

Public Confidentiality Expectations Regarding Data Linkage

Jennifer Hunter Childs, Casey Eggleston and
Aleia Clark Fobia

Center for Behavioral Science Research

U.S. Census Bureau

Disclaimers

The data discussed in this presentation were produced for strategic and tactical decision-making and exploratory research and not for official estimates.

Any views expressed are those of the author and not necessarily those of the U.S. Census Bureau.

Objectives

- Examine views towards administrative record use by:
 - Demographic characteristics
 - Attitudes towards the U.S. Federal Statistical System (decentralized)
- Model likelihood to support administrative record use as an alternative or supplement to survey data collection
- Hypothesis: Favorability towards use of administrative records for statistical purposes will be predicted by trust in the U.S. Federal Statistical System.

Overview

- Methods
- Big Picture
- Study 1: Source and type of information, question frames
- Study 2: Information for whom and privacy views
- Study 3: Attitudes towards combining data
- Conclusion



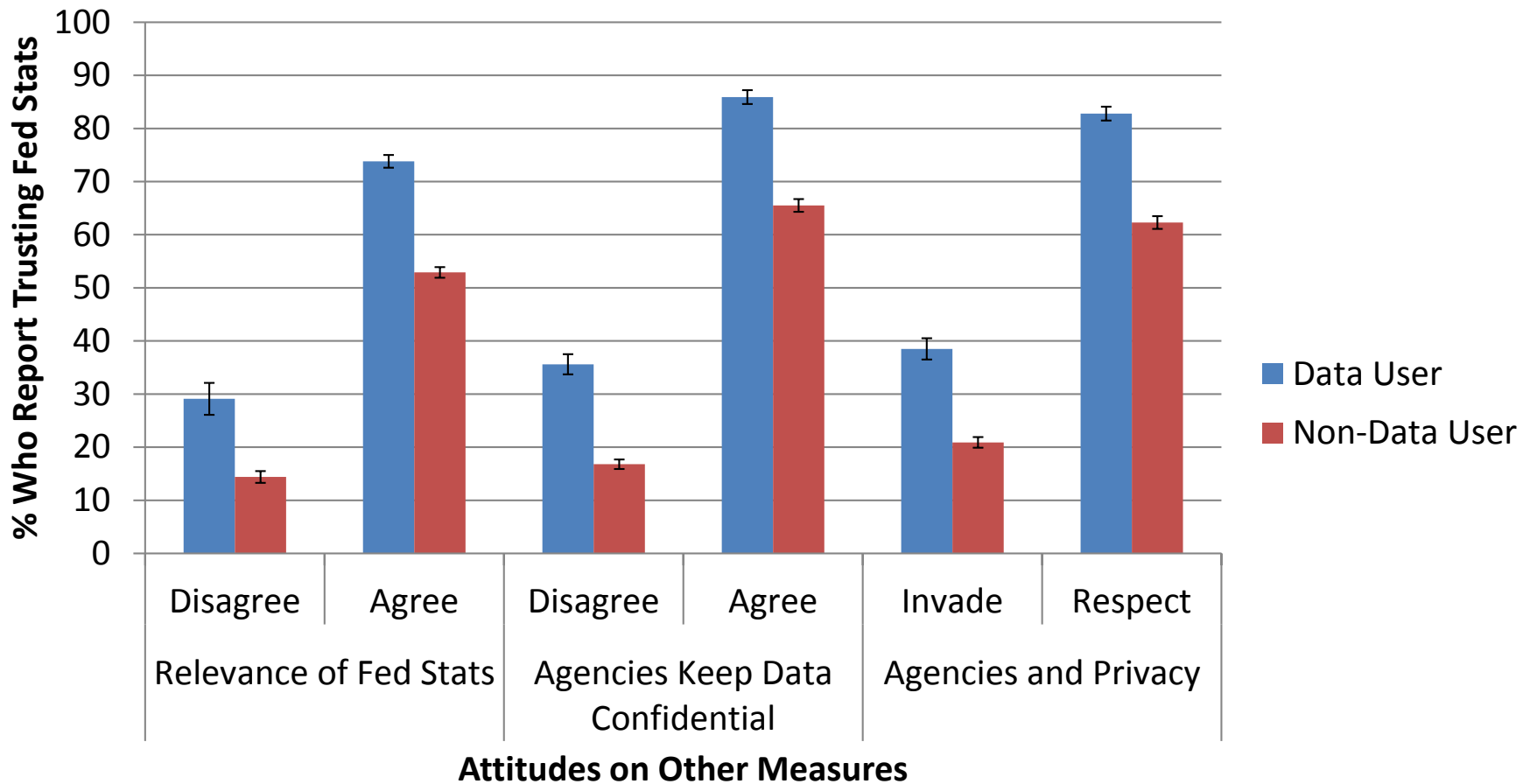
Public Opinion Poll

- Questions part of nationally representative Gallup Daily Tracking Survey of 100-200 people nightly
- Both landline and cell phone numbers
- Began data collection in February 2012 through July 2018 (with several brief breaks in data collection)
- Core and rotating questions
- AAPOR RR3: about 8-11% for each wave

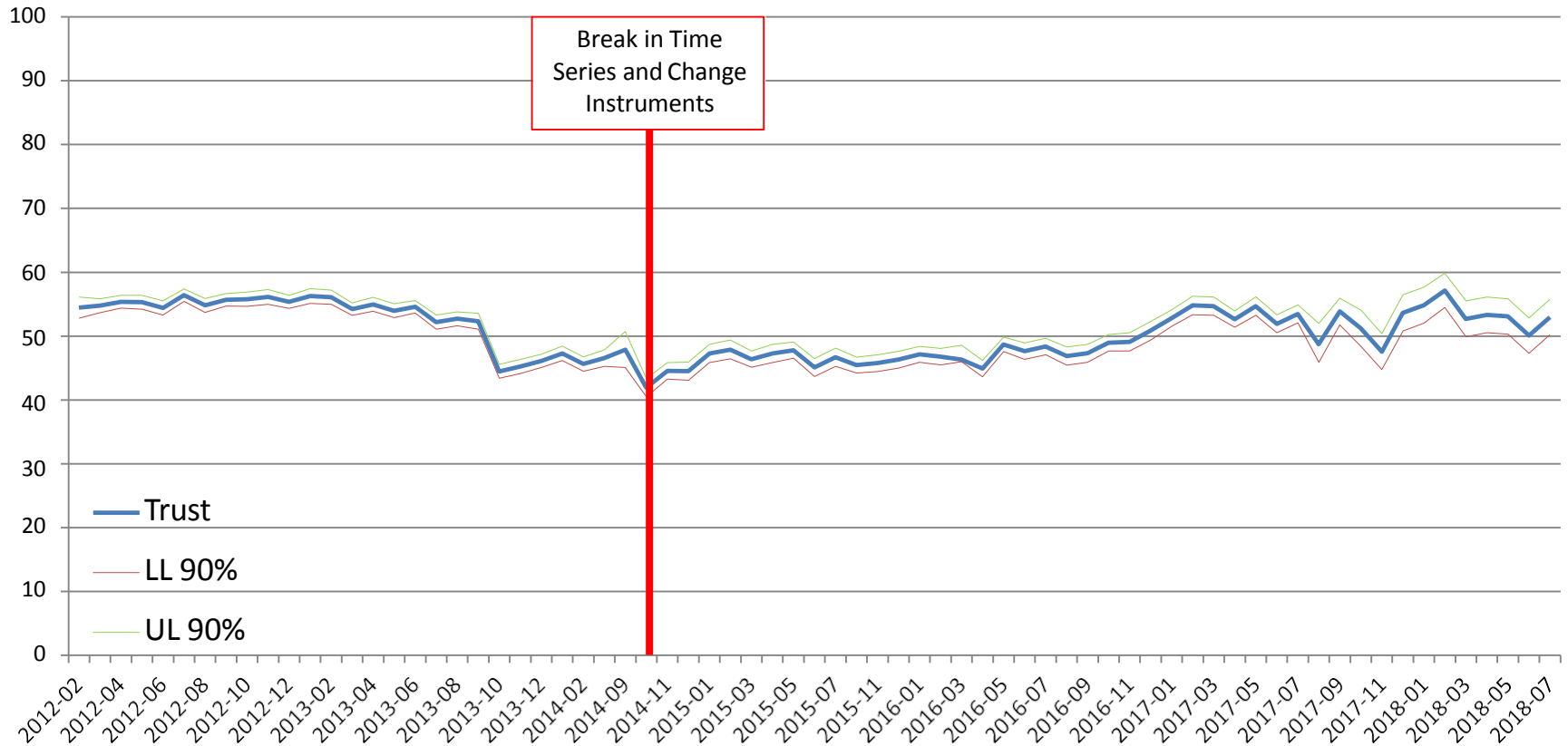
Public Opinion Poll Core Questions

- Awareness
 - Orienting respondents: Introduce federal statistics
 - Data user? (key knowledge variable)
- Attitudes towards federal statistics
 - Trust
 - Confidentiality
 - Relevance/Importance
 - Privacy

Effects on Trust by Data User

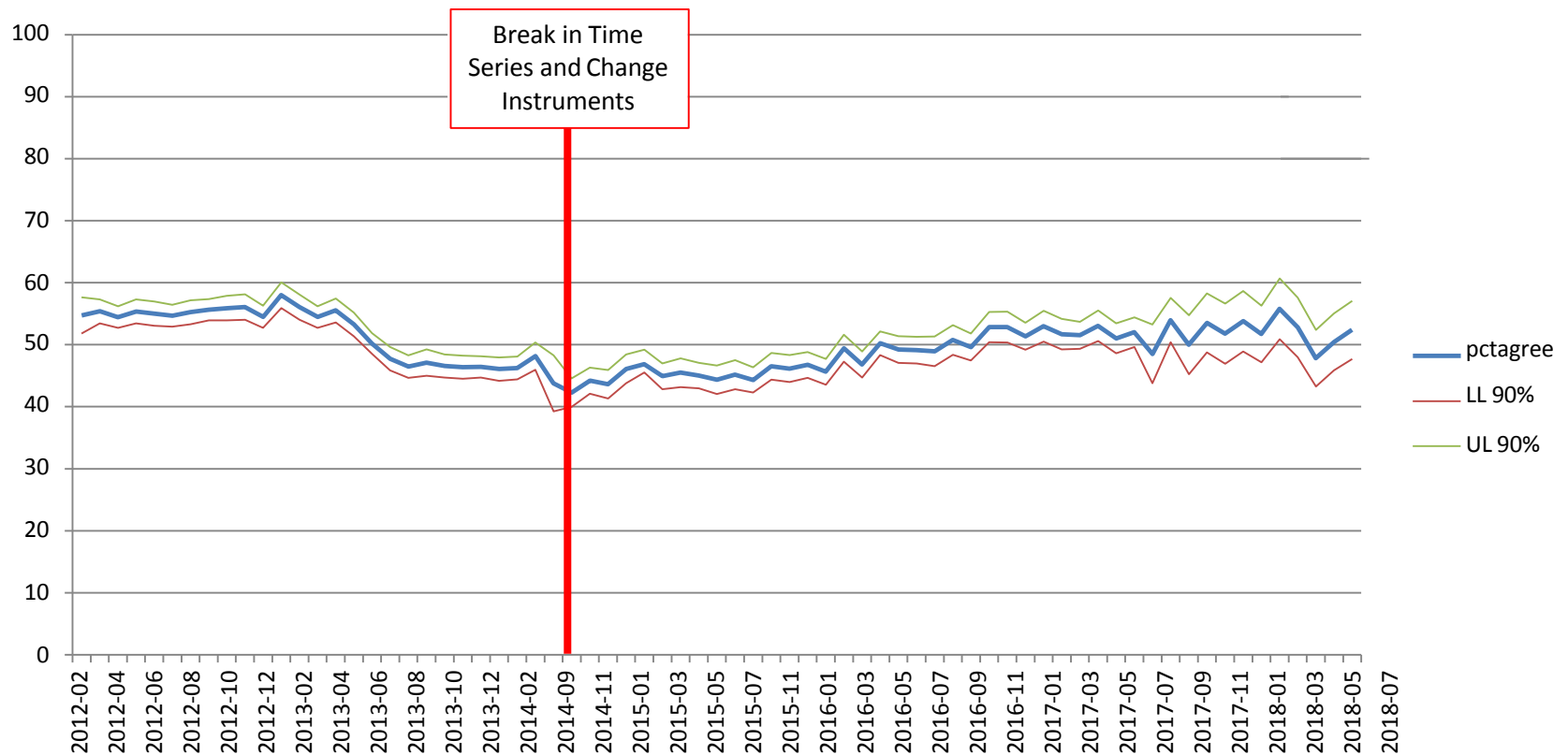


Reported Trust in Federal Statistical System over Time



* Change in instruments coincided with a 3.2% decrease in reported trust.

Reported Belief that Data are Kept Confidential in the Federal Statistical System



* Change in instruments coincided with a 4.8% decrease in reported belief.

Summary of Background

- Overall, knowledge of the statistical system, as well as using data, believing the data are relevant, kept confidential and that agencies respect privacy all correspond with increased trust in the statistical products.
- Trust in Federal Statistics and Belief in Confidentiality of Data declined for a time, but has returned to the previous level.

Hypothesis

- Favorability towards use of administrative records for statistical purposes will be predicted by trust in the Federal Statistical System.

Study 1

- Source of administrative records
- Type of administrative records
- Purpose (frame):
 - Cold ask
 - Save government money
 - Social good

Source/Type of Admin Rec

Source	Type			
	Generic Information	Name and Age	Income	Contact Information
Social Security Administration (SSA)	X	X	X	
Internal Revenue Service (IRS)		X	X	
Dept. of Motor Vehicles (DMV)				X
Company that provides publicly accessible info (public records)		X		X

Frame: “Cold Ask”

For the next census in 2020, if the Census Bureau got information that you already provided to the Social Security Administration, they may not need to ask you for this information on a questionnaire.

If you knew that this information was being obtained from the Social Security Administration only to produce statistics, and that your personal information would remain unavailable to the public, would you be strongly in favor of it, somewhat in favor of it, neither in favor nor against it, somewhat against it, or strongly against it?



Frames



Save Government Money

The 2010 Census cost over **[\$10 billion dollars / about \$100 dollars per household.]** Getting your information directly from the Social Security Administration could reduce the cost for the 2020 Census and save government money.

Social Good

Because some people choose not to respond to the Census, getting information from the Social Security Administration could help the Census Bureau get a better idea of population sizes informing **[where the community should build new schools, hospitals, roads, and firehouses./ how to distribute the seats in the U.S. House of Representatives.]**

Overall Effects

Compared to “Info” from Social Security Administration...

- Less desirable:
 - Any data from the Internal Revenue Service
 - Income data from Social Security Administration
- No difference for:
 - Getting name and age from Social Security Administration
 - Contact info from the Dept. of Motor Vehicles
 - Data from public records
- Both frames improved favorability by an average of about 15%

Predictors

- Attitudes much stronger predictors than demographics
- **Favoring** use of administrative records predicted by:
 - Knowledge of federal statistics
 - Trust in federal statistics
 - Belief that policy makers need good statistics to make decisions
 - Belief that federal statistical agencies keep information confidential

Predictors (cont.)

- Being **against** use of administrative records predicted by:
 - Less confidence in institutions (newspapers/universities)
 - Don't know/Refuse to report income
- Not significant: (Race, Sex, Income, Job status)

Why Are Respondents Against the Use Of Records?

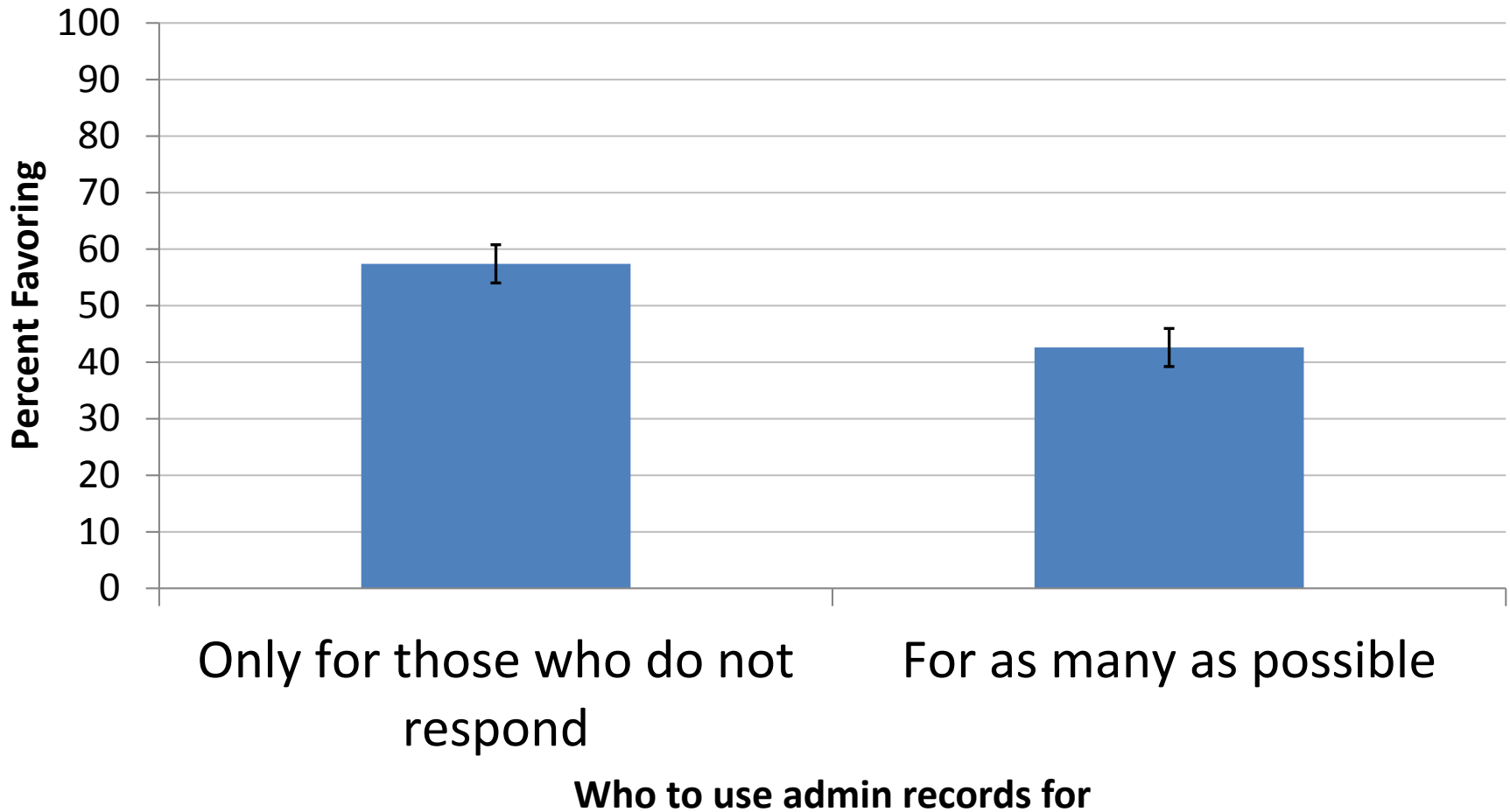
- Privacy/Confidentiality
- Don't trust the government/Government is too big
- Government has too much info about them
- Concern about an unauthorized person gaining access to their information (e.g., government's ability to protect data)



Study 2

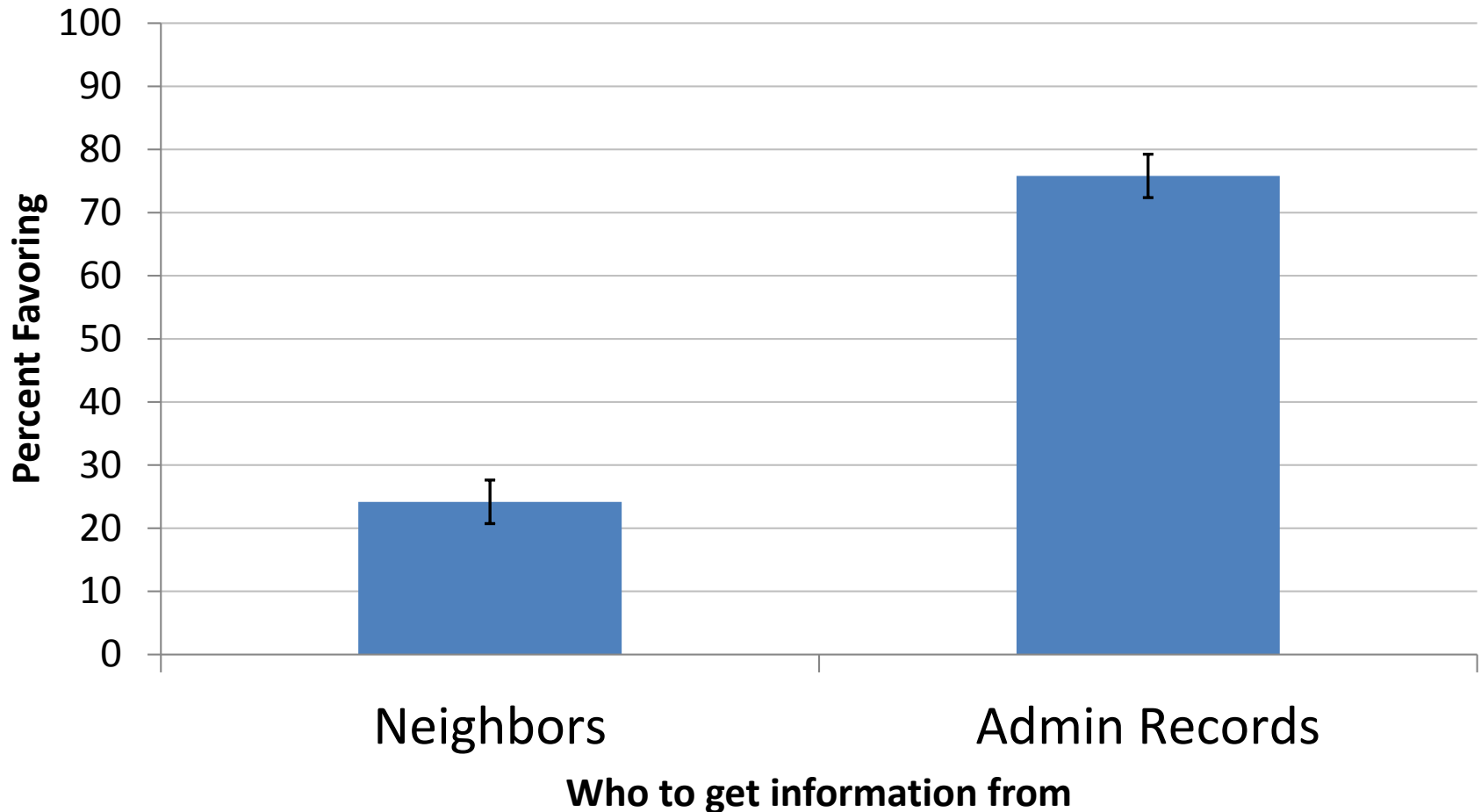
- For whom should we use administrative records in replacement of surveys?
 - For nonrespondents only, or everyone?
 - Proxies/Imputation: Neighbors or records?

To Save Money, Who Should Census Use Administrative Records for?



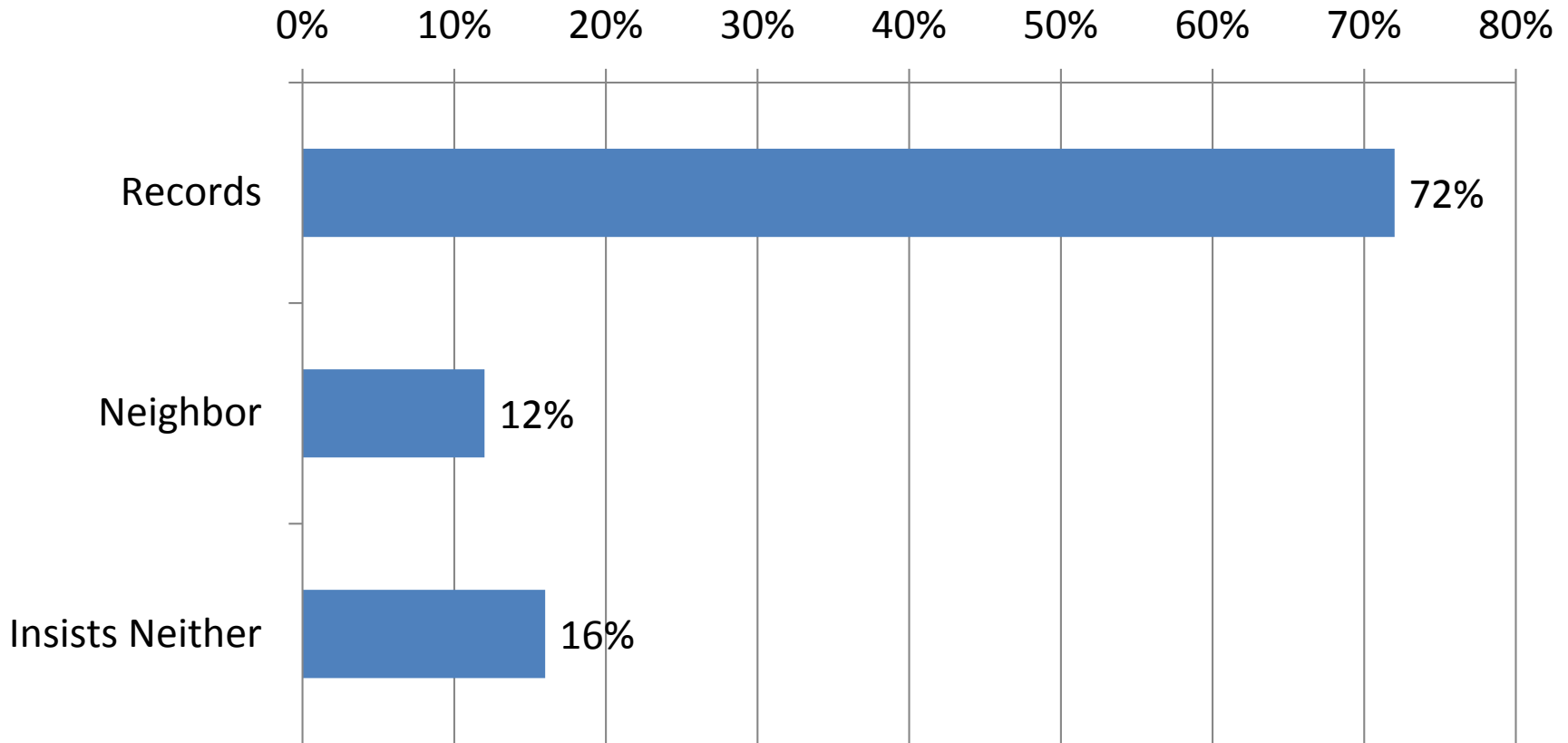
Source: Gallup Daily Tracking Survey, Sept. 2013-Oct. 2013. Not intended for point estimates

For Non-Responders, Should Census Go to Neighbors or Records?



Source: Gallup Daily Tracking Survey, Sept. 2013-Oct. 2013. Not intended for point estimates

Which is more respectful of your privacy? Neighbor vs. Admin Record



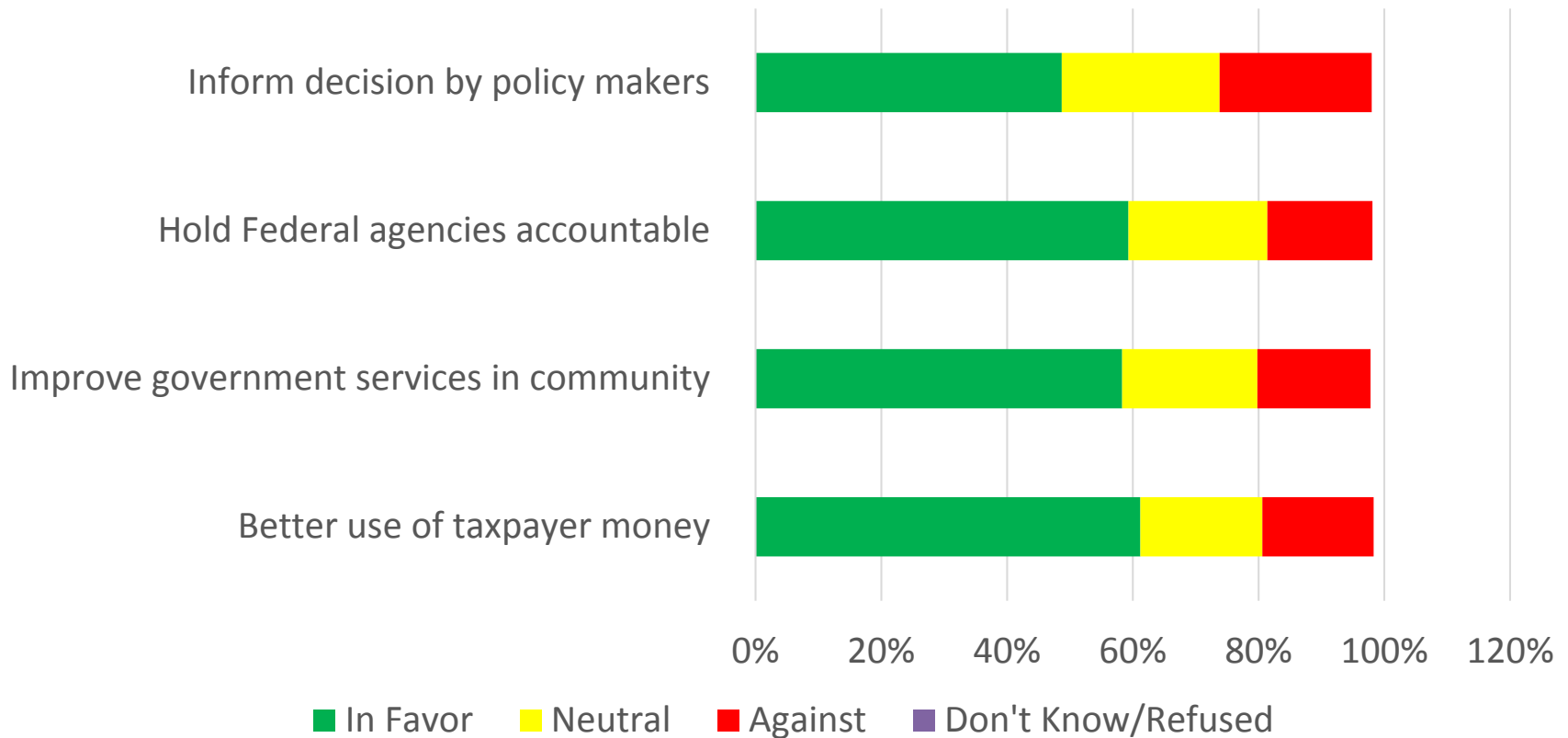
Predictors of Favoring Administrative Record Use

- Greater trust in the statistical system
- More educated respondents were more likely to prefer records over neighbors
- Presence of a child in the home predicts preferring more extensive administrative record use

Study 3

- Predicting public acceptance of combining data

Public's Favorability Towards Combining Data by Question Frame



$X^2 (9, N = 11572) = 94.0, p < .0001$

Odds Ratios from Logistic Regression Predicting Favorability Towards Combining Data

Predictor	Odds ratio	95% CI	
		Lower bound	Upper bound
Attitudes towards FSS			
Trust in statistical products (v. Not trust)	1.82 **	1.56	2.11
Agree/Neutral that federal statistics are relevant (v. Disagree)	1.81 **	1.56	2.1
Agree/Neutral trust in confidentiality (v. Disagree)	1.53 **	1.32	1.78
Beliefs about Privacy			
Responded to income question (v. Did not respond)	1.61 **	1.37	1.88
FSS respects privacy (v. Invades privacy)	1.44 **	1.25	1.66
Knowledge of Federal Statistics			
Data user (vs. Not a data user)	0.92	0.78	1.08

¹¹ We also ran the regression with “Neither Agree or Disagree” responses collapsed with the “Disagree” category. It did not alter the pattern of results for attitudes, beliefs, and knowledge, but the significance of some demographic predictors differed slightly.

Predictor	Odds ratio		95% CI	
			Lower bound	Upper bound
Demographics				
Black, non-Hispanic (v. White non-Hispanic)	0.67	**	0.55	0.83
Hispanic (v. White non-Hispanic)	0.65	**	0.52	0.84
Other non-Hispanic or 2+ races (v. White non-Hispanic)*	0.75	*	0.59	0.95
25-44 (v. 18-24)	0.69	*	0.51	0.93
45-64 (v. 18-24)	0.58	**	0.43	0.78
65+ (v. 18-24)	0.38	**	0.28	0.53
Some college (v. high school education or less)	1.39	**	1.19	1.64
Bachelor's degree (v. high school education or less)	1.37	**	1.14	1.64
Post-graduate education (v. high school education or less)	1.76	**	1.43	2.17
Female (v. Male)**	1.19	**	1.04	1.36
Married or in a domestic partnership (v. Single)	1.11		0.95	1.32
Separated, divorced, or widowed (v. Single)	1.05		0.85	1.3
Full-time (v. Not in Labor Force)	1.09		0.92	1.29
Part-time (v. Not in Labor Force)**	1.39	**	1.09	1.77
Unemployed (v. Not in Labor Force)	1.16		0.77	1.77
* p < .05, **p < .01, N=9999				

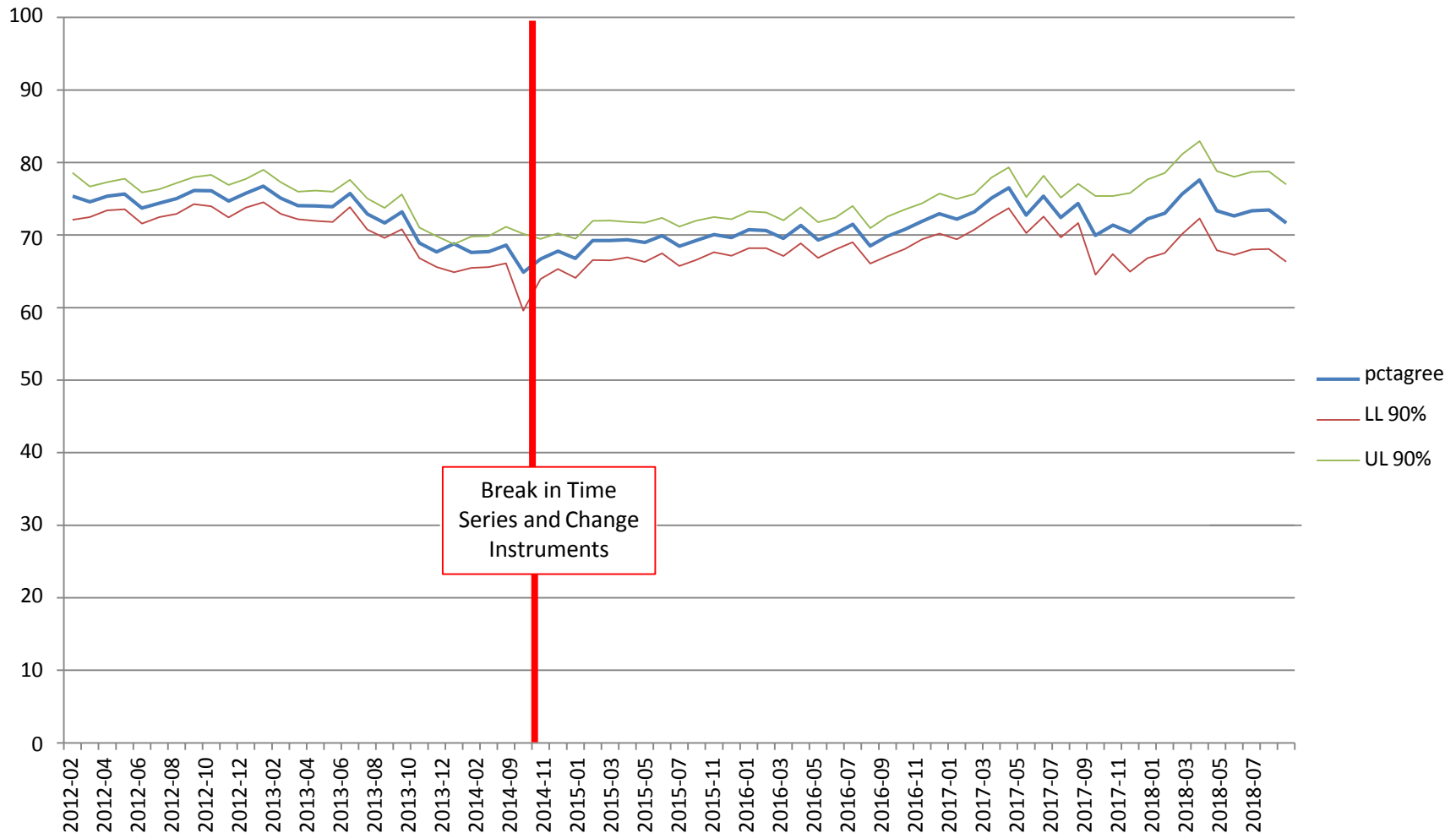
Conclusions

- Transparency is important.
- Need for a carefully planned communications strategy.
- Public wants to know how their info is protected, and in what circumstances administrative records would be used.

Jennifer Hunter Childs
jennifer.hunter.childs@census.gov

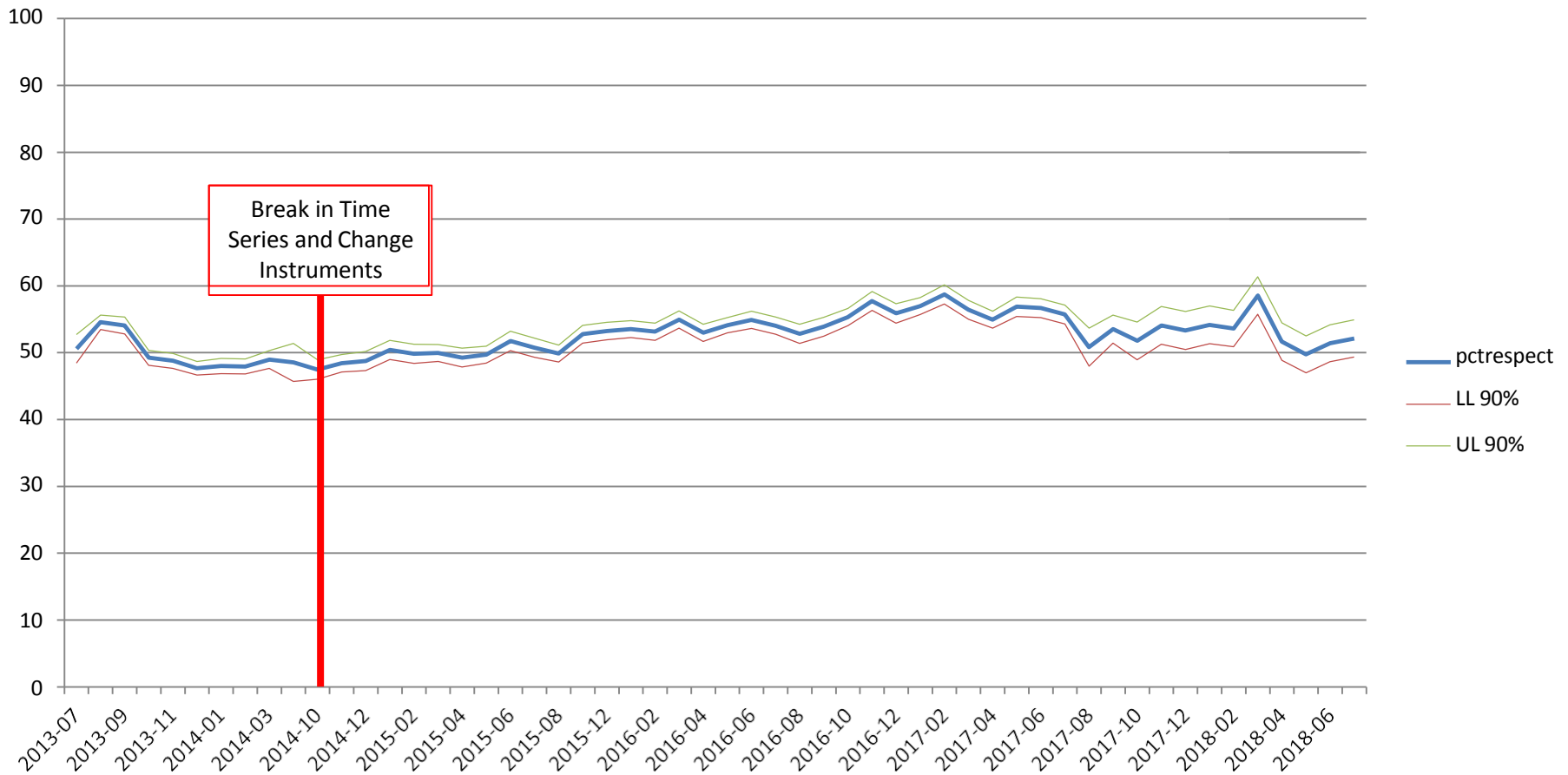
Additional Resources

Reported Belief that Policy Makers Need Statistics for Decision-Making



Break in Time Series and Change Instruments

Reported Belief that the Federal Statistical System Respects (Rather than Invades) Privacy



*No statistically significant difference after change in instrument.