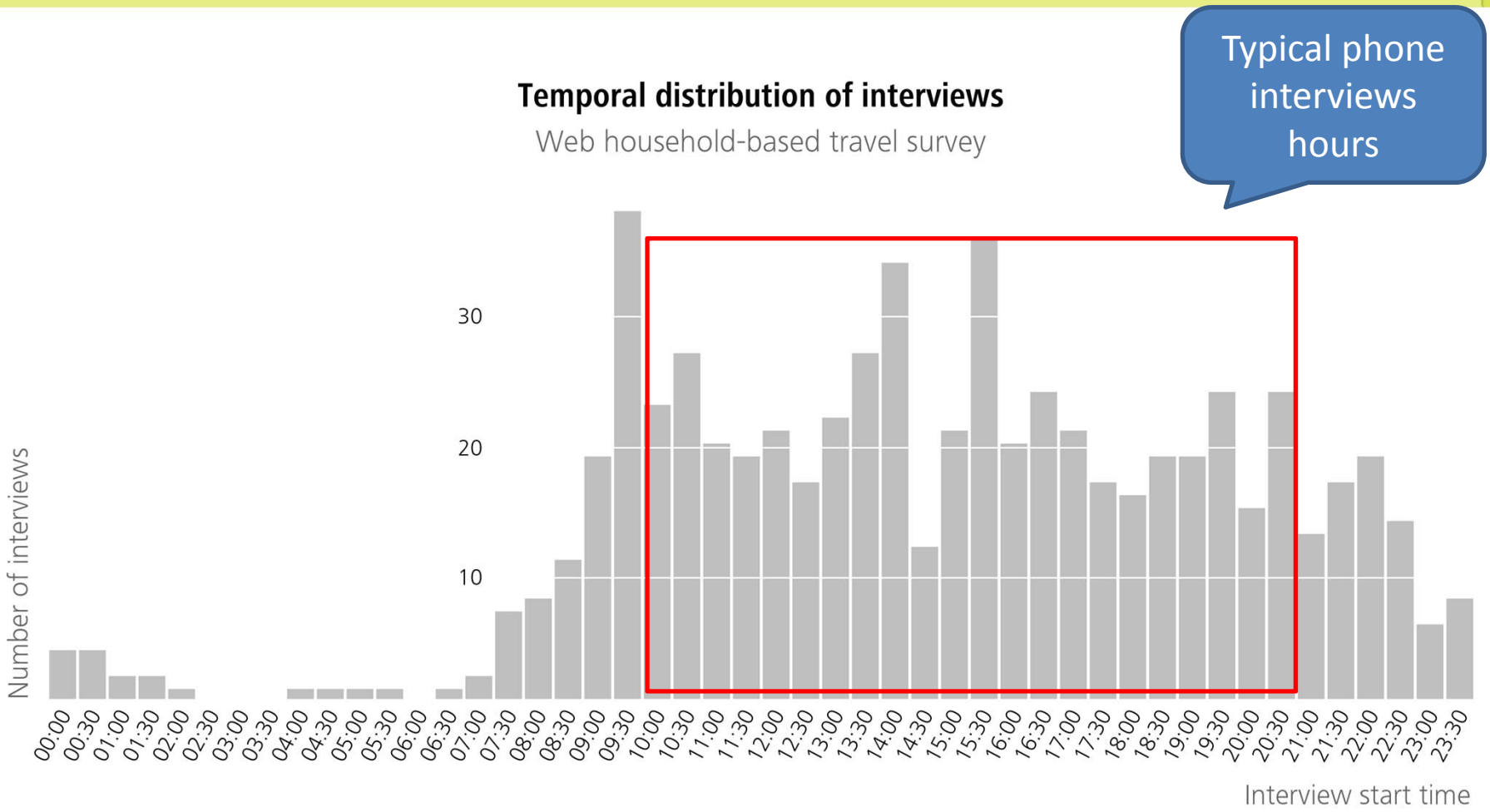
Survey	Mode	# Start. inter.	# Comp. inter.	Mean (min)	SD (min)	Q1 25%	Med.	Q3 75%
QC11	Letters	138	98	**23.7**	**13.4**	14.7	20.1	29.0
	Univ. residence	60	41	**11.4**	**7.2**	5.6	9.7	16.9
	Social net.	36	26	**12.7**	**6.4**	7.6	12.2	16.4
MTL13	Phone ref.	70	46	**23.0**	**13.1**	12.6	22.0	31.9
	website	599	443	**21.4**	**12.5**	12.8	19.0	26.6
Total (HH)	ALL	903	654	**20.9**	**12.6**	12.1	18.1	26.3

Interview duration (people questionnaire)

Survey	# Start. inter.	# Comp. inter.	Mean (min)	SD (min)	Q1 25%	Med.	Q3 75%
PY10	1,972	1,530 1,458 < 40 min.	**13.2**	**6.5**	8.7	11.6	16.1
PY11	1,929	1,673	**12.6**	**5.9**	8.5	11.4	15.4
UM11	7,948	6,501	**14.1**	**6.9**	9.4	12.8	17.4
TR11 cell	109	81	**11.9**	**7.5**	7.3	10.0	14.8
TR11 mailing	54	35	**13.0**	**7.2**	7.9	11.8	17.8
SH13 univer.	2,399	1,838	**13.0**	**5.5**	9.1	12.1	15.7
SH13 college	683	467	**13.7**	**6.5**	9.4	12.3	16.7
CM13	3,143	2,527	**13.8**	**7.1**	8.9	12.3	17.1
BX11	6,191	4,423	**13.6**	**7.3**	8.7	12.1	16.9
Total person-based*	24,428	19.075	**13.6**	**6.8**	9.0	**12.2**	16.7

Interview durations longer than 40 minutes are not included in the descriptive statistics for PY10 to limit the effect of outliers. In fact, the PY10 sample is not included in the interview duration models because timestamps' paradata for this survey was not precise enough to obtain genuine validated interview durations.

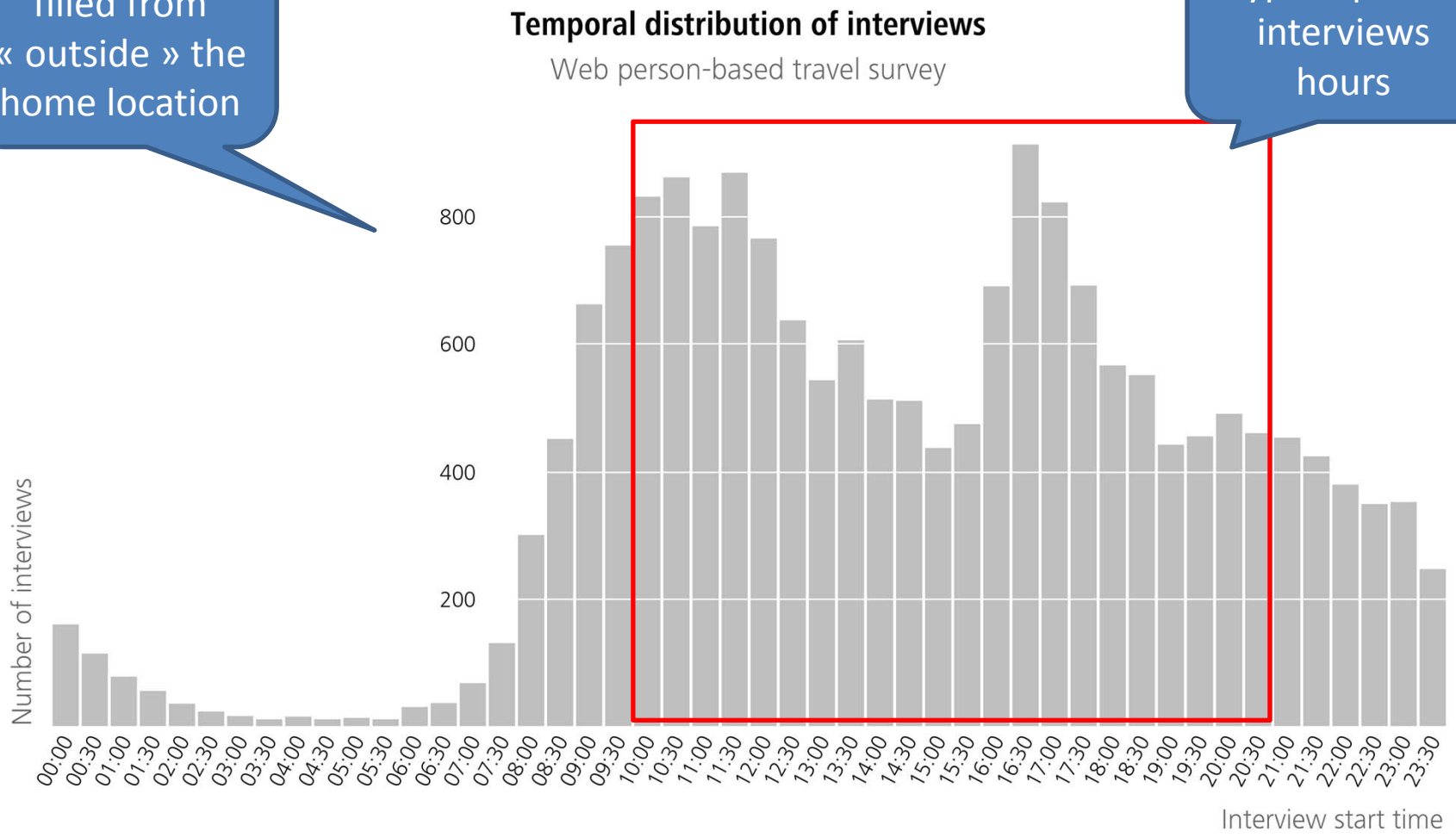
Temporal distribution of interviews (HH surveys)



Temporal distribution of interviews (people surveys)

Surveys can be filled from « outside » the home location

Typical phone interviews hours



() Impact of phone interview periods on sample composition

OD 2008	% full time workers (on total population)
Monday	42.53%
Tuesday	41.37%
Wednesday	40.72%
Thursday	41.17%
Friday	**45.30%**
Total	42.21%

Saturday calls = higher probability of workers being at home location
Impacts on travel behaviours of « Fridays »
Correction through the weighing process?
Would require exogenous data on workers

**** small samples ****

2011 – Trois-Riviere, Quebec

2013 - Sherbrooke

2012 PILOTE - Montreal regional household survey -

WEB VS PHONE – COMPARISON OF BEHAVIOURS

Main outcomes: Trois-Rivières + Québec (in regional HH survey settings)

Small samples

Trois-Rivières

- **PERSON** questionnaire
- 3 recruitment modes: cell phone numbers, dedicated sampling lot (mailings), dormitory of Cégep de Trois-Rivières
- **<u>146 completed questionnaires</u>**
- **86% of questionnaires** started have been completed

Québec

- **HOUSEHOLD** questionnaire
- 3 recruitement modes dedicated sampling lot (mailings), U Laval dormitory, social networks
- **<u>139 completed questionnaires</u>**
- **58% of questionnaires** started have been completed

Synthesis: comparative analysis behaviors

Phone vs Web (+ means more important in phone)	Trois-Rivières : respondent		Québec : all people that were reached	
	Diff.	Commentaires	Diff.	Commentaires
% non-mobiles	+	Only women	+	Men and women
Trip rate	-	Men and women	-	Men and women
Trip rate - WORK	-	↑↑ women 45-54 y.o.	-	↑↑ men 45-54 y.o.
Trip rate – STUDY	-	Only men	-	Men and women of 15-24 y.o.
Trip rate - OTHERS	-	Particularly women	-	Men and women
Daily km travelled	-	More important for some men segments	-	↑↑men 35-44 y.o.
Trips between 6h-8h – WORK	+		+	
Trips between 6h-8h – STUDY	-		+	

Synthesis: comparative analysis behaviors

Phone / Web	Trois-Rivières : respondent		Québec : all people that were reached	
	Diff.	Commentaires	Diff.	Commentaires
Trip length – Car driver	+	Small differences men and women	=	
Trip length – AP	-	Men	-	Small differences men and women
Trip length– WORK	+	Men and women	=	Slightly more important for women
Trip length – SHOPPING	+		+	Small differences for men and women
Car modal share	-		+	Active modes and transit higher in web

It seems people declare more trips in web-based surveys, namely those related to non-mandatory activities

May be related to the fact that the questionnaire

Main outcomes from these 2 web surveys (in regional OD survey settings)

- People-based questionnaire: interesting completion rates (85% of people who start the questionnaire will complete it); Lower for household questionnaire at app. 60%
- Some people participate in the survey at periods outside of typical calling hours = flexibility
- Samples reach are complementary in many areas (higher web sampling rates when lower phone sampling rates)
- It seems people declare more trips in web-based surveys, namely those related to non-mandatory activities;
  - May be related to the fact that the questionnaire asks to list all places visited « yesterday » and then the way this places were linked by trips

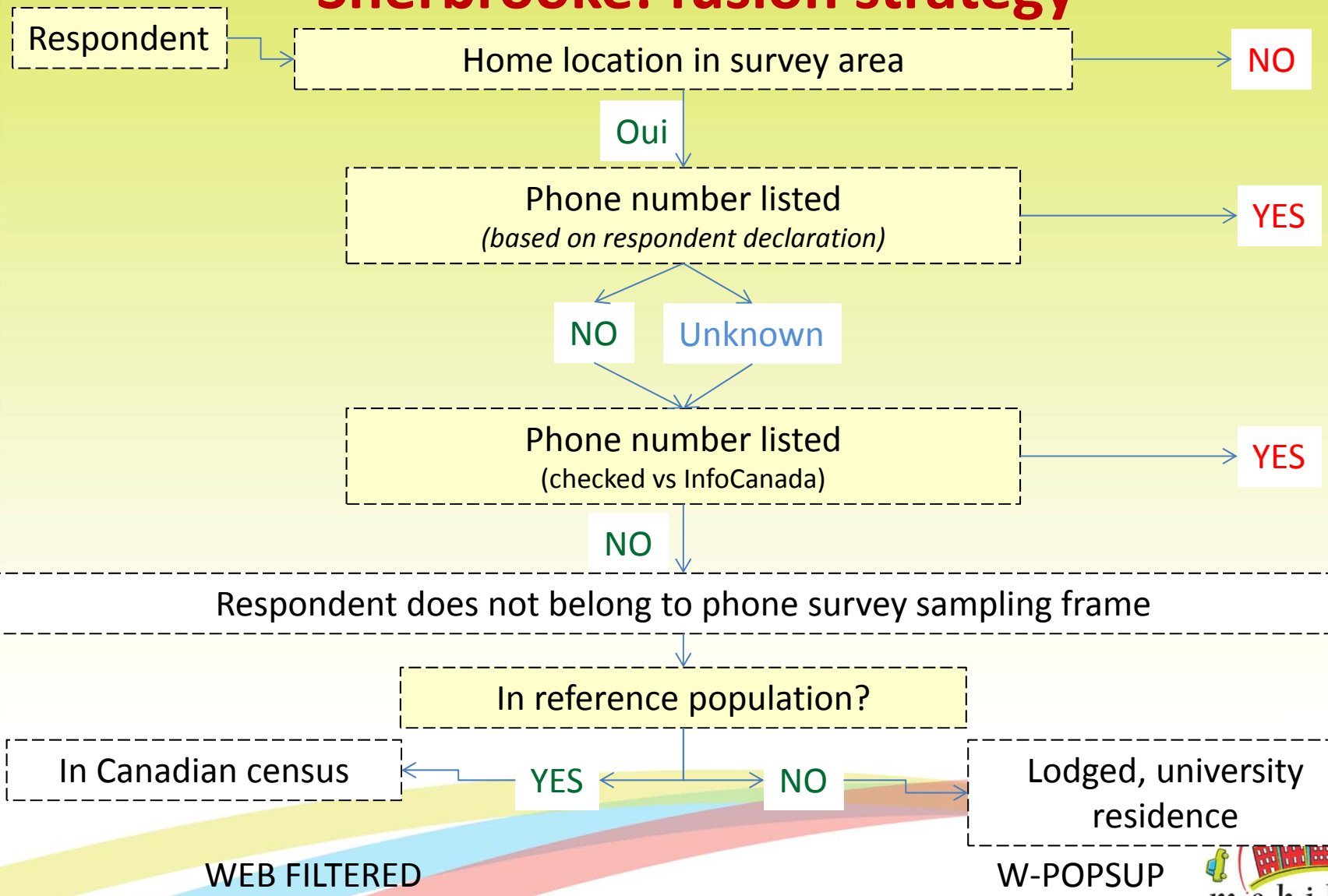
Sherbrooke

- **Objective**: web survey among educational institutions to reach young people typically not included in phone survey and combined web + phone survey results

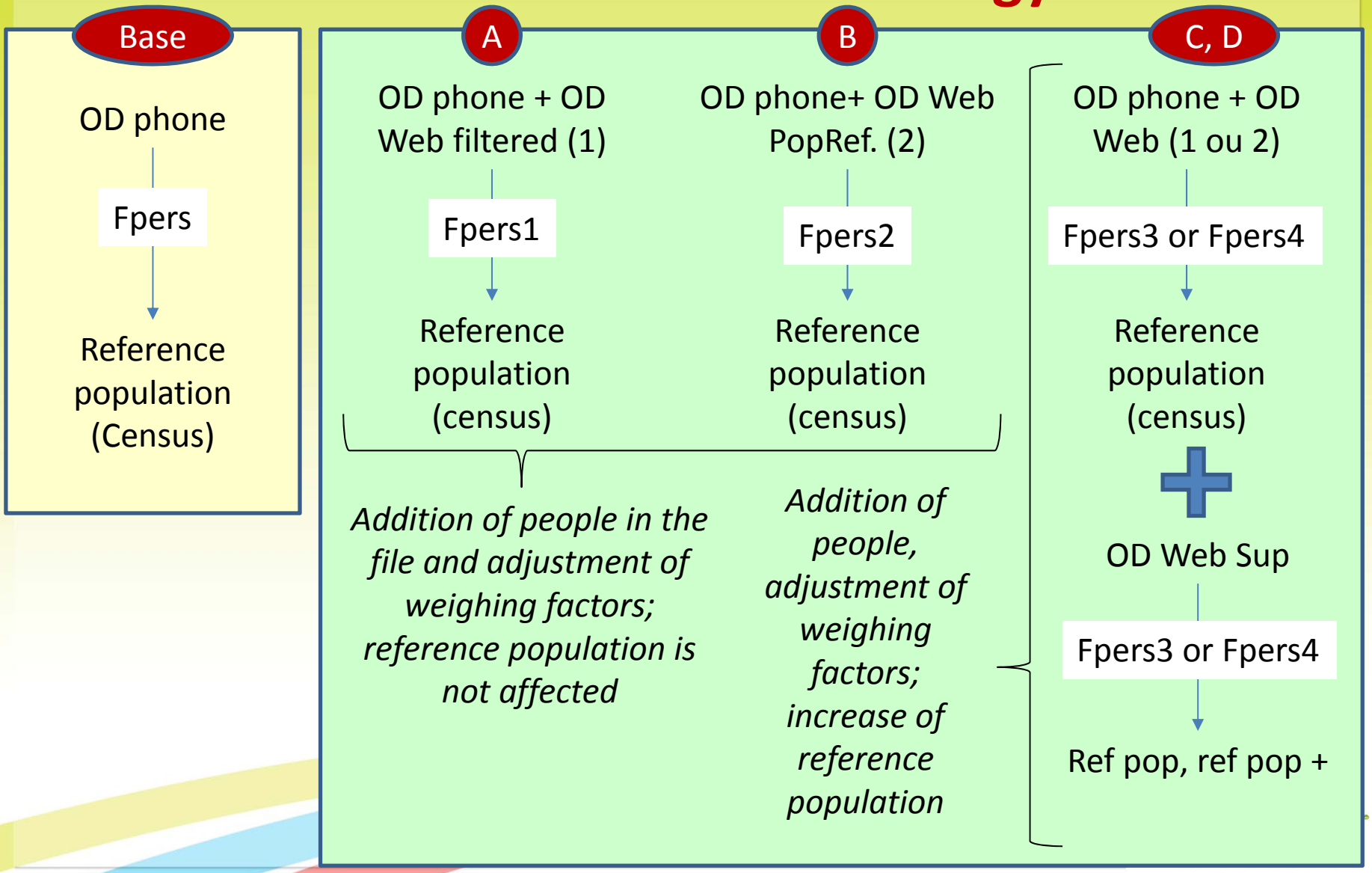
WEB SURVEYS

Source	Ref Pop	Completed	Response rate	OD area respondents
Séminaire de Sherbrooke	500	22	4,4%	20
UdeS	16187	1838	11,4%	1708
Champlain	1067	120	11.3%	109
Cegep	5753	337	5,8%	315
TOTAL	23507	2317	9,9%	2152

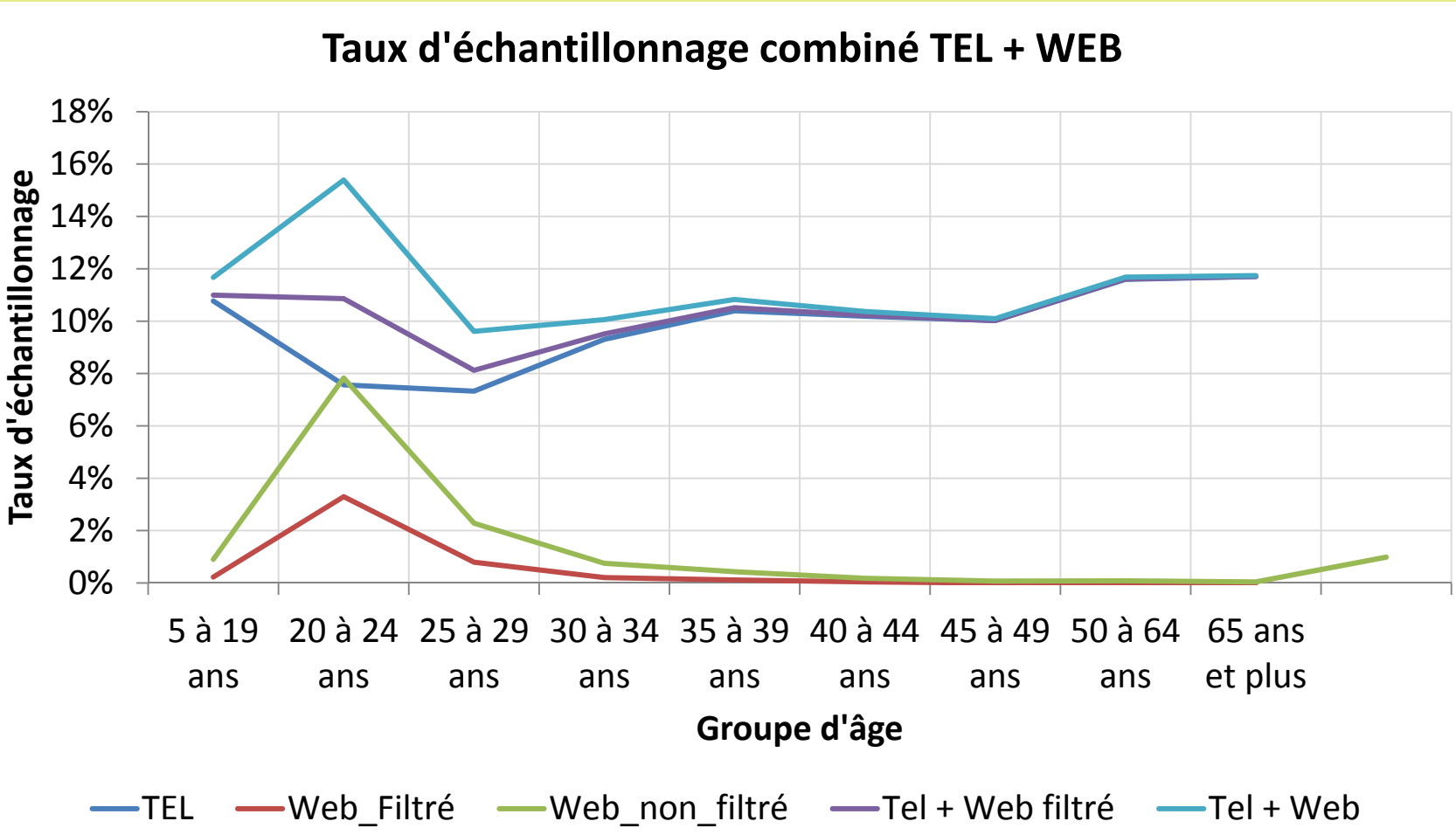
Sherbrooke: fusion strategy



Sherbrooke: Fusion strategy



Sherbrooke: impact of fusion on sampling rates



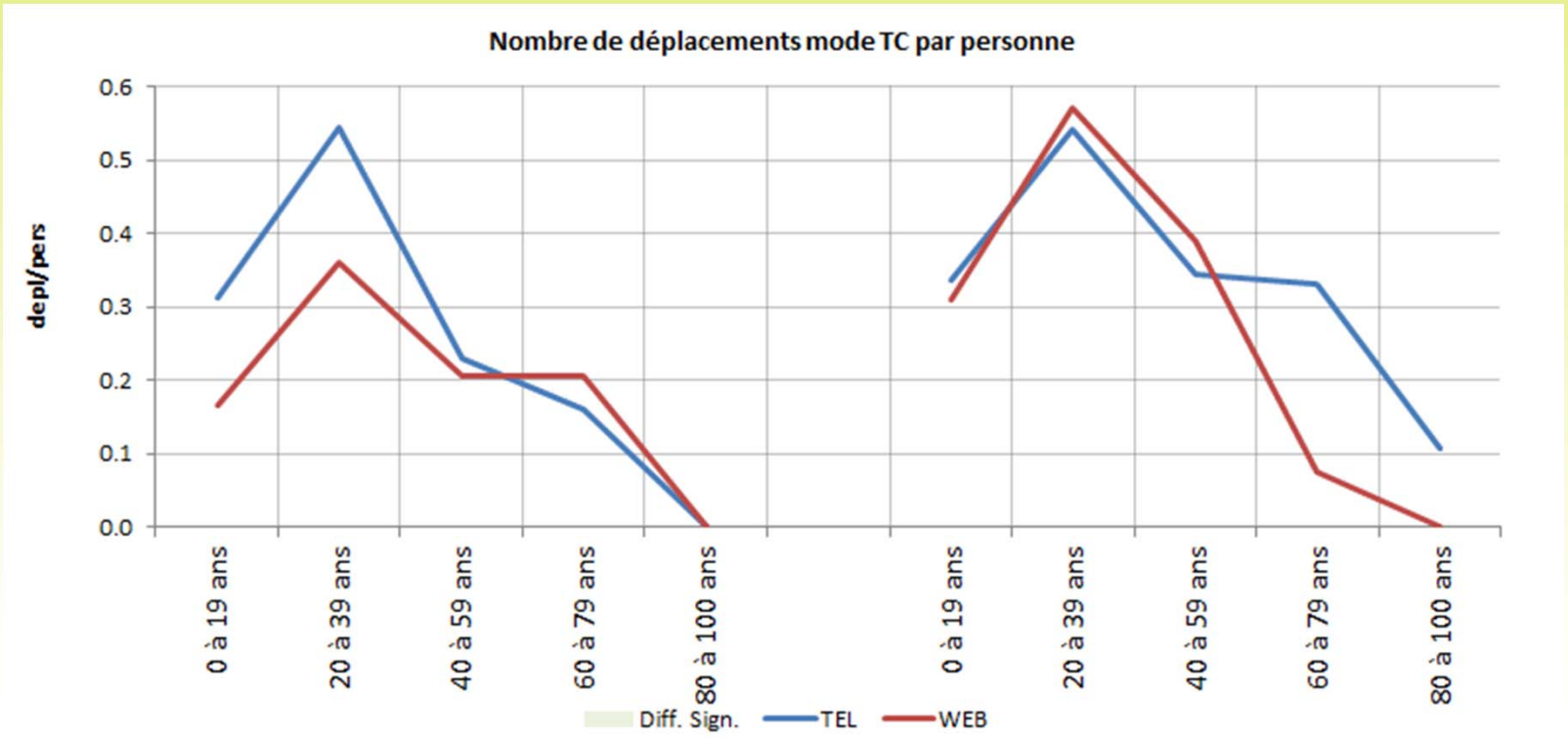
Montreal pilote survey

- Fall 2012
- 2000 letters sent (0.43% of reference population)
- 135 completed interviews
- 24.4 % of the completed households don’t have a landline
- Comparison with « continuous survey sample » of the same period

Montreal pilote – Comparison of sample behaviours 30 indicators + statistical test… small sample

Trip rate

Transit trip rate



Montreal pilote – Comparison of sample behaviours

- Again: trip rate is higher
- Similar modal shares;
- Lower proportion of « return home trips » in web survey (related to higher trip rates and more non-mandatory trips);
- More kilometers travelled during the day in web = f(more trips;
- Higher proportion of simple trip chains in phone surveys and higher proportion of mobile people doing only on simple trip chain per day.

RESPONDENT BIAS

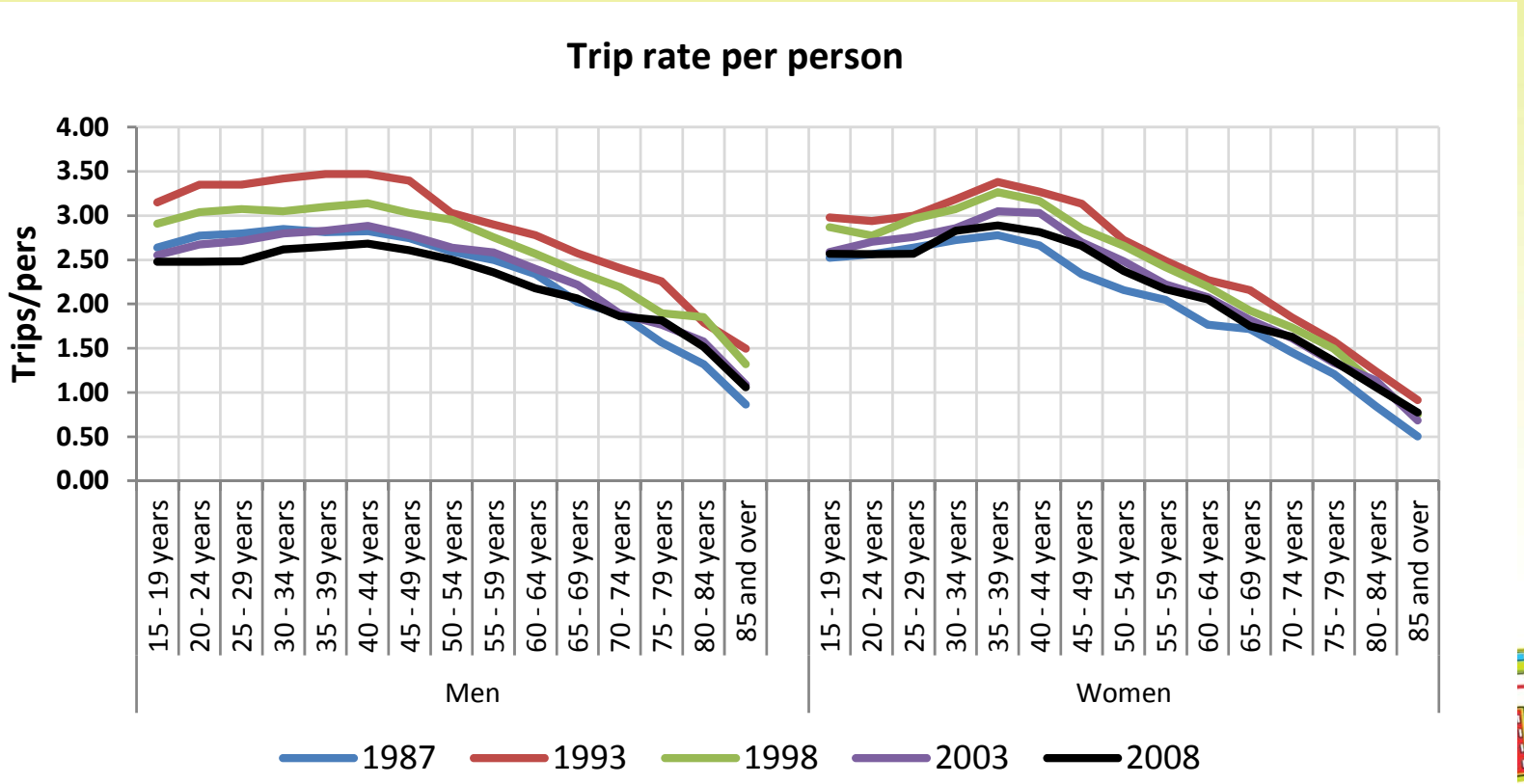
General context: self-respondent vs indirect participant

Survey	Self-respondents	Indirect participant	% Direct respondents
1987	53 177	84 188	38.7%
1993	61 988	98 526	38.6%
1998	65 227	98 848	39.8%
2003	58 000	81 527	41.6%
2008	66 124	90 596	42.2%

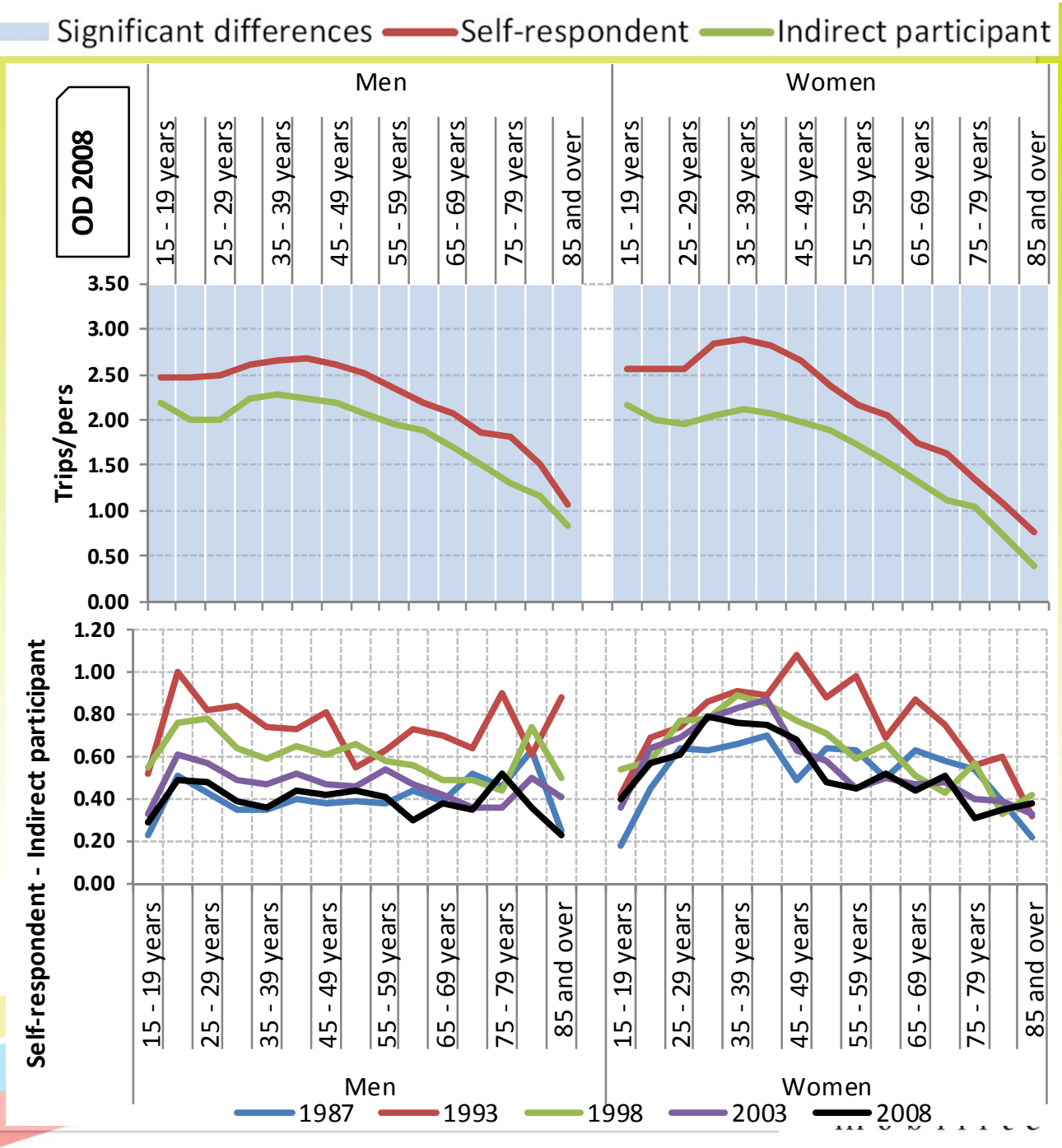
Context: trends

Influencing factors for proxy respondent bias

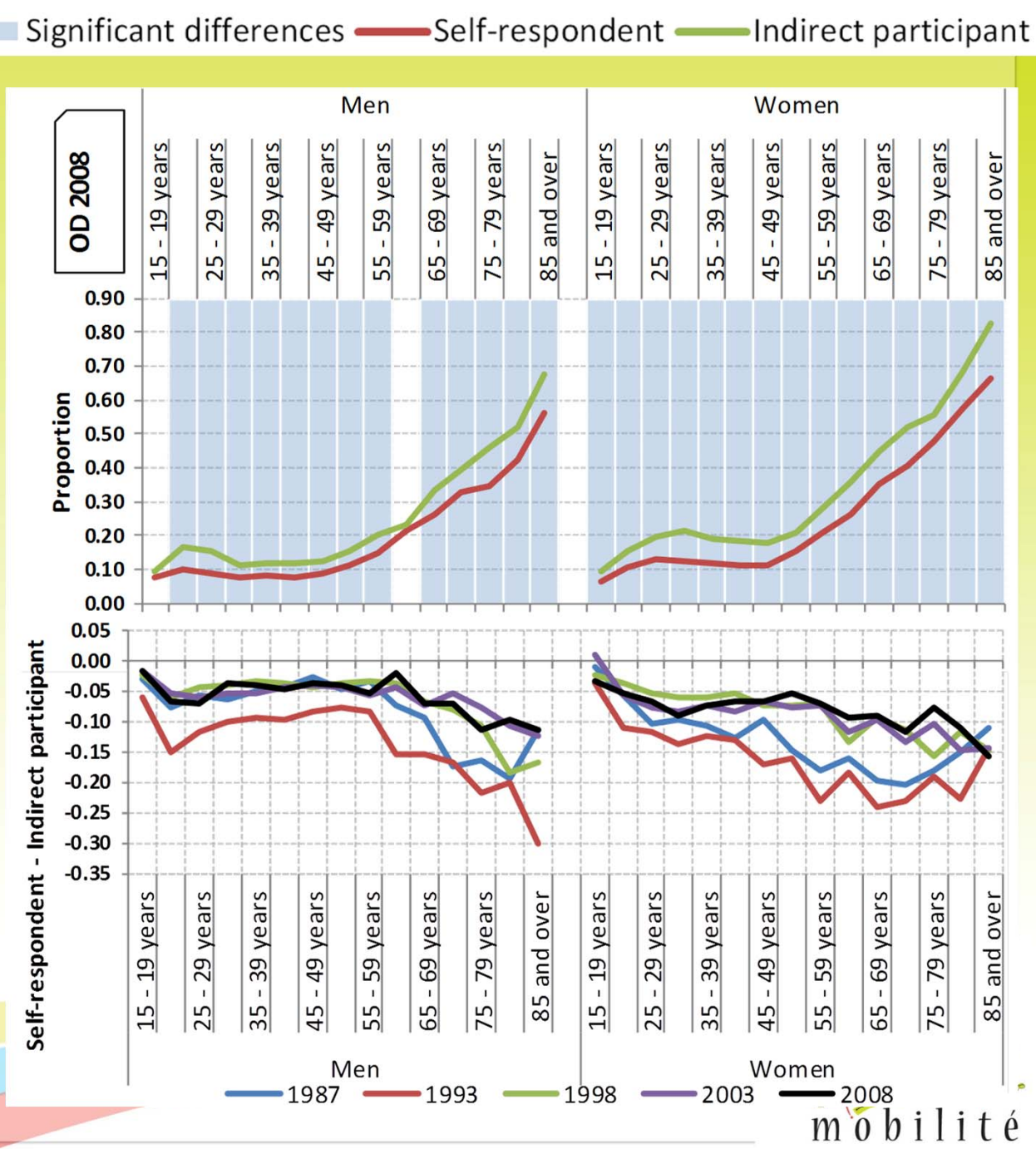
- Decline in household size since 1987 (from 2.56 in 1987 to 2.38 in 2008)
- Decreasing trips rates since 1993



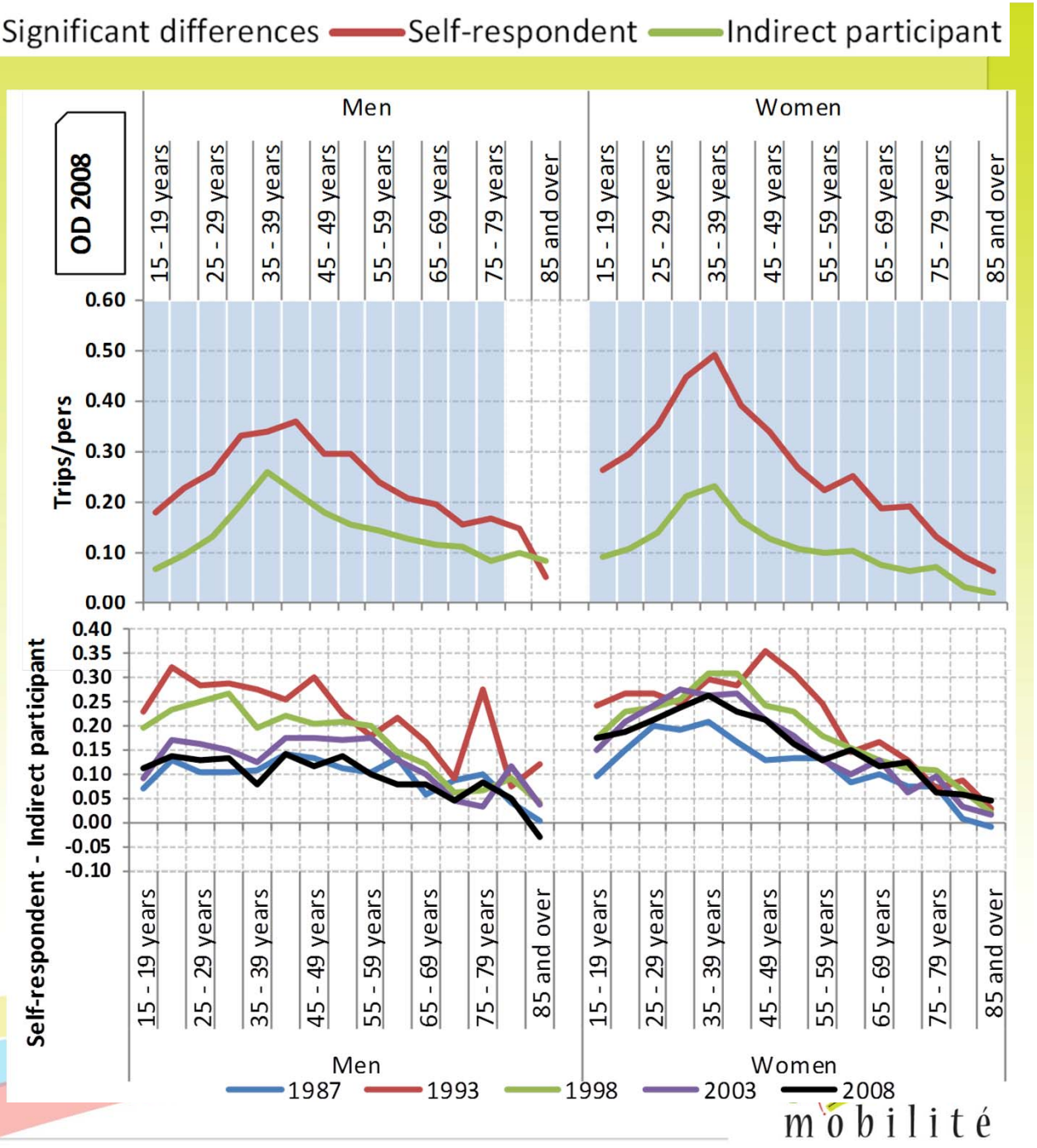
Trips per person per day



% of non-mobile people



Non-home based trips per person per day



What if??? Indirect participants actually behave like self-respondent?

Indicators	Additional	%
Trips per person per day		
All trips	756 659	10.7
WORK trips	22 133	1.4
SCHOOL trips	-16 886	-4.7
LEISURE trips	154 990	30.6
SHOPPING trips	167 496	28.0
Car-driver	450 841	10.8
Car-passenger	-107 042	-15.8
Public transit	236 178	20.5
Walking and Cycling	222 886	30.4
AM peak	6 203	0.3
Non-home-based	229 541	33.0
Non-mobiles	-89 597	-14.0

Controlling for age * gender * area
*** Impact of main occupation!

Understanding differences?

- Decomposition statistical method
- The difference between the two samples can be explained by two phenomena:
  - The composition of the population is not the same in both samples (Sample effect)
    - Example : higher proportion of workers among the indirect participants for instance.
  - The trip behaviors of the two samples are not the same (Coefficient effect)
    - Respondent bias or real differences in behaviors

Statistical decomposition method

Coefficient effect

Interaction effect : objective = close to 0.

	Taux		Différence		Composante					
Indicateurs	Non-Rep	Répondant	Diff	Diff sign	Échantillon	Coefficient	Interaction	Échan_Sign	Coeff_Sign	Inter_Sign
Nb_depl	2.43	2.88	-0.45	***	0.00	-0.42	-0.03		***	

Sample composition explains 0% of the difference

The coefficient effect explains 94% of the difference

Difference is significant

Decomposition method

- Variables included in the model :
  - Region of residence
  - Cohort and gender
  - Being a full-time worker
  - Household type
- Significant variables :
  - Home = Island of Montreal
  - Full-time worker (men and women)
  - Household of 2, 3 or 4 people

Statistical decomposition method

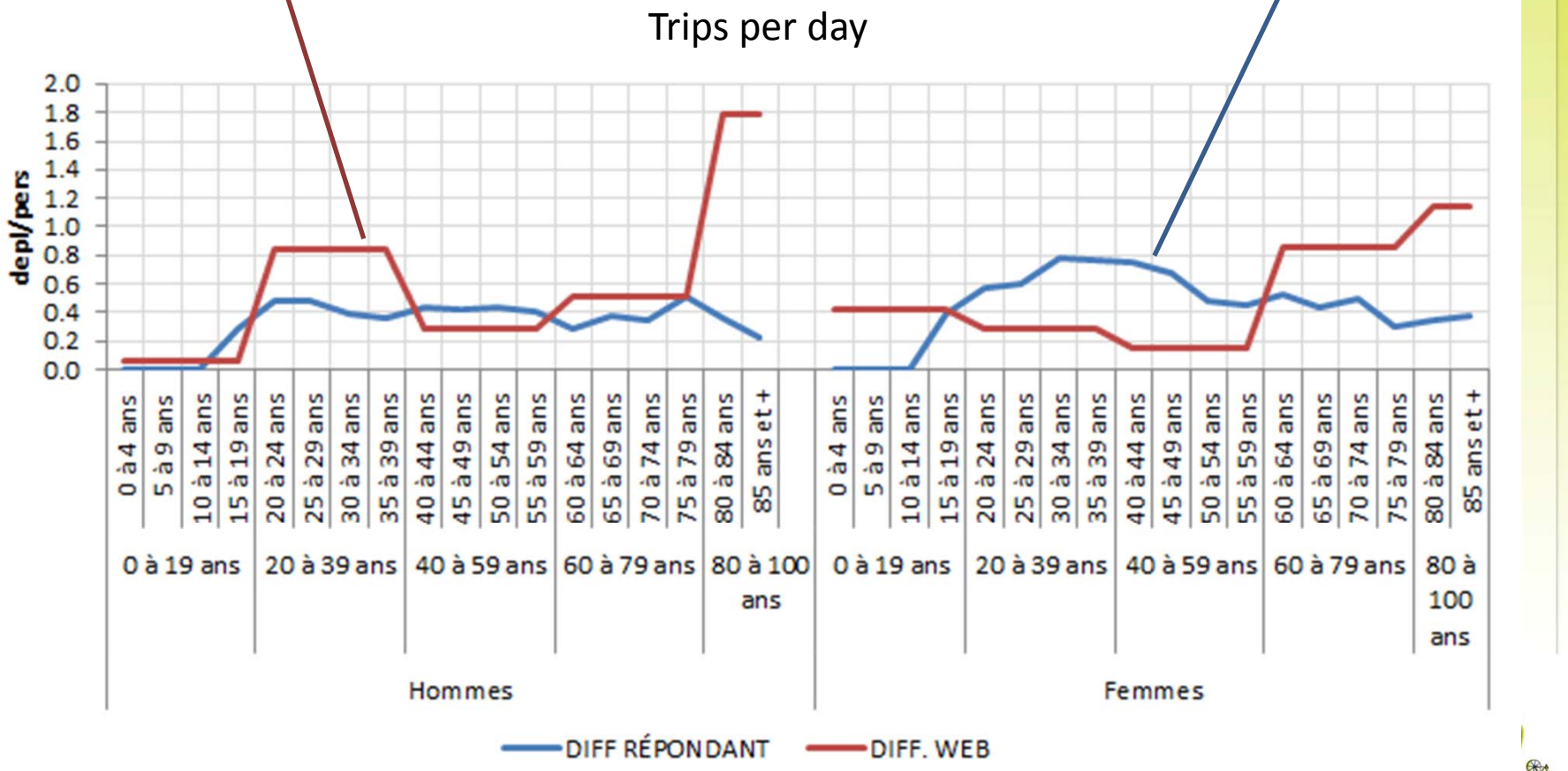
		Difference	Significant	% of difference explained: Composition	% of difference explained: Coefficient	% of difference explained: Interaction	Statistically significant: Composition	Statistically significant: Coefficient	Statistically significant: Interaction
All people	Trips per person	-0.37	***	-60.2%	150.2%	10.1%	***	***	***
	% non-mobile	0.01	***	-437.4%	587.0%	-49.6%	***	***	***
People who made at least one trip	Trips per person	-0.42	***	-11.6%	101.9%	9.7%	***	***	***
	Working trips per person	0.02	***	122.6%	-2.4%	-20.3%	***		
	School trips per person	0.12	***	78.2%	21.4%	0.4%	***	***	
	Leisure trips per person	-0.10	***	29.2%	87.0%	-16.3%	***	***	***
	Shopping trips per person	-0.18	***	45.7%	62.2%	-7.9%	***	***	***
	Other trips per person	-0.14	***	-13.7%	79.0%	34.7%	***	***	***
	Car-driver trips per person	-0.37	***	-24.9%	125.2%	-0.3%	***	***	
	Car-passenger trips per person	0.08	***	-7.7%	179.4%	-71.6%	**	***	***
	Public transit trips	-0.02	***	-108.5%	-25.7%	234.2%	***		***
	Walking trips	-0.15	***	47.4%	71.7%	-19.1%	***	***	***
	Am peak trips per person	0.05	***	174.0%	-11.7%	-62.3%	***		***
	Non-home-based trips per person	-0.17	***	-1.7%	97.6%	4.1%		***	*
	Distance per person trips per persor	0.82	***	97.1%	20.5%	-17.6%	***	**	**
	Activity duration per person (min)	132.60	***	48.6%	43.2%	8.1%	***	***	***

Confidence interval :*** 99%, ** 95%, * 90%

How does the difference between web and phone compares to the difference between self-respondent and indirect participant??

Absolute difference between web and phone

Absolute difference between self-respondent and indirect participants (phone)



Difference between web and phone is smaller if only respondents are compared

		Taux		Différence		
	Indicateurs	Web	Téléphone	Diff	% Diff/web	Diff sign
Répondant	Nb_depl	2.85	2.83	0.02	0.8%	
	NB_NMOB	0.16	0.13	0.03	18.7%	
	Nb_depl_mob	3.40	3.26	0.14	4.2%	
	Nb_travail_mob	0.72	0.62	0.10	13.8%	
	Nb_étude_mob	0.16	0.07	***0.10***	58.4%	**
	Nb_loisir_mob	0.31	0.30	0.01	3.2%	
	Nb_magasinage_mob	0.34	0.47	***-0.13***	-39.1%	*
	Nb_autre_mob	0.57	0.47	0.10	17.3%	
	Nb_AC_mob	1.91	1.74	0.18	9.3%	
	Nb_AP_mob	0.31	0.21	0.10	31.0%	
	Nb_TC_mob	0.53	0.47	0.06	11.5%	
	Nb_MAR_mob	0.38	0.53	-0.15	-41.1%	
	Nb_PAM_mob	0.74	0.70	0.04	4.8%	
	Nb_EXT_mob	0.79	0.59	0.19	24.6%	

Smaller differences and less significant

Respondents

DISCUSSION

Survey questionnaire

- What questions are essential + how to value available questions
- What is the sample size required for the expected analysis? Is this an opportunity to gather additional information (for smaller samples)
- What level of confidence should be given to proxy response and what questions (behaviors) are more sensitive to proxy bias

Web vs Phone

- Web is relevant for some population segments – need to make sure there is an appropriate survey mode for each segment
- Does not solve the sampling issue… and a very important one; landline list are not sufficient anymore
- Household web surveys are long and have higher dropoffs– Alternative: combine people and household surveys – probably provide multiple weights and recommendations on when to use which sample

Bias

- Proxy responses are an issue
- In web – more direct respondent hence differences are lower
- Issue related to sample composition – opportunity to monitor sample by demography, spatial location and main occupation?