Case 5

Introduction to Demographic Research Using Aggregated ACS Data for Ecological Regression:
Changes in County Poverty

Katherine Curtis
Adam Slez
Jennifer Huck

University of Wisconsin – Madison

Center for Demography & Ecology

Applied Population Laboratory

- Comparability of ACS with Census Long-Form

  – Variable Comparability (data & measures)

  – Sample Comparability (statistical inference)

- Focus on changes in relationships between county poverty rates and structural covariates

Sample

- Generalized standard error:

$$SE(\hat{Y}) = D * \sqrt{\hat{Y}\left(\frac{1}{R} - 1\right)\left(1 - \frac{\hat{Y}}{N}\right)}$$

- SE of an estimate (Y) is inversely related to R (*sampling fraction*) & N (*total population*), and positively related to D (*design factor*)
  - SE increases as R & N decreases and as D increases

- ACS is at a disadvantage for estimate reliability given the smaller sample size (compared to SF3)

Variable

- ## Sample Design Issues
  - **Poverty** is based on calendar year income (i.e., 1999) for SF3 and income during the past 12 months of a multi-year period for ACS

- ## Universe Issues
  - **Eligibility** surrounding the 2-month residency rule

  - **Underemployment** (male workers) reported for population age 16-64 in the ACS and 16+ in the SF3

- ## Suppressed Data Issues
  - **Race/ethnicity** is not reported for 274 of the 988 counties

Variable Selection

For Industry

Data Profile
• 65 "missing" cases

Detailed Tables (collapsed)
• 4 "missing" cases

Detailed Tables (uncollapsed)
• 963 "missing" cases
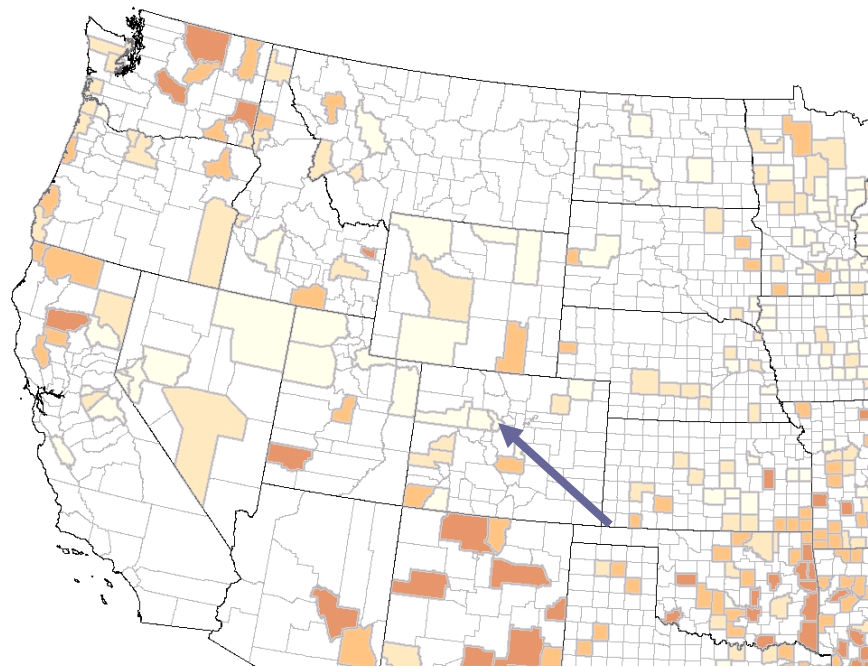
2005-2007 American Community Survey 3-Year Estimates

Select a Geographic
Summary Level

```
=== National Level ===
United States (010)
All States (040)
All Counties (050)
All Places (160)
All Metropolitan and Micropolitan Statistical Areas (310)
All Congressional Districts -110th (500)
```

Select a Download
Method
  ⦿ Selected Detailed Tables - up to 50 tables in pipe-delimited format
  ◯ Data Profiles
  ◯ Subject Tables

GO

All Counties 20-65k

N = 988
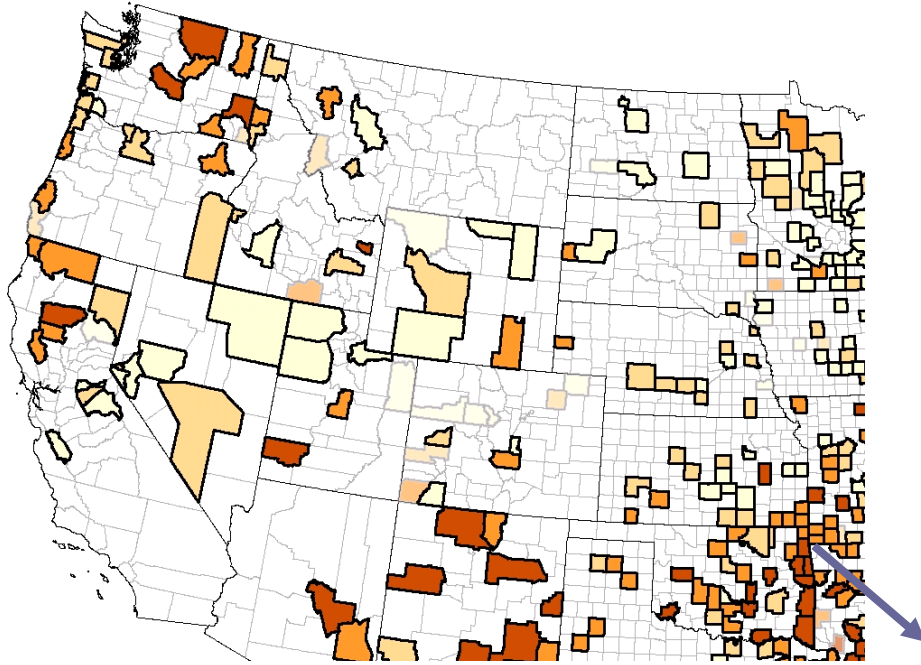
**Minus All
Suppressed Data**

N = 708

- Comparative analysis to examine the way differences in survey design influence results of a conventional ecological regression analysis

  - County poverty rates

  - 2000 SF3 & 2005-2007 ACS

  - Counties size 20,000 and 65,000

- Required Adjustments

1. Calculate margin of error for derived proportions
   – ACS New Compass Handbook for Federal Agencies, Appendix 3

2. Reduce sampling error
   – WLS (thanks Freddie!)

3. Address spatially correlated errors
   – Not the focus per se, but important for ecological analyses

- ## Data Access

  – American FactFinder > Download Center > *Data Profiles*
  – American FactFinder > Download Center > *Selected Detailed Tables*

- ## Variable Calculation

  – Use of different denominator (e.g., education)
  – Changing variable definitions (e.g., industry)
  – Create new variables (e.g., underemployment and commuter rates)

ACS versus SF3
County Poverty Rates

	Minimum	Maximum	Mean
2000 (SF3)	0.03	0.51	0.14
2005-2007 (ACS)	0.03	0.50	0.16
Δ Poverty Rate	-0.65	1.26	0.15

$$y = x\beta + \lambda Wu + \varepsilon$$

- **Spatial Error Model**
    - **y** is the county poverty rate
    - **x** is the set of structural covariates associated with poverty
    - β is the set of effects associated with these factors
    - λ measures the extent to which the spatial error in a county tends to be correlated with the spatial error in neighboring counties
    - **W** is a row-standardized matrix depicting the spatial relationship between counties
    - **u** is a measure of spatial error
    - ε is a measure of non-spatial error

ACS: Unadjusted versus Adjusted
Regression Analysis of
County Poverty Rates (log odds), (N=708)

	ACS					
	Unadjusted		**Population Adjusted**		**Residual Adjusted**	
	β	SE	β	SE	β	SE
Constant						
African American						
Hispanic						
More than High School						
Commuter						
Unemployment						
Underemployment						
Female-Headed HH						
Extractive Industry						
Professional Services						
Manufacturing						
Miscellaneous Services						

Rsq

* p < .05, ** p < .01, *** p < .001

Note: All variables are in proportions.

ACS versus SF3

Regression Analysis of County Poverty Rates (log odds) with Spatial Corrections, (N= 708)

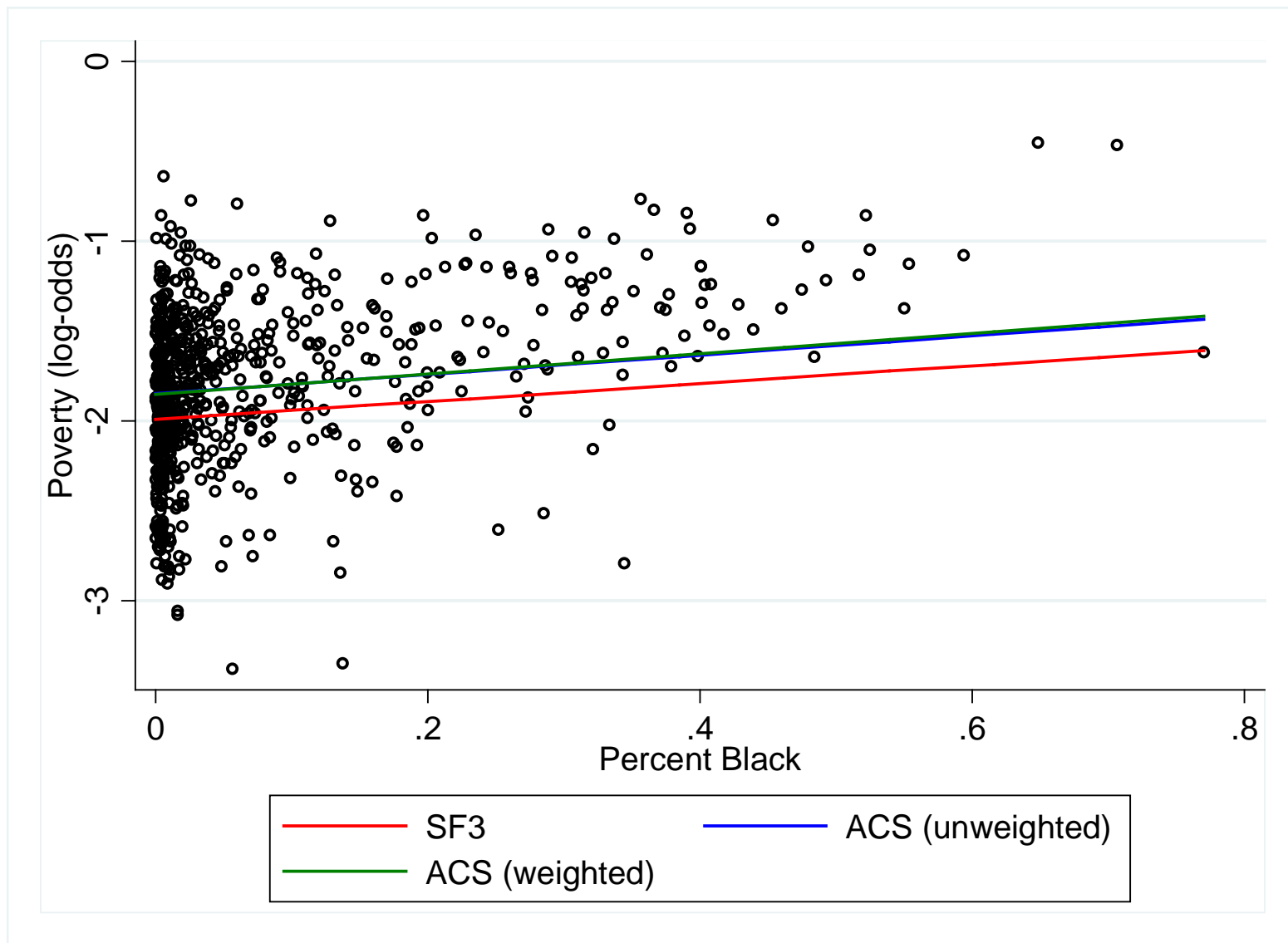|	SF3		ACS	
			Population Adjusted	
	β	SE	β	SE
Constant	-4.21 ***	0.18	-4.49 ***	0.17
African American	0.30 *	0.13	0.59 ***	0.13
Hispanic	0.25 *	0.11	0.21	0.12
More than High School	1.59 ***	0.15	1.86 ***	0.16
Commuter	0.35 ***	0.06	0.58 ***	0.07
Unemployment	5.42 ***	0.58	3.71 ***	0.56
Underemployment	3.43 ***	0.26	2.76 ***	0.20
Female-Headed HH	7.25 ***	0.90	3.01 ***	0.67
Extractive Industry	0.63	0.37	0.26	0.39
Professional Services	-0.71 *	0.31	0.44	0.28
Manufacturing	-0.83 ***	0.24	-0.31	0.23
Miscellaneous Services	-0.83	0.46	0.08	0.44
Lambda	0.46 ***		0.29 ***	

* p < .05, ** p < .01, *** p < .001

Note: All variables are in proportions.

- Necessary user practices:

  - Review variable definitions

  - Confirm variable universe

  - Calculate MOE for derived variables

  - Adjust standard errors for statistical inference

Table 1. Calculating a margin of error for a derived count and derived proportion, Sauk County, Wisconsin, ACS 2005-2007 Table B17001

Variable	Estimate		MOE
Population with income in the past 12 months below poverty level	5,256	±	731
Male:	2,132	±	359
Under 5 years	346	±	155
5 years	66	±	59
6 to 11 years	227	±	96
12 to 14 years	140	±	74
15 years	8	±	10
16 and 17 years	28	±	34
18 to 24 years	199	±	149
25 to 34 years	397	±	143
35 to 44 years	170	±	72
45 to 54 years	240	±	92
55 to 64 years	144	±	85
65 to 74 years	40	±	28
75 years and over	127	±	70
Female:	3,124	±	479
Under 5 years	231	±	104
5 years	29	±	26
6 to 11 years	340	±	149
12 to 14 years	142	±	76
15 years	35	±	29
16 and 17 years	184	±	119
18 to 24 years	409	±	160
25 to 34 years	434	±	180
35 to 44 years	395	±	118
45 to 54 years	237	±	91
55 to 64 years	122	±	52
65 to 74 years	114	±	86
75 years and over	452	±	149
Total population	57,154	±	124
Estimated proportion below poverty	0.092	±	0.013

ACS versus SF3
Regression Analysis of
County Poverty Rates (log odds), (N=708)

	SF3		ACS	
			Population Adjusted	
	β	**SE**	**β**	**SE**
Constant	-4.21 ***	0.21	-4.61 ***	0.18
African American	0.50 ***	0.13	0.57 ***	0.12
Hispanic	0.29 **	0.11	0.23 *	0.12
More than High School	1.84 ***	0.15	2.05 ***	0.16
Commuter	0.47 ***	0.07	0.64 ***	0.07
Unemployment	6.01 ***	0.65	3.82 ***	0.57
Underemployment	3.50 ***	0.29	2.86 ***	0.22
Female-Headed HH	6.70 ***	0.98	3.19 ***	0.70
Extractive Industry	-0.14	0.41	0.07	0.40
Professional Services	-0.99 **	0.35	0.39	0.29
Manufacturing	-1.28 ***	0.25	-0.38	0.24
Miscellaneous Services	-1.39 **	0.53	-0.20	0.47
Rsq	0.71		0.65	

* p < .05, ** p < .01, *** p < .001

Note: All variables are in proportions.