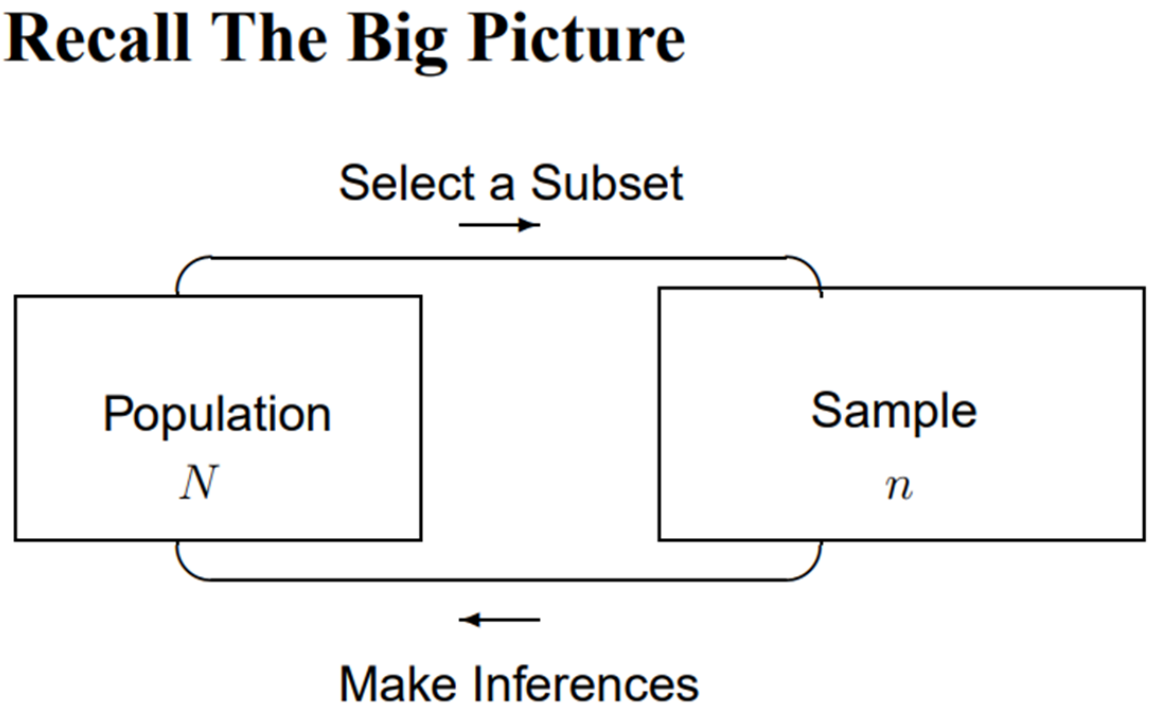
Research Methods (HDFS 3390),
Alan Reifman, Texas Tech University
Sampling (Selecting Participants)

- How can we describe a population of perhaps millions of people, without having to survey every one of them?
- Let's start with the "Big Picture" from this U. of Illinois PowerPoint show (a few pages in).



Carolyn J. Anderson (2005), "Sampling Distributions, the CLT, and Estimation"

Next, let's carefully define the differences between a population and a sample...

POPULATION	SAMPLE
All members of a defined group	Small subset of population
Population of Texas = All residents of Texas	Sample = 500 Texans (for example)
Mean (average) of population is known as a "true mean" or "true value" because it is based on all members	Mean of sample is estimate of true population mean (estimate is good if sample is representative)
True population mean symbolized by Greek letter Mu (μ)	Sample mean symbolized by "M"
What we often call "statistics" (e.g., the mean) are known as "parameters" in a population	"Statistic" is the applicable term for a sample
We virtually never know μ or other parameters because it's impractical to survey an entire population (U.S. Census does, but only once every 10 years)	Sample surveys are done every day

What Pollsters and Samplers are Trying to Do

- **The goal is to generalize from a relatively small sample (e.g., 1,000 people) to a larger population (e.g., U.S. adults), as in a political poll.**
- **In other words, we want to use the findings from the sample survey to characterize the larger population even though the researchers didn't interview every American adult (examples include saying that the average American adult devotes 3 hours per day to entertainment or determining the monthly unemployment rate from the Current Population Survey).**
- **Being able to generalize from a sample survey is also known as "external validity" (not to be confused with measurement validity).**
- **Sample must be representative of the larger population to have external validity.**
- **As described by Frank Newport, former polling director at Gallup, polls:**
  - **"...use the results of interviews with relatively small numbers of people to estimate what would have been found if it had been possible to interview every person in enormously large groups. Pollsters, in other words, aren't interested only in the opinions of the people actually interviewed. We're interested in using these sample data to estimate the opinions of much, much larger numbers of people" (p. xi).**
  - **"...although all humans are unique at some basic level, they are *not* unique when it comes to the broad and general characteristics in which pollsters are usually interested" (p. 76).**
  - **Newport, F. (2004). *Polling matters: Why leaders must listen to the wisdom of the people*. New York: Warner Books.**
- **As the authors of *The Numbers Game* describe sampling: "...only a few... are counted, assuming they are representative of the rest, then multiplied to the right size for the whole country" (p. 111)**

Two Most Prominent Random Sampling Techniques

SIMPLE RANDOM SAMPLE (SRS), like a lottery (example)

- **Random Digit Dialing (RDD) can be seen as a variation of SRS (in theory, all possible phone numbers could be listed, e.g., in Lubbock, (806) 797-0000, 0001, 0002, ..., 9999)**

SYSTEMATIC SAMPLE, "interval jumping," as shown in the next slide…

ADDITIONAL POINTS

- **SRS and systematic sampling require that you have a master list of everyone in the study population, technically known as a sampling frame.**
- **The sampling frame is similar to the population. However, whereas the population includes all members of a group (e.g., all residents of Dallas), a pollster using a list of all Dallas driver's-license holders as a sampling frame would miss people without licenses. This is known as coverage error (the sampling frame not "covering" the full population).**
- **SRS and systematic sampling will produce a sample in which every member of the sampling frame has an equal probability of selection. This is the common definition of a random sample.**
- **Advanced point for graduate students: The broader term "probability sampling" refers to when "every element of the population has a known probability of being included in the sample." Oversampling of groups that comprise a relatively small share of the population, when a researcher wants to study the group in depth, is OK. According to this document, "Probabilities of selection may be different for different groups, as long as they are known." Sample weighting (discussed in later slide) can be used to make a group's share of the sample match its share of the population, if desired (i.e., oversampled group can be weighted back down to original share of the population).**

Suppose we wanted to survey approximately 85 of the 435 members of the U.S. House of Representatives, at random.

When taking every 5th name, we need a random starting point between 1- 5.

	Representative	Party	State	District
1	Ackerman, Gary L.	D	NY	5th
2	*Aderholt, Robert B.*	R	AL	4th
3	Adler, John H.	D	NJ	3rd
4	*Akin, W. Todd*	R	MO	2nd
5	*Alexander, Rodney*	R	LA	5th
6	Altmire, Jason	D	PA	4th
7	Andrews, Robert E.	D	NJ	1st
8	Arcuri, Michael A.	D	NY	24th
9	*Austria, Steve*	R	OH	7th
10	Baca, Joe	D	CA	43rd
11	*Bachmann, Michele*	R	MN	6th
12	*Bachus, Spencer*	R	AL	6th
13	Baird, Brian	D	WA	3rd
14	Baldwin, Tammy	D	WI	2nd
15	*Barrett, J. Gresham*	R	SC	3rd
16	Barrow, John	D	GA	12th
17	*Bartlett, Roscoe G.*	R	MD	6th
18	*Barton, Joe*	R	TX	6th
19	Bean, Melissa L.	D	IL	8th
20	Becerra, Xavier	D	CA	31st
21	Berkley, Shelley	D	NV	1st
22	Berman, Howard L.	D	CA	28th
23	Berry, Marion	D	AR	1st
24	*Biggert, Judy*	R	IL	13th
25	*Bilbray, Brian P.*	R	CA	50th
26	*Bilirakis, Gus M.*	R	FL	9th

If all the D's (Democrats) were listed first, on lines 1 through 255*, followed by all the R's (Republicans), on lines 256 onward, this would be an example of *STRATIFICATION*. What might be an advantage of stratifying?

*Number of Democrats in the U.S. House as of May 25, 2010.

Sources: http://clerk.house.gov/member_info/olmbr.html (list); http://en.wikipedia.org/wiki/111th_United_States_Congress (party breakdown)

STRATIFICATION, the lining up of people with like characteristics, is often used together with systematic sampling, but it is a separate concept.

What's the Average Weight of the Texas Tech Football Team?

Demonstration of drawing random samples and using those sample means (statistic M) to estimate the true population mean (parameter μ) of player weights on the Texas Tech football team. We have computer generate random numbers, which we can use to select players by uniform numbers.*

Mu (μ) calculated from all players (see "*N*" in left-hand column); sample means calculated from 10 or 15 randomly selected players, according to this color scheme: 10 SRS, 15 SRS, ~10 Sys, ~15 Sys

Weight Intervals	200-210	211-215	216-220	221-223.9	224-226	227-230	231-233	234-237	238-242	243-247	248+
Spring '06 Class (2005 Roster, *N* = 90)							μ = 232.83	235.0	241.75		
Spring '07 Class (2006, *N* = 86)			219.38		224.25	227.38 227.81	231.77	μ = 236.85 236.89	238.4	246.5	
Fall '07 Class (2007, *N* = 91)				222.67 223.2	224 225.08	228.8 229.40	μ=231.54 231.4			243.71	
Spring '08 Class (2007, *N* = 91)	209.0		219.71		225.4 225.69		μ = 231.54	234.67 235.36		244.62	249.3
Spring '09 Class (2008, *N* = 87)	209.56	213.77 214.11	219.83			μ=229.33 227.67 227.75	231.5				
Spring '10 Class (2009, *N* =118)					μ = 224.67						
Summer I '10 (2009, *N* = 118)	208.3		217.2	223.2	μ = 224.67		233.4				
Fall '19 Class (2019, *N* = 116)				223.21	225.0	μ = 229.40 227.0	231.0				
Spring '21 Class (2020, *N* = 122)						μ = 228.3					
Spring '22 001 (2021, *N* = 121)							μ = 233.1				
Spring '22 002 (2021, *N* = 121)							μ = 233.1				

*Texas Tech's football roster, like other schools', sometimes lists more than one player for a given uniform number. Usually , only one guy will play in the games, with the other(s) only participating in practice on the "scout" team (emulating the plays of the coming week's opponent). For the Spring '06-Spring '09 classes, I kept only one player per uniform number (based on greater player experience or coin flip). Beginning Spring '10, I'm keeping all players and assigning extra players with duplicate numbers to new make-believe numbers (100, 101, etc.) and effective Fall '19, to unused numbers.

Computation of True Mean (μ) Weight of 2021 Texas Tech Football Team

https://www.calculatorsoup.com/calculators/statistics/descriptivestatistics.php

Descriptive Statistics:

Minimum	min =	155
Maximum	max =	350
Range	R =	195
Size	n =	121
Sum	sum =	28210
Mean	$\overline{x}$ =	233.140496
Median	$\tilde{x}$ =	220
Mode	mode =	220, 230, 195
Standard Deviation	s =	46.3327289

FALL 2021 TTU Football (Spring '22 Classes Combined)

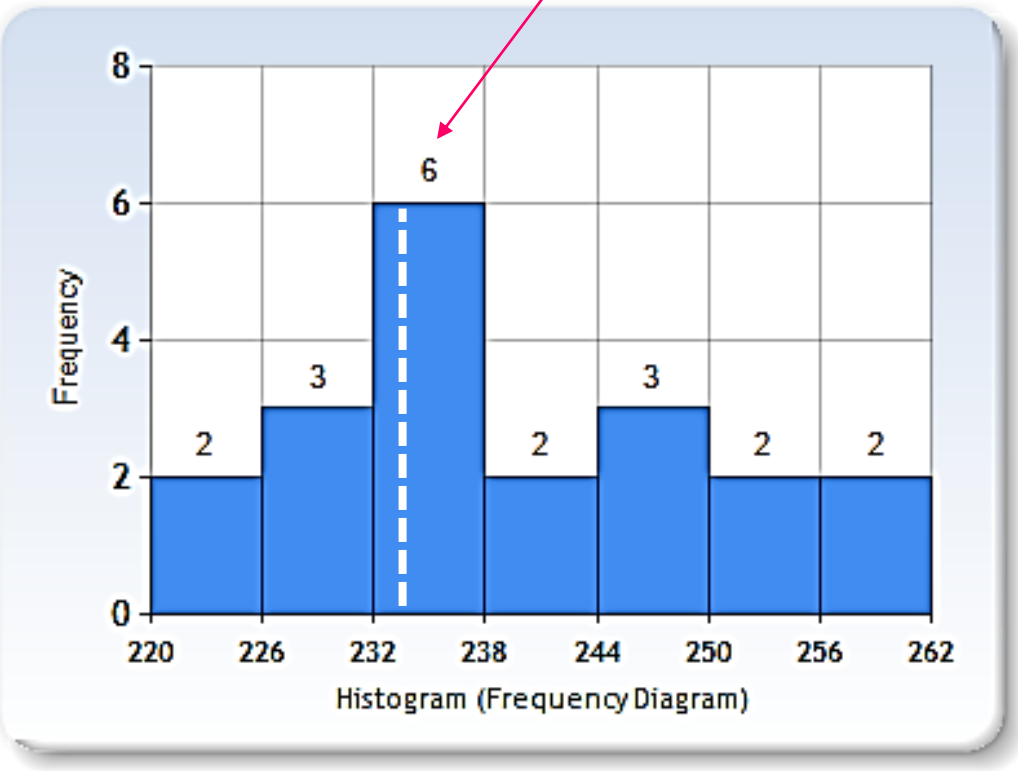
Taking the "Average of the Averages" from Everyone's 10-Player SRS Comes Very Close to the True Mean

Success! We have edited your histogram, and updated the frequency table and histogram information, based on the values you provided via the edit tool.

True mean (from entering all 121 players' weights) = μ = 233.1

Frequency Table	
Class	*Count*
220-225.9	2
226-231.9	3
232-237.9	6
238-243.9	2
244-249.9	3
250-255.9	2
258-261.9	2

Your Histogram	
Mean	238.125
Standard Deviation (s)	10.93507
Skewness	0.44014
Kurtosis	-0.49865
Lowest Score	220.5
Highest Score	259.5
Distribution Range	39
Total Number of Scores	20
Number of Distinct Scores	18
Lowest Class Value	220
Highest Class Value	261.9
Number of Classes	7
Class Range	6



Note that having your 10-player mean well below or well above the true value of 233.1 <u>in no way reflects anything done incorrectly</u>. Nor does a mean close to 233.1 signify anything done extraordinarily well. These are just random fluctuations.

https://www.socscistatistics.com/descriptive/histograms/

A little musical number to drive home these points...

SRS or Systematic

Lyrics by Alan Reifman

(May be sung to the tune of "Every Little Thing She Does Is Magic," Sting)

YouTube video of in-class performance (filmed by one of the students)

So you need to get a sample,
'Bout a thousand respondents, you'll find, makes sense,
There are two main ways to do it,
With a comprehensive list of residents,

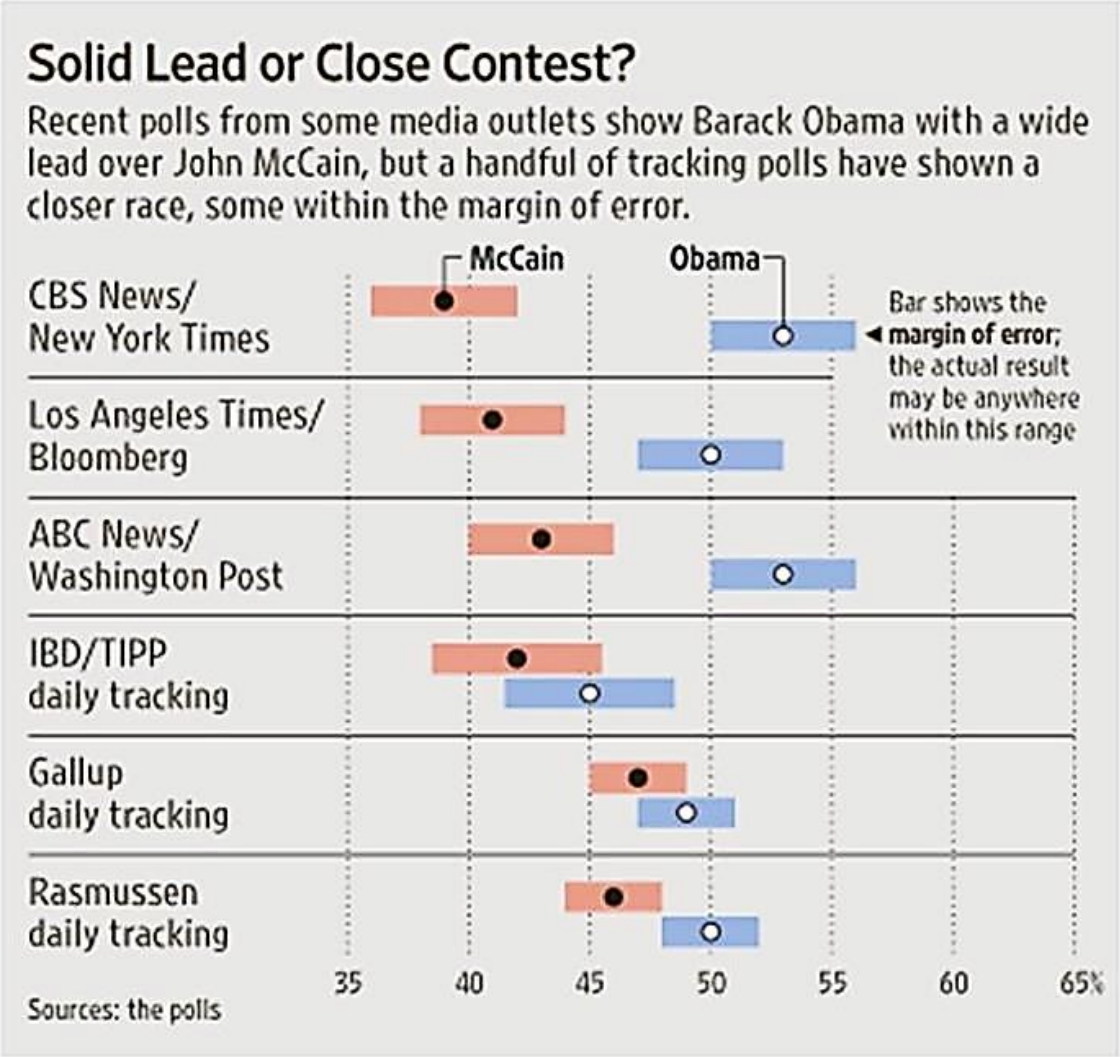
Interval jumping is systematic,
Like a lottery, is SRS,
Both of these techniques, give random samples,
To ensure rep-re-sen-ta-tive-ness,

All throughout the population,
An equal chance for each person, is the key,
Your results should be accurate,
That's within, plus or minus, the M-o-E,

Interval jumping is systematic,
Like a lottery, is SRS,
Both of these techniques, give random samples,
To ensure rep-re-sen-ta-tive-ness...

Margin of Error

Samples can often provide a good estimate of the true population value but will usually be off by a little. The margin of error (typically, +/- 3 percent away from the sample value) tells us that, if we were to have surveyed the entire population of interest, the true value would fall within a given range, with 95% confidence. Here's a nice example from the October 17, 2008, *Wall Street Journal*:



Margin of Error (Continued)

- **This AAPOR webpage provides an excellent description of how the margin of error relates to sample size and why a sample size of about 1,000 is "as good as it gets." Margin of error (or as the linked document calls it, "Margin of *Sampling* Error") depends not only on sample size, but also the breakdown of opinion in the survey (e.g., 50/50, 60/40, between two candidates). The AAPOR page appears to apply roughly to a 50/50 breakdown.**
    - **Article criticizing poll results based on sample size of only 165 respondents (which has a large MoE)**
- **The margin of error pertains only to random sampling error. Random sampling error occurs, for example, when a computer generates a set of phone numbers that is unintentionally slanted toward Democratic or Republican voters (like how a series of coin flips can yield slightly more or less than 50% heads, purely at random). This poll report from Harris (bottom of inked document) explains the distinction between sampling and other types of error:**
    - ***All surveys are subject to several sources of error. These include: sampling error (because only a sample of a population is interviewed); measurement error due to question wording and/or question order, deliberately or unintentionally inaccurate responses, nonresponse (including refusals), interviewer effects (when live interviewers are used) and weighting. With one exception (sampling error) the magnitude of the errors that result cannot be estimated...***
    - **In other words, there is no such thing as a +/- margin of error for question wording, inaccurate responses, etc. There is only a margin of error for random error in drawing the sample.**
- **Pew Research Center overview of margin of error (September 2016)**
- **Huffington Post's *Pollster* column further probes the concept of "margin of error," including whether it is ever justifiable to report a MoE for opt-in Internet surveys (2/3/2015).**

Further Illustration of Margin of Error:

Americans' Reaction to 9/11 Attacks (2001)

Poll Results

Let's look at question 4. Based on the information at the bottom of the linked document, the margin of error is listed as +/- 3 percent, with a sample size of roughly 1,000 (1,032 to be exact). In this sample, 58% supported a "long-term war to eliminate terrorist groups world-wide."

We don't know what percentage of the full population of American adults 18 and older (roughly 200 million people) would have supported then-President George W. Bush's position of a long-term war. The only way to know for sure would have been to survey all 200 million American adults, which was not practical.

What the sample survey tells us, however, is that if the entire American adult population had been surveyed, the percentage of the full population supporting Bush's position would likely (with 95% confidence) have ranged somewhere between ___ and ___ percent.

Ways to Increase Participation in a Survey:
Traditionally Thought to Improve the Accuracy of the Results*

- **Pre-notification (Census director explains purpose of sending advance letters, March 2010)**
- **Personalizing the request (i.e., "Dear [name of person]," instead of "Dear Resident“)**
    - **Results are actually mixed, with some studies showing personalization to increase participation, but others failing to find any difference. One chapter concludes that “… it appears that use of personalized salutations in invites may be helpful, particularly if invites can be sent from authoritative sources” (source).**
- **Sending a token of appreciation (e.g., a nice pen) in advance (group receiving pen had 5% higher participation rate than group not receiving pen).**
- **Cash rewards provided in advance.**
- **Callbacks (Census Bureau sends someone to your home if you haven't responded).**

However, giving incentives can actually be counter-productive under some circumstances. Jerold Pearson, the Director of Market Research for the Stanford Alumni Association, had found that women and people who donate to the university responded to surveys at higher rates than did men and non-donors, respectively. As summarized in this report (under "Incentives"), providing incentives had the effect of making women and donors participate in ever-higher numbers, but not men and non-donors. The conclusion was that "the incentive in fact *worsened* the sample composition by further increasing response among over-represented groups" (p. 2).

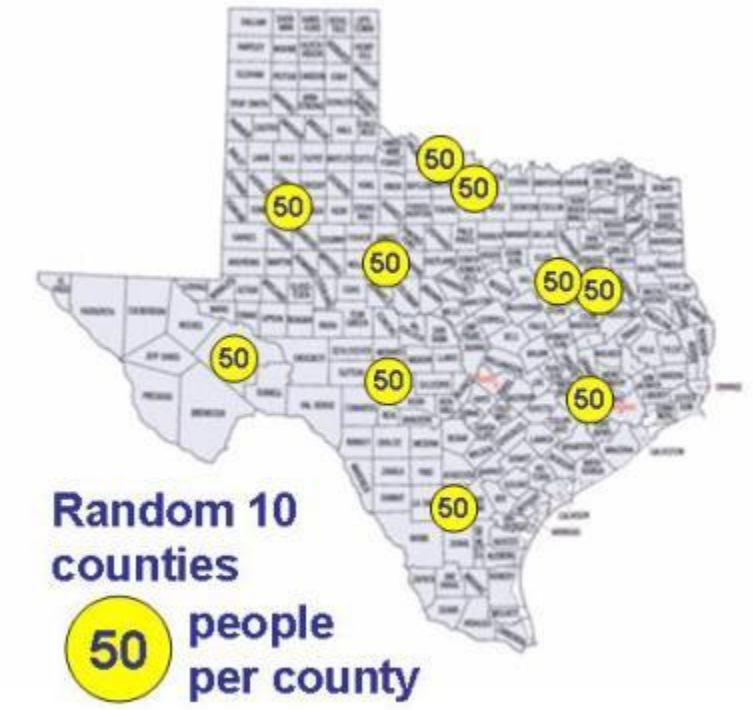
Some evidence, however, questions whether a low participation rate really is that harmful; also discussed here).

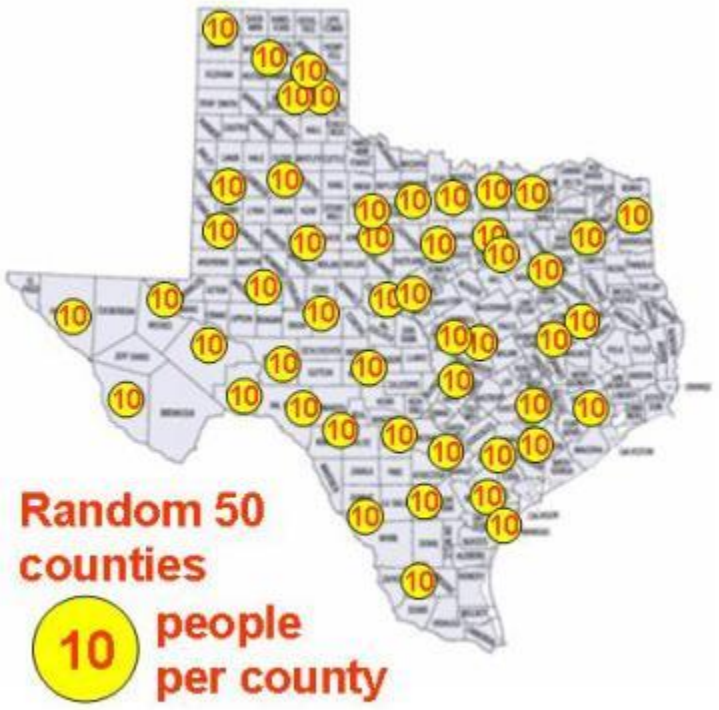
Multistage Cluster Sampling

For Drawing Random Samples without a Master List of Potential Participants

In MCS, you first sample "big units" where people are clustered (e.g., colleges, churches, counties). Then, within each selected cluster, you sample people ("small units").

EXAMPLE 1: Let's say we want to sample Texas counties to obtain a representative sample of state residents using MCS, and let's further say we only have the resources to survey a total of 500 Texans. Here are two possible strategies:

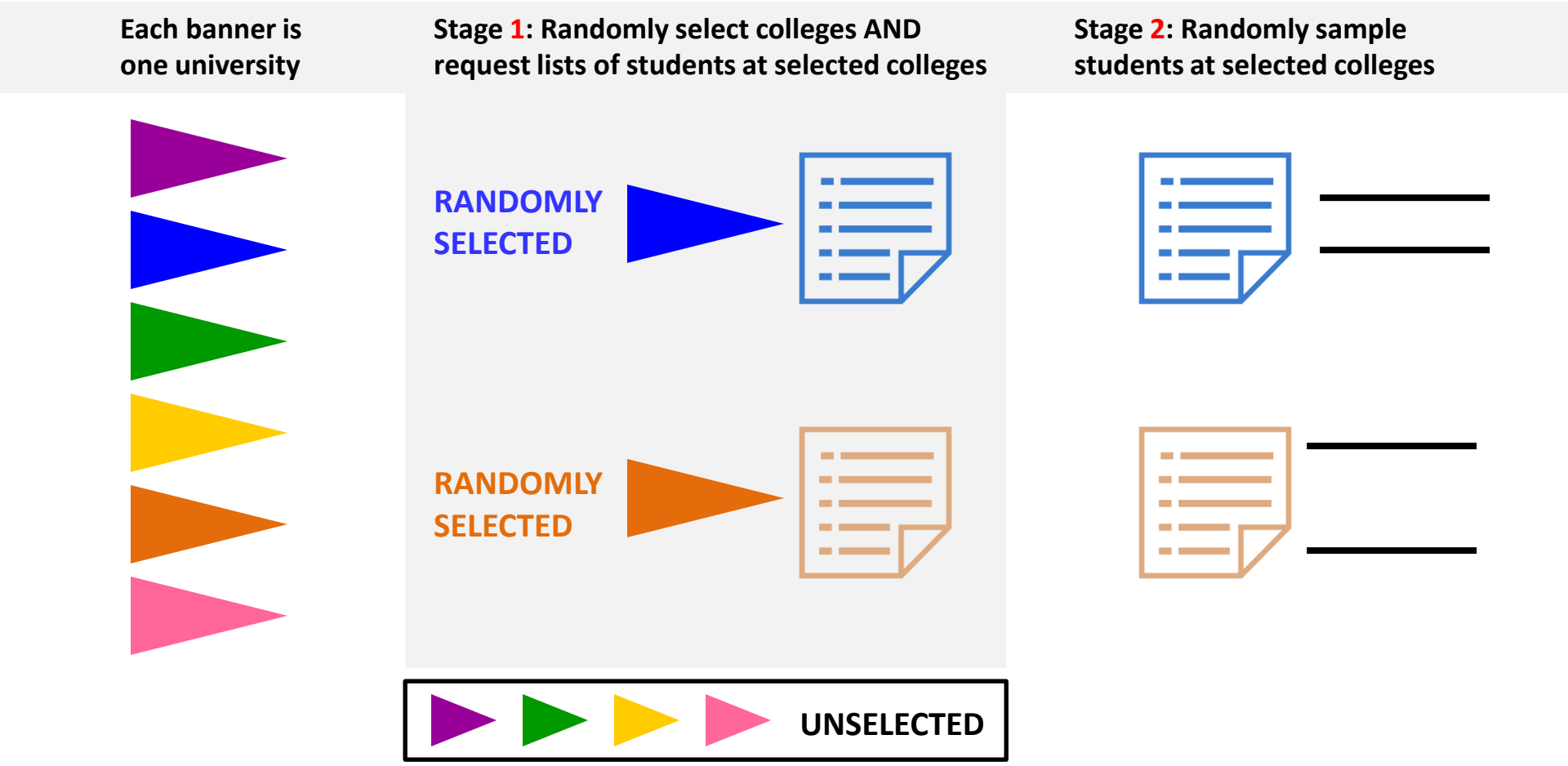




Babbie (p. 220) advises as follows: ***The general guideline for cluster design, then, is to maximize the number of clusters selected while decreasing the number of elements within each cluster.***

EXAMPLE 2: Harvard School of Public Health conducted the College Alcohol Study, a national survey of college-student binge drinking (Wechsler et al., 1994, vol. 272, *JAMA*)

- No central list of all American college students
- However, there are lists of colleges (~2,000 4-year colleges)
- Researchers first sampled colleges using American Council on Education list (195 colleges selected*)
- Then, asked each selected college to provide random sample of full-time undergraduate students*



*Further details for anyone interested: 140 colleges (72% of 195) participated; 28,709 students (~200 per college) were in the sample; 3,082 eliminated (incorrect address, withdrawal, leave of absence),leaving sample at25,627; 17,592 returned surveys (69%)

EXAMPLE 3: **Exit-polling on election days, to help the media project winners on election night and analyze possible reasons for the winning candidate's victory.**

- **Pollsters first randomly select precincts (clusters of voters or what I call "big units“**
- **Then randomly select exiting voters ("small units") to survey.**
- **This article provides an excellent overview of the process, and discusses how sampling error (i.e., margin of error) is wider in cluster sampling than in studies that sample directly from a comprehensive list (frame) of all people in the population.**

EXAMPLE 4: **Attempt by Johns Hopkins University researchers to estimate the number of civilian deaths in the Iraq War. The number the researchers came up with (as of October 2006) was 655,000 (note also the verification procedures used). However, a critic claims the number of clusters (neighborhoods) randomly sampled was too small, thus throwing off the results.**

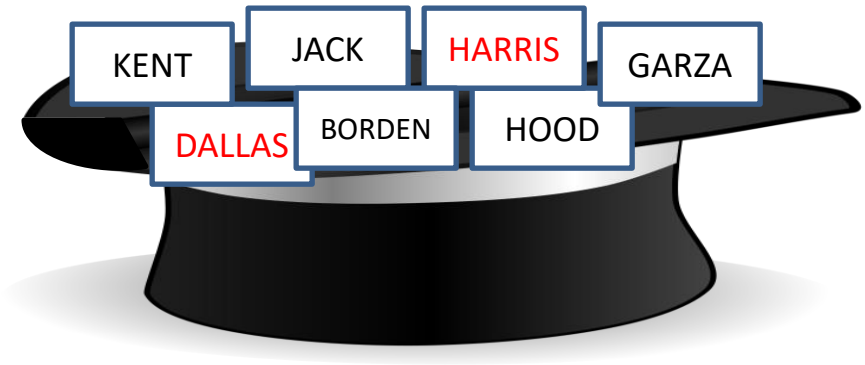
HYPOTHETICAL SITUATION: **How would you obtain a sample of fans of your favorite musical group or artist? Is there a comprehensive list of all fans of that group or artist? Where would fans of the group/artist be clustered?**

Probability Proportionate to Size (PPS) Sampling

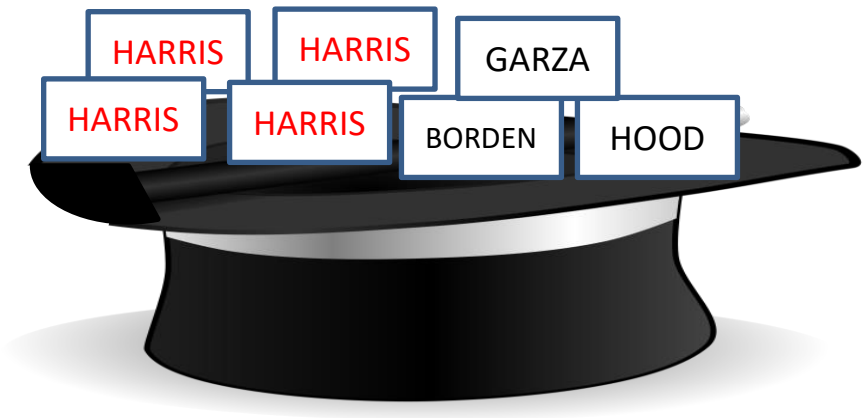
Used When Clusters are of Greatly Differing Sizes

- **Sticking with our example of using Multistage Cluster Sampling to draw a random sample of 500 Texans, we now know that we should randomly select 50 counties (either with SRS or Systematic) from the full set of 254 counties, and then randomly select 10 people from each selected county.**
- **Some metropolitan counties have very large populations, such as Bexar (San Antonio, 1.96 million), Dallas (2.6 million), Harris (Houston, 4.6 million), Tarrant (Fort Worth, 2.1 million), and Travis (Austin, 1.2 million), whereas some very small counties have only a few thousand people or even just a few hundred (list).**

- **If each county has an equal chance of being selected in the first stage (one card in the hat), larger counties may very well be excluded. With 50 counties being selected, each county would have a 50/254 chance (about 1-in-5) of making the cut.**



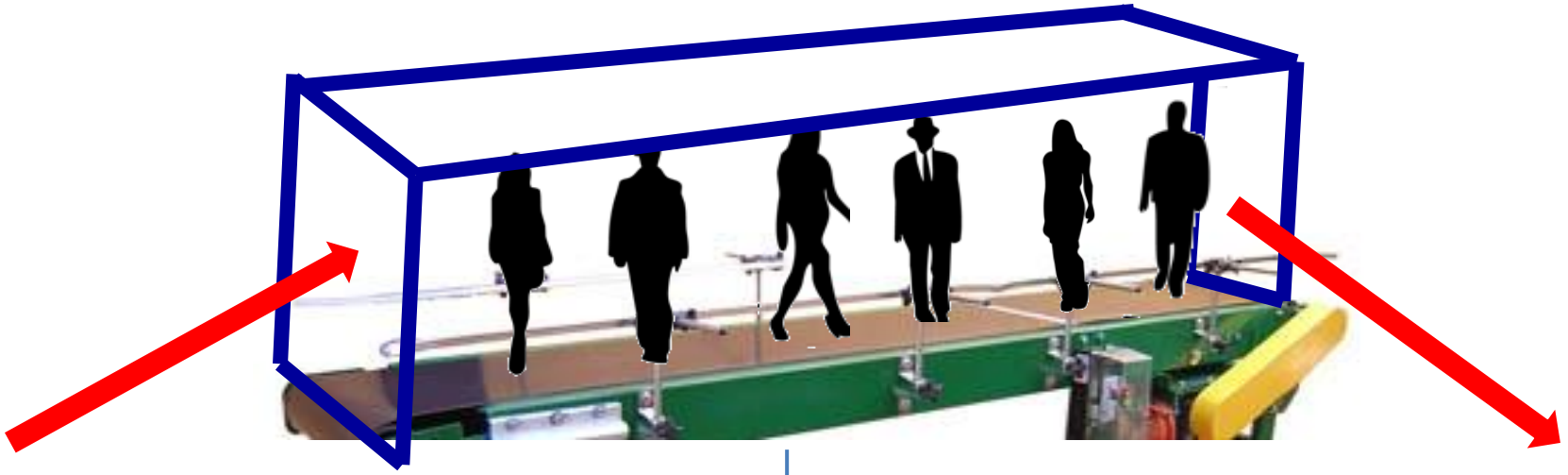
- **Because of their size and prominence, however, it could be argued that the large counties should be given a larger chance (more cards), proportionate to their size. After all, It would seem odd if a Texas sample included no one from Houston, Dallas, etc.**



Probability Proportionate to Size (Continued)

- **Giving some counties a greater number of cards in the hat than other counties appears to violate one of our established principles. Which one?**
- **After sampling counties, we would survey 10 residents per selected county. Thus, if Dallas County made it into the sample, each Dallas County resident's chance (10/2,600,000 = .000004) would be less than that of a resident of, say, Sutton County(10/4,000 = .0025), if Sutton County got in.**
- **Thus, in the long run, any individual resident of Dallas County and any individual resident of Sutton County have an equal chance of being selected. Dallas County, with its large population, has an increased chance as a county in the first stage compared to smaller counties. However, Dallas County residents, as individuals, have a lower chance of selection in the second stage than do residents of smaller counties. This balances everything out and satisfies the equal-probability principle.**

Metaphor of "Building a Sample" on an Assembly Line



FRONT END (also known as DESIGN focus)

- Things you can do from the beginning to enhance representativeness of sample.
- Sample randomly, giving all possible respondents an equal probability of selection.
- Random-digit-dial phone interviewing commonly used.
- Obtain a high response/participation rate.

IF EVERYTHING WORKED PERFECTLY AT THE FRONT END…

- … there'd be no need to repair things at the back end.
- However, response rates have dropped below 10%.
- Many pollsters have thus been shifting from traditional random, probability sampling (e.g., random-digit-dial phone surveys) to opt-in surveys (e.g., running online ads and taking anyone who wants to complete surveys).

BACK END (also known as MODEL focus)

- Online opt-in surveys unlikely to yield proportions of men and women, different ages, different racial-ethnic groups, etc., that match the U.S. Census parameters.
- Statisticians then come in after the fact to "make the nonrandom sample look like the population, [by using] weighting and modeling techniques" (link). See "synthetic sampling" in *NY Times* article.

ARTICLES ON OPT-IN/NONPROB. POLLING

- *New York Times* (2019, *REQUIRED CLASS READING*)
- *Washington Post* (2014)
- AAPOR Task Force on Non-Probability Sampling (2013)
- Pew Research Center (2016)

Sample Weighting

(a BACK-END process, AFTER collecting data, also known as POST-stratification)

- **Try to obtain a random sample on the front end, but if demographic groups are over- or underrepresented in your sample compared to the population (Census), then weight groups up or down to bring percentages into balance.**
- **Video on "Weighting and War on Error" (Elon University)**
- **Sample weighting involves adjusting the amount certain people will be counted after the data have been collected, to make the sample percentages match the true population values. Sort of like the "Honey, I Shrunk the Kids" and "Honey, I Blew Up the Kid" movies.**
- **As described in this 2014 article, "...if few people reply, the data have to be 'reweighted,' divided into clumps based on demographics like age, race, sex, and ethnicity, which are then scaled to match their shares of the population."**
- **This 2016 article gives an in-depth look at sample-weighting and the possible perils of doing so with very small subgroups: "A run of the U.S.C./[*Los Angeles Times*] poll, for instance, might have only 15 or so 18-to-21-year-old men. But for those voters to make up 3.3 percent of the weighted sample [of approximately 3,000], these 15 voters have to count as much as 86 people — an average weight of 5.7."**
- **Dr. Reifman's website on sample weighting (see numerical example at bottom).**

NOTE: Surveys can (and sometimes do) get highly accurate matches between their sample estimates and population true values (when true values are available) on the front end, *without* having to resort to sample weighting on the back end.

Practices of companies conducting or facilitating online surveys* (compiled by Dr. Reifman)	
Company	**Method**
American Life Panel (RAND Organization)	From program site: "The ALP is a nationally representative, probability-based panel of over 6000 members ages 18 and older who are regularly interviewed over the internet for research purposes. All data are available for free to researchers. The ALP is also a service for researchers to field their own questionnaires and experiments. We will work with you to program, field, and monitor your survey. Combine your data with over 400 previously collected surveys... "
AmeriSpeak (National Opinion Research Center/ University of Chicago)	"[The] initial offering will be a general population adult panel of 10,000 households across the country, with plans to expand in later years... AmeriSpeak households are selected initially from NORC's National Sample Frame... [which] is designed to provide over 99 percent sample coverage by supplementing the U.S. Postal Service Delivery Sequence File... NORC contacts sampled households by a variety of means to make the household aware of their eligibility to join AmeriSpeak. The contact methods are email (when available), U.S. mail, and telephone (when available)... AmeriSpeak panel members typically participate in AmeriSpeak web-based or phone-based studies two to three times a month." (methodology white paper)
Gallup Panel	Recruited through random digit dial or address-based sampling, to be representative (~60,000 members).
GfK (formerly Knowledge Networks)	Assembles large (~50,000) online panel through representative sampling (e.g., postal addresses, phone numbers), then randomly selects small subsamples for study purposes; occasionally uses opt-in component.
Google Surveys (formerly Google Consumer Surveys) OPT-IN	Visitors to websites (most commonly online news sites) encounter one or more survey items before they can read an article: "Surveys partially and temporarily block the content on each publisher's site. " The following quotes describe how respondents' demographic characteristics are accounted for: • "The mobile app is the simpler of the two panels: We ask users to self-report their age, gender, and zipcode when they sign up to use the app. We periodically ask these questions again to refresh the panelists' demographics in case they have changed." • "The publisher network uses inferred demographics, which means that we don't explicitly ask the panelists for their demographics. We do this to minimize the number of questions…, which offers a better respondent experience and encourages higher response rates. Any researcher who doesn't want to rely on inferred demographics can explicitly add demographic questions to their surveys... " • "Like many ads on the web, Google Surveys infers the age and gender of anonymous respondents based on browsing behavior and geography based on IP addresses" (from Google Surveys "White Paper," linked in left-hand column)
Mechanical Turk (Amazon) OPT-IN	"…relies on a large online panel of volunteers who complete short tasks, including surveys and experiments, for micropayments that Amazon administers."

Qualtrics Panels OPT-IN	Works through other companies such as SSI
Survey Monkey OPT-IN	Maintains a panel of respondents (known as the Audience): "We recruit from the diverse population of 30+ million people who complete SurveyMonkey surveys every month." Researchers pay to use the panel (includes free option of asking 10 questions and getting 100 responses).
Survey Sampling International (SSI) OPT-IN	"…participants are invited via banners, invitations and messaging of all types, but then go through rigorous quality controls before being included in any sample. SSI can potentially access anyone online via a network of relationships with websites, panels, communities and social media groups" (source)
Understanding America Study (U. Southern California)	From program site:"new panel... of approximately 2,000 households representing the entire United States. The study is an 'Internet Panel,' which means that respondents answer our surveys on a computer, tablet, or smart phone, wherever they are and whenever they wish to participate. The majority of the panel members have their own Internet access. The remaining panel members have been provided Internet access by USC... It is not possible to volunteer for joining the UAS. We randomly select people around the country using postal codes. "
YouGov OPT-IN	"…offers what it calls a 'matched sample,' using census data to guide the selection of participants from a large opt-in panel, and constructing weights that the company claims can be used to approximate a representative sample (Rivers, & Bailey, 2009)."

*Quotes are from this Ohio State document, unless noted otherwise; other sources are linked in the table.

Non-Probability Sampling

- Unlike sampling from a frame, where you know, for example, that each person has a 1-in-500,000 probability of selection, with non-probability sampling you don't know what anyone's probability of selection was
- Inexpensive, good for early-stage research, but limits generalizability and statistical inference from sample
- Types of non-probability sampling
  - Convenience
  - Quota; similar to opt-In (discussed above) with post-stratification weighting to meet demographic criteria
  - Judgmental/Purposive
  - Snowball

  (All of these methods discussed in this required class reading, except opt-in and snowball)

Quota Sampling (Hypothetical Example)

	Known proportions from "population" of TTU freshmen (total = 4180)		Suppose you wanted to conduct a survey of 500 TTU freshmen	
	Male	Female	Male	Female
Am Ind	14 (0.3%)	10 (0.2%)		
Asian	83 (2.0%)	58 (1.4%)		
Black	113 (2.7%)	99 (2.4%)		
Hispanic	335 (8.0%)	292 (7.0%)		
White	1677 (40.1%)	1499 (35.9%)		

Preserve same percentages in your 500-person sample, e.g., recruit 40 Hispanic males (8% of 500), and 180 white females (35.9% of 500)

Texas Tech U., Fall 2008 Entering Freshmen http://www.irim.ttu.edu/

Rare use of quota sampling in current polling (2014; scroll down to "Suffolk Poll Methodology Part I: Calling Quotas " when page opens)

Additional Websites Relevant to Sampling

General resources

- History of sampling and political polling (Mark Blumenthal, 2016)
- Random number generator
- American Association for Public Opinion Research
- Gallup Poll
- Polling Report
- Survey Practice (journal)
- Pew Research Center page on Sampling
- Inside the interviewing process in public-opinion polls
- "Four Innovative Methods to Recruit Community Samples of Young Adults" (mostly non-prob.)

Implications for sampling/polling of shift from landline to cell/mobile phones and other technologies

- Estimate of Wireless-Only U.S. Households (2018)
- Pros and Cons of Different Sampling/Polling Approaches (August 2016)
- Argument becoming stronger for calling only cellphones and no landlines plus a leading compiler of polling results will no longer recognize landline-only polls (August 2016)
- "2016 Could Mark Telephone Poll's Last Stand"
- "Why Polls Don't Work" (February 2016)
- Counterpoint: Polls' Demise "Greatly Exaggerated" (February 2016)
- Guidelines for "Dual Frame"(Landline and Cellphone) Telephone Samples (2014; see Figure 6 and beyond)
- Issue of "telephone number portability," how someone who has moved to a new city still has cell-phone area code from city where phone was purchased. "...no surveyor conducting a small area survey would attempt a New York City, Atlanta, or Birmingham area code and exchange in hopes of reaching a respondent in their target market of Mississippi." (June 2014)
- "Sorry, Wrong Number: Surveying Americans by phone is increasingly problematic. As a result, pollsters are trying to incorporate new – but imperfect – alternatives" (August 2012)

Evaluating the pollsters' accuracy

- Performance of polls in forecasting 2020 presidential vote (here and here)
- Performance of polls in forecasting 2016 presidential vote (brief overview of national polls; in-depth study of state poll errors)
- Performance of polls in forecasting 2012 presidential vote
- Performance of polls in forecasting 2008 presidential vote
- Performance of polls in forecasting 2004 presidential vote
- National Council on Public Polls: Performance of different polls in forecasting *true* vote in 1936-2000 elections

Census-related

- U.S. Census Bureau
- Article on the planning of the 2020 Census
- "Census Scope" website from U. of Michigan, depicting Census Data in colorful charts and graphs