# Challenges in Conducting Nonresponse Bias Analyses for Business Surveys: A Perspective from the U.S. Census Bureau

Katherine Jenny Thompson
Statistical Methods and Sample Design Staff
Office of Statistical Methods and Research for Economic Programs

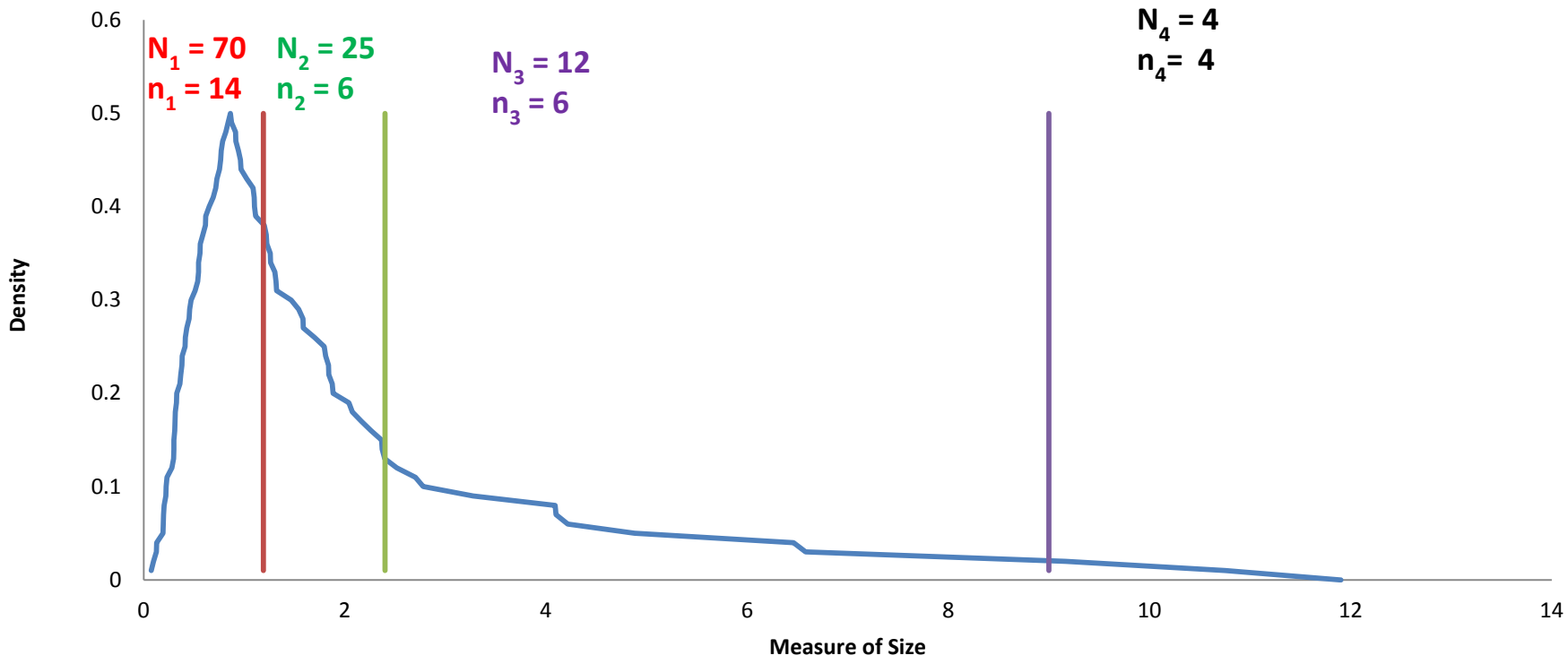
*Presented at the Fourth International Conference on Establishment Surveys*
*Invited Overview Lectures – Session 6*

United States Census Bureau

# SET THE SCENE

United States Census Bureau

# Fictional Stratified Business Population




United States Census Bureau

# Underlying Assumptions

| Stratum Number | Population Units ($N_h$) | Sampled Units ($n_h$) | Stratum Parameters |
|---|---|---|---|
| 1 | 70 | 14 | $\mu_1, \sigma_1$ |
| 2 | 25 | 6 | $\mu_2, \sigma_2$ |
| 3 | 12 | 6 | $\mu_3, \sigma_3$ |
| 4 | 4 | 4 | $\mu_4, \mu_5, \mu_6, \mu_7$ |

- $\mu_1 \neq \mu_2 \neq \mu_3 \neq \{\mu_4, \mu_5, \mu_6, \mu_7\}$ (or why stratify???)
- The "certainty" units in stratum 4 are unique
  - Self-representing
  - There aren't very many "large" units in the population
- The auxiliary variable used for stratification variable is positively correlated with the survey characteristic(s) of interest

United States Census Bureau

# Design Stage

| Population Units ($N_h$) | Sampled Units ($n_h$) | Stratum Parameters | Sampling Rate $n_h/N_h$ | Design Weight |
|---|---|---|---|---|
| 70 | 14 | $\mu_1, \sigma_1$ | 14/70 = 0.2 | 70/14 = 5 |
| 25 | 6 | $\mu_2, \sigma_2$ | 6/25 = 0.24 | 25/6 = 4.17 |
| 12 | 6 | $\mu_3, \sigma_3$ | 6/12 =0.50 | 12/6 = 2 |
| 4 | 4 | $\mu_4, \mu_5, \mu_6, \mu_7$ | 4/4 = 1.00 | 4/4 = 1 |

United States Census Bureau

# “Reality”/Challenges

| Population Units ($N_h$) | Sampled Units ($n_h$) | Respondent Units ($r_h$) | Sampling Rate $n_h/N_h$ | Response Rate $r_h/n_h$ |
|---|---|---|---|---|
| 70 | 14 | 7 | 14/70 = 0.2 | 7/14 = 50% |
| 25 | 6 | 3 | 6/25 = 0.24 | 3/6 = 50% |
| 12 | 6 | 6 | 6/12 =0.50 | 6/6 = 100% |
| 4 | 4 | 3 | 4/4 = 1.00 | 3/4 = 75% |

- Larger units very different from smaller units
- Larger units more likely to respond than smaller units
- “Representative” subsample?

United States Census Bureau

# In General

- Business populations are positively skewed
- Business data may not be positively skewed
  - Sales, Payroll, Employment
    - Positively skewed, non-negative by definition
  - Income
    - Real valued (bell shaped curve)
- Estimates of interest – generally TOTALS

United States Census Bureau

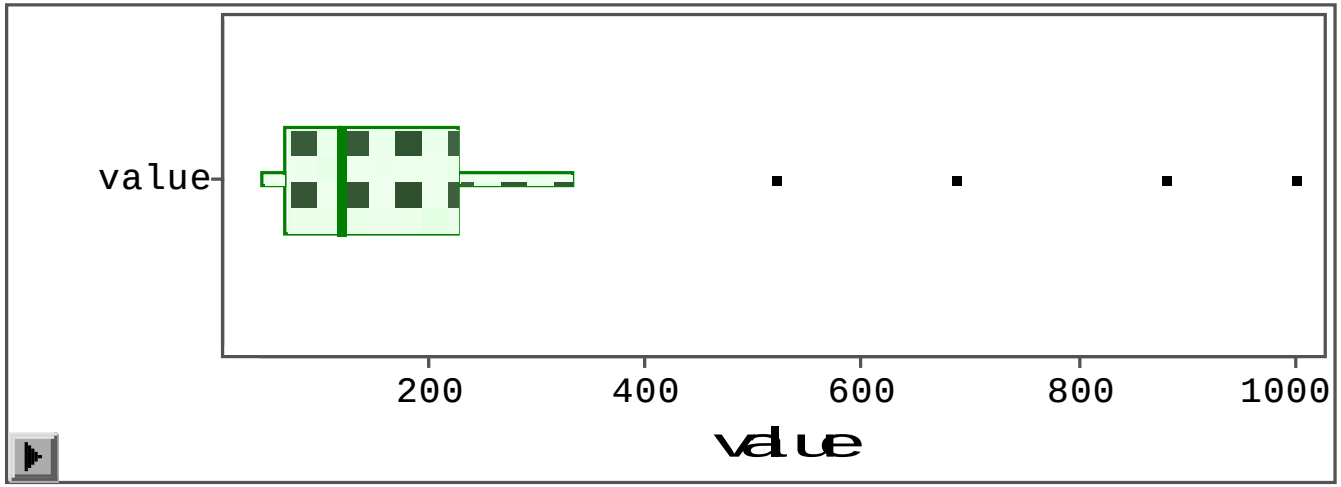
# Not All Units Are “Equal”

- Unit response rate
  - Unweighted proportion of responding units
  - One per program/survey

- Total Quantity Response Rate
  - weighted proportion of an estimate reported by responding units and from equivalent quality sources
  - 1 rate per key item/program (can be several)

United States Census Bureau

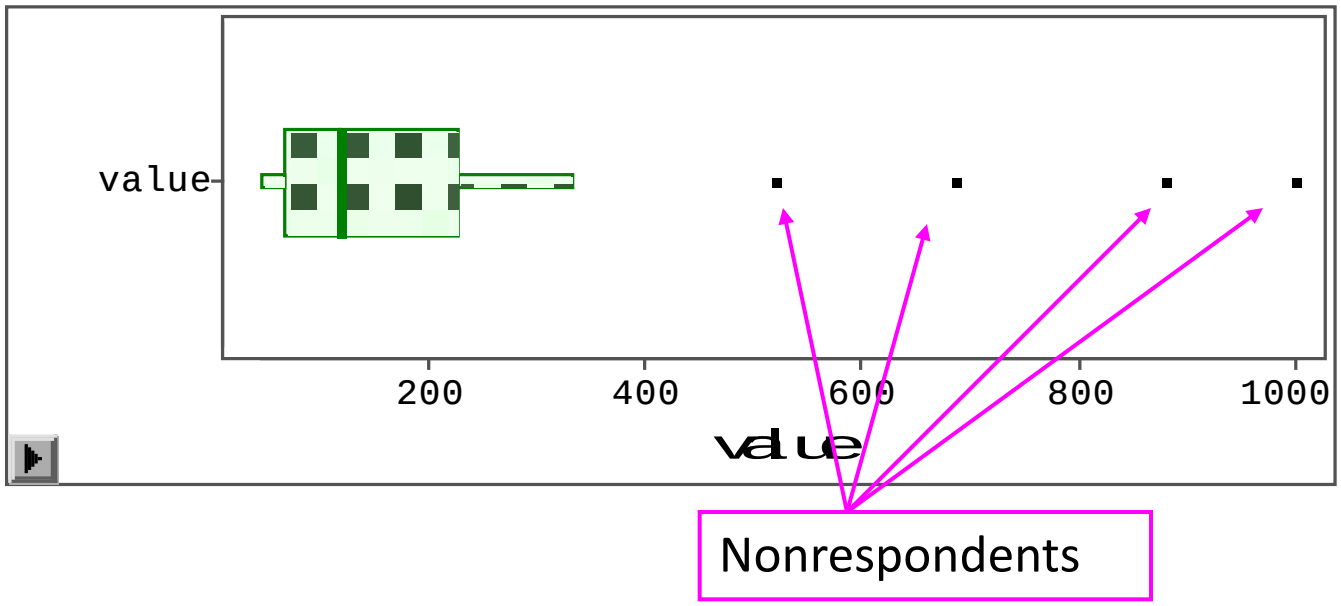
# Return to Fictional Example




- Skewed business population
- 30 sampled units
- 4 units provide ≈ 47% of TOTAL Value
  - Certainty units (census)
  - Noncertainty units (sampled)

United States Census Bureau

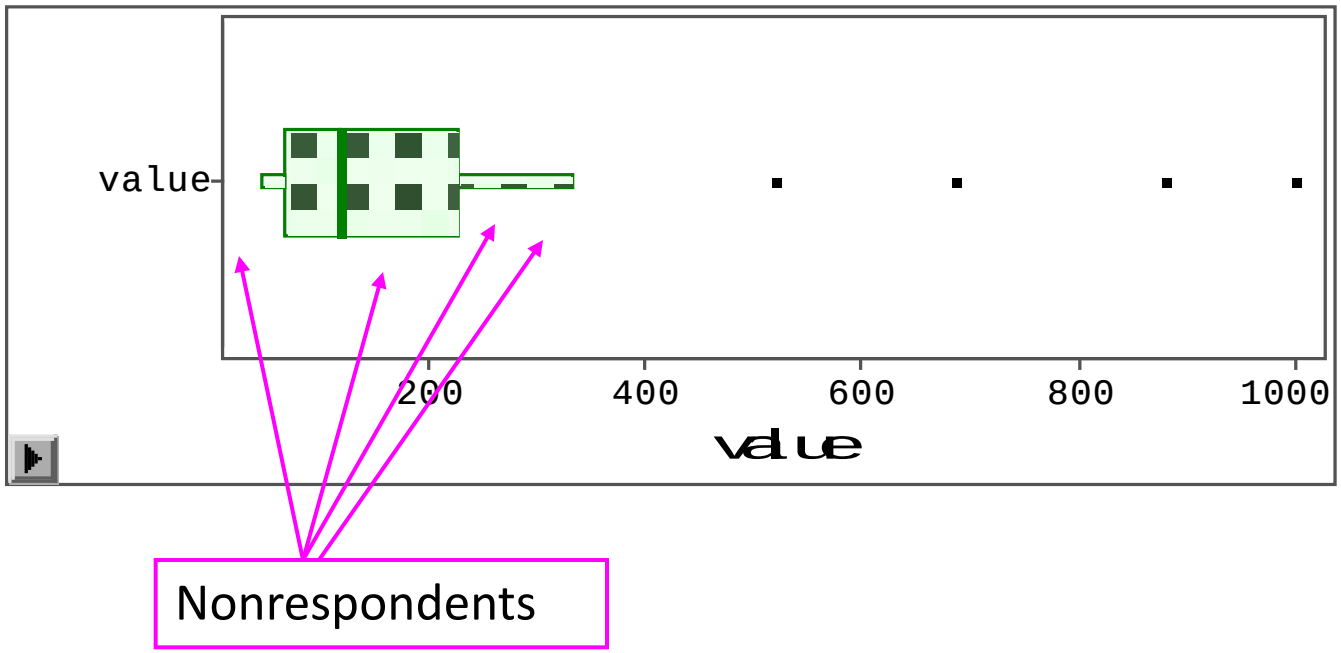
# Not All Units Are “Equal”




- Unit response rate = 26/30 ≈ 87%
- Total Quantity Response Rate (Value) = 3423/6507 ≈ 53%

United States Census Bureau

# Not All Units Are Equal




- Unit response rate = 26/30 ≈ 87%
- Total Quantity Response Rate (Value) = 6023/6507 ≈ 93%

United States Census Bureau

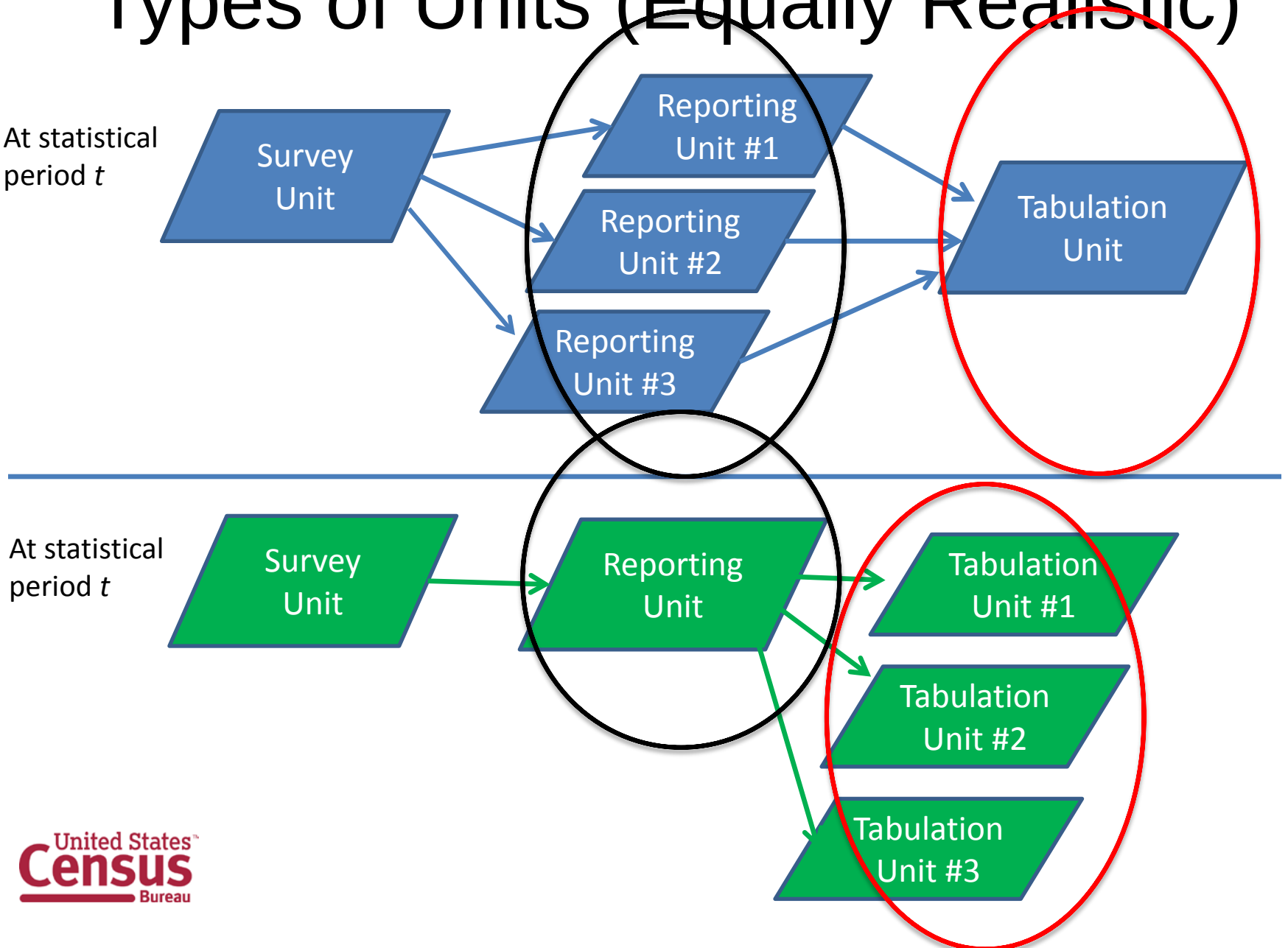
# Types of Units

| Type of Unit | Definition | Established By |
|---|---|---|
| Survey Unit | An entity selected from the underlying statistical population of similarly-constructed units | Frame |
| Reporting Unit | An entity from which data are collected | Sampled unit(s) providing the data |
| Tabulation Unit | An entity that houses the data used for estimation/tabulation | Program managers and/or methodologists |

Simplest Case


United States Census Bureau

# Types of Units (Equally Realistic)




United States Census Bureau

# One More Challenge

| Provide the number of employees working at this establishment on March 15, 2011 | |
|---|---|
| Category | Value |
| 1a. Production workers | <To Be Collected> |
| 1b. Other employees | "" |
| 1c. Total workers | "" |

Detail items

Total item

- Total items collected from each respondent
  - Administrative data may be available for some units
- Detail items
  - Depend on category (industry)
  - May not be available from respondent

United States Census Bureau

Getting More Formal

# NONRESPONSE BIAS ANALYSIS CONSIDERATIONS

United States Census Bureau

# The Issues

- Administrative Standards/Requirements

- Statistical Issues
  - Data quality
  - Representatives of respondent set
    - Response mechanism
    - Mitigation strategies
      - Behavioral protocols
      - Statistical adjustments
  - Impact of nonresponse bias on estimates

United States Census Bureau

# Nonresponse Bias (Total)

SAMPLE (*n* units)

$$\hat{Y} = N\bar{y}_n = N\left(\sum_{i=1}^{n} y_i / n\right)$$

RESPONDENTS (*r* units)

$$\hat{Y}_r = N\left(\sum_{i=1}^{r} y_i\right) = \frac{N}{n}\left(r\bar{y}_r\right)$$

Observed (respondent data)

NONRESPONDENTS (*nr* units)

$$\hat{Y}_{nr} = N\left(\sum_{i=r+1}^{n} y_i\right) = \frac{N}{n}\left(nr(\bar{y}_{nr})\right)$$

Not observed (nonrespondent data)

United States Census Bureau

# Some (Simplified) Formulae

- Bias of unadjusted total (respondent data)

$$B = N(\bar{y}_r) - N(\bar{y}_n)$$

$$= N\left\{(\bar{y}_r) - \left[\left(\frac{r}{n}\right)(\bar{y}_r) + \left(\frac{nr}{n}\right)(\bar{y}_{nr})\right]\right\}$$

$$= N\left\{\frac{nr}{n}(\bar{y}_r - \bar{y}_{nr})\right\}$$

- Unbiased if
  - Complete response
  - Mean value per respondent = mean value per nonrespondent $\Rightarrow$ Items may be affected by nonresponse bias differently

United States Census Bureau

# Stochastic View

$$E(\bar{y}_r - \bar{y}_n) = \frac{\sigma_{yp}}{\bar{p}}$$

Where $\sigma_{yp}$ is the covariance between the survey variable, *y*, and the response propensity, *p*

- What mechanisms produce the covariance?

Source: Tucker, Dixon, and Cantor (2007). Measuring the Effects of Unit Nonresponse in Establishment Surveys. Third Conference on Establishment Surveys (ICES III)

United States Census Bureau

# ARE YOU READY TO CONDUCT A NONRESPONSE BIAS ANALYSIS?

United States Census Bureau

# The “Structured” Detective Work

1. Determine the potential for nonresponse bias
2. Examine the extent of nonresponse bias
3. Understand the response mechanism
4. Mitigate the nonresponse bias

Repeat, repeat, repeat…

United States Census Bureau

Administrative requirements…

# DETERMINE THE POTENTIAL FOR NONRESPONSE BIAS

United States Census Bureau

# In Case You Weren't Listening to Brian

**OMB Standard 1.3 (Survey Response Rates)**

- Agencies must design the survey to achieve the highest practical rates of response…to ensure that survey results are representative of the target population so that they can be used with confidence to inform decisions.
- ***Nonresponse bias analyses must be conducted if unit or item response rates suggest the potential for bias to occur.***

United States Census Bureau

# More Definitions and Concepts

- Respondent
- Reported Data
- Equivalent Quality, but Not Reported Data
- Required Data Items

United States Census Bureau

# What is a Respondent?

A respondent is an *eligible* unit for which

- an attempt was made to collect data;
- the unit belongs to the target population;
- the unit provided *sufficient data* to be classified as a response.

[Source: Census Bureau Standard: Response Rate Definitions, Version 1.0.]

United States Census Bureau

# Reported Vs. Equivalent Quality

- **Reported Data** are directly received from the reporting unit for the survey questionnaire in the data collection period.

- **Equivalent Quality (but Not Reported) Data** are indirectly received from the reporting unit or the tabulation unit in the data collection period

United States Census Bureau

# Response Rates (Refresher)

- **Unit Response Rate (URR)** – the rate of the total unweighted number of "responding" units to the total number of sampled units eligible for tabulation.
  - 1 rate per program

- **Total Quantity Response Rate (TQRR)** – the weighted proportion of a key estimate reported by responding units and from equivalent quality sources.
  - 1 rate per key item/program (can be several)

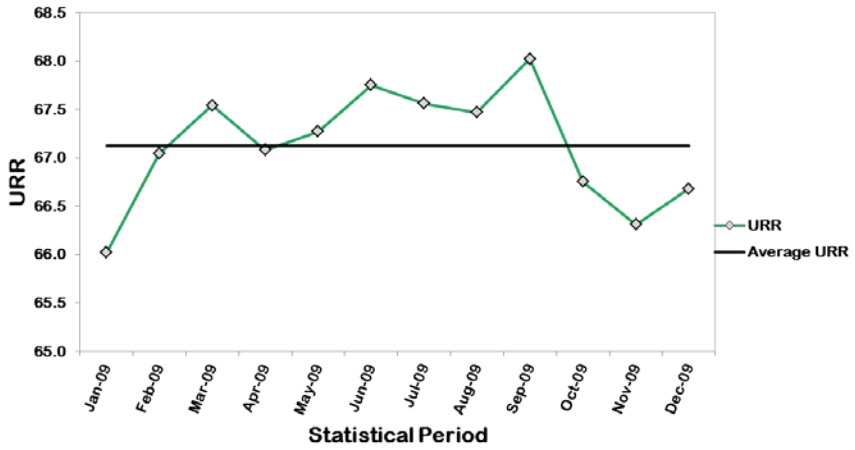
United States Census Bureau

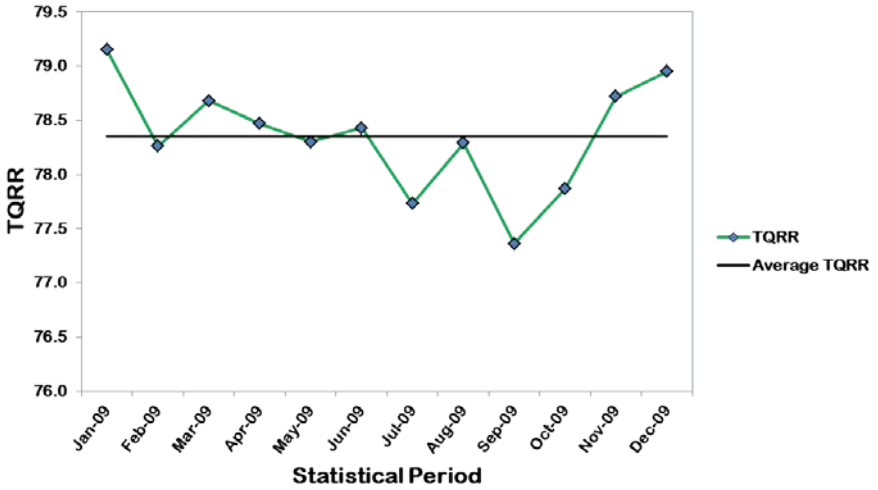
# Response Rate Analyses

- Compute the rates
  - For program
  - Different (meaningful) subgroups
    - Sector/Industry
    - Certainty/Noncertainty status
- Analyze the rates
- *Monitor the rates over time*

United States Census Bureau

# URR and TQRR (Sales)

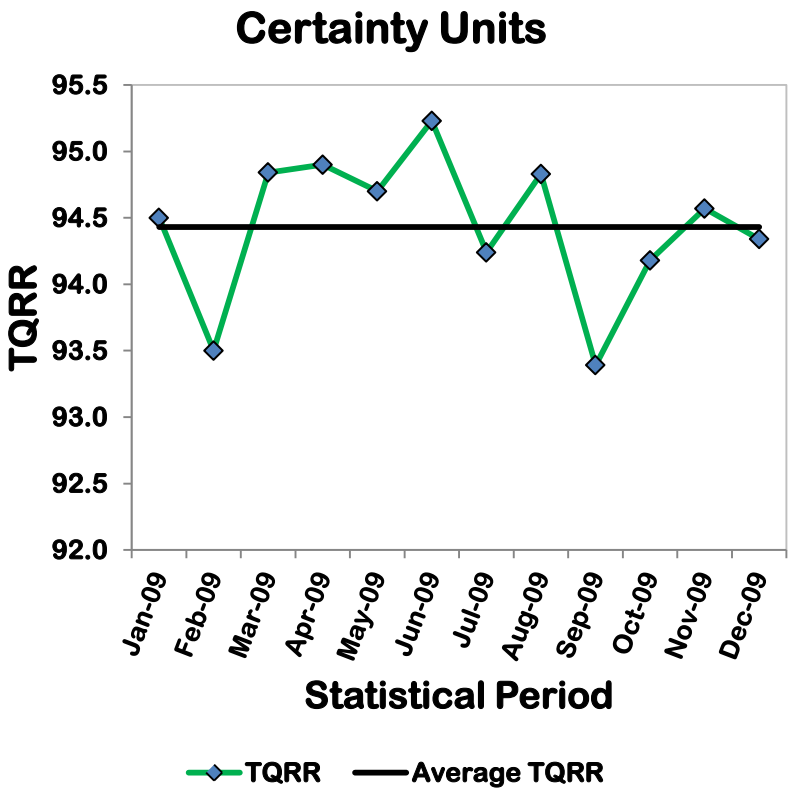


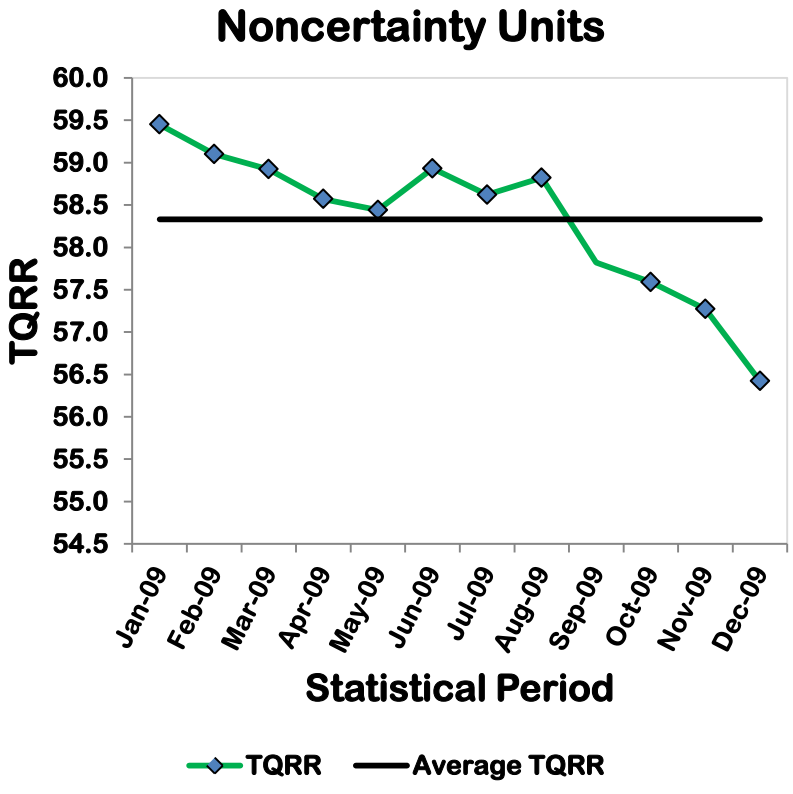




Source: Monthly Retail Trade Survey

United States Census Bureau

# TQRR for Sales: By Certainty and Noncertainty Status




United States Census Bureau

Source: Monthly Retail Trade Survey

Response Rates as Process Control Measures

# DIGRESSION #1

United States Census Bureau

# Response Rates as Quality Measures

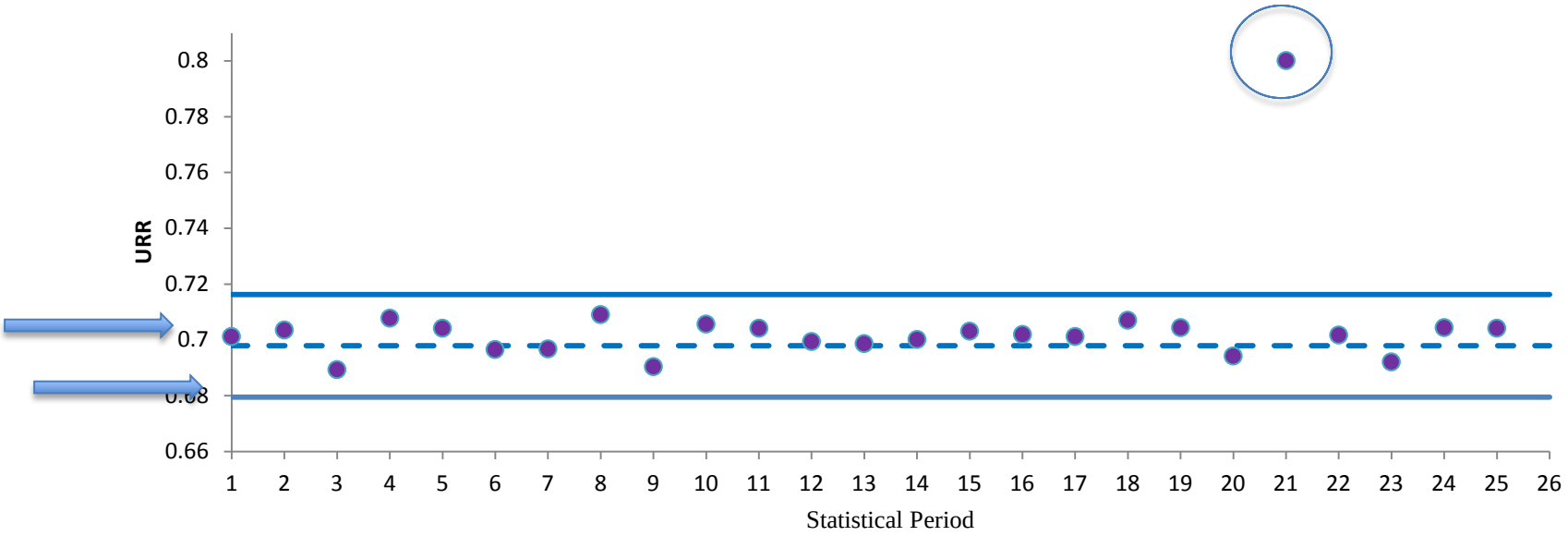
- Assess the state of the process
- Understand the response rate process capability
- Monitor over time

González, Y. and Oliver, B. (2012). Producing Control Charts to Monitor Response Rates for Business Surveys in the Economic Directorate of the U.S. Census Bureau. FCSM Research Conference.

Thompson, Katherine J. and Oliver, B. (2012). Response Rates in Business Surveys: Going Beyond the Usual Performance Measure . Journal of Official Statistics.

United States Census Bureau

# P-chart for URR (Fictional Series)




- Time-series plot of URR
- Centerline (process average)
- Control Limits

United States Census Bureau

Understanding the Response Mechanism

# DIGRESSION #2

United States Census Bureau

# Response Mechanisms

- Ignorable
  - Model can “explain” the response mechanism
  - Auxiliary variable (covariate) that “explains” nonresponse is not the studied variable(s)
  - Nonresponse can be “ignored” *after the estimates are adjusted with respect to the model*

- Nonignorable
  - Probability of response depends on studied variable (nonresponse systematic, not random)
  - Difficult (if not impossible) to correct through adjustment

United States Census Bureau

# Ignorable Response Mechanism #1: Missing Completely at Random (MCAR)

- Uniform
- Probability of response equal and independent for all units in sample
- Not terribly realistic

United States Census Bureau

# Ignorable Response Mechanism #2: Missing at Random (MAR)

- Probability of response depends on auxiliary variable, not directly related to characteristic(s) of interest

- Probability of response differs by adjustment cell
  - Adjustment cells may or may not be strata
  - Same probability of response for all units in an adjustment cell

United States Census Bureau

# Ignorable Response Mechanism #3: Covariate – Dependent Missing

- Probability of response depends on auxiliary variable, not characteristic(s) of interest (more general formulation)
  - $P^{(}M_i \mid Y,X) = P(M_i \mid X)$
- Probability of response differs by unit
- Probability of response **predicted** by level of auxiliary variable

United States Census Bureau

# Options (Ignorable Unit Nonresponse)

- Reweighting
  - Divide sample into weighting cells
  - Increase respondents' weights to represent sample

- Imputation
  - Divide sample into imputation cells
  - Create complete records

United States Census Bureau

# Notation (Reweighting)

Adjustment factor  Adjustment Cell  Sampling weight

$$w_i^{p*} = f_i^{\,p} \times w_i$$

$f_i \geq 1$ if $J_i^p = 1$ (unit $i$ responded)

$f_i = 0$ if $J_i^p = 0$ (unit $i$ did not respond)

$$\hat{Y}_{adj} = \sum_p \sum_{i \in p} f_i^{\,p} J_i^p w_i y_i = \sum_p \sum_{i \in p} w_i^{p*} y_i$$

United States Census Bureau

# Notation (Imputation)

Imputation factor  Auxiliary variable (covariate)

$$y_i^{p*} = f_i^{p*} \times x_i$$

$$\hat{Y}_{imp} = \sum_p \sum_{i \in p} w_i y_i J_i^p + \sum_p \sum_{i \in p} w_i y_i^{p*} (1 - J_i^p)$$

United States Census Bureau

# Reweighting (Mitigation Strategy 1)

**Advantages**

- Statistically valid if response model is correct
- Easy to correctly compute variance estimates of totals
- Preserves multivariate relationships between items

**Disadvantages**

- Can increase the variance
- Not valid if response set is not a representative random sample

United States Census Bureau

# Imputation (Mitigation Strategy 2)

**Advantages**

- Can develop “best” predictive model for each item *y*
- Can set a hierarchy of imputation methods to attempt in expected order of reliability
- Ratio imputation model parameters can be B.L.U.E.

**Disadvantages**

- Predictive model may not be very good
- Does not preserve multivariate relationships
- Difficult to correctly compute variance estimates (especially if composite imputation is used)

United States Census Bureau

# Adjustment Cells

- Auxiliary variables (categorical)
  - Available data for all sampled units
  - "Sufficient" respondents in each category
  - May be recoded continuous variables
- Hope
  - Auxiliary variable is correlated with response propensity; **and**
  - Auxiliary variable is correlated with outcome

United States Census Bureau

# Adjustment Effects on Mean or Totals

| Auxiliary Variable Association with Response (Propensity) | | Auxiliary Variable Association with Outcome (Prediction) | |
|---|---|---|---|
| | | Low | High |
| | Low | Bias:--<br>Variance:-- | Bias:--<br>Variance: ↓ |
| | High | Bias:--<br>Variance: ↑ | Bias: ↓<br>Variance: ↓ |

Little, R.J. and Vartivarian, S. (2005). Does Weighting for Nonresponse Increase The Variability of Survey Means? *Survey Methodology*, **31**, pp. 161-168.

United States Census Bureau

# Certainty Units

Proposition: Nonresponding certainty units always cause some nonresponse bias

- Unique quality → guaranteed inclusion in sample
- Nonignorable nonresponse

United States Census Bureau

# At Last…the Promised Topic

- This talk is titled “Challenges in Conducting Nonresponse Bias Analyses for Business Surveys: A Perspective from the U.S. Census Bureau”

- Insights into “tried and true” (aka) methods
  - Cites older analyses (before 2010)
  - Nothing “cutting edge”

United States Census Bureau

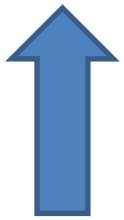
Studying internal variation within the data collection

# COMPARE RESPONDENTS AND NONRESPONDENTS ON FRAME VARIABLES

United States Census Bureau

# Adjustment Effects on Mean or Totals

| Auxiliary Variable(s) Association with Response (Propensity) | | Auxiliary Variable Association with Outcome (Prediction) | |
|---|---|---|---|
| | | Low | High |
| | Low | Bias:--<br>Variance:-- | Bias:--<br>Variance: ↓ |
| | High | Bias:--<br>Variance: ↑ | Bias: ↓<br>Variance: ↓ |



United States Census Bureau

# Response Propensity Analysis

- Are categorical variables used to form adjustment cells predictive of unit nonresponse?
  - Logistic regression analysis
- Could other variable be used to form adjustment cells?
- Do response propensities differ between cells?
  - Assess by comparing unit response rates

United States Census Bureau

# Are Categorical Variables Used To Form Adjustment Cells Predictive Of Nonresponse?

Model:

$$\log\left(\frac{P(I_i^p = 1)}{1 - P(I_i^p = 1)}\right) = \beta^p X_i^p + \varepsilon_i^p$$

Test:

$$H_0 : \beta^p = 0$$

$$H_A : \beta^p \neq 0$$

United States Census Bureau

# Logistic Regression Models

- Want to reject the null hypothesis
  - Propensity model holds
- Need to account for complex survey design
  - PROC SURVEYLOGISTIC®
  - Noncertainty units only
- Minimum cell size requirements (challenge!)
  - Actual sample size
  - Effective sample size

United States Census Bureau

# Annual Capital Expenditures Survey

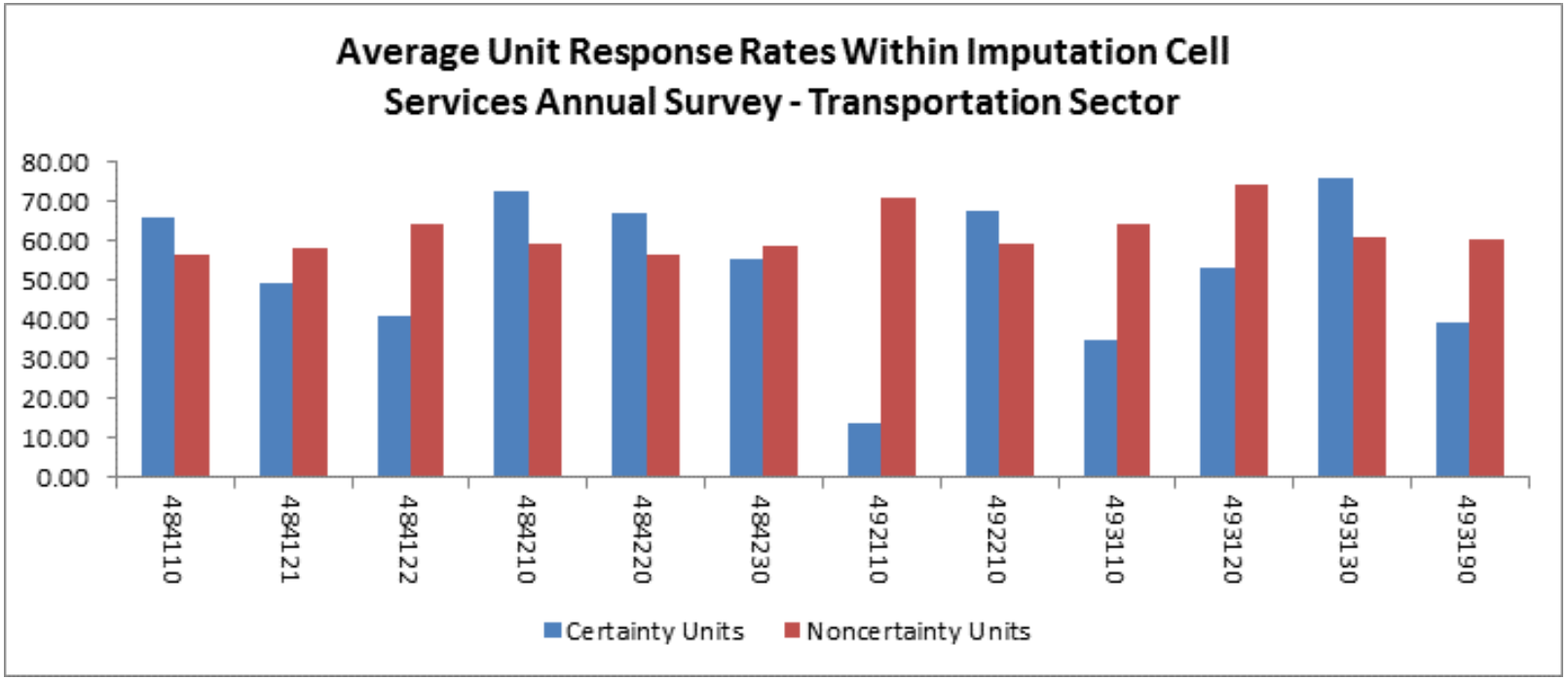
Largest companies

Smallest companies

| Percentage of Weighting Cells Where Propensity Model Held | | | | | |
|---|---|---|---|---|---|
| Size Class Stratum | Survey Year | | | | |
| | 2002 | 2003 | 2004 | 2005 | 2006 |
| 2A | 84.1 | 78.3 | 84.1 | 80.3 | 82.7 |
| 2B | 78.2 | 65.6 | 72.1 | 73.3 | 78.8 |
| 2C | 76.0 | 52.0 | 69.0 | 63.3 | 70.0 |
| 2D | 44.7 | 27.9 | 34.1 | 31.5 | 30.3 |

United States Census Bureau

# Do Response Propensities Differ Between and Within Cells?




United States Census Bureau

# Adjustment Effects on Mean or Totals

| Auxiliary Variable(s) Association with Response (Propensity) | | Auxiliary Variable Association with Outcome (Prediction) | |
|---|---|---|---|
| | | Low | High |
| | Low | Bias:--<br>Variance:-- | Bias:--<br>Variance: ↓ |
| | High | Bias:--<br>Variance: ↑ | Bias: ↓<br>Variance: ↓ |

United States Census Bureau

# Response Model Assessment: Prediction Model Evaluation

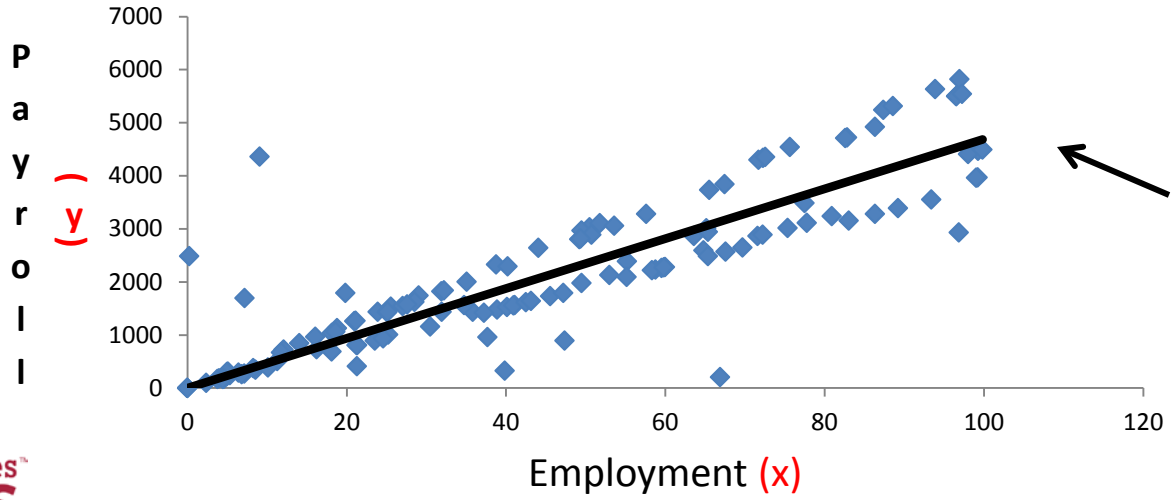
Model:

$$Y_i = \beta^p X_i^p + \varepsilon_i^p$$

Test:

$$H_0 : \beta^p = 0$$

$$H_A : \beta^p \neq 0$$




$$\hat{\beta}^p = \frac{\sum_{i \in p} y_i}{\sum_{i \in p} x_i}$$

United States Census Bureau

# Prediction Model/Regression Analysis

- Want to reject the null hypothesis
- WLS linear regression no intercept model

**Trend** $y_{ti}^{*} = \beta y_{t-1.i} + \varepsilon_{ti}, \ \varepsilon_{ti} \sim (0, y_{t-1.i}\sigma^2)$

**Auxiliary** $y_{ti}^{*} = \beta x_{ti} + \varepsilon_{ti}, \ \varepsilon_{ti} \sim (0, x_{ti}\sigma^2)$

- Minimum cell size requirements (challenge!)
  - Actual sample size
  - Effective sample size

United States Census Bureau

# Prediction Model/Regression Analysis

- Need to account for the complex survey design
  - PROC SURVEYREG®
  - Noncertainty units only
- Want model to be highly predictive
  - Model $R^2$ = measure of “predictive power” of auxiliary variable on item of interest
- Adjustment cell as “instrumental variable”

United States Census Bureau

# Example: ACES

| Percentage of Weighting Cells Where Prediction Model Held | | | | | |
|---|---|---|---|---|---|
| Size Class Stratum | Survey Year | | | | |
| | 2002 | 2003 | 2004 | 2005 | 2006 |
| 2A | 93.8 | 93.1 | 92.6 | 93.3 | 94.8 |
| 2B | 85.3 | 89.9 | 84.1 | 78.9 | 88.7 |
| 2C | 77.0 | 74.6 | 73.6 | 80.8 | 74.6 |
| 2D | 56.1 | 45.9 | 56.3 | 52.4 | 58.7 |

Largest companies

Smallest companies

United States Census Bureau

# Are the Respondents a Random Subsample Within Adjustment Cell?

- Assessing the assumption of “ignorable” nonresponse
- Need auxiliary variables available for all sampled units
  - Proxy variables for characteristic(s) of interest
- Adjustment cells may not be strata
  - Must account for differential sampling

United States Census Bureau

# Are the Respondents a Random Subsample Within Adjustment Cell?

- Are the respondent-based means different from the full-sample means?
- *Are respondents systematically different from nonrespondents?*

United States Census Bureau

# Examining the "Balance"

$$\mathrm{H}_0 : \mu_R^p = \mu_n^p$$

$$\mathrm{H}_\mathrm{A} : \mu_R^p \neq \mu_n^p$$

- Are the respondent-based means different from the full-sample means?
- Measured on auxiliary variables

United States Census Bureau

# Examining the "Distance"

$$\mathrm{H}_0 : \mu_R^p = \mu_{NR}^p$$

$$\mathrm{H}_\mathrm{A} : \mu_R^p \neq \mu_{NR}^p$$

- Are respondents systematically different from nonrespondents?
- Measured on auxiliary variables

United States Census Bureau

# Examining the “Distance”

$$\mathrm{H}_0 : \mu_R^p = \mu_{NR}^p$$

$$\mathrm{H}_\mathrm{A} : \mu_R^p \neq \mu_{NR}^p$$

- Two-sample t-tests on frame variable
  - Performed within adjustment cell *p*
  - Cross-sectional and longitudinal analysis

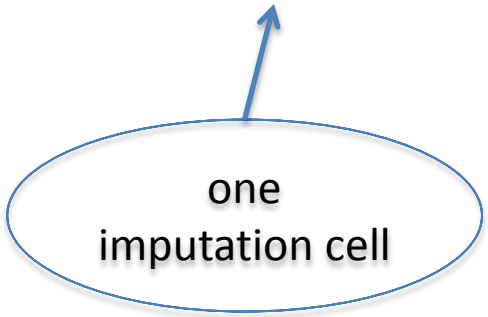
# T-tests

- Want to fail to reject the null hypothesis
- Need to account for complex survey design
  - Weighted unbiased estimates
  - Complex survey variances
  - Noncertainty units only
- Need “sufficient” observations in adjustment cell
  - Respondents and Nonrespondents
  - Challenge for
    - “Large” unit strata (few units to begin with)
    - High nonresponse strata (large variances)

United States Census Bureau

# Quarterly Services Survey (QSS)

| | | 2004 | | | | 2005 | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Sector** | | **Q1** | **Q2** | **Q3** | **Q4** | **Q1** | **Q2** | **Q3** | **Q4** |
| **51** | **Total Imputation Cells** | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 12 |
| | **Different Means($\mu_R \neq \mu_{NR}$)** | 0 | 2 | 1 | 0 | 2 | 1 | 1 | 1 |




United States Census Bureau

# Cross-sectional Analysis

| | | 2004 | | | | 2005 | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Sector** | | **Q1** | **Q2** | **Q3** | **Q4** | **Q1** | **Q2** | **Q3** | **Q4** |
| **51** | **Total Imputation Cells** | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 12 |
| | **Different Means($\mu_R \neq \mu_{NR}$)** | 0 | 2 | 1 | 0 | 2 | 1 | 1 | 1 |

- Test whether the number of cells with different respondent and nonrespondent means is larger than expected
- No evidence of systematic difference in mean receipts

United States Census Bureau

# Longitudinal Analysis

| | | 2004 | | | | 2005 | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Sector** | | **Q1** | **Q2** | **Q3** | **Q4** | **Q1** | **Q2** | **Q3** | **Q4** |
| **51** | **Total Imputation Cells** | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 12 |
| | **Different Means($\mu_R \neq \mu_{NR}$)** | 0 | **2** | **1** | 0 | **2** | **1** | **1** | 1 |

- Examine the cells with significantly different means across quarters
- Identify cells that exhibit “consistent” differences
  - Imputation cell 512000T (5 of 8 quarters)

United States Census Bureau

# Chi-Squared Tests for Independence

| | Respondents | Nonrespondents | |
|---|---|---|---|
| $0 < \text{weight} < P_{33}$ | $r_{11}$ | $r_{12}$ | $r_{1\bullet}$ |
| $P_{33} \leq \text{weight} < P_{66}$ | $r_{21}$ | $r_{22}$ | $r_{2\bullet}$ |
| $P_{66} \leq \text{weight}$ | $r_{31}$ | $r_{32}$ | $r_{3\bullet}$ |
| | $r_{\bullet 1}$ | $r_{\bullet 2}$ | $r_{\bullet\bullet}$ |

- PROC SURVEYFREQ®
  - Want to fail to reject null hypothesis
- Noncertainty units only

United States Census Bureau

# Sensitivity Analysis

- Presented techniques require auxiliary variables
  - Available for all sampled units
  - Correlated with characteristics
- Other options (sensitivity analysis)
  - Quantile regression (Tucker and Dixon, 2007)
  - “Influence” functions – examine contribution of each weighted observation to tabulated total

United States Census Bureau

# A FEW OTHER APPROACHES: SOME CONSIDERATIONS

United States Census Bureau

# Comparison to Other Estimates – Benchmarking to Other Programs

- Alternative estimates may or may not be independent
- Definitions between programs may not agree (caution advised)
- Big assumption that other set of estimates is “superior” to studied program with respect to sampling and/or measurement errors

United States Census Bureau

# Benchmarking Survey Estimates to those from Another Data Source

- Compare linked microdata
  - Same units, other programs
  - Administrative data (where feasible)
  - Same cautions apply with respect to definitions for items
- Compare early and late respondents on key estimates
  - Not studied extensively for business surveys

United States Census Bureau

We Did It, Now What?

# MAKING NONRESPONSE BIAS ANALYSIS EFFECTIVE

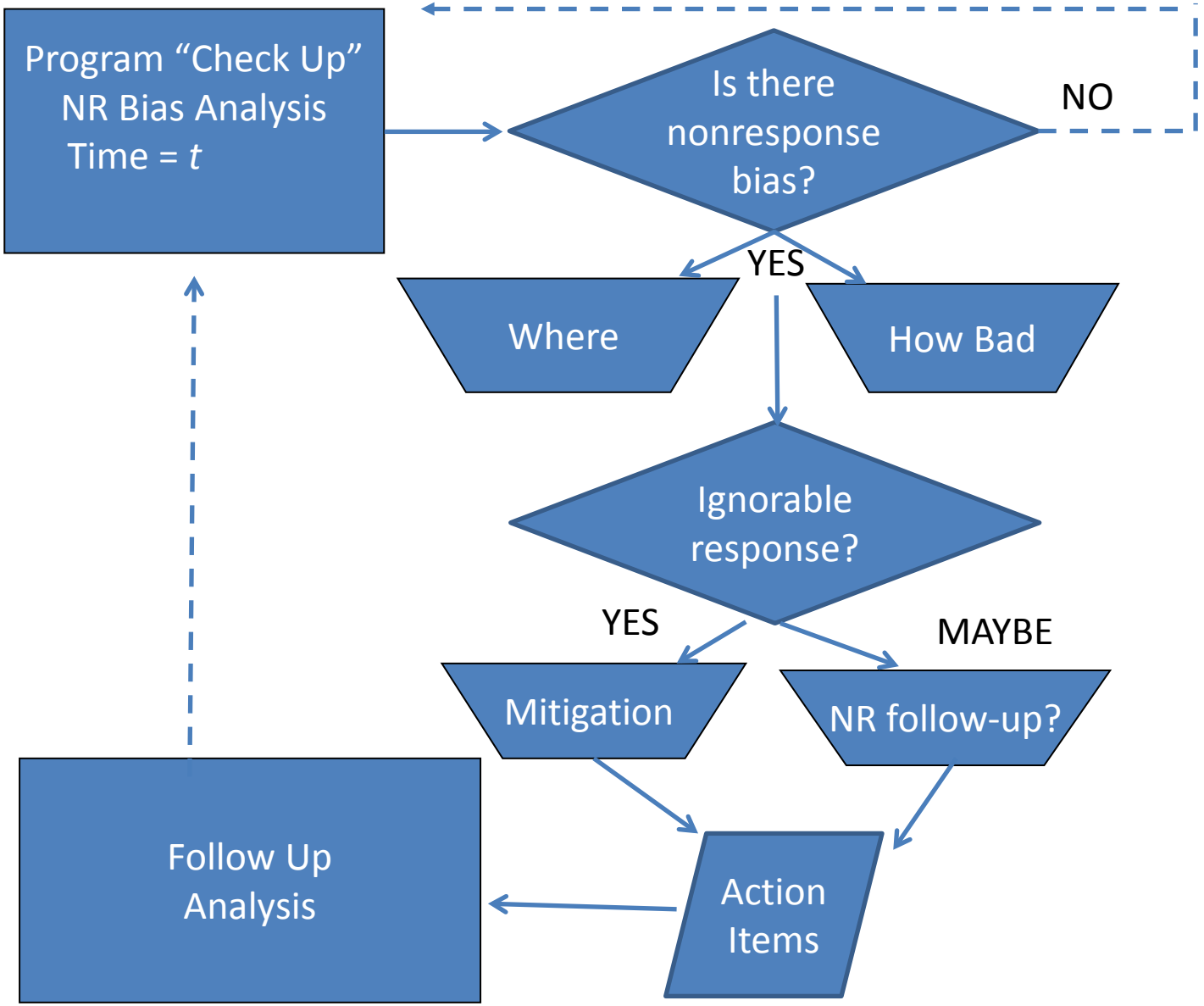
United States Census Bureau

# Official Disclaimer

*The purpose of this presentation is to inform interested parties of research and to encourage discussion of work in progress. Any views expressed are those of the author and not necessarily those of the U.S. Census Bureau.*

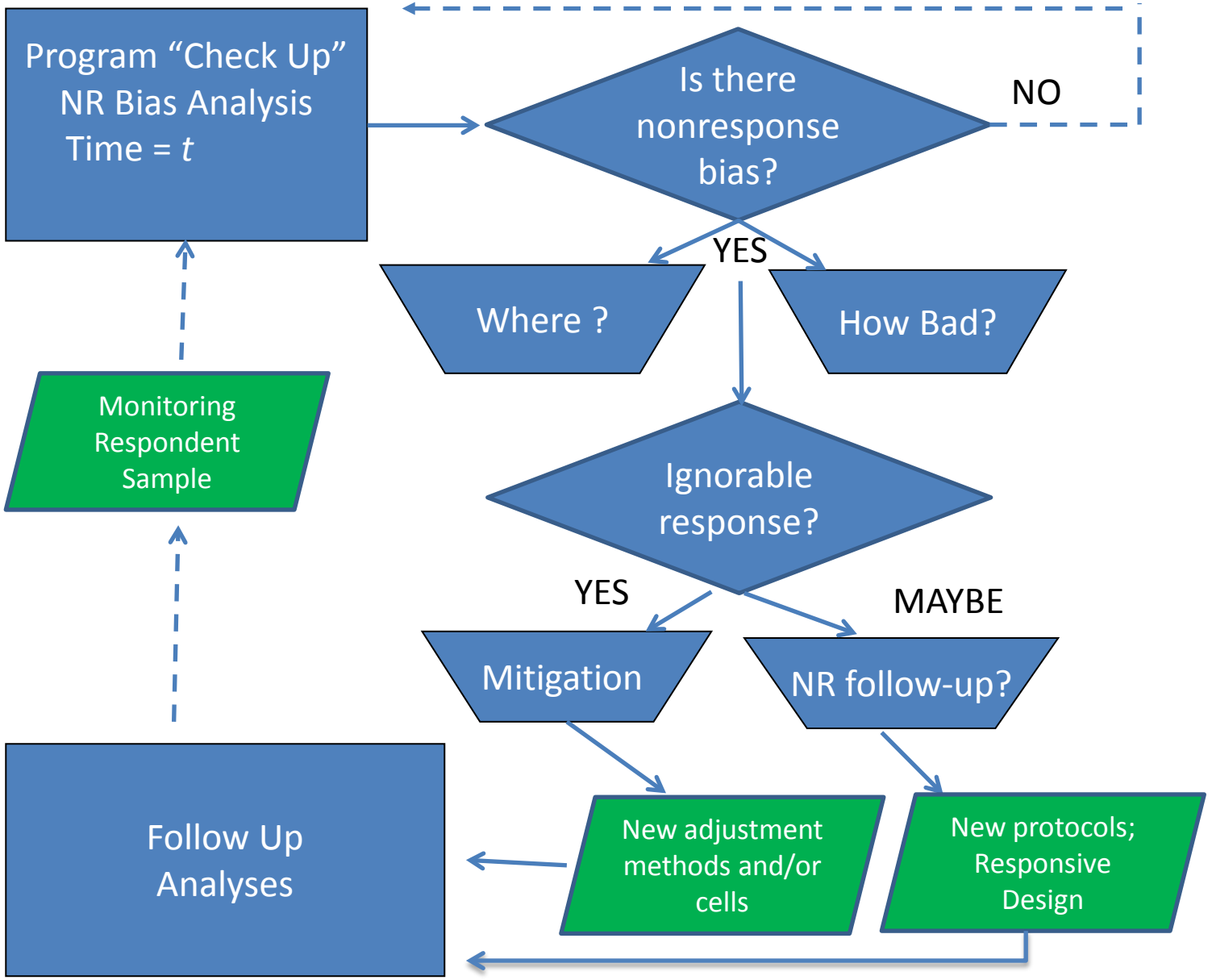
United States Census Bureau

# Objectives of NR Bias Analysis




United States Census Bureau

# Useful Outcomes of NR Bias Analysis




United States Census Bureau

# Being Pro-Active

- Monitor respondent sample during collection to avoid nonresponse bias
  - P-charts of unit response rates
  - Contingency table analysis
  - R-indicators
- Especially for nonignorable nonresponse
  - **Study** response contact strategies
    - Qualitative – focus groups, company visits
    - Quantitative – using paradata
  - Responsive designs

United States Census Bureau

# The Beginning…

- **Initial Analysis** (Nonresponse Bias Study)
  - Action items
- **Corrective measures**/process changes
  - Assess over time using the same analysis tools
- **Monitor** stable/in-control process
  - Collaborative effort between methodologists and subject-matter experts

Katherine.J.Thompson@census.gov

United States Census Bureau