

# Using Large Data Sets

Mary Kay Rizzolo, PhD

Institute on Disability and Human  
Development, UIC

## What data set are you using?

- Using public data to do secondary analysis – an example from my own dissertation...



## *Statement of the Problem*

- ❑ 4.6 million persons in the U.S. have intellectual or developmental disabilities (I/DD)
- ❑ Approximately 10% live in out-of-home settings
  - ❑ **Community settings for 1-15 persons**  
(group homes, supported living, foster care)
  - ❑ **Institutional settings for 16+ persons**  
(state institutions, nursing homes)
- ❑ The states have great latitude in deciding which settings to use

## ***Statement of the Problem (continued)***

- ❑ In 1987, the Nursing Home Reform Amendments (OBRA '87) were passed which addressed the inappropriate placement of individuals with mental disabilities in nursing facilities
- ❑ States were required to identify all persons with MI or I/DD who were inappropriately placed in nursing facilities and move them to appropriate settings

## ***Statement of the Problem (continued)***

- ❑ Although the number of persons with developmental disabilities in nursing homes has declined since OBRA, every state in the U.S. continues to utilize this type of setting as a residential option for this population
- ❑ Most states also utilize state-operated institutions for persons with developmental disabilities



## *Research Questions*

What variables are associated with state utilization of **nursing facilities** for persons with developmental disabilities...

In 1991?

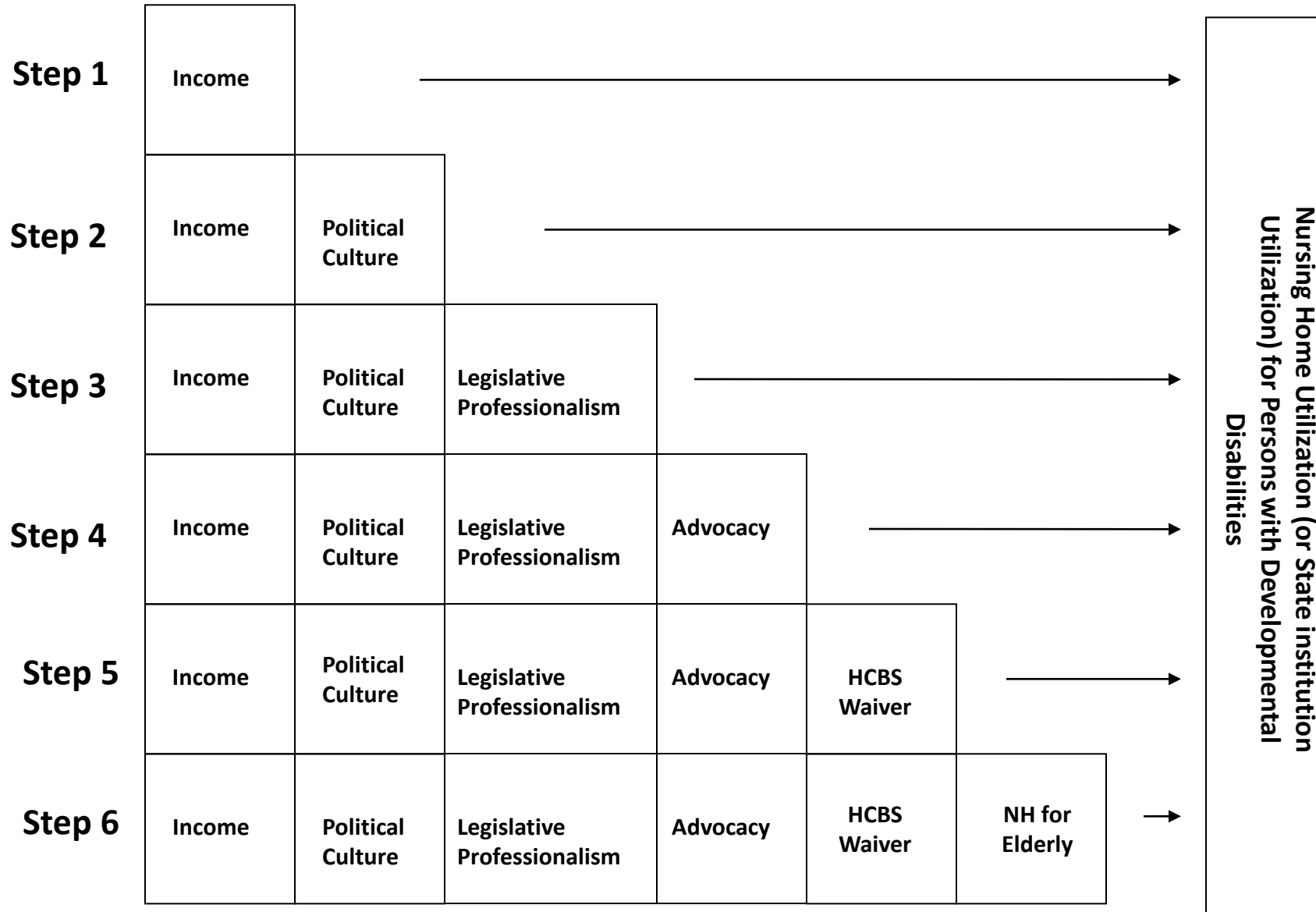
In 2000?

What variables are associated with state utilization of **state-operated institutions** for persons with developmental disabilities...

In 1991?

In 2000?

# ***A Political-Socioeconomic Model***



# The Independent Variables



<b>Variables</b>	<b>Definition</b>	<b>Source</b>	<b>Rationale for Including in Model</b>
State Wealth	Personal income per capita of state general population	Bureau of Economic Analysis; U.S. Census Bureau	Significant predictor of deinstitutionalization of state DD institutions in 1977 (Sigelman et al., 1981)
Advocacy	Membership per capita in the state Arc chapter	The Arc of the United States	Significant predictor of spending for community DD services (Braddock, 2002a, 2002b; Braddock & Fujiura, 1991; Braddock et al., 1998; Braddock et al., 2000)
HCBS Waiver (DD)	State spending per capita for HCBS Waiver services for persons with developmental disabilities	The State of the States in Developmental Disabilities Project, The University of Colorado	Significant predictor of spending for community DD services (Braddock et al., 2000, 2002; Braddock, 2002a, 2002b)
Nursing Facility Use for General Population 65+	Total nursing facility utilization per capita of state general population 65+	The Department of Social and Behavioral Sciences, University of California, San Francisco	
Political Culture	Measure of a state's political culture: traditionalistic, individualistic, or moralistic	Sharkansky, 1969	Significant predictor of deinstitutionalization of state DD institutions in 1977 (Sigelman et al., 1981) and deinstitutionalization of facilities for juvenile offenders (Downs, 1976)
Legislative Professionalism	Measure of the capacity of a state's legislature	Squire, 1992, 2000	Significant in Sigelman et al. (1981) model

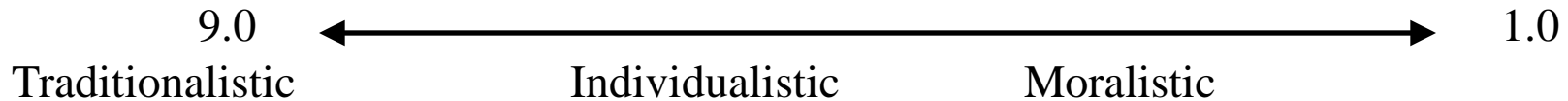
## ***Political Culture (independent variable)***

- Elazar (1966) conceived notion of **political culture**
  - "Attitudes, values, beliefs, and orientations that individuals in a society hold regarding their political system." (Riley, 2004)
- Elazar developed three sub-groups (**Moralistic, Individualistic, & Traditionalistic**) of American culture based on state histories, voting data, field work, and direct observation.
- These groups differed in their views regarding political participation, the Government's role in society, and the initiation of new programs.



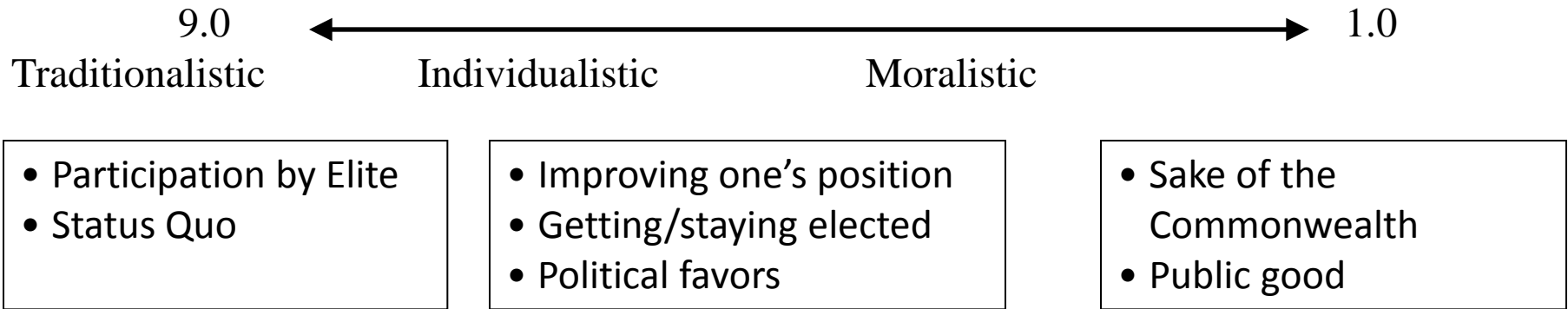
## *Political Culture*

Shankansky (1969) created a measurable scale based on Elazar's concept of political culture



- The Shankansky (1969) scale was used in my dissertation (sometimes an “updated” version of a variable from a previous model may be used)

# Political Culture (independent variable)



Arkansas	9
Mississippi	9
Georgia	8.8
South Carolina	8.75
Alabama	8.57
North Carolina	8.5
Tennessee	8.5
Oklahoma	8.25
Louisiana	8
Virginia	7.86
Florida	7.8
Missouri	7.66
Kentucky	7.4
West Virginia	7.33
Texas	7.11
Delaware	7
Maryland	7
New Mexico	7
Indiana	6.53

Arizona	5.66
Ohio	5.16
Nevada	5
<b>Illinois</b>	<b>4.72</b>
Pennsylvania	4.28
New Jersey	4
Wyoming	4
Kansas	3.66
Nebraska	3.66
New York	3.62
Massachusetts	3.6
California	3.55

Connecticut	3
Montana	3
Rhode Island	3
South Dakota	3
Idaho	2.5
Maine	2.33
New Hampshire	2.33
Vermont	2.33
Iowa	2
Michigan	2
North Dakota	2
Oregon	2
Utah	2
Wisconsin	2
Colorado	1.8
Washington	1.66
Minnesota	1

## ***Hierarchical Regression Analysis for Variables Predicting State Institution Utilization for Persons with Developmental Disabilities (2000)***

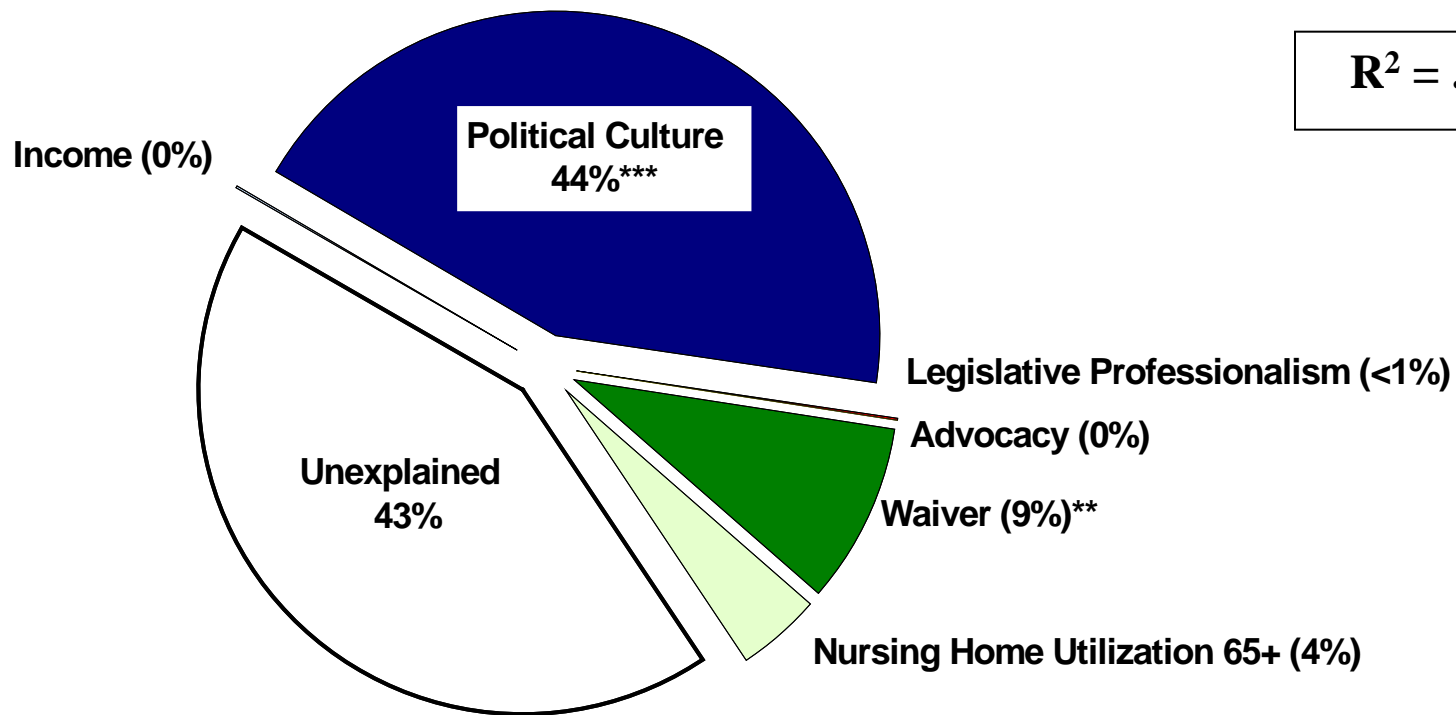
Variable	B	SE B	$\beta$	$R^2$	$\Delta R^2$
Step 1				.001	.001
Income	5.99E-07	.000	.029		
Step 2				.440	.439***
Income	5.96E-06	.000	.290*		
Culture	.024	.004	.712***		
Step 3				.441	.001
Income	6.30E-06	.000	.306*		
Culture	.024	.004	.711***		
LegProf	-.025	.091	-.035		
Step 4				.441	.000
Income	6.29E-06	.000	.306*		
Culture	.024	.004	.711***		
LegProf	-.025	.095	-.035		
Advocacy	2.69E-06	.000	.001		
Step 5				.531	.09**
Income	1.01E-05	.000	.491**		
Culture	.018	.004	.537***		
LegProf	-.082	.090	-.116		
Advocacy	9.82E-05	.000	.046		
Waiver	-.003	.001	-.415**		
Step 6				.572	.042
Income	1.00E-05	.000	.487**		
Culture	.018	.004	.536***		
LegProf	-.060	.088	-.085		
Advocacy	7.57E-05	.000	.035		
Waiver	-.003	.001	-.482**		
NH_65	1.598	.800	.218		

\* $p < .05$ . \*\*  $p < .01$ .

By including my dissertation as a project in my advisor's federally funded research center, I was able to secure minimal funding which I used to hire a statistician.

# 2000 State Institution Utilization

## Explained Variance in 2000 State-Operated Institution Model



47,372 persons with DD in state-operated institutions (48 states)

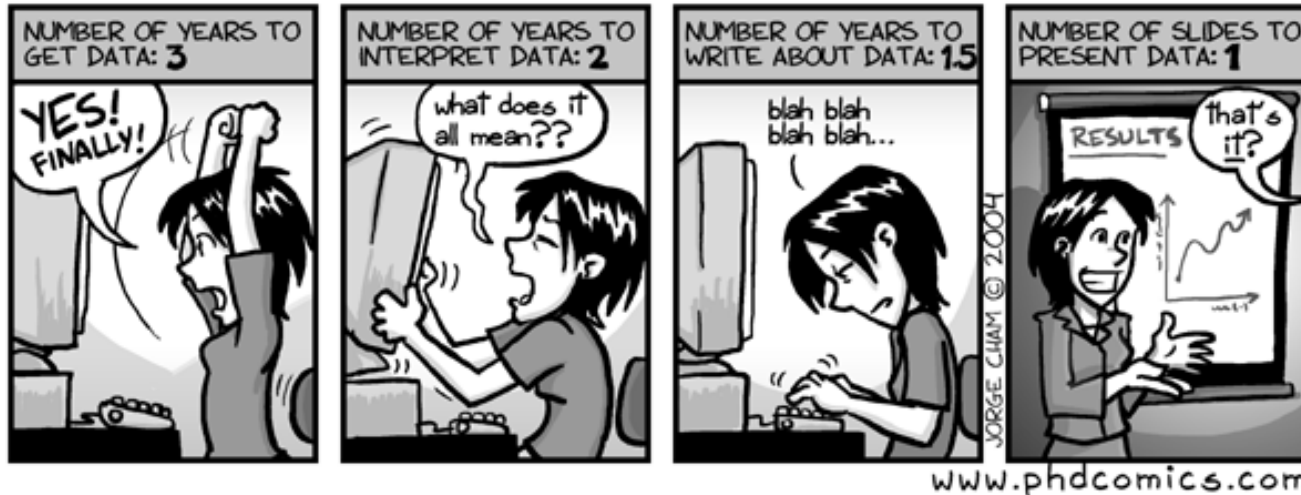
$F(1,41) = 9.14, p = .000$

Adjusted  $R^2 = .51$

\*\*  $p < .01$ . \*\*\*  $p < .001$ .

# What are the benefits of using existing data as opposed to doing collection on your own?

## DATA: BY THE NUMBERS



### 1. Time Commitment

– In the words of Glenn Fujiura’s wife: “Don’t be famous, be done!”

- ### 2. Institutional Review Board
- much easier to get approval to do secondary data analysis than primary data collection -- especially with “vulnerable populations” – review your IRB’s policies on educational research to see which types of studies can be exempt from review (or expedited).

***Your thesis/dissertation will not (usually) define you but the process should arm you with the tools you need to conduct further research***



JORGE CHAM © THE STANFORD DAILY

[phd.stanford.edu/comics](http://phd.stanford.edu/comics)

# ***Resources to find other public data sets...***

## **Sometimes data sets are analyzed and summarized by others**

- US Census Bureau (website link entitled “Data Access Tools”  
<http://www.census.gov/main/www/access.html> - they have software available for download to access survey data

- Behavioral Risk Factor Surveillance System (CDC)

User can select demographic or geographic variables

For example if you choose disability...

Choose a question to view its data.

- [Adults who are limited in any activities because of physical, mental, or emotional problems](#)
- [Adults with health problem\(s\) that requires the use of special equipment](#)

**Sometimes you can download entire data sets using SPSS or SAS or other statistical software to analyze “records”**

Surveys measure “disability” in their own ways (though some are now trying to coordinate to provide comparability across surveys). The survey you choose will depend on your research questions. Few surveys provide information specifically on people with I/DD.

- State of the States in Developmental Disabilities – revenue, spending and programmatic trends in each of the states and DC over a 33 year period

- Create a Chart option on website

- <http://sos.arielmis.net/index.php/interactive-charts>

- Total, community, and institutional fiscal effort for I/DD services

- Community spending as a % of total I/DD spending

- % of total out-of-home placements in settings for 6 or fewer

- % of total statewide I/DD caregivers supported by state DD agency

- More information available in tables throughout book

- National Institute on Disability and Rehabilitation Research (NIDRR)
  - Chartbook on Disability  
<http://www.infouse.com/disabilitydata/disability/index.php>
  - Chartbook on Mental Health and Disability  
<http://www.infouse.com/disabilitydata/mentalhealth/index.php>
  - Chartbook on Women and Disability  
<http://www.infouse.com/disabilitydata/womendisability/index.php>
  - Chartbook on Work and Disability  
<http://www.infouse.com/disabilitydata/workdisability/index.php>

Access to Disability Data *An InfoUse Project*  
 funded by the National Institute on Disability and Rehabilitation Research

Printer-friendly page

## Chartbook on Disability in the U.S.

### Section 1: Prevalence of Disabilities

#### 1.4. How many people have mental retardation?

Because it is difficult to measure in the population, there is only general agreement on how many people have mental retardation. While estimates of the percent in the U.S. population range from .67% to 3%, the most accepted estimates find approximately 1% of the noninstitutionalized population has mental retardation. This translates to about 2.5 million noninstitutionalized people having mental retardation. Estimates vary by age and definition used.

Additionally, in 1995, there were an estimated 346,659 people with mental retardation in residential settings and institutions: 33,943 are in nursing homes; 62,028 are in state institutions; 37,311 reside in private institutions with 16 or more residents; and 213,377 live in other community facilities. Another estimated 1,400 persons with mental retardation are in the non-mental retardation/developmental disabilities units of psychiatric hospitals.

In 1995, there were an estimated 1.3 million people of all ages reporting mental retardation as a cause of activity limitation (NHIS). People, age 15 and over, reporting mental retardation as a cause of disability in the SIPP in 1992 totaled 501,000.

**The number of people with mental retardation causing an activity limitation goes down with age.**

Age Group	Percentage
under 18	15.8%
18-44	3.1%
45-69	0.4%
70-84	0.1%
85 and over	0.1%

Source: LaPlante 1995; McNeil 1993; Prouty and Lakin, 1996  
 Survey: NHIS, 1992; SIPP, 1992

# ***Resources to find other public data sets...some examples***

- **Disability Statistics Center (University of California, San Francisco)**
  - **Finding disability data on the web:**  
[http://dsc.ucsf.edu/main.php?name=finding\\_data](http://dsc.ucsf.edu/main.php?name=finding_data)
- **U.S. Census Bureau**
  - **How Disability Data is Collected**  
<http://www.census.gov/hhes/www/disability/datacollection.html>
- **Disability Statistics (Cornell University)**
  - <http://www.disabilitystatistics.org/faq.cfm>

- **Center for Studying Disability Policy (Mathematica)**
  - **Disability Data in National Surveys** [http://www.mathematica-mpr.com/publications/PDFs/disability/Data\\_National\\_Surveys.pdf](http://www.mathematica-mpr.com/publications/PDFs/disability/Data_National_Surveys.pdf)
    - Overview of Existing Disability-Related Survey Data
      - Ongoing National General Population Surveys
      - Surveys on Health, Disability, Aging, and Long-Term Care
      - Surveys on Youth and Education
      - Other Surveys
      - New Survey Efforts



## Disability Data in National Surveys (continued) [http://www.mathematica-mpr.com/publications/PDFs/disability/Data\\_National\\_Surveys.pdf](http://www.mathematica-mpr.com/publications/PDFs/disability/Data_National_Surveys.pdf)

- **Appendix B: Detailed Survey Summaries**

**Information on:**

“General purpose, sponsoring agencies, sampling universe, sample size, regional and state-level estimates, longitudinal data, frequency of data collection and year(s) fielded, physical health, sensory, and communication disability measures, mental health and disability measures, cognitive and intellectual disability measures, other health and disability measures, topic areas, links to administrative data, **potential uses for disability-related research**, sources of information about the survey.”

- For example, on page B-20... the “B&B\* can be used to investigate educational outcomes for students with physical, emotional, or learning disabilities. Researchers can compare employment rates, occupations, and starting salaries for people with and without disabilities. The survey data can also be used to explore predictors for pursuing graduate or professional degrees.”

\*The Baccalaureate and Beyond Longitudinal Study