Mixed Mode Surveys

Professor Edith D. de Leeuw

PPSM Short Course Series

Bremen 8 May 2009

Introduction:

Instructor

Participants

Topic

Instructor

- **Edith Desiree de Leeuw** is a professor of survey methodology at the Department of Methodology & Statistics of the University of Utrecht. She is a fellow of the Netherlands Interuniversities Joint Institute for Psychometrics and Sociometrics (IOPS), associate editor of *Journal of Official Statistics (JOS)* and member of the editorial board of *Sociological Methods and Research, Field Methods*, and *MDA*. She has edited books on methodology and statistics, including the recently published International Handbook of Survey Methodology, and has more than 20 years experience in survey research. See also http://www.xs4all.nl/~edithl/

Acknowledgements

- Bill Blyth
- Don Dillman
- Martin Frenkel
- Joop Hox
- Michael Link
- Lars Lyberg

Modes of Data Collection

- Respondent Participation Required
  - Interviewer-administered Modes
    - Face-to-face interviews (PAPI or CAPI)
    - Telephone Interviews (PAPI or CATI)
  - Self-administered
    - Postal or mail survey
    - Self-administered with interviewer present
      - PAPI or CASI
    - Web or Internet Survey
    - IVR (Interactive Voice Response)

Modes of Data Collection 2

- No respondent participation involved
  - Enumerator needed
    - Direct Observation of Behaviour
    - Counting (e.g., traffic), etc
  - No enumerator needed
    - Administrative Records
    - Scanning, data mining
    - Electronic Data Exchange (EDI)
    - TV-usage, ‘people meters’

Participants' Experience

- Hands-on Experience?
  - Face-to-face interviews (PAPI or CAPI)
  - Telephone Interviews (PAPI or CATI)
  - Postal or mail survey
  - Web or Internet Survey
  - IVR (Interactive Voice Response)
  - Mixed Mode Surveys

Mixed Mode Survey

- Combine two or more communication modes
  - Contact
  - Data collection
- Contact
  - Screening or convincing
    - ESS allows for appointments made by telephone, Actual interview face-to-face
- Data Collection
  - Nonresponse follow-up by different method
  - SAQ-module during face-to-face interview
  - Web + telephone (PPSM)

A New Trend

"Mixed mode surveys, that is, surveys that combine the use of telephone, mail, and/or face-to-face interview procedures to collect data for a single survey project are occurring with increasing frequency. A second, or in some cases even a third, method to collect data for a single survey is being used throughout the world…. Indeed, mixed mode is becoming one of the survey buzz words of the late 20th century"

Dillman & Tarnai, 1988

- Important issues in mixed mode identified by Dillman & Tarnai are a.o.
  - Data comparability
  - Questionnaire construction and pretesting

Mixed-Mode the Norm

"In general, data collection systems do not consist of one mode only, since mixed-mode surveys are the norm these days."

Biemer & Lyberg, 2003

"An emerging new breed of survey software is starting to make this

[combine CATI/CAWI]

possible"

Macer, 2004

"Mixed-Mode: The only fitness regime."

Blyth, 2008

Why Mix Modes?

- Increase in Online Surveys
  - Coverage
  - Special groups
- Response/nonresponse problems
  - Effort to increase response
  - Investigating bias
- Increase in International Surveys
  - Different tradition in countries
  - Different coverage

Mixed Mode Surveys

- Which Mode to Choose
  - Web, telephone, face-to-face, mail?
- Which Mix to Chooses
  - Face-to-face / Telephone
  - Mail / Telephone?
  - Internet / Mail?
  - CATI / CAWI?
  - ???
- Why? Consequences?

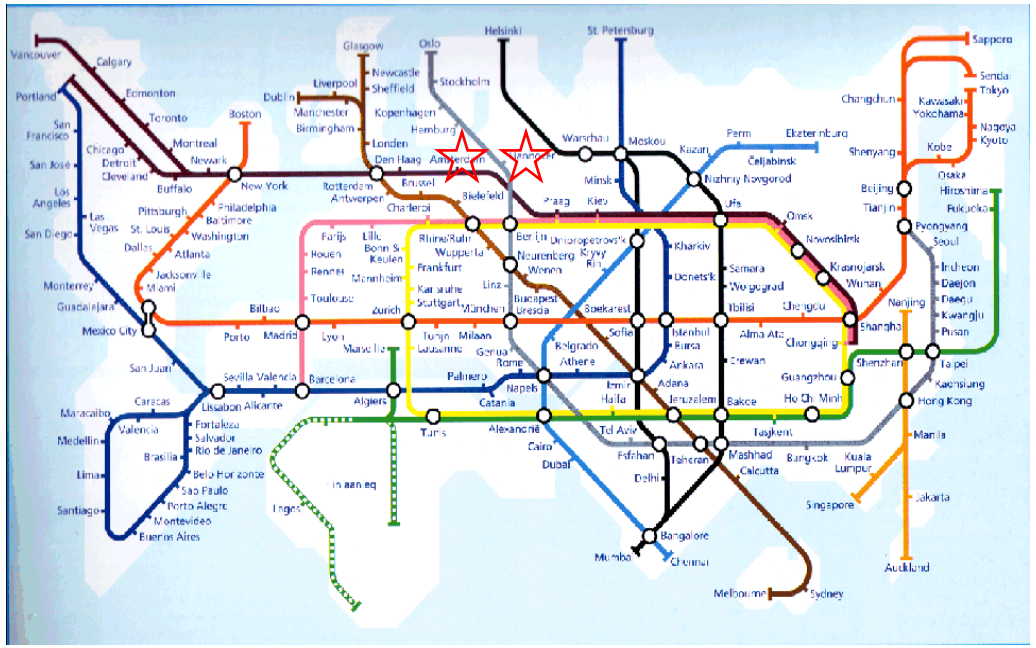
Terminology

- Mixed Mode
- Multi Mode
- Multiple Mode
  - Often used interchangeably
- Mixed Mode
  - Any combination of **survey** data collection methods (modes)
  - In any part of the data collection process

Note: Term mixed methods used in qualitative studies

Mixed Mode Surveys
A Road Map

- New adventurous continent
  - Mixed Mode Territory
- How to plan the trip
  - Where to visit?
  - What to pack?
- Short course
  - Road map
  - Itinerary



Why Mix Modes? Total Survey Error Perspective

Why Mixed-Mode?

Choosing the Optimal Data Collection Method

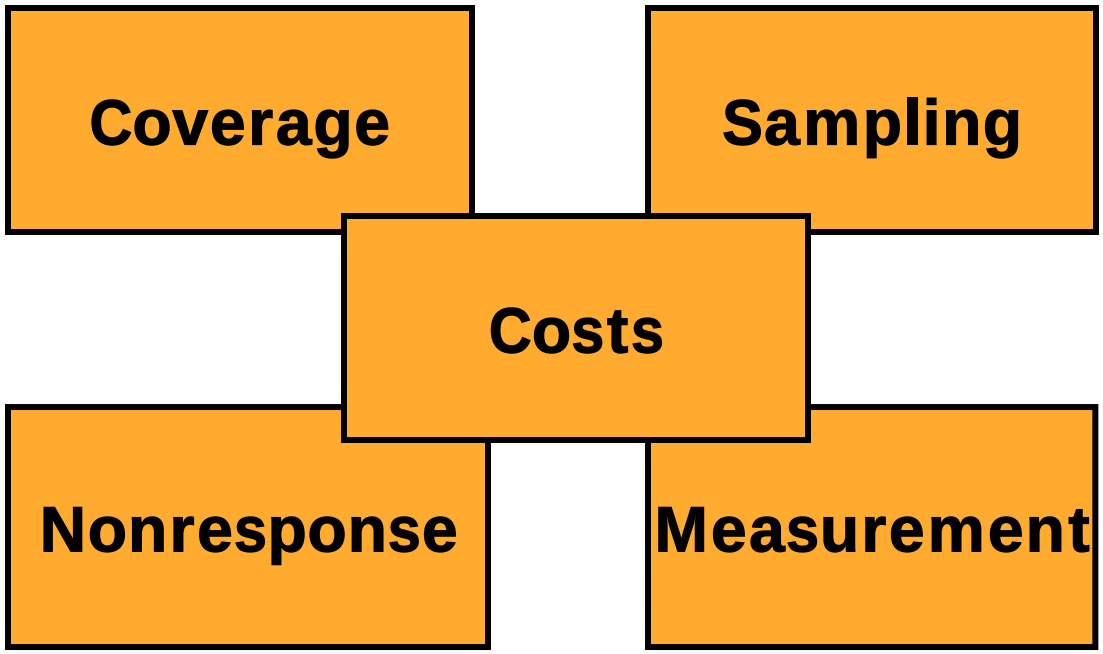
- Best data collection procedure given
  - Research question
  - Population
- Reduce total survey error
- Respect survey ethics/privacy
- Within available time
- Within available *budget*

Biemer & Lyberg, 2003

Best *Affordable* Method

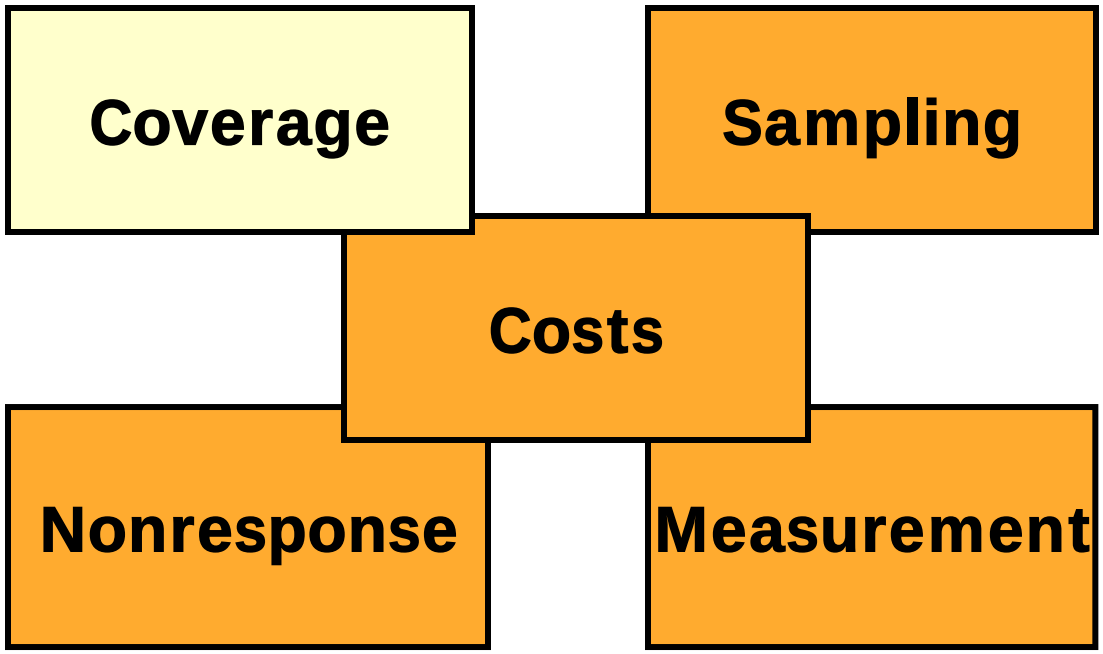
- Mixed-mode explicit trade-off
  - Survey Errors
  - Costs
- Example: Nonresponse follow-up
  - Mail survey
  - Telephone follow-up
  - Face-to-face for sub-sample of remaining nonrespondents

Costs & Errors



Groves, 1989

Survey Errors



Coverage Error

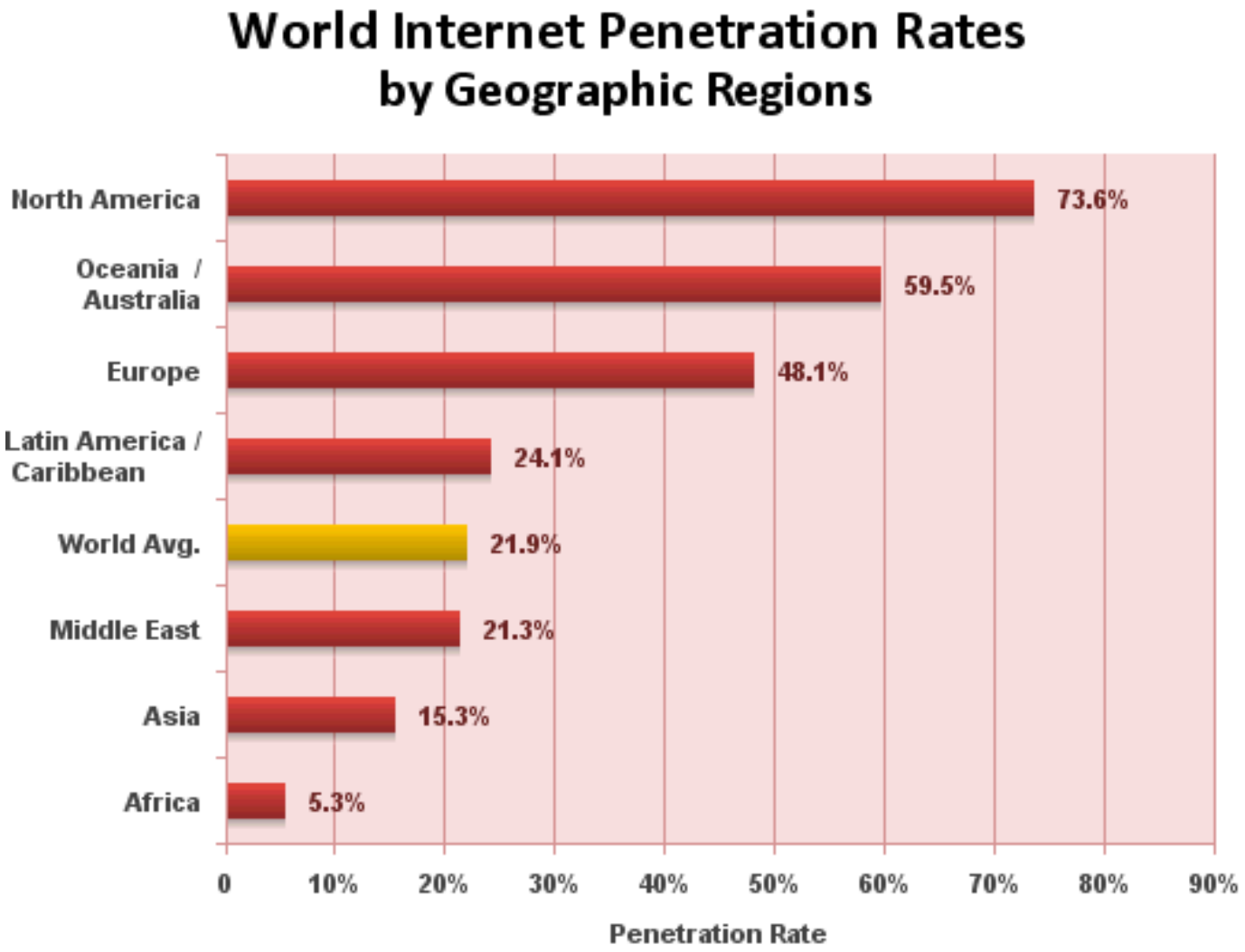
- Sampling frame must include *all* units of population of interest (once), if not:
- Frame Coverage Errors
  - Non-sampling error
- Errors arising from construction of sampling frame
  - Omissions
  - Erroneous inclusions
  - Duplications

Coverage Error and Mode

- Sampling frame must include *all* units of population of interest (once)

- Are all intended covered by mode or is there danger of undercoverage?
  - Telephone
    - Telephone penetration
      - Landlines vs mobile (cell) phones
  - Web
    - Internet penetration differs per country

Web Surveys and Coverage



Source: Internet World Stats - www.internetworldststs.com/stats.htm

Mid year 2008

Miniwatts Marketing Group

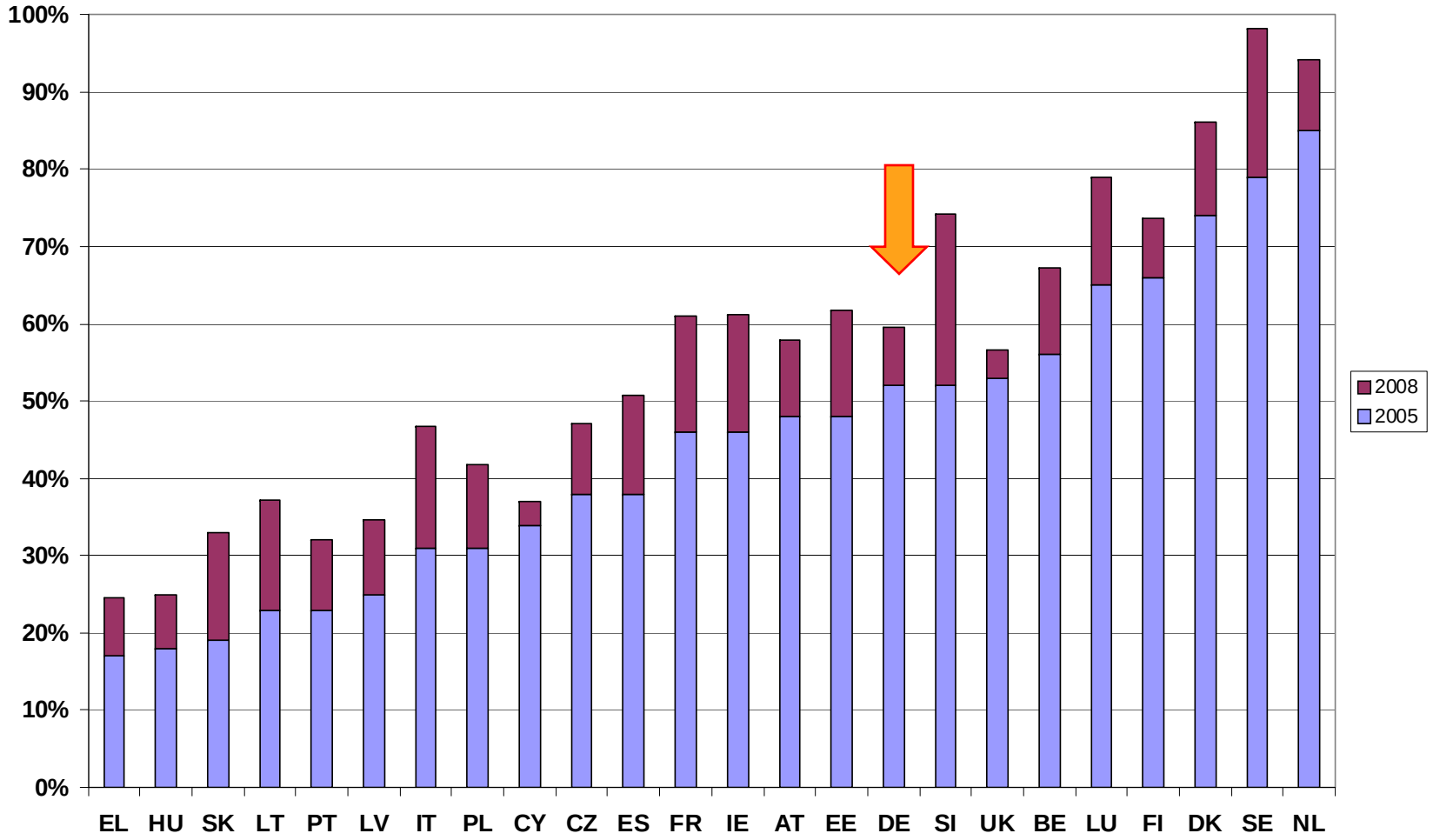
Europe Diverse Picture

- Euro Barometer Data: face-to-face survey!
- Internet Access at Home Adults 15 + (2008)
  - Holland: 86%, Sweden: 83%, Denmark: 81%, Finland: 73%
  - **Germany: 58% overall**
    - Former West Germany: 61%
    - Former East Germany: 48%
  - Romania: 11 %, Turkey 15%, Bulgaria 21%
- Positive trend over time: Growth 2005 to 2008
  - Between 1% (Holland) and 21%(Bulgaria0→21%)
    - 10% (UK: 53→63; Estonia 48→58%)

% Individuals with Internet Access at Home

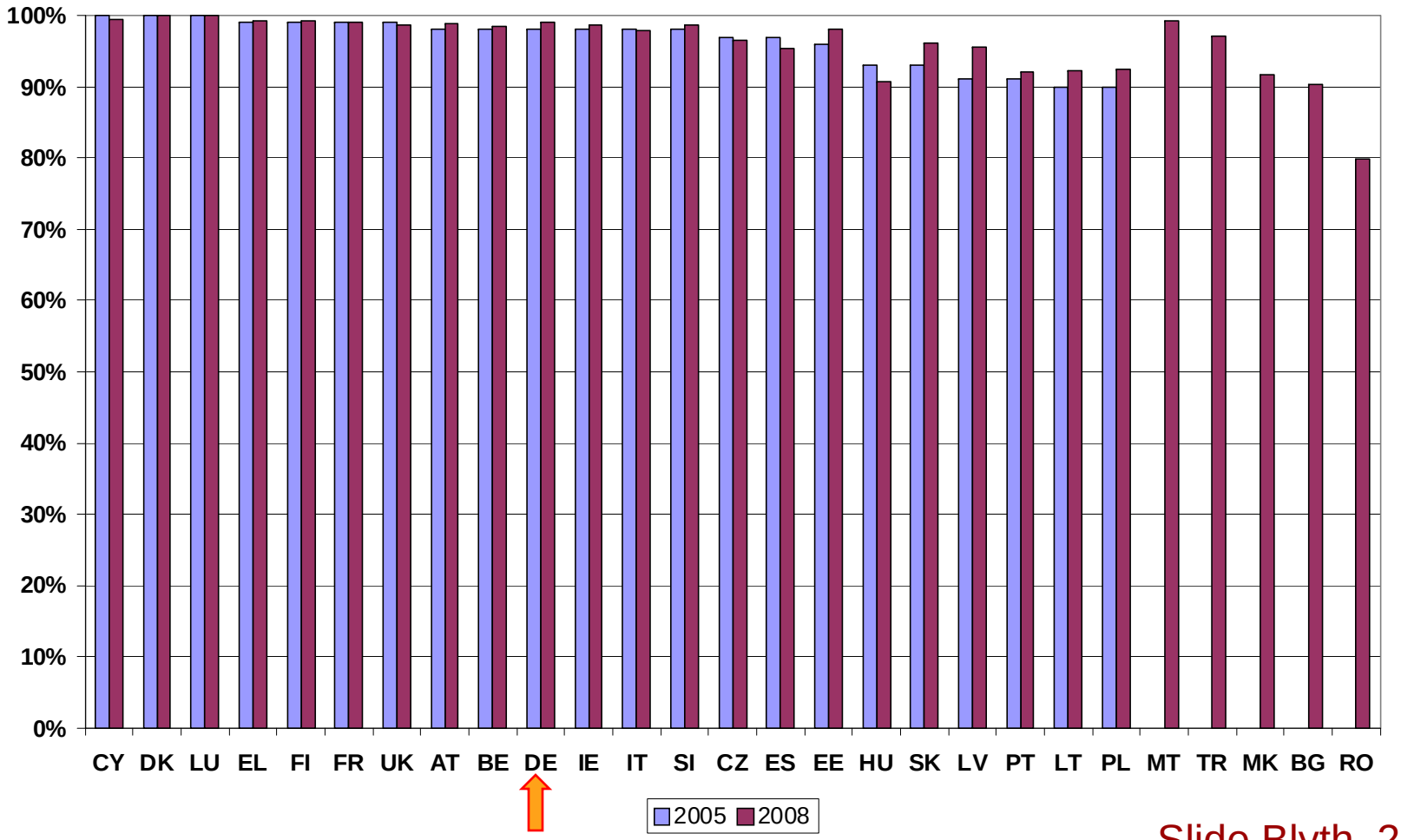
Slide Blyth, 2008

Source: Eurobarometer 2005 & 2008



% Individuals with a Telephone (of any sort)

Source: Eurobarometer 2005 & 2008

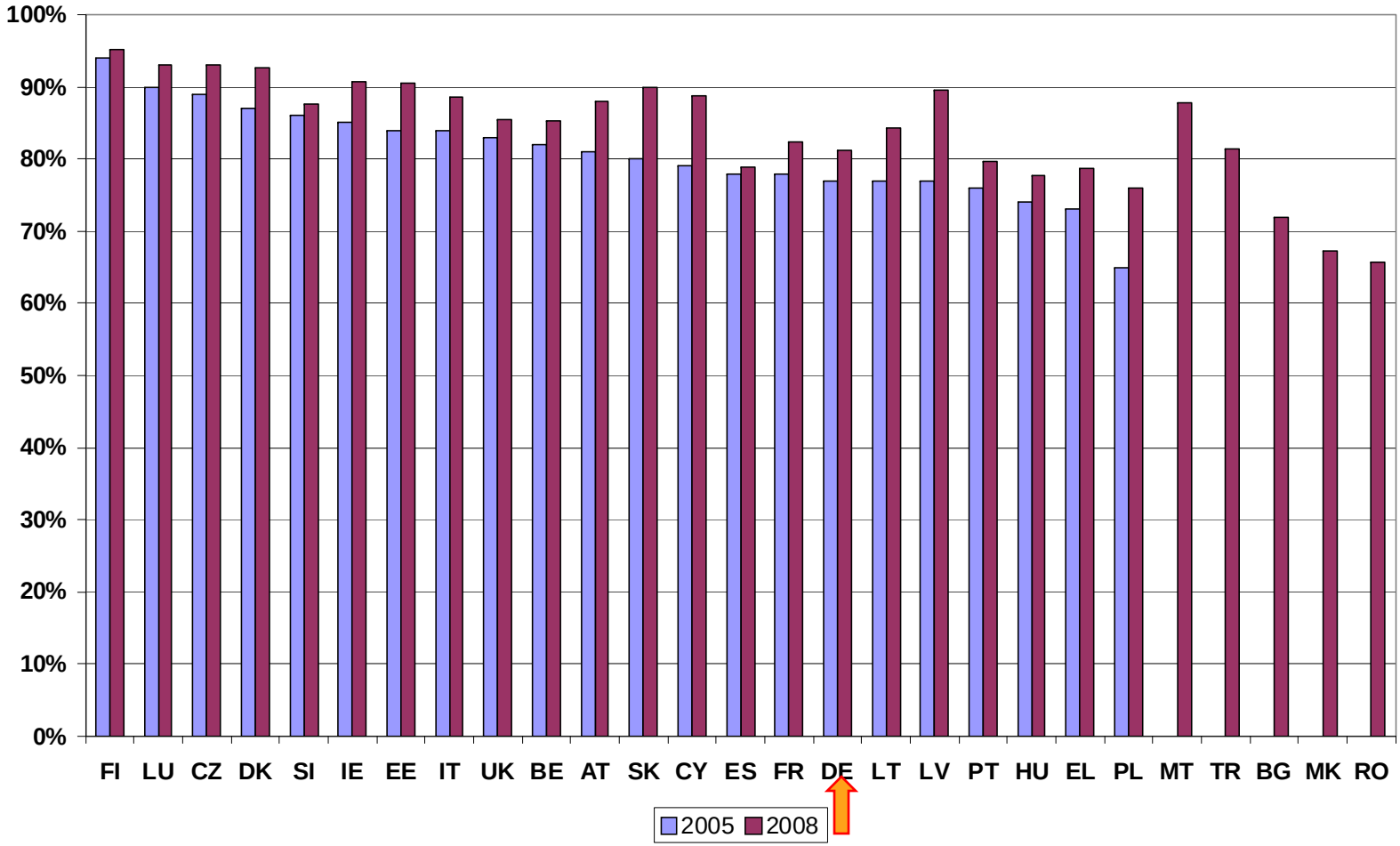


Slide Blyth, 2008

% Individuals with a Mobile (Cell) Phone

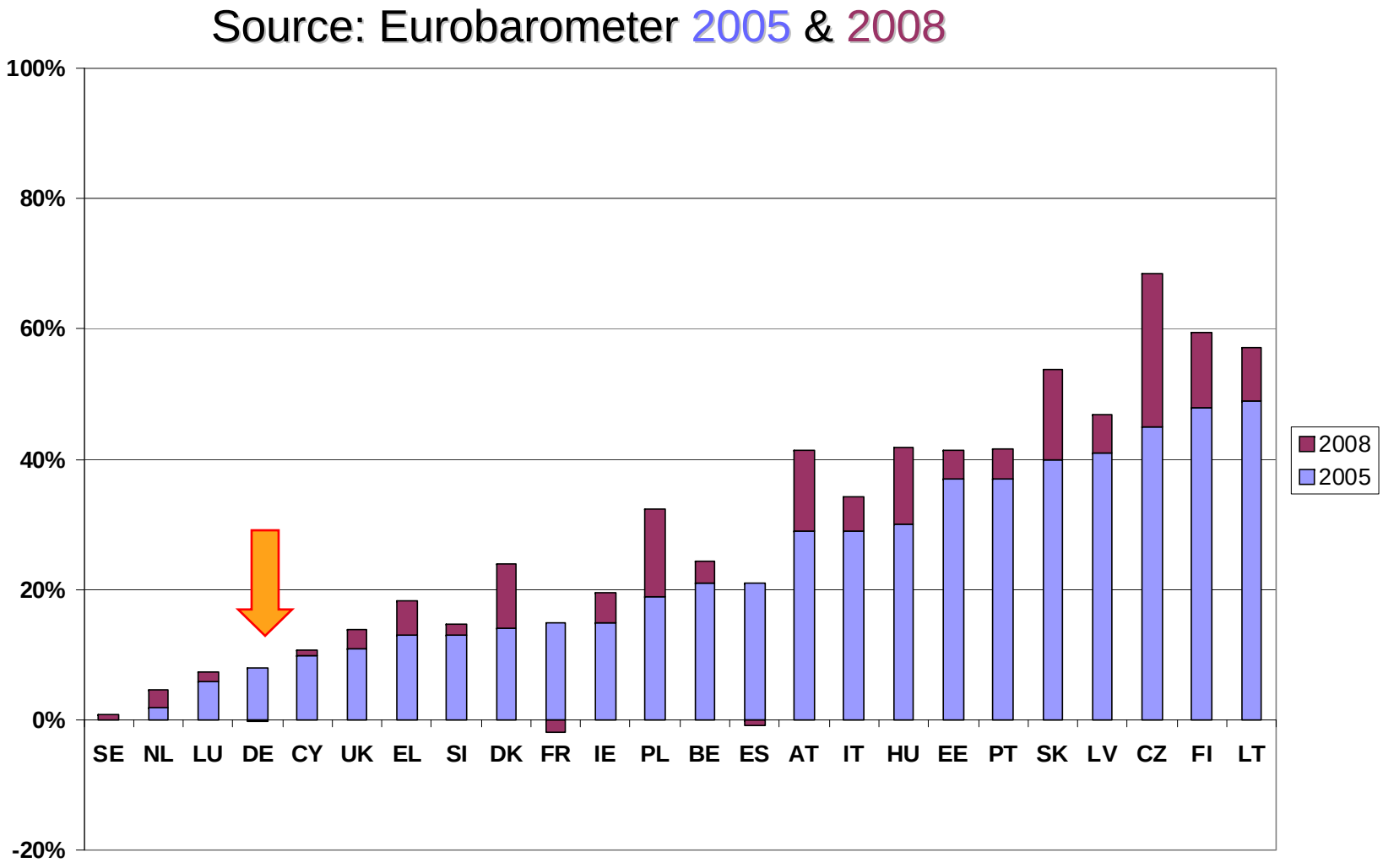
Slide Blyth, 2008

Source: Eurobarometer 2005 & 2008



% Individuals Mobile only No Fixed (land)line

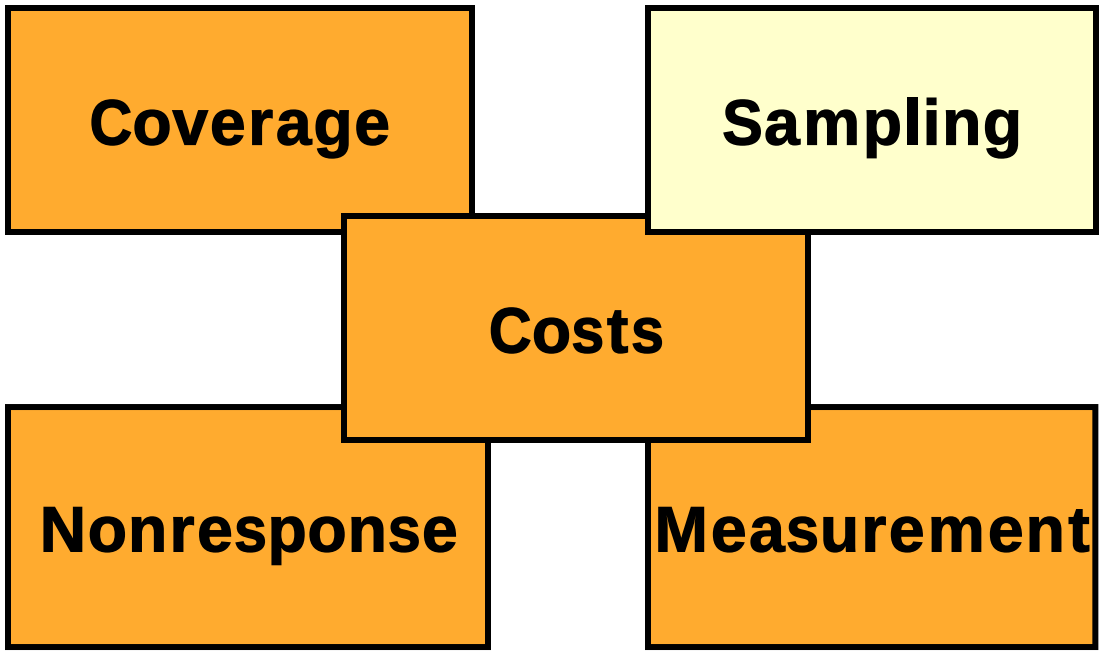
Slide Blyth, 2008



Country Codes

- **BE** Belgium, **BG** Bulgaria, **CZ** Czech Rep, **DK** Denmark
- **DE** Germany, **D-W** West Germany, **D-E** East Germany
- **EE** Estonia, **EL** Greece, **ES** Spain, **FR** France
- **IE** Ireland, **IT** Italy, **CY** Cyprus, **LV** Latvia, **LT** Lithuania
- **LU** Luxemburg, **HU** Hungary, **MT** Malta, **NL** Netherlands
- **AT** Austria, **PL** Poland, **PT** Portugal, **RO** Romania
- **SI** Slovenia, **SK** Slovakia, **FI** Finland, **SE** Sweden
- **UK** United Kingdom, **HR** Croatia, **TR** Turkey

Survey Errors



Sampling Error

- Occurs because only a sample of the population is invited to participate in a survey in stead of the total population
  - Statistic of interest is computed on sample
- Provided a ***probability*** sample is used
  - Each element in the population has a **known** and non-zero probability of selection from the sampling frame
    - Provide protection against selection bias (e.g. self-selection)
    - Give a means of quantifying sampling error

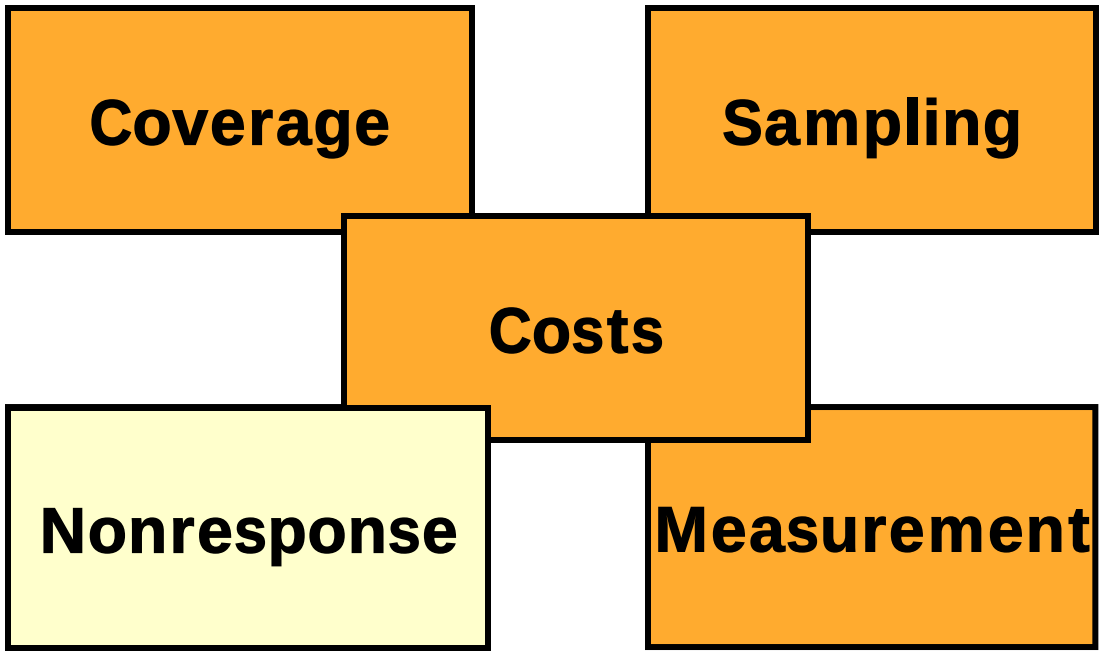
Sampling Error 2

- Statistic of interest is computed on sample
- Statistical procedures to accommodate this
  - Standard error, p-value, statistical tests, etc
- Standard software assumes Simple Random Sampling
- But there are more complicated sampling schemes (e.g., stratified, cluster)
- This needs more complicated statistics
  - Multilevel analysis, Sudaan, etc

Types of Samples

- Information from whole population, that is, NO sampling: A Census ☺
- Probability Sampling
  - Random selection, random sampling
  - Allows for statistical theory, inference ☺
- Non probability Sampling
  - Selection probability unknown
    - No known probabilities of inclusion in sample
  - No statistical theory
    - No p-values, no margins of error ☹

Survey Errors



Non-response Error

- Nonresponse / Missing Data
  - Inability to obtain data on all questionnaire items from all persons: *Missing Data*
- Unit non-response
  - Whole unit fails to provide data
    - Sampling unit, data collection unit, or analysis unit

Non-response Error 2

- Quantification (Non) Response Figures
  - Response Rate, Refusal Rate, etc

- Standardization response figures
  - WWW.AAPOR.ORG section survey methods, standards and best practice
    - RDD telephone, in person household, mail and internet surveys of specifically named persons
  - WWW.ESOMAR.ORG professional codes and guidelines for guidelines on internet research including non response categories that have to be reported

Nonresponse Internationally

- International Comparison Official Statistics
  - Longitudinal data statistical offices around the world
  - Internationally nonresponse increased over time, both noncontact and refusal
  - Countries differ in overall response rate
  - In general a negative trend over time
  - Speed of increasing nonresponse differ from country to country
    - Source De Leeuw & De Heer (2002)

Table 10.1 Response overview European Social Survey 2002/2003

country	eligible sample units	res-ponse rate	non-contact rate	refusal rate		incapacity rate[a]		un-known[c]	net sample
				total	by proxy[b]	total	language		
Austria	3736	60.41	12.04	25.67	3.27	1.87	0.00	0.00	2257
Belgium	3207	59.21	8.11	24.70	5.61	7.98	1.87	0.00	1899
Czech Republic	3139	43.33	11.63	20.01	4.81	5.32	0.13	19.72	1360
Denmark	2229	67.56	3.77	23.60	0.18	5.07	0.00	0.00	1506
Finland	2732	73.21	3.88	19.03	1.76	3.88	0.62	0.00	2000
France	3488	43.09	14.68	38.53	33.03	3.70	0.86	0.00	1503
Germany	5242	57.13	8.34	26.12	1.87	8.28	1.11	0.13	2919
Greece	3208	79.99	2.74	16.08	10.35	1.18	0.22	0.00	2566
Hungary	2412	69.86	6.84	13.93	0.00	9.37	0.00	0.00	1685
Ireland	3174	64.46	9.80	20.26	8.13	5.48	0.00	0.00	2046
Israel	3520	70.99	5.77	21.76	8.38	1.48	0.51	0.00	2499
Italy	2761	43.72	4.13	43.86	37.02	8.29	0.04	0.00	1207
Luxemburg	3535	43.93	11.46	44.61	0.00	0.00	0.00	0.00	1552
Netherlands	3484	67.85	2.70	24.37	0.00	3.16	1.32	1.92	2364
Norway	3132	65.01	2.65	24.87	0.51	7.47	1.05	0.00	2036
Poland	2881	73.24	2.08	19.75	4.82	4.93	0.00	0.00	2110
Portugal	2196	68.81	2.82	25.82	22.22	1.23	0.09	1.32	1511
Slovenia	2154	70.52	5.06	17.27	4.22	4.97	0.00	2.18	1519
Spain	3249	53.22	11.05	32.10	11.45	3.32	0.49	0.31	1729
Sweden	2878	69.46	4.07	20.95	0.00	5.52	1.84	0.00	1999
Switzerland	6097	33.48	2.80	54.81	24.95	8.91	5.04	0.00	2040
UK	3696	55.52	4.92	33.17	11.39	5.09	1.33	1.30	2052

a Respondent mentally or physically unable to cooperate throughout the fieldwork period; language barrier; respondent unavailable throughout the fieldwork period for other reasons.

b Refusal to cooperate by other person within the household or at the address, possibly within knowing whether the refuser is the target sample person.

c Percentage of sample units not accounted for.

Source: NSD (2004)

From Ineke Stoop, 2005

Non-Response Error 3

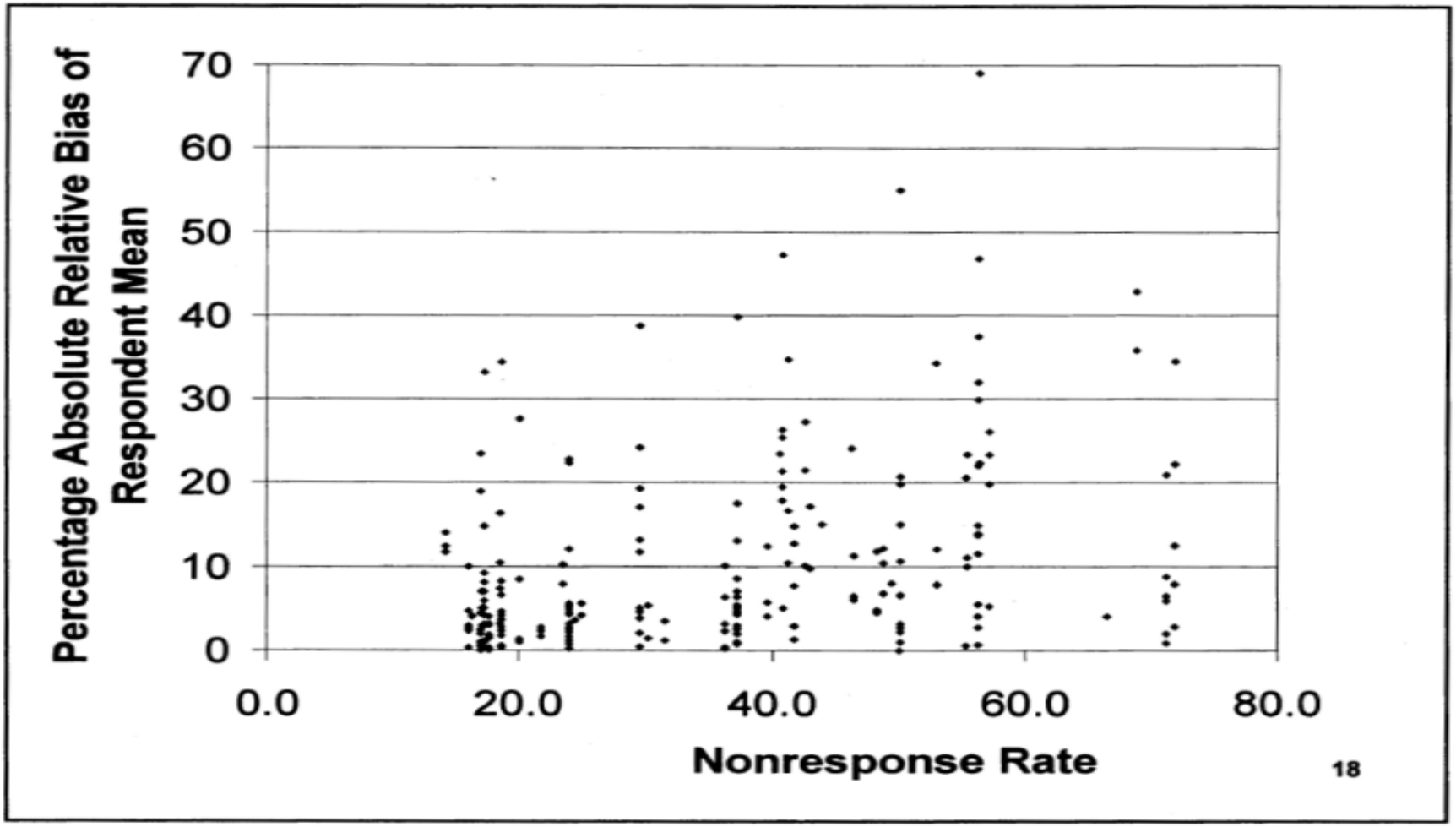
- Beware Nonresponse Error is more than nonresponse!

- Nonresponse error
  - I. Nonresponse occurs
  - II. Respondents and non-respondents differ on variable of interest (key variable study)

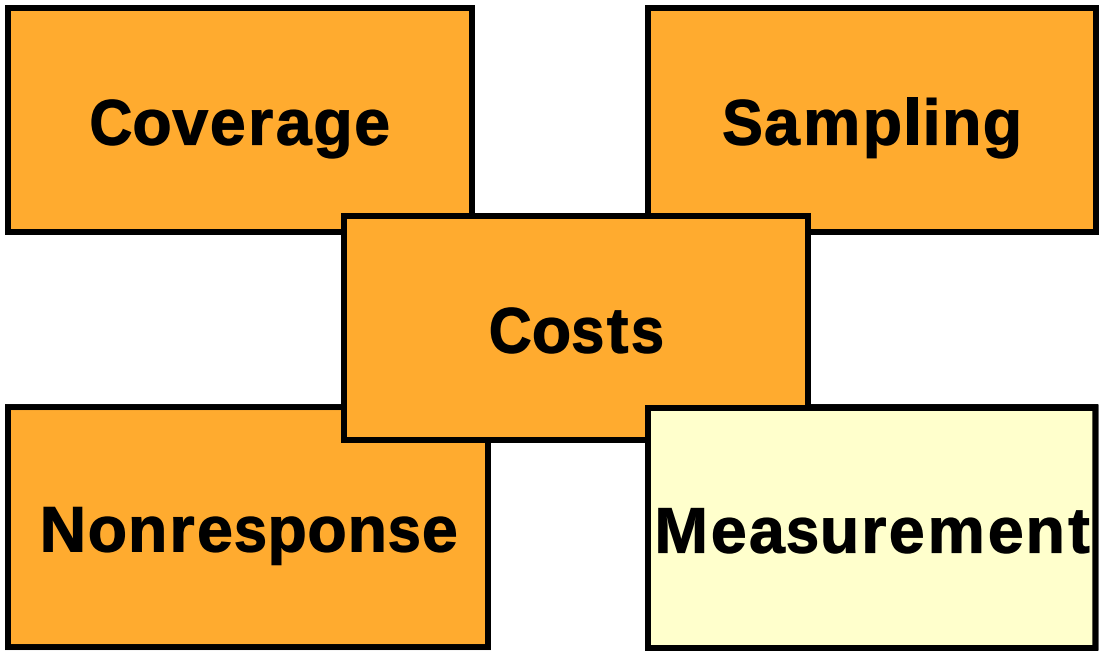
- Nonresponse figures as such uninformative
  - High nonresponse but little or no error (or vice versa)
  - Need nonresponse studies / diagnostics

Nonresponse Rate vs. Bias

Figure 2. Percentage absolute relative nonresponse bias of 235 respondent means by nonresponse rate from 30 different methodological studies (Groves, 2006 POQ)



Survey Errors



Measurement Error

- Measurement Error
  - Nonsampling error or error of observation.
- Measurement errors are associated with the data collection process itself.
- Measurement error occurs when a respondent's answer to a question is inaccurate,
  - In other words when answer departs from the 'true' value

Measurement Error 2

- Measurement errors are associated with the data collection process itself
- There are three main sources of measurement error:
  - Questionnaire
  - Respondent
  - Method of data collection
- When interviewers are used for data collection, the interviewer is a fourth source of error

Mixed Mode Surveys

A Solution to Problems Illustrated with Some Special Cases

How do Modes Differ?

- Practical advantages & disadvantages
  - Personal needed, time, equipment, etc
- Differences coverage
- Differences sampling
- Different cost structure
- Differences measurement

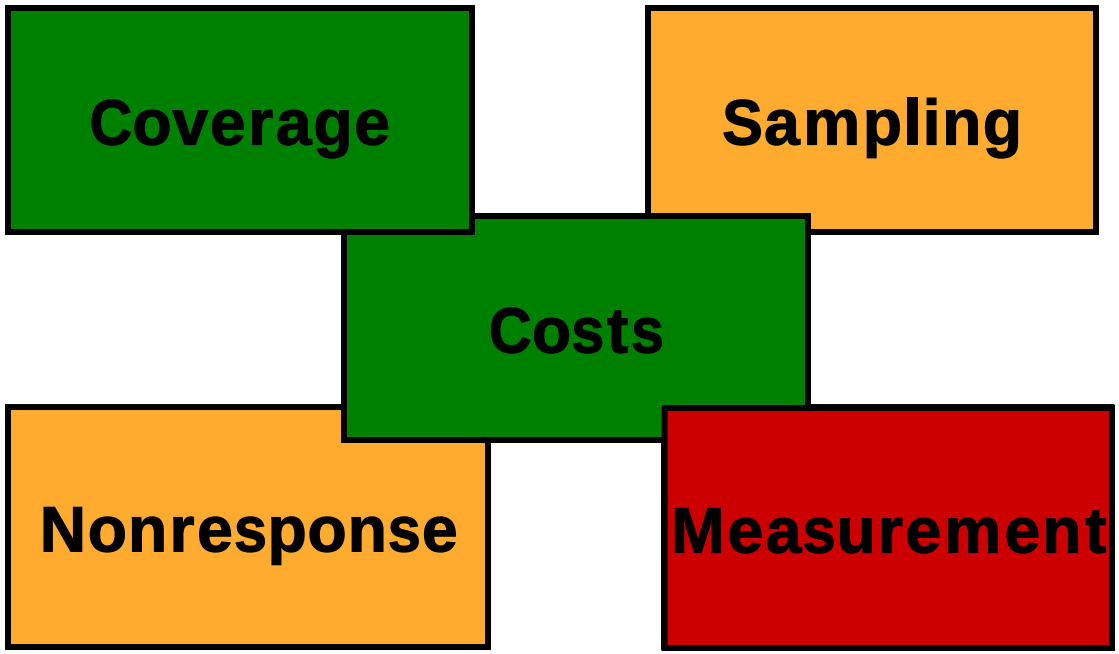
Known Coverage Problems

- **Face-to-face coverage:**
  - Available household lists not complete
  - Need to manually count and list
- **Telephone coverage:**
  - Households with no telephones
  - Cell phone only households
  - No directory of cell phone numbers (country specific)
  - Number portability and erosion of geographic specificity
- **Mail coverage:**
  - Depends on good mailing list.
    - Country specific
    - Registers or postal lists
- **Email coverage:**
  - No systematic directory of addresses

Note: general population coverage problems

Solution Web Coverage

Concurrent Mixed Mode

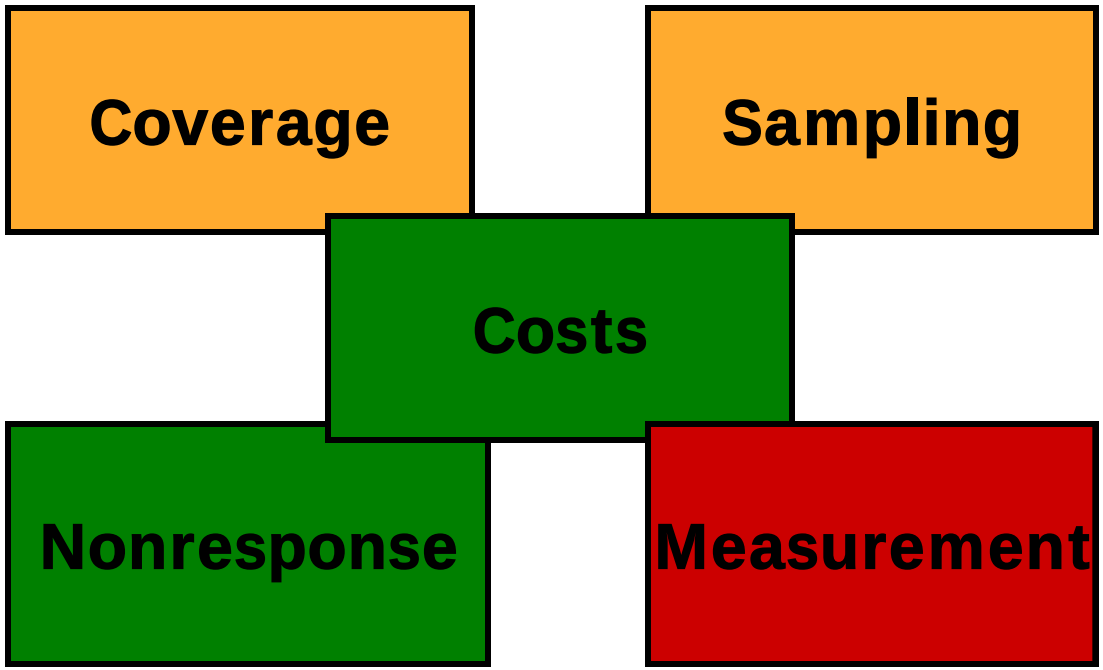


Declining Response Rates

- Response rates decreasing significantly in the last 20 years.
  - Decline has occurred for most types of surveys— particularly telephone and in-person interviews
  - Evidence of trends for mail surveys not as clear
- Web surveys are too new to provide good trend data.
  - But, in general lower than comparable mail surveys
  - And other modes
    - Lozar Manfreda et al, 2008; Shih & Fan,2008
- Increase in nonresponse is a global problem
- No **single** or clear explanation for these trends.
  - Several theories on nonresponse

Nonresponse Solution

Sequential Mixed Mode



Sequential Mixed Mode

- Sequential
  - Different modes for successive phases of interactions (contact phase, data collection phase, follow-up phase)
    - Screen or contact by phone,collect data by face-to-face interview
  - Different modes in sequence during data collection phase
    - American Community Survey
      - Mail, telephone, face-to-face

Balden, 2004

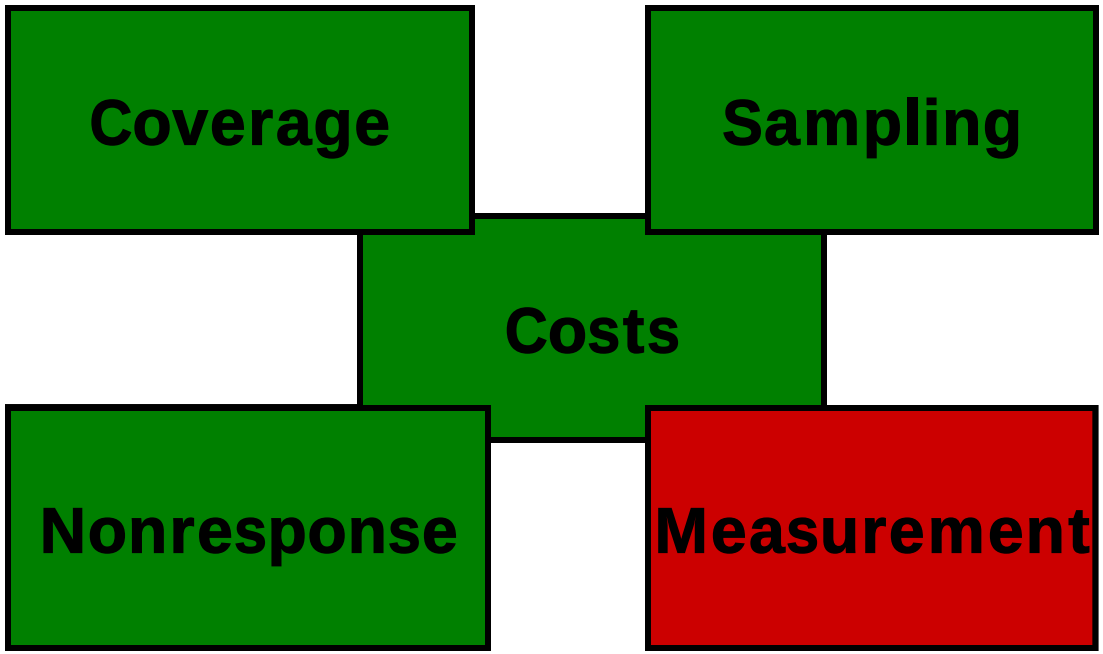
American Community Survey

- Sponsor: U.S. Census Bureau
- Target population: Households in U.S.
  - 2.9M addresses sampled
- Focus: social, housing, & economic characteristics
- Frame: Census Master Address File
- Modes (sequential):
  - Mail
  - Telephone follow-up
  - In-person follow-up
- Field period: 3 months
- Response rates: 97.3% (for 2005)
  - 1.9M interviews completed

Frankel & Link, 2007

International Surveys

Concurrent Mixed Mode



Concurrent Mixed Mode

- Concurrent
  - Multiple modes are used simultaneously for data collection: implemented at same time
    - Asthma awareness study
      - Invitation postcard offering choice
    - Establishment and business surveys
    - Dual frame surveys
    - International surveys

Balden, 2004

To Mix or Not to Mix

- Mixing modes has advantages, but
  - Will the answers provided by respondents differ by mode?
  - Can data that are collected through different modes be combined in one study?
    - Cross-sectional?
    - Longitudinal?
  - Can data that are collected through different modes be compared over studies or countries?

Taxonomy of Mixed Mode Surveys I Multi(ple) Mode Contacts

Multi Mode Survey Systems

- Survey more than data collection
- Communication with Respondent
  - Contact Phase
    - Pre-notification
    - Screening
    - Invitation
      - PPSM
  - Data collection
  - Follow-up
    - Reminders

Terminology Revisited

- Multiple Mode In General
  - Communication with Respondent
    - Contact Phase
      - Pre-notification
      - Screening/Invitation
    - Follow-up
      - Reminders
- Mixed Mode
  - Data collection itself
    - Data collected with more than one mode
      - At same time (concurrent)
      - One after another (sequential)

One Survey System: Multiple Modes of Communication

- Example Nielsen media research
  - Multiple modes of contact in 7 steps
    1. Pre-recruitment postcard
    2. Recruitment phone call
    3. Advance postcard announcing diary
    4. Diary survey package
    5. Reminder postcard
    6. Reminder phone call
    7. Reminder postcard

Bennett & Trussel, 2001
Trussell & Lavrakas, 2004

Contact Phase

Advance Notification/Screening Different Mode from Data Collection

- Rationale
  - Correct sampling frame
  - Raise response rate
  - Enhance legitimacy and trust
  - Send incentive in advance
- Effect on Quality
  - Reduce coverage and nonresponse error
  - No threats to measurement if data collection itself is in **single**-mode (= data are collected with one method only)

Contact Phase

Invitation in Different Mode from Data Collection Itself

- Why?
  - Reduce coverage and nonresponse error
- Effect on measurement
  - No threats to measurement if data collection itself is in one **single-mode**
    - Telephone invitation for IVR
    - Nielsen media study: data collection diary
  - Potential threats if data collection is **multiple-mode**
    - Postcard invitation for Web/CATI offering choice of mode to respondent

Follow-up Phase

Reminder(s) in Different Mode from Data Collection Itself

- Rationale
  - Simple reminder, such as postcard, short telephone call, etc has low costs
  - Raise response rate
- Effect on Quality
  - Reduce nonresponse error
  - If pure reminder (data collection ***single***-mode) no threats to measurement

Follow-up Phase

Reminder(s) in Different Mode from Data Collection Itself + Questions

- Rationale
  - Simple reminder, such as short telephone call, has low costs
  - Raise response rate
  - At low additional cost ask extra questions
- Effect on Quality
  - Reduce nonresponse error
  - If reminder plus additional questions, then ***multiple***-mode
    - Part of data collection different from main mode
      - Threat to measurement

In Sum: MM &Contact Phases

- Second or third method for **interaction** with respondent
- No actual data are collected with additional modes (e.g. only encouraging people to respond)
  - Data quality enhanced ☺
  - Generally more costly
  - More complicated logistics
- Second or third mode for **data collection** too (e.g., some respondents by one, other by second mode (mail with telephone follow-up)
  - Potential threats to data integrity ☹

Taxonomy of Mixed Mode Surveys II: Measurement Error

Data Collection Phase

- Mixed mode for interaction with respondent in Contact Phase and Follow-up Phase mostly Win-Win situation!
- More complicated in Data Collection Phase
  - Threats to data integrity
  - One big exception: win-win situation in mixing interview mode with SAQ for sensitive questions
    - **Some questions by one mode, other questions by another mode, but same procedure for all respondents**
    - E.g. CAPI/CASI
      - Interviewer guidance for non sensitive questions
      - More privacy, less social desirability sensitive questions

Implications Mixed Mode in Data Collection Phase

- Potential Risk
  - Introducing mode effects in data set
- Result:
  - Increasing measurement error
- However:
  - Reduction of other errors
    - E.g., Coverage / nonresponse
- Careful consideration needed
- Careful design for optimal mixed mode

Recap Sequential vs. Concurrent Data Collection Phase

- **Sequential**
  - Different modes in sequence during data collection phase
    - Example: American Community Survey
      - Mail, telephone, face-to-face
    - Example: LFS Sweden
      - Longitudinal face-to-face, telephone
- **Concurrent**
  - Multiple modes are used simultaneously for data collection: implemented at **same time**
    - Example: Asthma awareness study
      - Invitation postcard offering choice of modes
    - Example: Using more private method for sensitive questions

Data Collection Phase: Concurrent Mixed Mode 1

- Multiple modes implemented at same time
  - For **sub set** of questions only
- Reduce Social Desirability Bias
  - Sensitive questions in more 'private' mode
    - CAPI - (A)CASI mix
    - Telephone - IVR (or T-CASI) mix
    - Face-to-face – paper SAQ mix
  - Example: US National Survey on Drug Use and Health (NSDUH)
- Win-win situation ☺

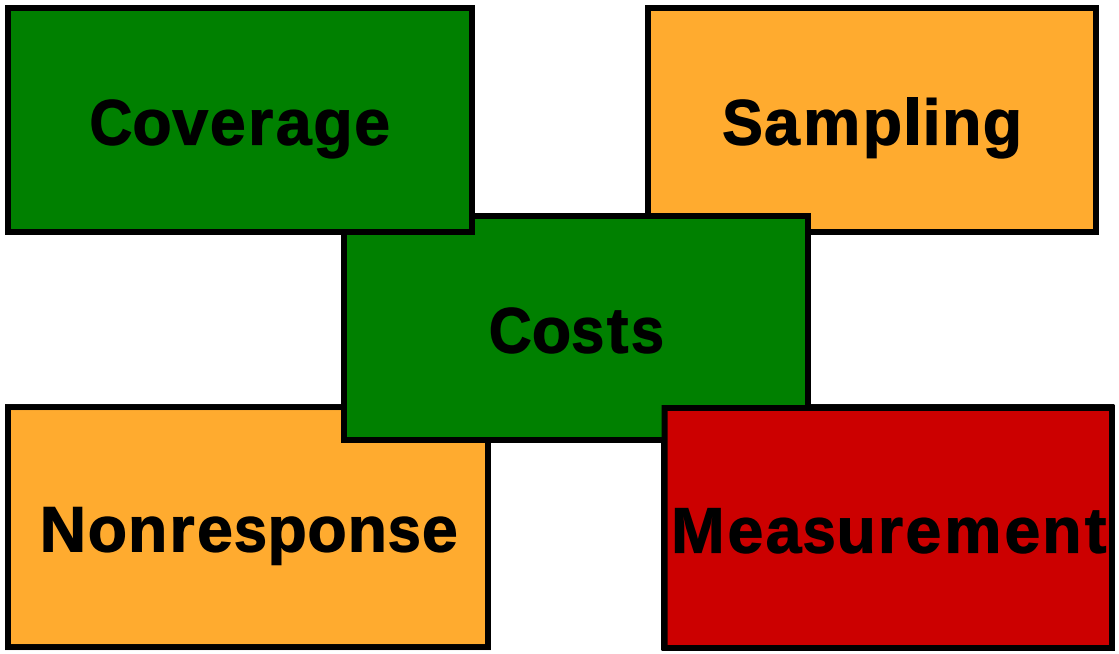
- **Warning:** Beware of concurrent mixed mode for **total** questionnaires when sensitive topics are studied!!!
  - Different groups get different modes

Data Collection Phase: Concurrent Mixed Mode 2

- Multiple modes implemented at same time
  - For **all questions**, full questionnaire, **one population**
- Reducing Coverage Error at reasonable costs
  - Dual frame sampling
- Dangers concurrent mixed-mode
  - Measurement differences
    - E.g., social desirability, recency effects
    - Often, difficult to entangle as (self-)selection and mode effect are confounded
    - PPSM: random allocation makes it possible to study mode effects
- Reduced coverage error at the price of increased measurement error

Remember Web Coverage

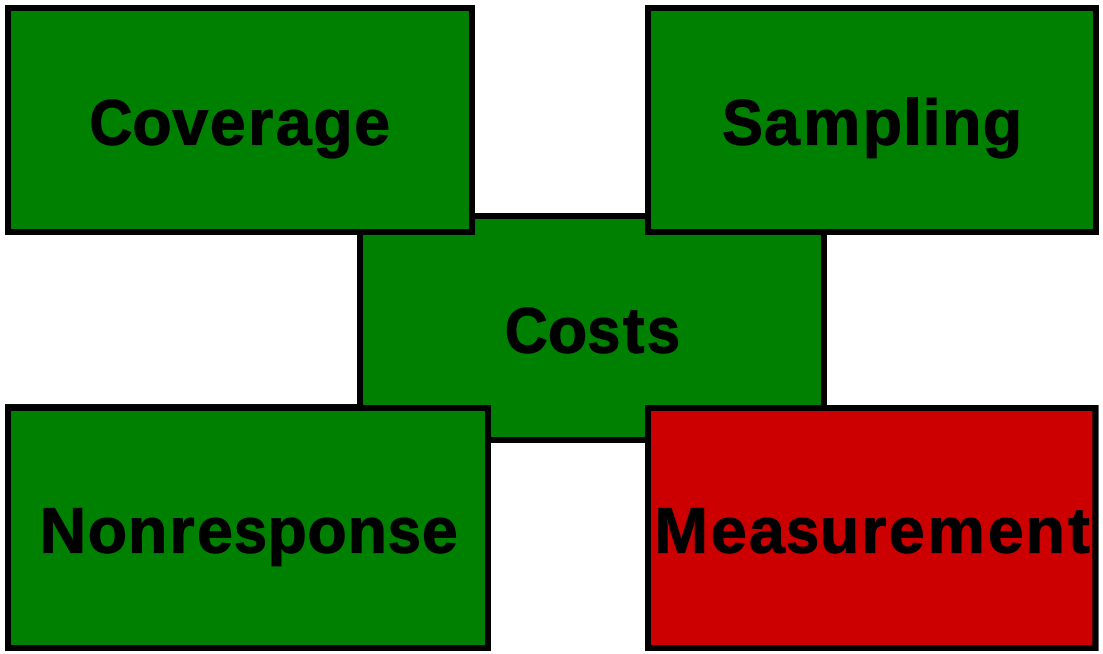
Concurrent Mixed Mode



Data Collection Phase: Concurrent Mixed Mode 3

- Multiple modes implemented at same time
  - For **all questions**, full questionnaire
- Different **populations** offered **different mode**
  - International Surveys
  - Regional comparisons
  - Multiple cultures
- Practical considerations
  - Only way to get data for certain groups/countries
  - Example ISSP
- Danger: measurement error
  - Population and mode effects confounded

International Surveys



Data Collection Phase : Concurrent Mixed Mode 4

- Multiple modes implemented at same time
  - For **all questions**, full questionnaire, one population
  - **Respondent is offered choice of mode**
- Rationale: be client centered in order to reduce nonresponse and save costs
- Dangers
  - Measurement differences confounded with self-selection groups
  - Higher nonresponse in household surveys!!!
    - 1-9% Dillman (2008).
    - More effective in establishment surveys by official statistics
      - Need more empirical data

Respondents Viewpoint:
Offering A Choice Makes Life More Difficult

- Researcher's viewpoint
  - Client centered to reduce nonresponse
  - Respondent friendly, establish good-will
- BUT Respondent's viewpoint is different
  - More information to read and process
    - Higher 'costs' in social exchange
  - Increased cognitive burden
    - Two decisions to make in stead of one
      - From "will I participate" to "will I participate and what method do I want to use"
      - Harder task so simplest thing is opt-out
    - May concentrate on choice, not on survey
      - Distracts from message and arguments on why to cooperate
        - Weakens saliency

Sequential Mixed Mode Nonresponse Reduction

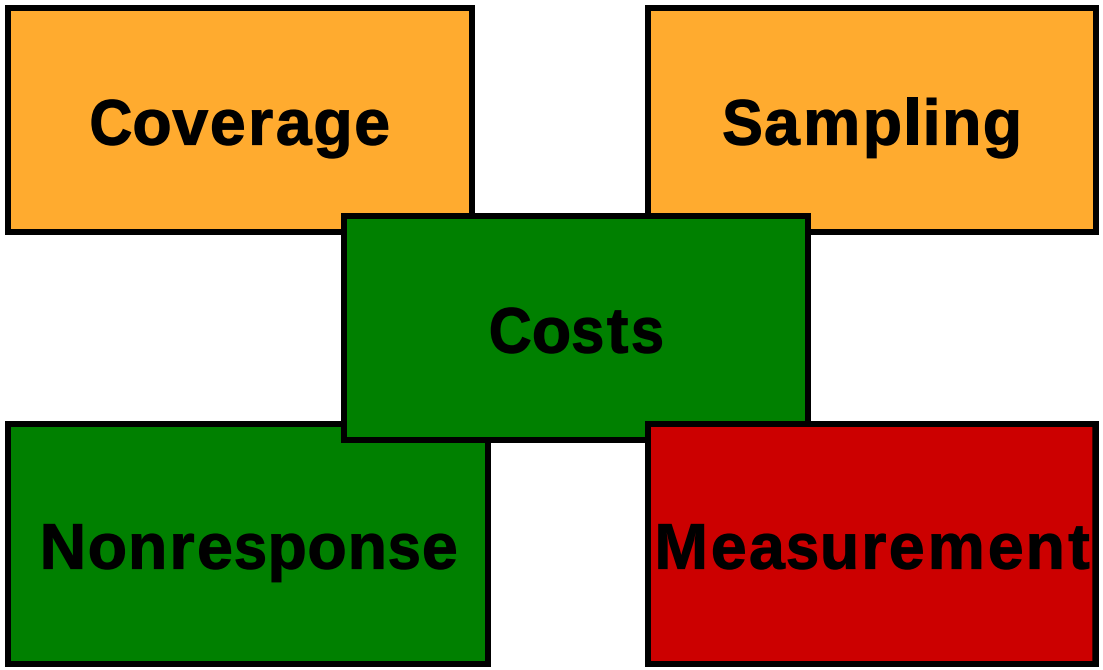
- Sequential mixed-mode approach may be more effective than giving respondents a choice
- Sequential for nonresponse reduction better than concurrent
- But, concurrent a good solution for coverage problems
- Sequential approach for nonresponse reduction
  - Different modes in sequence during data collection phase
    - Example: American Community Survey
      - Mail, telephone, face-to-face

Data Collection Phase: Sequential Mixed Mode 1

- Multiple modes implemented in sequence one time period / cross-sectional study
  - Successful for nonresponse reduction
    - Inexpensive mode first main mode
    - More expensive mode as follow-up
  - Potential measurement error
- Beware for using data of sequential mixed mode for assessment of nonresponse bias
  - Mode and nonrespondents may be confounded
  - Ideally assessment of nonresponse bias by small sub sample same mode

Nonresponse Solution

Sequential Mixed Mode



Data Collection Phase: Sequential Mixed Mode 2

- Multiple modes implemented in sequence, one sample
- Multiple time points / longitudinal study
  - Cost reduction and practical considerations
    - More expensive mode
      - Selection and screening for panel
      - Base-line study
    - Next waves less expensive study
  - Labor force survey many countries
    - Face-to-face first wave, later waves telephone
  - Web panel selection by face-to-face or phone,
  - Sometimes ABBAB design
    - A=Mode 1, B=Mode 2
    - Example NESTOR study on aging Holland
  - Potential measurement error
    - Time and mode may be confounded

In Sum Mixing Modes Data Collection Phase

- Some questions by one mode, other questions by another mode, but the same procedure for all respondents
  - Sensitive questions by SAQ, rest other mode
    - Data quality enhanced ☺
  - Selection and biographics first wave longitudinal/panel by one mode, next waves other modes
    - Mostly win-win
    - Beware of confounding mode vs. time effects longitudinal 😐
- Some respondents by one mode, other respondents by another
  - Nonrespondents by another mode
  - Offering choice of mode
  - Cross-national
    - Potential threats to data integrity ☹

In Sum: Problems MM Data Collection Phase

- Incomparability
  - Different subgroups different modes (e.g. nonresponse follow-up, or telephone survey in city A, mail survey in city R)
    - Confounding
      - Are groups different (more crime in R)
      - Or is it mode effect (e.g., Soc. Des.)
  - Different samples, different modes (e.g.,comparative research, international)
    - More social isolation in country X than Y or different survey methods (& associated social desirability)?

In Sum: Types of Mixed Mode

- Two major distinctions:
  - Different contact methods are used in different survey phases (e.g., recruitment, screening, questionnaire administration, etc.)
    - Mostly win-win situation, no threat to measurement if data collection is done in one single mode
  - Different methods used for data collection
    - Concurrent mixed mode:
      - Offer two or more modes at same time
    - Sequential mixed mode
      - Assign modes sequentially during life of the survey

To Mix is to Design

- Mixing data collection modes has advantages in reducing noncoverage and nonresponse errors, but
- Mixing methods may enhance measurement errors
- So,

I. Design for Mixed Mode Surveys
   - Design equivalent questionnaires!

II. If possible, measure potential mode effects

III. Adjust

Diagnosis/Adjustment: Design for Mix

Build in overlap	**Method 1**	**Method 2**
Group X	**Main Data Collection**	**Some Data**
Group Y	**Some Data**	**Main Data Collection**

Why and How Modes Differ

Self-Administered vs. Interviewer-Guided

Visual vs. Aural

Media-related customs

Modes & Measurement

- Measurement error occurs when a respondent's answer to a question is inaccurate (departs from the "true" value)
- Modes vary in terms of:
  - Interviewer versus self-administered questionnaires
    - Interviewer impact
  - Stimuli / manner in which survey question is conveyed to respondent (and response is recorded)
    - Information transmission
  - Knowledge about mode, usage, social customs
    - Media related factors

How Modes Differ

Overviews: De Leeuw 1992, 2005 and Dillman & Christian, 2005

- Empirical Evidence **Interviewer Impact**
  - More social-desirability in interview
    - E.g., drinking, fraud
    - More open in self-administered modes
  - More positive in interview
    - Less lonely, better health in interview
  - More acquiescence in interview
    - Tendency to agree
    - Easier to agree than disagree with another person
  - Less missing data/more detailed answers open questions in interview
    - In general interviewer probes help

How Modes Differ 2

- Empirical Evidence **Medium Impact**
  - Few systematic studies (Overviews De Leeuw, 1992, 2005) indicate **advantage of self-pacing**
    - Self-administered more consistent answers
    - SAQ higher psychometric reliability on scales
      - BUT all Paper SAQ vs. interview!
- Internet as medium still different (cf. Krug, 2006)
  - Multi-tasking
  - Scanning
  - Satisficing (close enough in stead of optimal)

Internet as Medium

- Empirical Evidence Medium Impact
  - Hardly any systematic studies
    - Satisficing (less differentiation in web, Fricker et al, 2005)
  - Psychological testing
    - Equivalence when no time pressure (De Leeuw et al, 2003)
  - Conveying sincerity of purpose and trust more difficult
  - **More research needed on nonresponse**
    - Web on average 11% lower (meta-analysis Lozar Manfreda, et al, 2008)
  - Research needed on **response to sensitive** questions
    - Influence of SPAM
    - Trustworthiness web
      - Panel should have advantage vs. one time web survey
      - Existing relationship vs one-time

How Modes Differ 3

- **Information transmission**: visual vs aural; spoken vs written vs typed; question by question or blocks (page)
- Some evidence recency effect in telephone surveys
  - More often last offered answer category is chosen
- Context and order effects less likely in self-administered (paper) than interview
  - Overview / segmentation
    - No empirical studies including web surveys
- Visual presentation & design & quality
  - Growing body of evidence that respondents use all information including visual cues to decide what answer they are going to report
    - Cf Dillman, 2007; Toepoel, 2008; Couper 2009

Good news, but….

- Literature reports that there are some mode difference
  - Not large
    - Except for more sensitive questions
  - But….
    - All empirical evidence is based on
      - Well conducted experiments
        - Controlling/adjusting population differences
        - **Equivalent questions and questionnaires!**

Lesson Learned

- To minimize mode effects one should:
  - Control/adjust for population differences
    - E.g., More younger, higher educated in web and more elderly, lower educated phone
  - Use equivalent questions and questionnaires!
    - Ensure measurement equivalence

Questionnaire Design

Traditional Designs for Specific Modes and the Implications for Mixed-Mode Surveys

Traditional Design F2F

- Face-to-face: Visual + Aural
  - Show cards with answer choices
    - Long lists of answers, long scales with each point labelled
  - Pictures may be used
  - Open-ended questions on wide variety of topics
    - Trained interviewers are carefully instructed to probe in order to get detailed and complete information
  - No opinion etc not explicitly offered, but accepted when given. Interviewers often trained to accept ‘no answer’ only after a standard ‘probe’
  - Transitional texts to guide interviewer and respondent to next block of questions

Traditional Design Tel

- Telephone: Aural only
  - Shorter answer scales (2-5 point scales)
  - Often only anchored end-points
    - On a scale from 1 to 5 with 1 being not at all satisfied and 5 being completely satisfied
  - Visual analogue questions
    - Imagine a ladder with 7 steps
    - Imagine a thermometer with a scale from 0 to 100
  - Unfolding for longer scales
    - Satisfied, dissatisfied or somewhere in the middle
    - Completely, mostly, somewhat (dis)satisfied

Traditional Design Tel2

- Telephone design
  - Difference with face-to-face
    - In general breaking up questions in parts to accommodate loss of visual channel
  - Like face-to-face
    - Open-ended questions and probes
    - No opinion / no answer not explicitly offered
      - But is accepted after probe by well-trained interviewer

Traditional Design Postal

- Mail survey: Visual only, no interviewer present
  - In general, no breaking up of questions in parts
  - But, use longer list of response categories in stead
  - Fully labelled scales
  - Check all that apply instead of yes/no answers
  - Only ‘no answer’ when person skipped question, in stead of interviewer coded ‘refused, do not know, no opinion’
  - Go back and forth: more context available
  - Use illustrations / visuals

Example Mail vs Telephone

Mail

- Is the home in which you live
  - Owned free & clear
  - Owned with a mortgage
  - Rented
  - Occupied under some arrangement

Telephone

- Do you own or rent a home?
  - Follow-ups accordingly, e.g. when owns a home
    - Do you have a mortgage or is it owned free and clear

Example Face to Face

Face-to-face using show card with response categories

- Is the home in which you live
  - Owned free & clear
  - Owned with a mortgage
  - Rented
  - Occupied under some arrangement

Or when quick doorstep only aural

- Do you own or rent a home?
  - Follow-ups accordingly, e.g. when owns a home
    - Do you have a mortgage or is it owned free and clear

Traditional Design Web

- Web survey:
  - Visual only, but audio potential
  - No interviewer, but intelligent computer system
- Many similarities with mail
- Differences
  - More sequential offering of questions
  - Check all that apply almost standard format
    - Radio buttons (but…)
    - Evidence Christian et al (2008) check-all-that apply not optimal
  - Grids often used for groups of questions
    - What is best visual design?

Traditional Design Web2

- Web survey many similarities with mail plus some additional features, such as,
  - Visual only, but audio potential
  - No interviewer, but intelligent computer system
- Also many differences in question design
  - Special formats
    - Slider bars
    - Drop down menus
  - Open questions influenced by box size, dynamic space (cf Dillman)

Visual Illustrations

- Visual Illustrations are attractive
- May motivate respondent
  - Cover mail survey positive on response (e.g. Dillman's work)
- Help question/word meaning
- BUT: May influence respondent's answer!!
  - Example "How often do you eat out "
  - Illustration 1: couple, candlelight, rose in vase
  - Illustration 2: picture of MacDonald
    - Visuals/illustrations and their influence (Couper, 2007,2009)

Designing for Mixed-Mode Two Cases

Naively Mixing Enhances Measurement Errors

- Different modes have a tradition of different formats, and question format has effect on response distribution
- Consequence: Designers may routinely enhance **unwanted** mode effects in mixed-mode survey
  - E.g. unfolding in one mode, full presentation of all response options in other mode
- What to do?
  - **Design Questionnaire for Mixed-Mode**
    - How?

Design for Mix

- Two Situations:
  - One main method that accommodates the survey situation best
    - Main method is used to maximum potential
    - Other methods auxiliary
      - Example: Nonresponse follow-up
  - Truly multiple mode design
    - All modes are equally important
      - Example: PPSM, International surveys, Longitudinal studies, Respondent is offered choice

Design for Optimal Mix 1

- One Main Method, other methods auxiliary (cf Biemer&Lyberg 2003)
- Identify main method
  - Use main method optimal and to its maximum potential
  - Auxiliary methods designed **equivalent**
    - To avoid measurement error
    - May be perhaps sub-optimal for auxiliary method
      - Example: less response categories
- Note: Dillman et al (2009) coined this ‘mode-enhancement-construction’

Example LFS

- Longitudinal face-to-face & telephone
- Identify main method
  - Main method not necessary first method
    - Main method **telephone**
    - **Face-to-face** auxiliary from longitudinal point of view
- Main design for telephone interview

cf. Biemer & Lyberg, 2003

Example LFS cont

- Design **longitudinal** questions for telephone use
  - Not full potential *face-to-face* used in face-to-face interview
    - No visuals, no show cards
    - Shorter scales, unfolding
    - Open questions
- Design **one-time** recruitment questions for face-to-face use (full potential visual)
- Ensure **data integrity** for longitudinal use!

One Main Method

- Telephone with Face-to-Face Mixes
  - If telephone main method
    - Relatively easy to design mix optimally
      - Interviewer assistance in both modes
      - Do not use the ‘extra’ visual channel in face-to-face
  - If face-to-face main method
    - Absence of visuals makes it more complicated
      - Carefully balance pro and cons
        - Optimize for one? (preferred-mode specific design, aka mode-enhancement construction)
        - Or use ‘uni-mode’ design?
    - Implement a small experiment within one mode if possible!

One Main Method 2

- **Self-Administered Questionnaires and Interviewer Mixes**
- SAQ or Interview Main Method?
  - Complexity of questionnaire
    - Big issue in mixes with paper-mail not in mixes interview with web
  - Are visuals essential?
    - Face-to-face in mix may accommodate visuals, phone does not
      - CAWI-CATI may have problems, CAWI-CAPI not
  - Sensitive questions
    - Social desirability differences, interviewer influence
  - Is interviewer probing essential or not?
    - Paper mail problems, but web can emulate some probes
      - NCES example

Example NCES

- RTI surveys for National Center for Educational Statistics
  - TSMII-paper Wine et al at www.rti.org
  - Original studies were done by telephone
  - Switch to Web with telephone follow-up
    - Highly Internet savvy population
    - So web good choice, but…

Example NCES 2

- Switch to Web with telephone follow-up
  - But, researcher did not want to lose advantages interviewer
    - (Non)Response conversion
    - Clarification, solving inconsistencies, coding, etc
  - Blend best features of both modes

Example NCES 3

- Start with web survey ‘enriched’
  - Offer incentive for early completion
  - Help desk with
    - Specially **trained telephone interviewers**
  - Telephone prompts by phone by trained interviewers help-desk instead of standard e-mail reminders
    - Directly or on answering machine
    - Reminding of early completion incentive

Example NCES 4

- Questionnaire equivalence
  - Field tested
    - Some adaptation to web questionnaire
    - To make situation more equivalent to telephone interview
  - Changes in web-questionnaire
    - No answer option equivalence with interview
      - Continue button in stead of explicit ‘no answer’
      - But generic pop-up after 3 consecutive no answers to remind of importance
      - Key-items redisplayed with tailored text
    - Sensitive questions and SAQ
      - Revision finance items to be less sensitive
  - Help text designed for web also helped interviewers

Truly Multiple Mode Surveys: Modes are Equivalent

Three Approaches in Design

Modes Are Equivalent

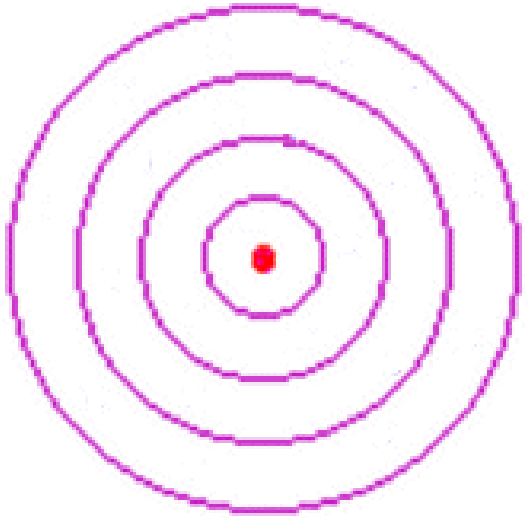
- Three schools of thought
  - Method Maximization
    - Optimize each mode ***separately***
  - Unified Mode Design or Uni-mode design
    - Provide the same stimulus (question format) in each mode, same ***offered*** stimulus
  - Generalized Mode Design
    - Purposively constructing questions to be different to achieve cognitive equivalence, same ***perceived*** stimulus
    - This can be seen as a sophisticated form of mode specific design (cf Dillman et al 2009)

I. **Method Maximization**

- Optimize each method **individually** as far as possible
  - If one method has an extra use it
- Rationale
  - Reduces overall error
  - Best of all possible worlds
- Assumption
  - Same concept is measured in both modes but with different accuracy only
  - Differences between methods only affect random error! (*no systematic bias*)

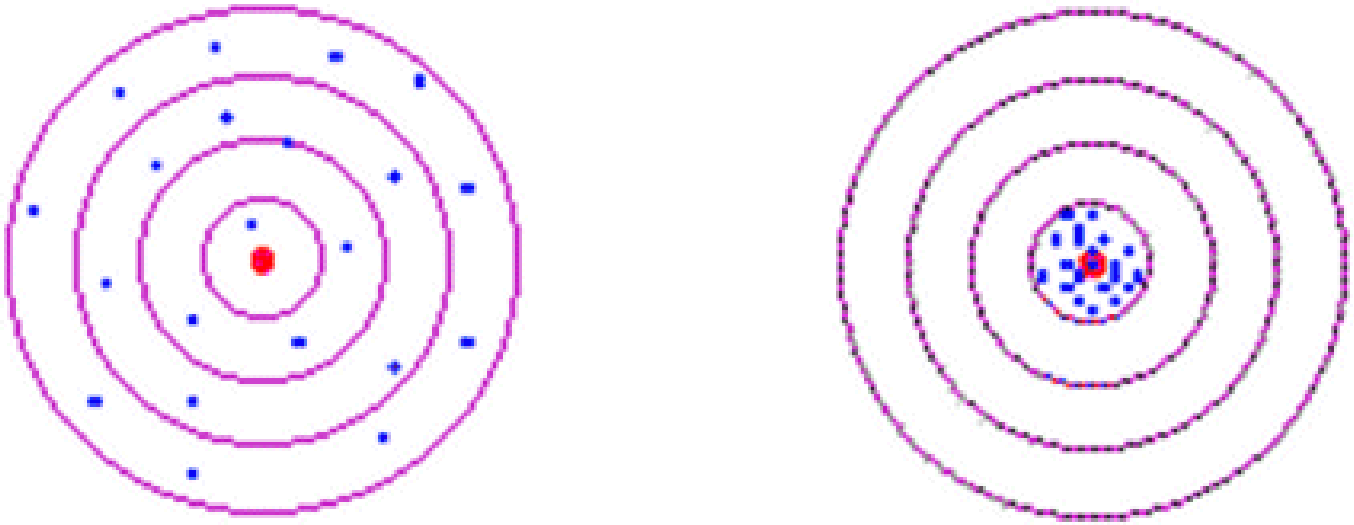
Reliability and Validity

- Imagine an English County Fair and ….

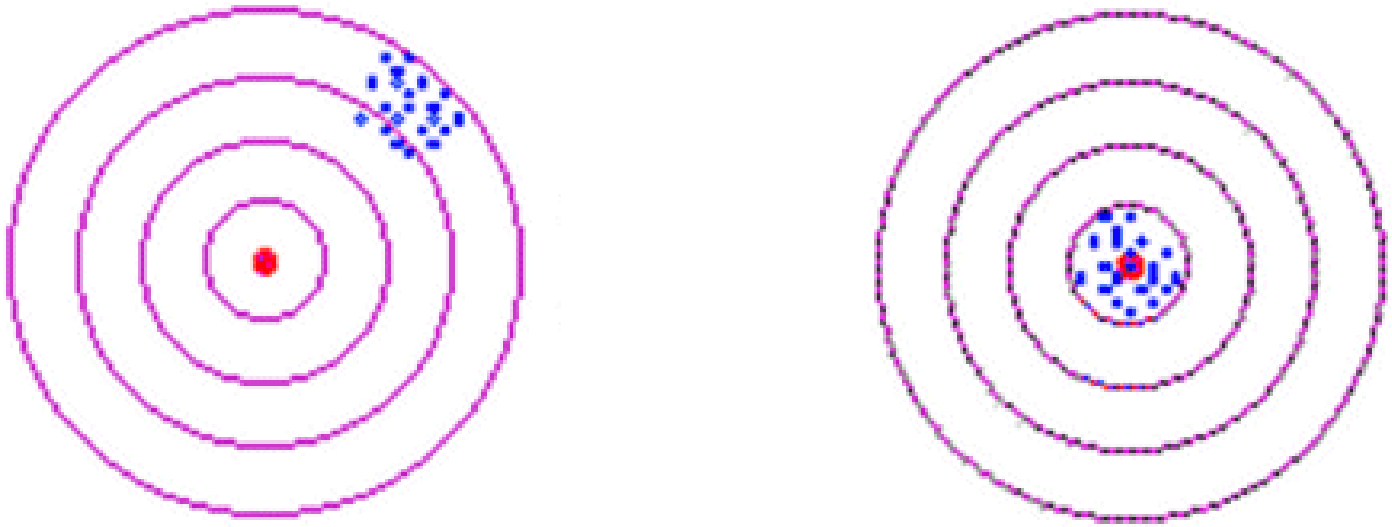


Variance vs Bias or Systematic Error

Low vs. High Reliability



Low vs. High Validity



Method Maximization continued

- Optimize each method individually
- Beware of Assumptions:
  - Differences between methods only affect random error!
  - M_1: $T+e_1$ $\quad e_1 \neq e_2$
  - M_2: $T+e_2$ $\quad e_1, e_2$ *random*
- But is this feasible?
- How about systematic error, bias?
  - **Danger of question format effects**
    - Example: check al that apply vs.. yes/no
    - Example: offer all response categories vs unfolding
- Burden of proof on designer

II. Unified Mode Design

- To minimize data integrity problems Dillman (2000) proposes UNI-mode design for all modes
- **Uni-mode design.** From **unified** or **uniform** mode design; designing questions and questionnaires to ***provide the same stimulus*** in all survey modes in order to reduce differences in the way respondents respond to the survey questions in the different modes.
  - Write and present questions the same or almost the same
  - Same offered stimulus in each mode
- How to do this, see Dillman (2000, 2006)

Uni Mode Design continued

- Unified or UNI-mode design for All Modes
  - Avoid inadvertently changing the basic question structure across modes in ways that change the stimulus.
  - Make all response options the same across modes and incorporate them into the stem of the survey question.
  - Reduce the number of response categories to achieve mode similarity.

(Dillman 2000, 2006, Chapter 6)

Uni Mode Design cont

- Unified or UNI-mode design for all modes
  - Use the same descriptive labels for response categories instead of depending on people's vision to convey the nature of a scale concept.
  - Develop equivalent instructions for skip patterns that are determined by answers to several widely separated items.
  - Reverse the order in which categories are listed in half the questionnaires to avoid recency/primacy effects
  - Evaluate interviewer instructions carefully for unintended response effects and consider their use for other modes.

(Dillman 2000,2006, Chapter 6)

Uni Mode Design cont

- Dillman, 200, 2006, chapter 6:

  "*Avoid question structures that unfold* "

- Comment:
  - Comes from paper mail survey-outlook.
  - One can and may unfold in both modes in CAWI-CATI design
  - Or in CAPI-CATI
  - Or in … any mix (but not in a mix with PAPI: paper mail)

Example UNI Mode Design

Mail, Telephone and Face-to-face interview

- Early attempt De Leeuw 1992, chap 4, p 37
  - http://www.xs4all.nl/~edithl/pubs/disseddl.pdf
- Response options the same across modes
- Same descriptive labels for response categories
- Reduced number of response categories
  - Maximum 7 pushing the limit for phone
  - But, used show cards in face-to-face
    - Equivalent with visual presentation mail
- Used simple open questions
- Interviewer instructions and instructions in mail questionnaire equivalent

In Sum: Uni-mode Design

- Designing for Mixed modes
  - Unified (uni-) mode questions to reduce mode effects
    - Question format
    - Response format
    - Instruction
- Uni-mode design for Mixed modes also implies
  - Besides questionnaire development
  - Questionnaire lay-out
  - Implementation procedures

Meeting the Challenge of Mixed-Mode Surveys

- Unified (uni-) mode questions to reduce mode effects
- Take it one step further, then designing questionnaire ….
- Do not think in traditions.
  - Example 1: RTI-Wine et all, 2006
    - Use telephone interviewers after training for web survey help-desk and for reminders

Do Not Think in Traditions

- Example
  - Question about employment
    - In CATI split up in two questions
      - Were you working for either pay or profit during the week of XXX?
      - If ‘no’ follow-up question: Were you not working for any of the following reasons: you were a student on paid work study, you were self-employed and did not get paid that week, you were on vacation from work or travelling while employed, you were on paid sick leave, personal leave or other temporary leave, or you were on a job that did not pay but had other benefits
      - Follow-up question was to make sure the respondent counted these experiences as employment

Do Not Think in Traditions continued

- Question about employment
  - To improve understanding CATI split up in two questions
    - Were you working for either pay or profit during the week of XXX? + follow-up question
    - Follow-up question was to make sure the respondent counted these experiences as employment
- Paper/Web traditionally no follow-up question
    - Paper/Web often incorporate the definition of employed in an instruction
  - But people do not read instructions and definitions carefully
  - Follow-up questions perhaps be better solution?

Meeting the Challenge of Mixed-Mode Surveys 2

- Step:1 unified (uni-) mode questions to reduce mode effects
- Step 2: Do not think in traditions
- Step 3: From unified to an integrated mode design?

III. Generalized Mode Design

- From unified (uni) mode design to an integrated, generalized mode design

- **Generalized mode design.**
  - Purposively constructing questions and questionnaires to be different in different modes with the goal of ***achieving cognitive equivalence*** of the perceived stimuli, thereby resulting in equivalent answers across modes.
  - The ***same*** offered stimulus is not necessarily the same ***perceived*** stimulus!

Generalized Mode Design continued

- Prerequisites integrated, generalized mode design
  - Designer understands
    - How differences between modes affect the question-answer process
    - How they affect the way respondents *perceive* the question, process the information and select and communicate the response
  - Designer does not think in traditions
  - Burden on the researcher to demonstrate that different questions elicit equivalent responses.

Understand What Happens

- To reach cognitive equivalence
  - Check with respondents, because Medium May Change Meaning
  - Same worded question can be perceived differently in different modes
    - Wording may be adequate in one medium and awkward in another

- Example Pierzchala et al, 2003
  - CATI: Are you {name}?
  - Web: Is your name {name}?

Example Generalized Mode Design

- Christian, Dillman & Smyth (2005)
  - CATI
    - When did you start attending WSU, when did you get your degree, etc
    - Interviewer probed for desired format
  - First Web-design
    - Same question text “When did you start attending WSU” ,“ When was your degree granted”, etc
    - With write in standard programming: mmyyyy

Date Degree Granted: ____ / ____ (MM/YYYY)

    - Too many errors

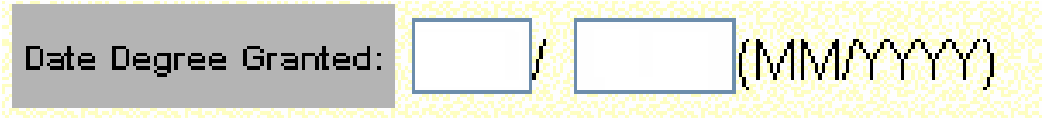
Example continued Generalized Mode Design

- In CATI
  - Interviewer probed for desired format
  - Interviewer is intelligent system
- Web Can Be Intelligent System too!
  - Try to emulate interviewer
- Christian, Dillman & Smyth (2005)
  - http://survey.sesrc.wsu.edu/dillman/papers/Month%20Year%20Technical%20Report.pdf
  - Redesigned size boxes, graphical language, symbols
    - Equivalence needed more than the same question wording!

Example continued Generalized Mode Design

- Web Can Be Intelligent System too!
  - Try to emulate interviewer
  - Christian et al (2005) redesigned size boxes, graphical language, symbols

In stead of



  - Equivalence needed more than the same question wording

Meeting the Mixed Mode Challenge Needs Investment

- Designing for Mixed modes
  - Questionnaire construction
  - Sampling
- Embedding small experiments / mode comparisons
  - Provide data to estimate mode effect
- Using these data for empirically based adjustment
  - Weighting
  - Propensity score adjustment

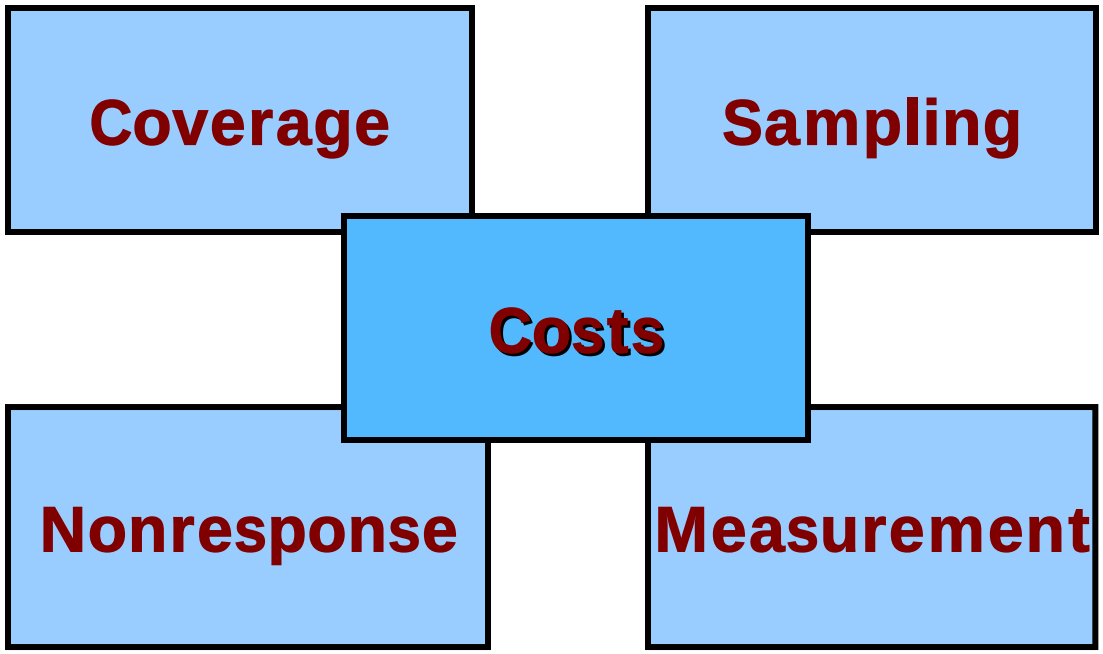
Embedded Experiments and/Adjustment

Build in overlap	**Method 1**	**Method 2**
Group X	**Main Data Collection**	**Some Data**
Group Y	**Some Data**	**Main Data Collection**

Logistics
Mixed Mode Surveys

Why Mixing Modes?

Balance Costs & Errors



Consequences

Mixed Mode Strategy

Coverage	**Costs**
Nonresponse	**Measurement**
Sampling	**Logistics**

Logistics

- Need for field reports
  - Not much literature
  - But, increasing
    - E.g., US Bureau of the Census website
- Lessons learned
  - Mixed mode studies
    - Past mode changes
    - International studies
    - Software developers
  - Emerging experience
    - Conference papers
    - E.g., methodological mixed-mode projects UK, Holland, USA

Main Issues

- In-House Expertise
- Communication
- Implementation and Timing
- Data Processing
- Quality Control
- Para information
- Cost

- Many of these issues well-known in International and Comparative Research
  - See for example CSDI International Workshop on Comparative Survey Design and Implementation &Cross-cultural survey guidelines at http://ccsg.isr.umich.edu/

Total Quality Approach

- Total survey design
- Document information
- Disseminate information
- Information on:
  - Survey process & quality
  - Methodologies
  - Software
  - Para data

Future

"Survey organizations are going to have to change dramatically in some ways in order to do effective surveys as we bring new technologies online and still use our other technologies where they work"

Dillman, 2000

Mixed Mode Surveys

- Survey research history of changes
  - Forced by changes in society and technology
  - Increased knowledge
- Remember first face-to-face interviews
  - Short & simple questions
  - Later one of the most flexible methods ☺
- Mixed mode has many challenges
  - We will meet those and learn ☺ ☺ ☺

Suggested Readings

- Introduction Mixed-Mode
  - Edith D. De Leeuw (2005). To mix or not to mix data collection methods in surveys. JOS, Journal of Official Statistics, 21,2, 233-255 (also available on www.jos.nu)
- On quality, data collection, writing questions, online research, mixed-mode, analysis
  - De Leeuw, Hox & Dillman (2008). International Handbook of survey Methodology. New York: Lawrence Erlbaum/Psychology Press, Taylor and Francis Group

Suggested Websites

- Don A. Dillman's homepage
  - http://survey.sesrc.wsu.edu/dillman/
    - Papers
- Edith de Leeuw homepage
  - http://www.xs4all.nl/~edithl/
    - Additional material and readings accompanying the International Handbook of survey methodology
- University of Michigan-ISR
  - http://ccsg.isr.umich.edu/
    - Guidelines for best practice in cross-cultural studies
- US Bureau of the Census
  - http://www.census.gov/srd/mode-guidelines.pdf
    - Guidelines on developing demographic questions for administration in different modes

Suggested Websites 2

- Journal of Official Statistics
  - http://www.jos.nu
    - All articles that appeared in JOS online available, simple search menu
- Survey Practice (AAPOR)
  - http://www.surveypractice.org
    - Short practical articles
    - Autumn 2009, theme propensity score adjustment
- European Survey Research Association
  - http://www.surveymethodology.eu/journal/
    - Online journal on survey methodology