

# Using Data from Secondary Sources and Existing Statistics

# Existing Statistics

“In existing statistics research, a researcher locates previously collected information, often in the form of government reports or previously conducted surveys, then reorganizes or combines the information in new ways to address a research question. Locating sources can be time consuming, so a researcher needs to consider carefully the meaning of what he or she finds.

# Existing Statistics

"Frequently, a researcher does not know whether the information of interest is available when he or she begins a study. Sometimes, the existing quantitative information consists of stored surveys or other data that a researcher reexamines using various statistical techniques. Existing statistics research can be used for exploratory, descriptive or explanatory purposes, but it is most frequently used for descriptive research."

-Neuman, p. 21

# Existing Statistics

Existing Statistics are a useful source of quantitative data for statistical analysis

Existing Statistics allow a researcher to remain objective, since she never contacts the subjects of the study; however, she is also limited to using data in its existing format

Existing Statistics are widely used in the social sciences

# Appropriate Research Questions for Existing Statistics

Existing Statistics Research is appropriate for research questions that:

- Involve large data sets from diverse sources, e.g. cross-national research
- Require longitudinal data, particularly from the past
- Involve well-developed concepts

# Existing Statistics

Using existing statistics requires a researcher to locate indicators that match theoretical constructs of interest. The researcher must carefully create an argument that links the indicators to the constructs.

# Sources for Existing Statistics

U.S. Census Bureau

Decennial Census

Current Population Survey

General Social Survey (Univ. of Chicago)

Statistical Abstract of the United States

United Nations Statistical Yearbook

Inter-University Consortium for Political  
and Social Research (Univ. of MI)

# Issues Using Existing Statistics

- (1) Finding appropriate data for a given research question, or using data that are inappropriate
- (2) Overemphasis on empirical data at the expense of theory or understanding
- (3) The “fallacy of misplaced concreteness,” or the tendency to quote very precise statistics in order to give the impression of scientific rigor



# Validity and Reliability

## Validity Problems:

- (1) Definitions
- (2) Proxy variables
- (3) Poor data collection

## Reliability Problems:

- (1) Changing definitions/data collection techniques over time
- (2) Difference in data collection in different governments/agencies