Introduction to Survey Design

Dr. Mark Rouleau
Social Sciences Department
Environmental and Energy Policy Program
What is a survey and why would I ever want to do one?
What is a survey and why would I ever want to do one?
A Proper Survey

• **Purpose:**
  • To produce a statistic or set of statistics (quantitative or numerical descriptions that concisely summarize a certain aspect of a particular group of interest)

• **Method:**
  • The main way to collect data is to ask people questions
  • The main analysis technique uses statistical tests to interpret people’s answers
  • The main analysis objective is to infer from a sample to a larger population
What is a survey and why would I ever want to do one?
What is a survey and why would I ever want to do one?
Reason 1: To efficiently describe the characteristics of a particular group (e.g., unemployment rate, public opinion, customer satisfaction)
Reason 2: To figure out why people think and behave the way that they do
What is the best way to design a survey?
There is NO one best way!
What are all the ways I can go wrong when designing a survey?
There are many!
Total Survey Error

- **Construct**: $\mu_i$
- **Measurement**: $Y_i$
- **Response**: $y_i$
- **Edited Response**: $\tilde{y}_{ip}$
- **Survey Statistic**: $\bar{y}_{prw}$

- **Target Population**: $\bar{Y}$
- **Sampling Frame**: $\bar{Y}_f$
- **Sample**: $\bar{Y}_s$
- **Respondents**: $\bar{Y}_r$
- **Post-Survey Adjustments**: $\bar{Y}_w$
Total Survey Error
Total Survey Error
Total Survey Error
Total Survey Error

**MEASUREMENT**

- **Construct** $\mu_i$
- **Measurement** $Y_i$
- **Response** $Y_i$
- **Edited Response** $\tilde{Y}_{ip}$

**REPRESENTATION**

- **Target Population** $\bar{Y}$
- **Sampling Frame** $\bar{Y}_s$
- **Sample** $\bar{Y}_s$
- **Respondents** $\bar{Y}_r$
- **Post-Survey Adjustments** $\bar{Y}_w$

**Survey Statistic** $\bar{Y}_{prw}$
Total Survey Error

1. **Construct** $\mu_i$
2. **Measurement** $Y_i$
3. **Response** $Y_i$
4. **Edited Response** $\tilde{Y}_{ip}$
5. **Survey Statistic** $\hat{Y}_{prw}$

**Target Population** $\bar{Y}$

**Sampling Frame** $\tilde{Y}$

**Sample** $\tilde{Y}_s$

**Respondents** $\tilde{Y}_r$

**Post-Survey Adjustments** $\tilde{Y}_w$
Total Survey Error
Total Survey Error

- **Target Population**: $\bar{Y}$
- **Sampling Frame**: $\bar{Y}$
- **Sample**: $\bar{Y}_s$
- **Respondents**: $\bar{Y}_r$
- **Post-Survey Adjustments**: $\bar{Y}_w$

**Survey Statistic**: $\bar{Y}_{prw}$

**Coverage Error**

- **Edited Response**: $\bar{Y}_{ip}$
- **Response**: $\bar{Y}_i$
- **Measurement**: $\bar{Y}_i$
- **Construct**: $\mu_i$
Total Survey Error

- Target Population $\bar{Y}$
- Sampling Frame $\bar{Y}$
- Sample $\bar{Y}_s$
- Respondents $\bar{Y}_r$
- Post-Survey Adjustments $\bar{Y}_w$

- Survey Statistic $\bar{Y}_{prw}$
- Edited Response $\bar{Y}_{e}$
- Response $\bar{Y}_i$
- Measurement $\bar{Y}_i$
- Construct $\mu_i$

Coverage Error
Total Survey Error

- **Survey Statistic** $\bar{Y}_{prw}$
  - **Edited Response** $\bar{Y}_{ip}$
  - **Response** $Y_i$
  - **Measurement** $Y_i$
  - **Construct** $\mu_i$

- **Target Population** $\bar{Y}$
- **Sampling Frame** $\bar{Y}_f$
- **Sample** $\bar{Y}_s$
- **Respondents** $\bar{Y}_r$
- **Post-Survey Adjustments** $\bar{Y}_w$

- **Coverage Error**
- **Sampling Error**
Total Survey Error

Target Population $\bar{Y}$

Sampling Frame $\bar{Y}_s$

Sample $\bar{Y}_s$

Respondents $\bar{Y}_r$

Post-Survey Adjustments $\bar{Y}_w$

Coverage Error

Sampling Error

Survey Statistic $\bar{Y}_{prw}$

Edited Response $\bar{Y}_{ip}$

Response $\bar{Y}_i$

Measurement $\bar{Y}_i$

Construct $\mu_i$
Survey Statistic

\[ \hat{Y}_{prw} \]

Edited Response

\[ \hat{Y}_{ip} \]

Response

\[ \hat{Y}_i \]

Measurement

\[ \hat{Y}_i \]

Construct

\[ \mu_i \]

Target Population

\[ \hat{Y} \]

Sampling Frame

\[ \hat{Y}_s \]

Sample

\[ \hat{Y}_r \]

Respondents

\[ \hat{Y}_r \]

Post-Survey Adjustments

\[ \hat{Y}_w \]

Total Survey Error

Coverage Error

Sampling Error

Nonresponse Error
Total Survey Error

Survey Statistic: \( \hat{Y}_{prw} \)

Edited Response: \( \hat{Y}_{ip} \)

Response: \( \hat{Y}_{i} \)

Measurement: \( Y_i \)

Construct: \( \mu_i \)

Target Population: \( \bar{Y} \)

Sampling Frame: \( \bar{Y}_s \)

Sample: \( \bar{Y}_s \)

Respondents: \( \bar{Y}_r \)

Post-Survey Adjustments: \( \bar{Y}_w \)

Coverage Error

Sampling Error

Nonresponse Error

Adjustment Error
Total Survey Error

Target Population \( \bar{Y} \)

Sampling Frame \( \bar{Y}_s \)

Sample \( \bar{Y}_s \)

Respondents \( \bar{Y}_r \)

Post-Survey Adjustments \( \bar{Y}_w \)

Coverage Error

Sampling Error

Nonresponse Error

Adjustment Error
Total Survey Error

1. Target Population $\bar{Y}$
2. Sampling Frame $\bar{Y}_s$
3. Sample $\bar{Y}_s$
4. Respondents $\bar{Y}_r$
5. Post-Survey Adjustments $\bar{Y}_w$

Coverage Error

Sampling Error

Nonresponse Error

Adjustment Error

Coverage Error

Survey Statistic $\bar{Y}_{prw}$

Edited Response $\bar{Y}_{ip}$

Response $\bar{Y}_i$

Measurement $\bar{Y}_i$

Construct $\mu_i$

Coverage Error
Survey Statistic

Edited Response

Total Survey Error

Coverage Error

Construct $\mu_i$

Measurement $Y_i$

Response $Y_i$

Edited Response $\tilde{Y}_{ip}$

Target Population $\bar{Y}$

Sampling Frame $\bar{Y}_f$

Sample $\bar{Y}_s$

Respondents $\bar{Y}_r$

Post-Survey Adjustments $\bar{Y}_w$

Survey Statistic $\bar{Y}_{prw}$

Coverage Error

Sampling Error

Nonresponse Error

Adjustment Error
Survey Statistic

Edited Response $\hat{\gamma}_{ip}$

Response $\gamma_i$

Measurement $Y_i$

Construct $\mu_i$

Target Population $\bar{Y}$

Sampling Frame $\bar{Y_f}$

Sample $\bar{Y_s}$

Respondents $\bar{Y_r}$

Post-Survey Adjustments $\bar{Y_w}$

Total Survey Error

Coverage Error

Measurement Error

Sampling Error

Nonresponse Error

Adjustment Error
Total Survey Error

- Construct: $\mu_i$
- Measurement: $Y_i$
- Response: $Y_i$
- Edited Response: $\tilde{Y}_{ip}$
- Target Population: $\tilde{Y}$
- Sampling Frame: $\tilde{Y}_s$
- Sample: $\tilde{Y}_s$
- Respondents: $\tilde{Y}_r$
- Post-Survey Adjustments: $\tilde{Y}_w$
- Survey Statistic: $\tilde{Y}_{prw}$

Error Components:
- Coverage Error
- Sampling Error
- Nonresponse Error
- Adjustment Error
- Measurement Error
- Processing Error
Survey Statistic $\hat{Y}_{prw}$

**Coverage Error**

**Measurement Error**

**Processing Error**

**Target Population** $\bar{Y}$

**Sampling Frame** $\bar{Y}$

**Sample** $\bar{Y}_s$

**Respondents** $\bar{Y}_r$

**Post-Survey Adjustments** $\bar{Y}_w$

**Total Survey Error**

**Construct** $\mu_i$

**Measurement** $Y_i$

**Response** $y_i$

**Edited Response** $\hat{y}_{ip}$
Any Questions?