

Comparing Estimates of Rare Events in the ACS to Administrative Records

KATHERINE NESSE, KANSAS STATE UNIVERSITY

MALLORY RAHE, OREGON STATE UNIVERSITY

ANDI IRAWAN, UNIVERSITY OF ILLINOIS, URBANA-CHAMPAIGN

ACS DATA USERS CONFERENCE

MAY 13, 2015

Right now, three-quarters of the fastest-growing occupations require more than a high school diploma. And yet, just over half of our citizens have that level of education.

We have one of the highest
high school dropout rates
of any industrialized nation.

Pres. Obama, 2009 State of the Union

Data Sources



Dropouts = students enrolled in Fall – students enrolled in Spring

$$\text{Dropout rate} = \frac{\text{Dropouts}}{\text{Students enrolled in Fall}}$$

Data Sources

American Community Survey

$$\text{Dropout Rate} = \frac{\text{Population 16–19 Not Enrolled in School, no degree}}{\text{Population Ages 16–19}}$$

Table B14005: Sex by School Enrollment by Educational Attainment By Employment Status for the Population 16 to 19 Years

Sources of error

- Statistical
 - Sampling
 - Nonsampling
- Conceptual



Statistical (sampling)

Percent of confidence intervals that contain the administrative record of the average dropout rate 2008-2012.

	Total number	Percent Not in CI
Illinois		
Counties	99	45%
School Districts	352	22%
Oregon		
Counties	35	51%
School Districts	132	24%

Source: 2008-2012 American Community Survey, table B14005; ISBE School Report Cards; ODE Dropout Records

Statistical (sampling)

Percent of confidence intervals that contain zero.

	Total number	Percent contain 0
Illinois		
Counties	99	14%
School Districts	352	52%
Oregon		
Counties	35	23%
School Districts	132	54%

Source: 2008-2012 American Community Survey, table B14005; ISBE School Report Cards; ODE Dropout Records

Statistical (sampling)

The coefficient of variation for estimates of the dropout rate.

	CV 0 to 9.9	CV 10 to 29.9	CV 30 to 59.9	CV 60 and over
Illinois				
Counties	2%	30%	54%	14%
School Districts	0%	10%	38%	52%
Oregon				
Counties	0%	43%	34%	23%
School Districts	0%	13%	34%	53%

Source: 2008-2012 American Community Survey, table B14005; ISBE School Report Cards; ODE Dropout Records

Statistical (nonsampling)

- Incomplete responses
- Response rate
- Interviewer bias
- Clerical errors

Person 2 (continued)

29 a. **LAST WEEK, did this person work for pay at a job (or business)?**
☐ Yes → SKIP to question 30
☐ No – Did not work (or retired)

30 b. **LAST WEEK, did this person do ANY work for pay, even for as little as one hour?**
☐ Yes
☐ No → SKIP to question 35a

30 **At what location did this person work LAST WEEK?** If this person worked at more than one location, print where he or she worked most last week.

a. **Address (Number and street name)**
[]

If the exact address is not known, give a description of the location such as the building name or the nearest street or intersection.

b. **Name of city, town, or post office**
[]

c. **Is the work location inside the limits of that city or town?**
☐ Yes
☐ No, outside the city/town limits

d. **Name of county**
[]

e. **Name of U.S. state or foreign country**
[]

f. **ZIP Code**
[]

31 How did this person usually leave home LAST WEEK?
[]

32 How many people, including this person, usually rode to work in the car, truck, truck, or van" in question 31. Otherwise, SKIP to question 33.
Person(s) []

33 What time did this person usually leave home to go to work LAST WEEK?
Hour [] Minute []
☐ a.m.
☐ p.m.

34 How many minutes did it usually take this person to get from home to work LAST WEEK?
Minutes []

35 a. **LAST WEEK, was this person a job?**
☐ Yes
☐ No → SKIP to question 39a

39 a. **Did this person receive any pay for work last week?**
☐ Yes
☐ No → SKIP to question 39a

K Answer questions 35 – 38 if this person did NOT work last week. Otherwise, SKIP to question 39a.

Statistical (nonsampling)

The imputation rate for school enrollment. (Counties)

Illinois Counties				Oregon Counties			
	Metro	Nonmetro	Cumulative		Metro	Nonmetro	Cumulative
1 % or less				1 % or less		4%	3%
1.01 to 2%	3%	17%	12%	1.01 to 2%		8%	8%
2.01 to 3%	44%	44%	57%	2.01 to 3%	73%	60%	72%
3.01 to 5%	50%	33%	96%	3.01 to 5%	27%	28%	100%
5.01 to 7%	3%	3%	99%	5.01 to 7%			100%
more than 7%		2%	100%	more than 7%			100%
Total Counties	36	63	99	Total Counties	11	25	36

Source: 2008-2012 American Community Survey, table B99141 Imputation rate of school enrollment for the population 3 years and older

Statistical (nonsampling)

The imputation rate for school enrollment.

	Illinois School Districts				Oregon School Districts		
	Metro	Nonmetro	Cumulative		Metro	Nonmetro	Cumulative
1 % or less	1%	3%	2%	1 % or less	6%	12%	9%
1.01 to 2%	10%	17%	15%	1.01 to 2%	17%	19%	27%
2.01 to 3%	36%	34%	50%	2.01 to 3%	35%	34%	61%
3.01 to 5%	46%	38%	93%	3.01 to 5%	33%	29%	92%
5.01 to 7%	5%	4%	97%	5.01 to 7%	8%	4%	98%
more than 7%	2%	4%	100%	more than 7%	1%	3%	100%
Total Districts	196	156	352	Total Districts	84	101	185

Source: 2008-2012 American Community Survey, table B99141 Imputation rate of school enrollment for the population 3 years and older

Statistical Uncertainty

- Large for rare characteristics
- Larger for small areas



Photo: Frédéric Poirot, "Your own destiny", via www.flickr.com, used under the creative commons license

Conceptual Uncertainty

- Who is a dropout?
 - Time of year
 - Time away from school
- Who is the population?
 - Age
 - Location



Photo: butterbits, "Dropout Bike Club", via [www.flickr.com](https://www.flickr.com/photos/butterbits/), used under the creative commons license

Conceptual Uncertainty

Comparison of the ACS-based dropout rate under different definitions.

Numerator	Denominator	Mean	St.D.	Min	Max
Illinois					
Students not enrolled in Spring	Students enrolled in fall	2.3%	1.5	0.0%	8.9%
16-19 not enrolled, not graduated	population 16-19 enrolled in school	7.9%	9.1	0.1%	79.5%
16-19 not enrolled, not graduated	Population 16-19	6.1%	5.6	0.1%	41.1%
Oregon					
Students not enrolled in Spring	Students enrolled in fall	2.8%	1.8	0.0%	11.7%
16-19 not enrolled, not graduated	population 16-19 enrolled in school	11.4%	10.9	0.4%	83.1%
16-19 not enrolled, not graduated	Population 16-19	8.4%	4.3	1.0%	38.6%

Source: 2008-2012 American Community Survey, table B14005; ISBE School Report Cards; ODE Dropout Records

Error matters to public policy

- Error is not solved by a larger sample size
- Data needs to be viewed in its full complexity
- ACS should augment policy decisions, not drive them

Comparing Estimates of Rare Events in the ACS to Administrative Records

KATHERINE NESSE, KANSAS STATE UNIVERSITY
KNESSE@K-STATE.EDU

MALLORY RAHE, OREGON STATE UNIVERSITY
MALLORY.RAHE@OREGONSTATE.EDU

ANDI IRAWAN, UNIVERSITY OF ILLINOIS, URBANA-CHAMPAIGN
IRAWAN1@ILLINOIS.EDU