ACS 5-Year Data:
A First Look at the First Release

Ken Hodges

May 11, 2011

Objectives

- NOT
  - A comprehensive review
  - Or a final verdict on ACS 5-year data
- Rather
  - One user's first look at ACS 5Y data
  - Focus on small area data
  - What looks good?
  - What looks . . . curious?

ACS 5-Year Data

- Released December 2010
- Data collected 2005-2009
- Excitement
  - Long awaited data for small areas
  - Small cities and towns
  - Census tracts and block groups
- Apprehensions
  - Accuracy of small area data?
  - Will users embrace ACS?
  - Or brace for large margins of error (MOE)?

The “Small” ACS Sample

- ACS sample smaller than long form sample
  - Even after 5 years
- 2000 Long Form
  - Total HU = 115,904,641
  - Sample HU = 18,345,474 (15.8 pct)
- 2005-2009 ACS
  - Total HU = 127,699,712
  - HU Sample = 14,450,288 (11.3 pct)
  - Unwgted HU = 9,658,438 (7.6 pct)
- What about block groups?

The “Small” ACS Sample

Block Groups by Number of Unweighted Housing Units

Unweigted Units	N	Pct
Missing (no ACS)	1,533	0.7
Missing (w/ACS) (1 or 2)	801	0.4
3 – 9	2,982	1.4
10 – 19	24,527	11.7
20 – 49	**115,865**	**55.5**
50 – 99	48,002	23.0
100 – 199	13.303	6.4
200 – 499	1,711	0.8
500 or more	73	0.0
Total	**208,797**	**100.0**

The “Small” ACS Sample

- 93 pct of BGs – fewer than 100 responses
- 70 pct of BGs – fewer than 50 responses

- Some MOEs will be large
- MOEs are important
- But not a measure of actual error
  - Can be misleading
  - Need more than MOEs to judge ACS

Households by Type and Size

- Look at “Households by Type and Size”

Total Households

- **Family 2-persons**
- **Family 3-persons**
- **Family 4-persons**
- **Family 5-persons**
- **Family 6-persons**
- **Family 7+ persons**
- **Nonfamily 1 person**
- **Nonfamily 2 persons**
- **Nonfamily 3 persons**
- **Nonfamily 4 persons**
- **Nonfamily 5 persons**
- **Nonfamily 6 persons**
- **Nonfamily 7+ persons**

Households by Type and Size

- Stable distributions
- Predominant pattern
  - Few with 6 or 7+ people
- Suspect data evident independent of MOE
- Also on decennial census
  - Allows comparisons vs. complete count

Households by Type and Size

Pct of BGs with Margin of Error Greater Than Cell Value

HH Type & Size	Pct MOE GT Cell Value
Total Households	1.1
Family 2-persons	8.4
Family 3-persons	25.4
Family 4-persons	34.6
Family 5-persons	65.4
Family 6-persons	90.6
Family 7+ persons	67.9
Nonfamily 1 person	10.3
Nonfamily 2 persons	76.7
Nonfamily 3 persons	99.0
Nonfamily 4 persons	99.7
Nonfamily 5 persons	100.0
Nonfamily 6 persons	100.0
Nonfamily 7+ persons	100.0

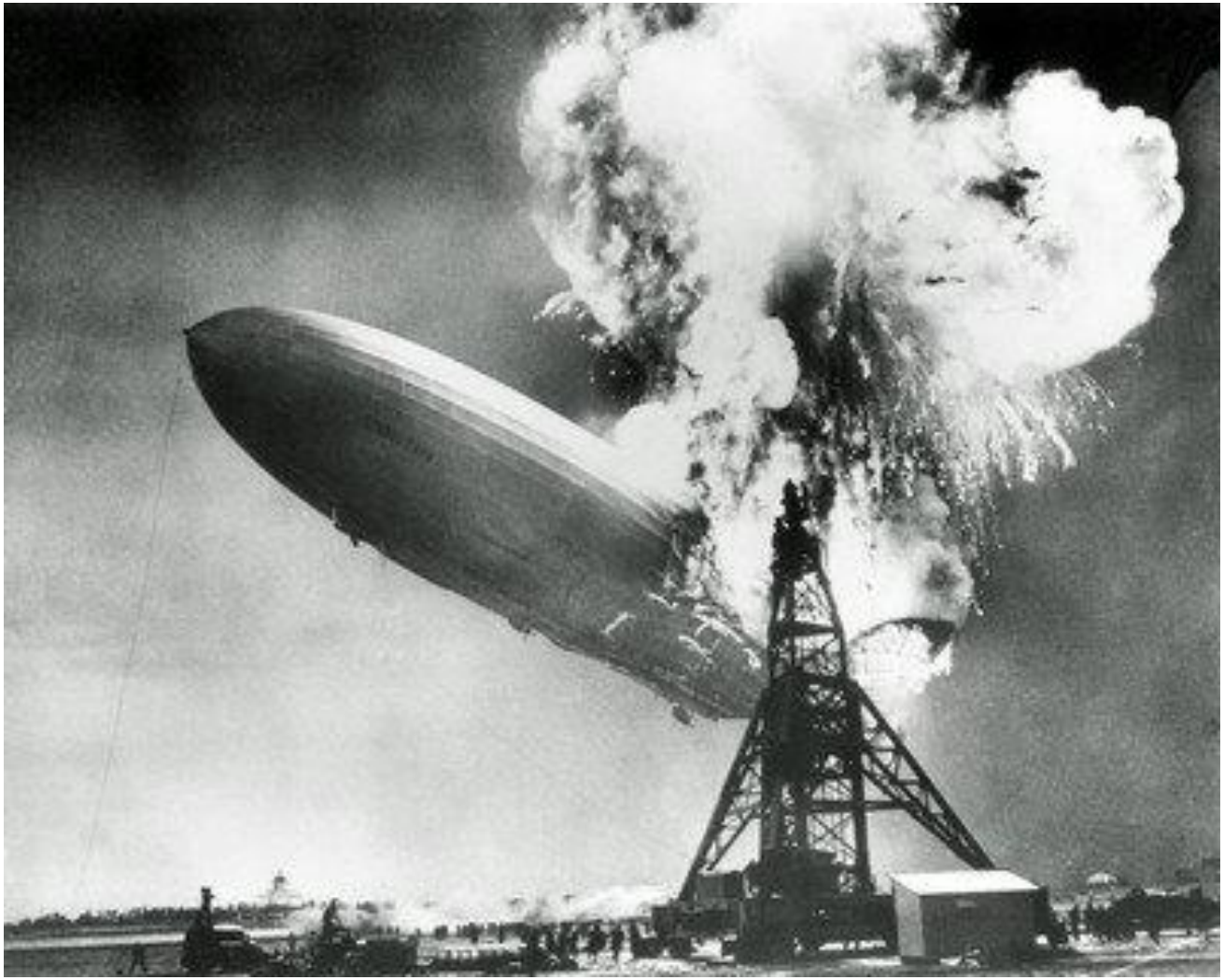
How does this make ACS look?

Large Margins of Error

Large Margins of Error

Large Margins of Error

Large Margins of Error



Large Margins of Error

Horror

Large Margins of Error

- Let's move away from the ledge

- Look at some data

Households by Type and Size

Pct of BGs with MOE GT Cell Value – and Pct Cell Value = 0

HH Type & Size	Pct MOE GT Cell Value	Pct Cell Value = 0
Total Households	1.1	**0.7**
Family 2-persons	8.4	**1.0**
Family 3-persons	25.4	**4.8**
Family 4-persons	34.6	**8.0**
Family 5-persons	65.4	**22.9**
Family 6-persons	90.6	**52.3**
Family 7+ persons	67.9	**67.7**
Nonfamily 1 person	10.3	**2.0**
Nonfamily 2 persons	76.7	**30.4**
Nonfamily 3 persons	99.0	**82.9**
Nonfamily 4 persons	99.7	**93.1**
Nonfamily 5 persons	100.0	**97.8**
Nonfamily 6 persons	100.0	**99.3**
Nonfamily 7+ persons	100.0	**99.6**

Households by Type and Size

San Diego County, CA Tract 99.01

HH Type & Size	Estimate	MOE	Implied Range
Total Households	8	**+/- 14**	0 to 22
Family 2-persons	8	**+/- 14**	0 To 22
Family 3-persons	0	**+/- 132**	0 to 132
Family 4-persons	0	**+/- 132**	0 to 132
Family 5-persons	0	**+/- 132**	0 to 132
Family 6-persons	0	**+/- 132**	0 to 132
Family 7+ persons	0	**+/- 132**	0 to 132
Nonfamily 1 person	0	**+/- 132**	0 to 132
Nonfamily 2 persons	0	**+/- 132**	0 to 132
Nonfamily 3 persons	0	**+/- 132**	0 to 132
Nonfamily 4 persons	0	**+/- 132**	0 to 132
Nonfamily 5 persons	0	**+/- 132**	0 to 132
Nonfamily 6 persons	0	**+/- 132**	0 to 132
Nonfamily 7+ persons	0	**+/- 132**	0 to 132

Households by Type and Size

Interesting that . . .

- MOE range extends to negative values
  - 90 pct confident between -132 and +132
- MOE is constant for “0” cells
  - Implies uncertainty same for all “0” cells
  - “0” as unlikely for “2 person family” as for “7 person non-family”?
- 90 pct confident “total HH” is between 0 and 22
- 90 pct confident “nonfam 7+” is between 0 to 132

Households by Type and Size

- MOE does not measure actual error
- Does not tell us
  - This is a good estimate
  - This is a bad estimate
- Some problems evident without MOE
- For example:
  - Large numbers of HHs with 7+ persons
- Let's check some example block groups

Households by Type and Size

HH Type & Size	482659606006	270619802001	170770109002
Unweighted HU	30	151	19
Total Households	535	401	271
Family 2-persons	98	123	75
Family 3-persons	81	22	46
Family 4-persons	16	42	18
Family 5-persons	122	12	0
Family 6-persons	0	6	0
Family 7+ persons	18	4	0
Nonfamily 1 person	63	86	46
Nonfamily 2 persons	47	29	0
Nonfamily 3 persons	0	0	0
Nonfamily 4 persons	0	0	0
Nonfamily 5 persons	0	0	0
Nonfamily 6 persons	0	0	0
Nonfamily 7+ persons	90	77	86

Households by Type and Size

BG 17 077 0190.00 2

HH Type & Size	ACS 2005-09	2000 SF1	2000 SF3
Total Households	271	273	260
Family 2-persons	75	75	83
Family 3-persons	46	54	45
Family 4-persons	18	30	29
Family 5-persons	0	18	18
Family 6-persons	0	8	0
Family 7+ persons	0	7	16
Nonfamily 1 person	46	71	50
Nonfamily 2 persons	0	7	19
Nonfamily 3 persons	0	2	0
Nonfamily 4 persons	0	0	0
Nonfamily 5 persons	0	1	0
Nonfamily 6 persons	0	0	0
Nonfamily 7+ persons	**86**	**0**	**0**

What's going on here?

- Too many "nonfam 7+"
  - Maybe 1 captured by ACS sample
  - Weighted up to 86

- But why weighted so high?
  - Nonfam 7+ is rare
  - Many BGs with 1 or 2. But none captured by ACS
  - ACS shows "0"
  - Where ACS does capture a "7+" HH
  - Have to weight extra
  - Compensate for BGs with 1 or 2 but show "0"
  - Otherwise national total is very low

What's going on here?

- Could improve accuracy of BG data
  - Reduce weight
  - Show fewer "Nonfam 7+"
  - But this would decrease accuracy for large areas
- Irony
  - Error in individual BGs can improve accuracy of aggregations
  - Reducing BG error can increase error of aggregations
- Remember
  - BG data intended for use in aggregations

What's going on here?

- Clearly some unrealistic estimates
  - But in relatively few block groups
- Most appear reasonable
- And some are clear improvement
  - Over aging 2000 Census data
- Consider some examples

More Typical Example

BG 09 001 0207.00 2

HH Type & Size	ACS 2005-09	2000 SF1	2000 SF3
Total Households	383	316	286
Family 2-persons	52	100	90
Family 3-persons	124	59	68
Family 4-persons	87	56	55
Family 5-persons	9	17	13
Family 6-persons	0	11	15
Family 7+ persons	0	1	0
Nonfamily 1 person	64	62	30
Nonfamily 2 persons	47	9	7
Nonfamily 3 persons	0	0	0
Nonfamily 4 persons	0	0	0
Nonfamily 5 persons	0	0	0
Nonfamily 6 persons	0	1	0
Nonfamily 7 persons	**0**	**0**	**8**

ACS Improvement

BG 51 059 4222.00 1

HH Type & Size	ACS 2005-09	2000 SF1	2000 SF3
Total Households	**681**	**6**	**6**
Family 2-persons	182	1	0
Family 3-persons	142	2	6
Family 4-persons	175	1	0
Family 5-persons	15	1	0
Family 6-persons	26	0	0
Family 7+ persons	39	0	0
Nonfamily 1 person	76	0	0
Nonfamily 2 persons	17	1	0
Nonfamily 3 persons	9	0	0
Nonfamily 4 persons	0	0	0
Nonfamily 5 persons	0	0	0
Nonfamily 6 persons	0	0	0
Nonfamily 7 persons	0	0	0

ACS and Rapid Growth

- Block Group 51 059 4222.00 1
  - The former Lorton Prison
  - 6 households in 2000
  - 921 households in 2010
- NOTE:
  - ACS improves over 2000 (more current)
  - Reflects dramatic growth in HHs
  - From 6 to 681
  - Even though ACS “not about counts”
- Consider another dramatic growth area

ACS and Rapid Growth

Tract 08 031 0041.05

HHs by Income	ACS 2005-09	MOE	2000 SF3
Total Households	**2,901**	+/- 139	**3**
Less than $10,000	117	+/- 80	0
$10,000 - $14,999	53	+/- 52	0
$15,000 - $19,999	32	+/- 27	0
$20,000 - $24,999	60	+/- 57	0
$25,000 - $29,999	28	+/- 28	0
$30,000 - $34,999	35	+/- 29	0
$35,000 - $39,999	64	+/- 53	0
$40,000 - $44,999	46	+/- 33	0
$45,000 - $49,999	111	+/- 65	0
$50,000 - $59,999	157	+/- 66	0
$60,000 - $74,999	126	+/- 64	0
$75,000 - $99,999	398	+/- 88	0
$100,000 - $124,999	402	+/- 109	0
$125,000 - $149,999	390	+/- 115	0
$150,000 - $199,999	463	+/- 95	0
$200,000 or more	419	+/- 100	0

ACS and Rapid Growth

- Tract 08 031 0041.05
  - Denver's former airport ("Stapleton")
  - Now "Stapleton" community
  - 3 households in 2000
  - 4,092 households in 2010
- NOTE:
  - ACS shows 2,901 households for 2005-2009
  - Improved income estimates
  - 2000 SF3 had no income distribution
  - ACS capturing major household growth

Curious Findings

- ACS often better than MOEs suggest
- Can provide significant value
- Still some curious findings

Curious Findings

- 2005-2009 ACS in 2000 census geography
- But ACS and 2000 census tract/BGs do not always match
- Example: Bibb County, AL

Census 2000 tracts	ACS Tracts
100.00	100.01
101.00	100.02
102.00	100.03
	100.04

Census Bureau Explanation

- For 18 Counties
- ACS inadvertently produced for 2010 tracts and BGs
  - Documented at this link
  - http://www.census.gov/acs/www/data_documentation/geography_notes/

Curious Findings

- Government units have priority over statistical geographies
- ACS samples
  - Stronger in small towns than tracts and BGs
- Following slide:
  - Ratio (ACS interviews / ACS households)
  - Rough measure: Percent of HHs interviewed
  - By number of households
  - For both Place and Block Group geographies

Government vs. Statistical Geography

Mean ratio (unweighted HU/HHs) by N of HHs

Places

HHs	N	Mean Ratio
All	24,727	20.4
1-49	1,562	47.2
50-99	2,197	35.0
100-199	3,166	32.3
200-499	4,892	22.8
500-999	3,757	17.5
1000 +	9,153	8.0

Block Groups

HHs	N	Mean Ratio
All	206,463	9.3
1-49	465	24.0
50-99	1,833	13.9
100-199	11,463	10.8
200-499	109,133	9.9
500-999	65,622	8.4
1000 +	17,947	7.1

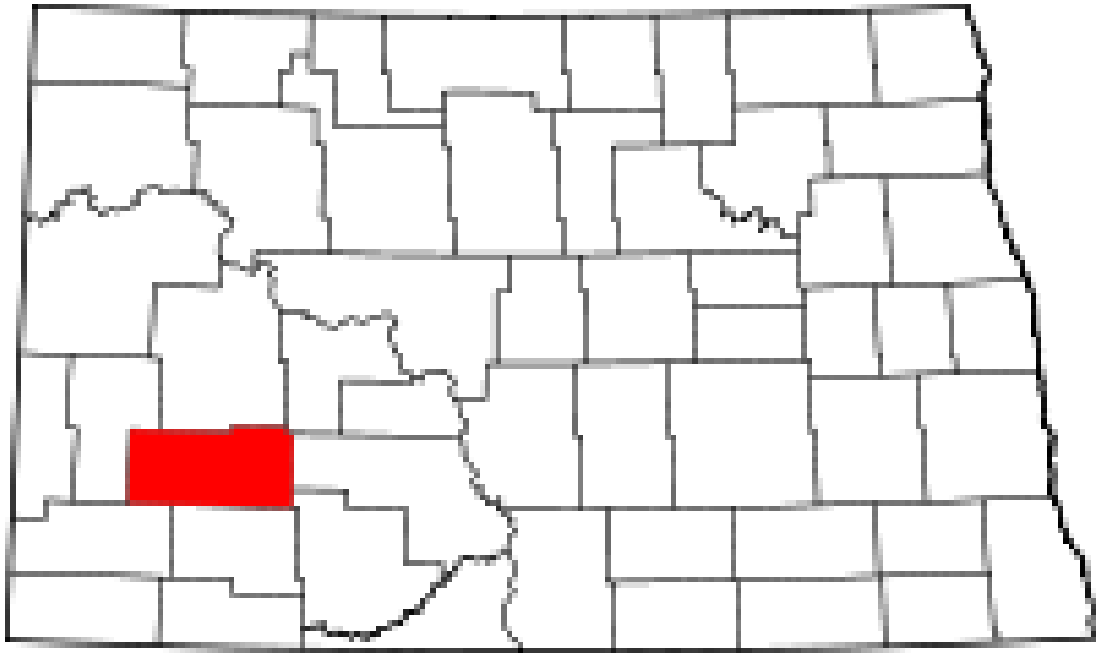
Government vs. Statistical Geography

- Cities and towns get stronger samples
- Even if very small
  - And there are many of them
- Did you know . . .
  - More Places than BGs with 1-49 HHs
  - More BGs than Places with 1,000+ HHs
- Let's look at an example . . .

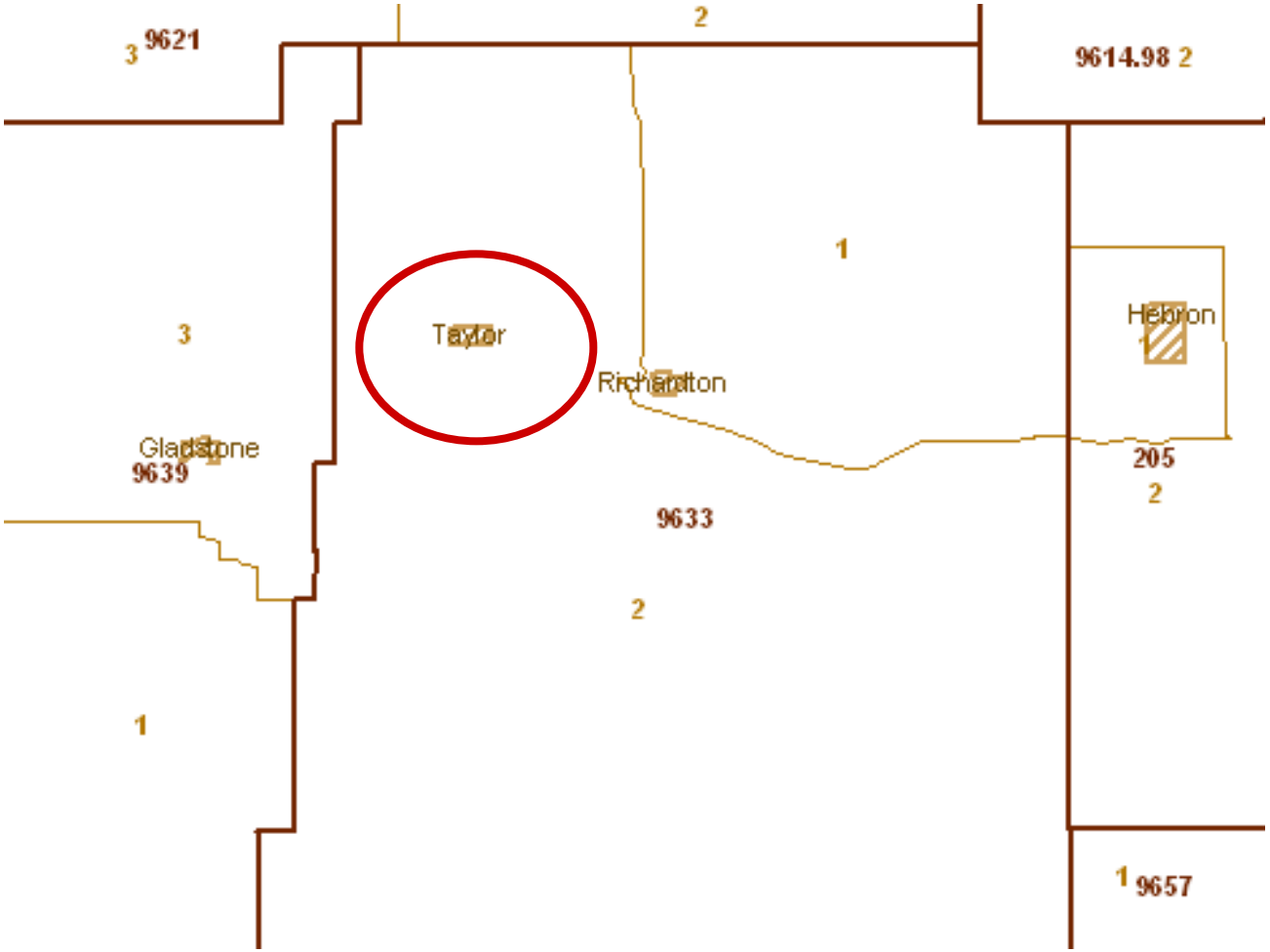
“Taylor City” North Dakota

- Small town in Stark County
- Home of the annual Taylor Horse Fest

Taylor, North Dakota



Taylor, North Dakota



Taylor, North Dakota

Taylor, North Dakota

- Taylor “city” is a small
- 2010 Census
  - Population = 148
  - Households = 75
  - Housing Units = 96

Taylor, North Dakota

- But has a pretty good ACS sample

- Unweighted HU = 31
  - 32 pct of 2010 HU
  - 41 pct of 2010 HH

Taylor, North Dakota

Households by Type and Size

HH Type & Size	ACS 05-09	MOE	2000 SF3	2000 SF1
Total Households	82	+/- 17	67	65
Family 2-persons	35	+/- 12	24	23
Family 3-persons	10	+/- 9	8	8
Family 4-persons	0	+/- 93	2	8
Family 5-persons	2	+/- 3	3	4
Family 6-persons	15	+/- 14	2	1
Family 7+ persons	0	+/- 93	0	0
Nonfamily 1 person	20	+/- 10	23	20
Nonfamily 2 persons	0	+/- 93	2	1
Nonfamily 3 persons	0	+/- 93	0	0
Nonfamily 4 persons	0	+/- 93	3	0
Nonfamily 5 persons	0	+/- 93	0	0
Nonfamily 6 persons	0	+/- 93	0	0
Nonfamily 7 persons	0	+/- 93	0	0

Taylor, North Dakota

Population by Race and Ethnicity

	ACS 05-09	MOE	2000 Cen	2010 Cen
Total Population	215	+/- 88	150	148
White	215	+/- 88	150	146
Black or African American	0	+/- 93	0	0
American Indian or AK Native	0	+/- 93	0	1
Asian	0	+/- 93	0	0
Native HI and Other Pac Islander	0	+/- 93	0	0
Some Other Race	0	+/- 93	0	1
Two or More Races	0	+/- 93	0	0
Hispanic Origin	0	+/- 93	0	1

Taylor, North Dakota

Housing Units by Year Structure Built

	ACS 05-09	MOE	2000 SF3
Total Housing Units	88	+/- 17	88
Built 2005 or later	0	+/- 93	na
Built 2000 – 2004	12	+/- 16	na
Built 1990 - 1999	0	+/- 93	4
Built 1980 – 1989	14	+/- 12	9
Built 1970 – 1979	8	+/- 6	17
Built 1960-1969	10	+/- 9	4
Built 1950 – 1959	4	+/- 4	11
Built 1940 – 1949	4	+/- 4	9
Built 1939 or earlier	36	+/- 14	34

Taylor, North Dakota

Occupied Housing by Heating Fuel

	ACS 05-09	MOE	2000 SF3
Total Occupied Units	82	+/- 17	66
Utility gas	71	+/- 17	54
Bottled, tank, or LP gas	2	+/- 4	3
Electricity	9	+/- 7	9
Fuel oil, kerosene, etc.	0	+/- 93	0
Coal or coke	0	+/- 93	0
Wood	0	+/- 93	0
Solar energy	0	+/- 93	0
Other fuel	0	+/- 93	0
No fuel used	0	+/- 93	0

Taylor, North Dakota

Occupation: Civilian Employed Population 16+

	ACS 05-09	MOE	2000 SF3
Civilian employed pop 16+	114	+/- 52	73
Management, Prof, and related services	31	+/- 17	24
Service occupations	4	+/- 4	7
Sales and office occupations	57	+/- 36	22
Farming, fishing, and forestry occup.	0	+/- 93	4
Construction, extraction, maint, & repair	11	+/- 8	2
Production, transpt, & material moving	11	+/- 9	14

Conclusions

- ACS 5-year data are a mixed bag
- Limitations
  - Relatively small sample
  - Large margins of error
  - Some extreme values for rare characteristics
- Promise
  - Data often better than MOEs suggest
  - Especially for small cell values
  - ACS adds value in rapid growth areas
  - Small town data might be pretty good

Conclusions

Bottom Line

- Too early to pass final judgment on ACS
- Value depends on one's application
- At this early stage
  - ACS good enough to merit user support and advocacy
  - But not so good that it can absorb cuts in funding
- We would miss the ACS if it were eliminated

Thank You

Ken Hodges
ken.hodges@nielsen.com