

The Decennial Census and the ACS: Looking Back and Looking Ahead

**Robert Groves and
Frank Vitrano**

U.S. Census Bureau

Presentation to:

Population Reference Bureau

March 16, 2011

Agenda

- Hints about the Quality of the 2010 Census
- 2020 Planning – Organization
- American Community Survey (History & Future Review)
- 2020 Census – Early Research
- Operational Design Alternatives
 - Establishing Where to Count
 - Enumeration
 - IT and Operational Infrastructure
- Policy and Communication Challenges
- Early Planning Highlights

Concepts of Quality in a Census

1. Distance from an Ideal Outcome
2. Similarity to an Alternative
3. Soundness of Process of Data Collection

No Two Uses of Census Counts have the Same Quality

- Total population count of the US
- Total population count of each of the 50 states
- Total population count of each of the 8 million census blocks
- Total African-American, Hispanic, Asian population count of each of the 8 million blocks

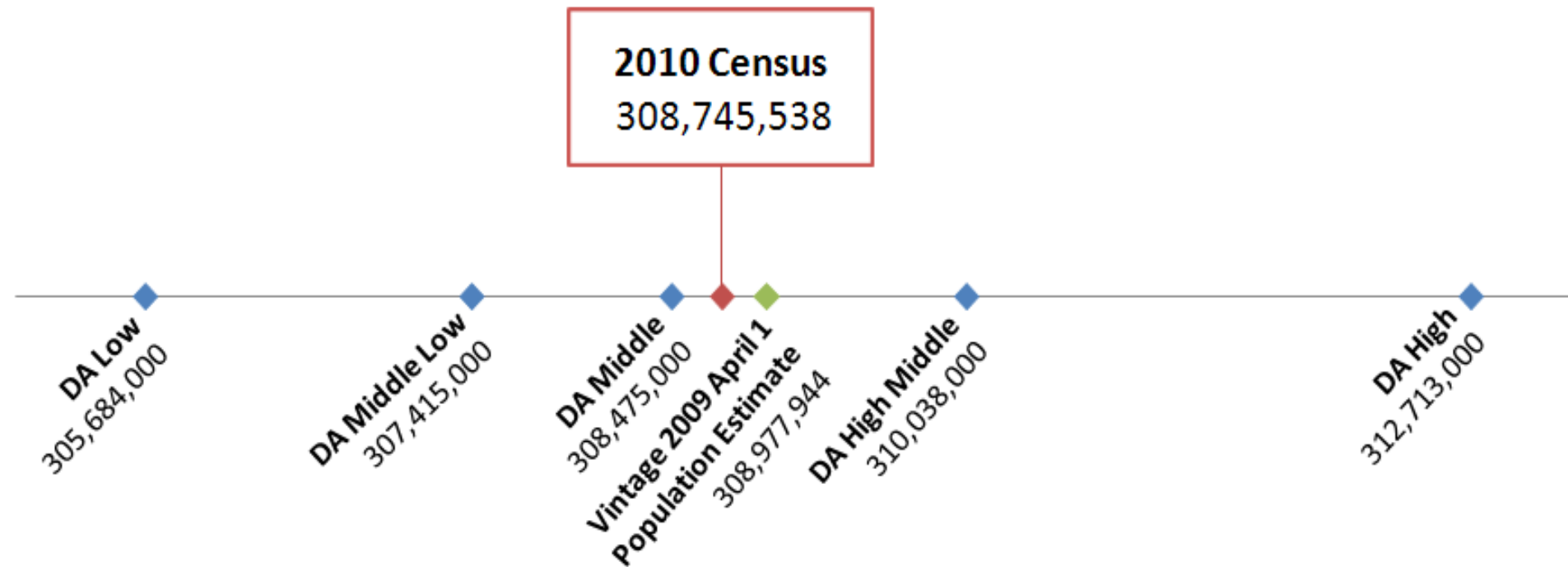
1. Distance from an Ideal Outcome

- Common evaluation with known truth
 - Percentage correct on examinations
 - Percentage unscheduled downtime for computer systems
- Ideal for all population enumeration use:
 - “Count every resident once and only once and in the right place”
- For a census we have no measured ideal

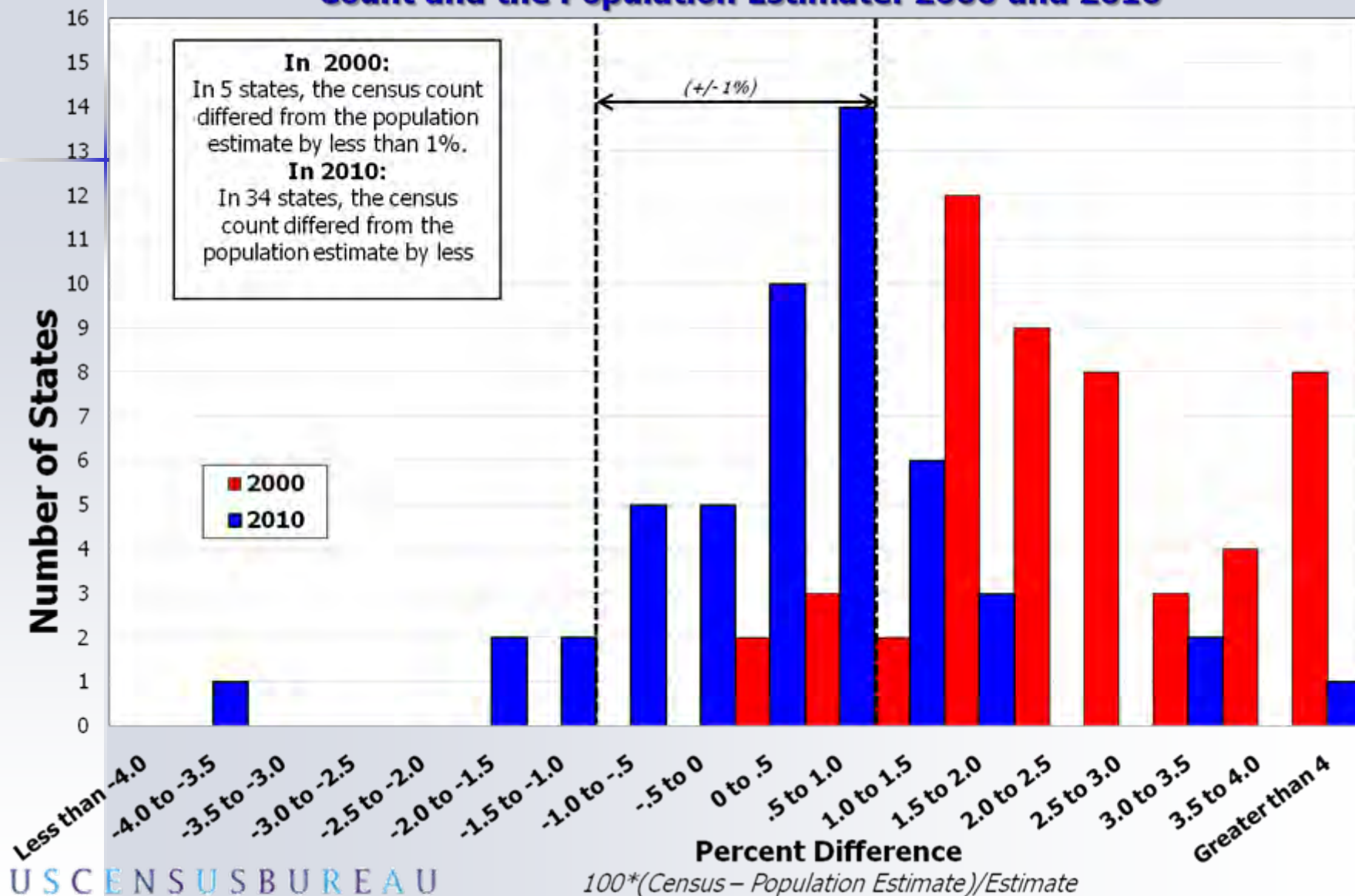
2. Similarity to an Alternative

- Demographic Analysis (Fall, 2010)
 - Vital registration system supplies births, deaths
 - Emigration and immigration from records and diverse other sources
- National totals for age by gender by race/ethnicity
- Weaknesses
 - Undocumented immigration
 - Ethnicity measurement on records

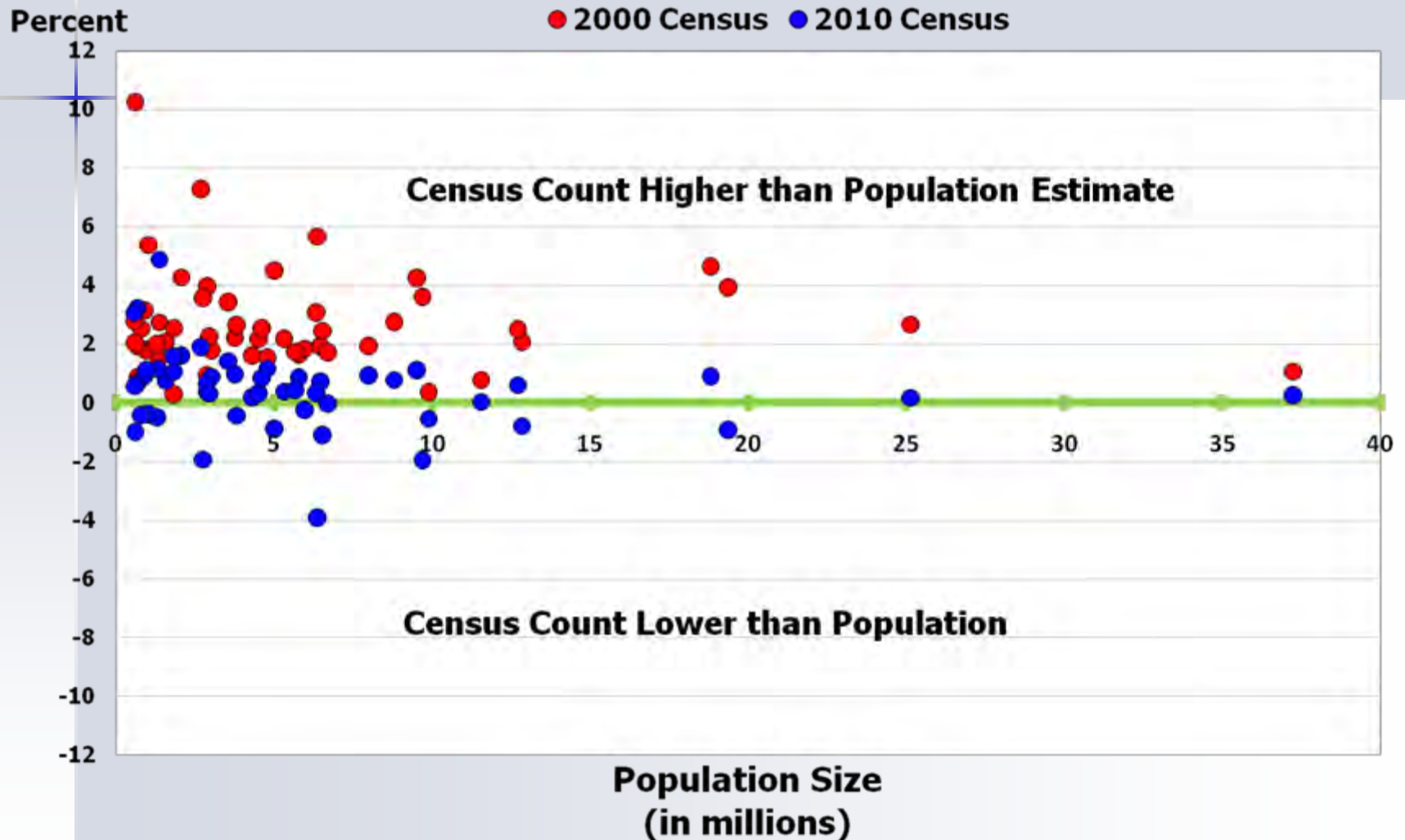
2010 Census, Demographic Analysis and Population Estimates



Distribution of States by Percent Difference Between the Census Count and the Population Estimate: 2000 and 2010



Percent Difference Between the Census Count and the Population Estimate by Population Size of State: 2000 and 2010



2. Similarity to an Alternative (contd.)

- Post-enumeration sample survey (2012)
 - Matches and nonmatches to census
- Estimates of matches and nonmatches for states and larger areas
- Weaknesses
 - Subject to sampling variance
 - “correlation bias”
 - Recall error re April 1 status

Preliminary Results from the Post-Enumeration Survey

- Unweighted housing unit match of sample to master address file (2010 – 96.5%; 2000 – 91.4%)
- Percent of sample cases correctly enumerated (2010 – 96.3%; 2000 – 89.9%)
- Percent of sample cases found duplicated in census (2010 – 0.79%; 2000 – 1.05%)

3. Soundness of Process of Data Collection

Feature	Evaluative Indicator Relative to 2000
Local government participation in Update of Census Addresses	Percentage of governments is lower; percentage of population represented is similar
Continuous update of Address List	Total housing units closer to independent estimate; fewer deleted listings (4 million vs 6 million in 2000); fewer additions in followup phase (635K vs 689K)
Communications campaign	Similar awareness, intention to participate



Real-time Evaluation – Before Mailout/Mailback

- Daily RDD survey asking about 2010 Census
- Analysts monitoring data
- Asking questions about whether key predictors of form completion are progressing positively across all groups

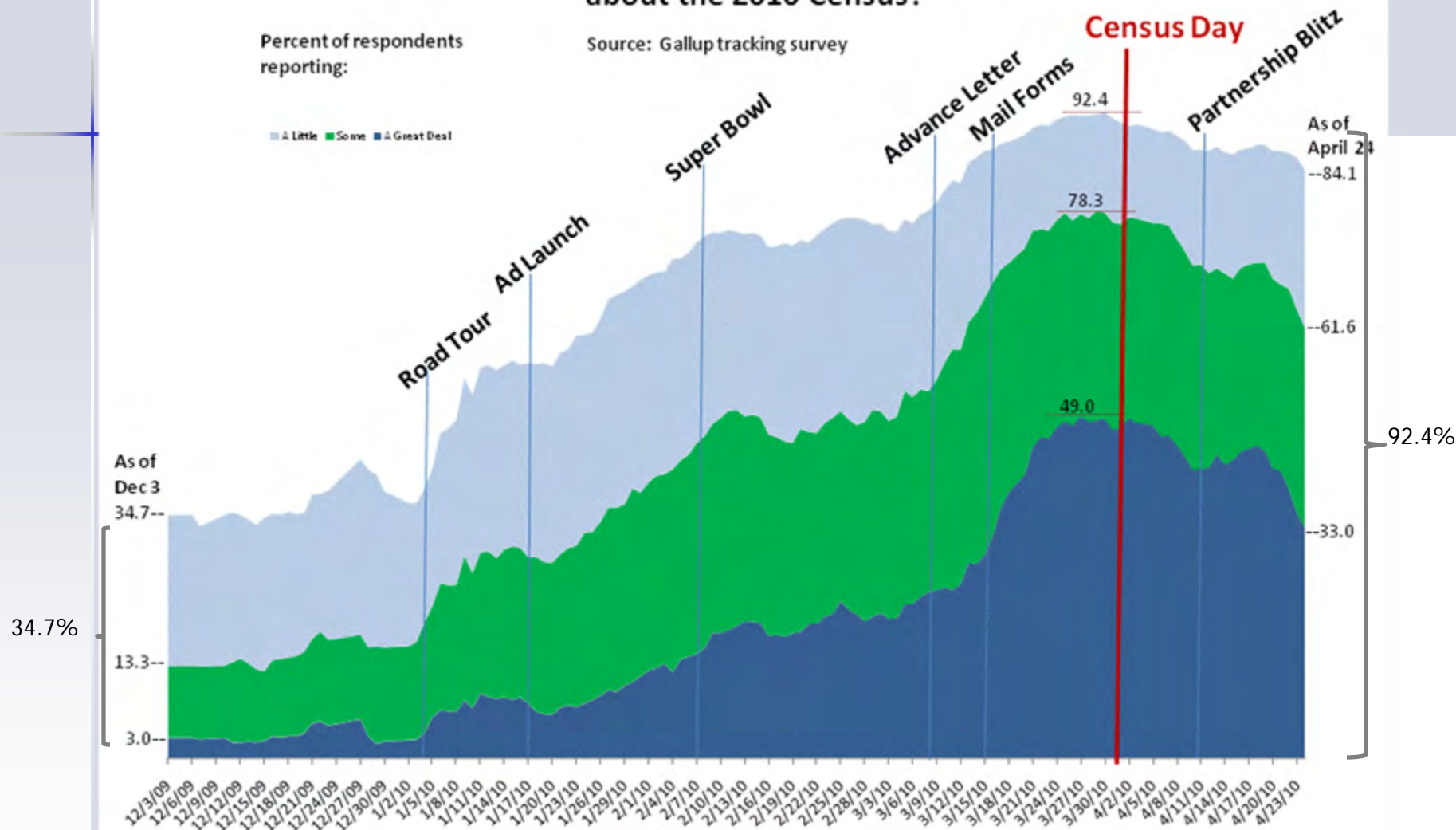
The Census is Coming

How much have you seen or heard recently – within the last week or so
– about the 2010 Census?

Percent of respondents reporting:

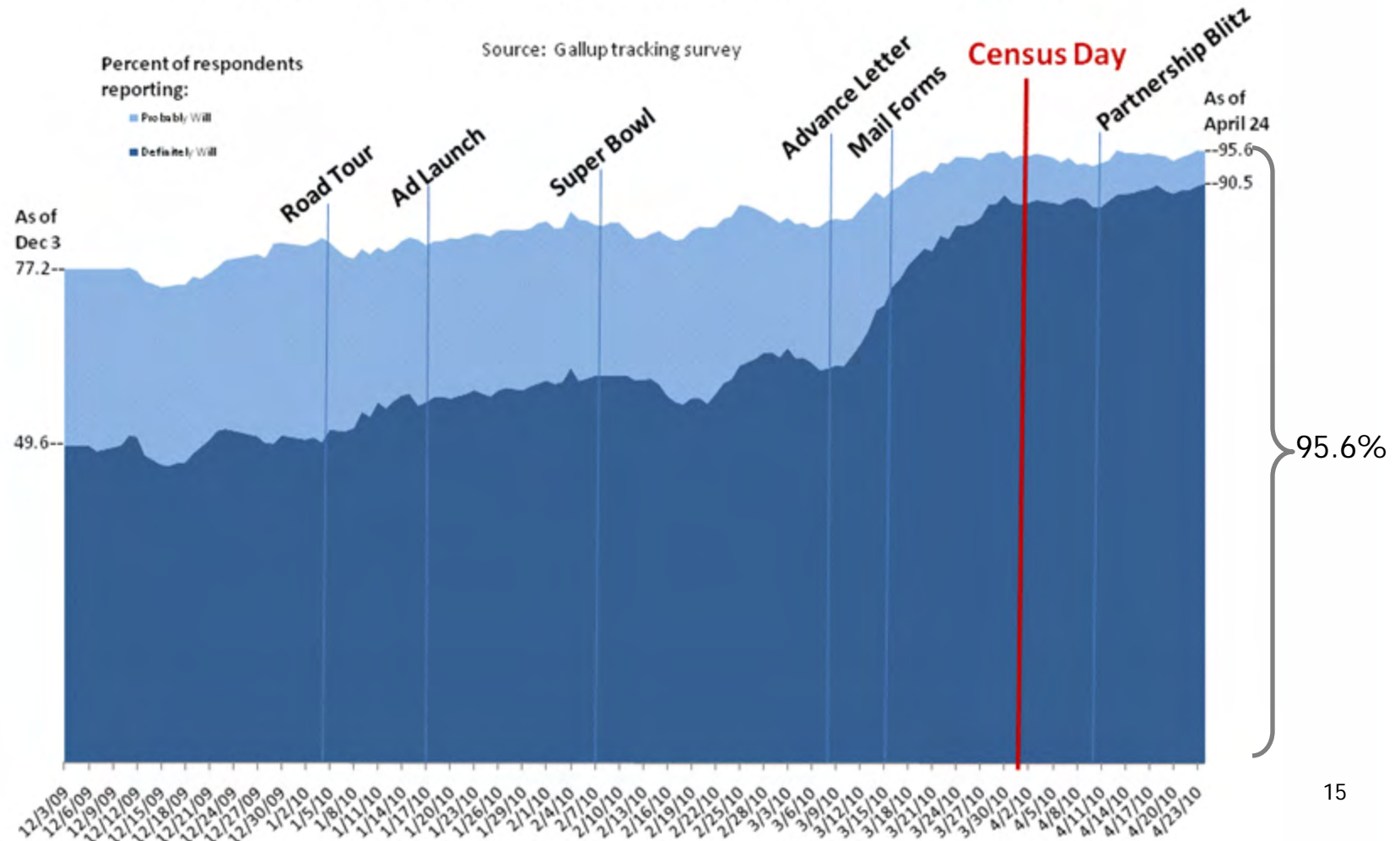
Source: Gallup tracking survey

■ A Little ■ Some ■ A Great Deal



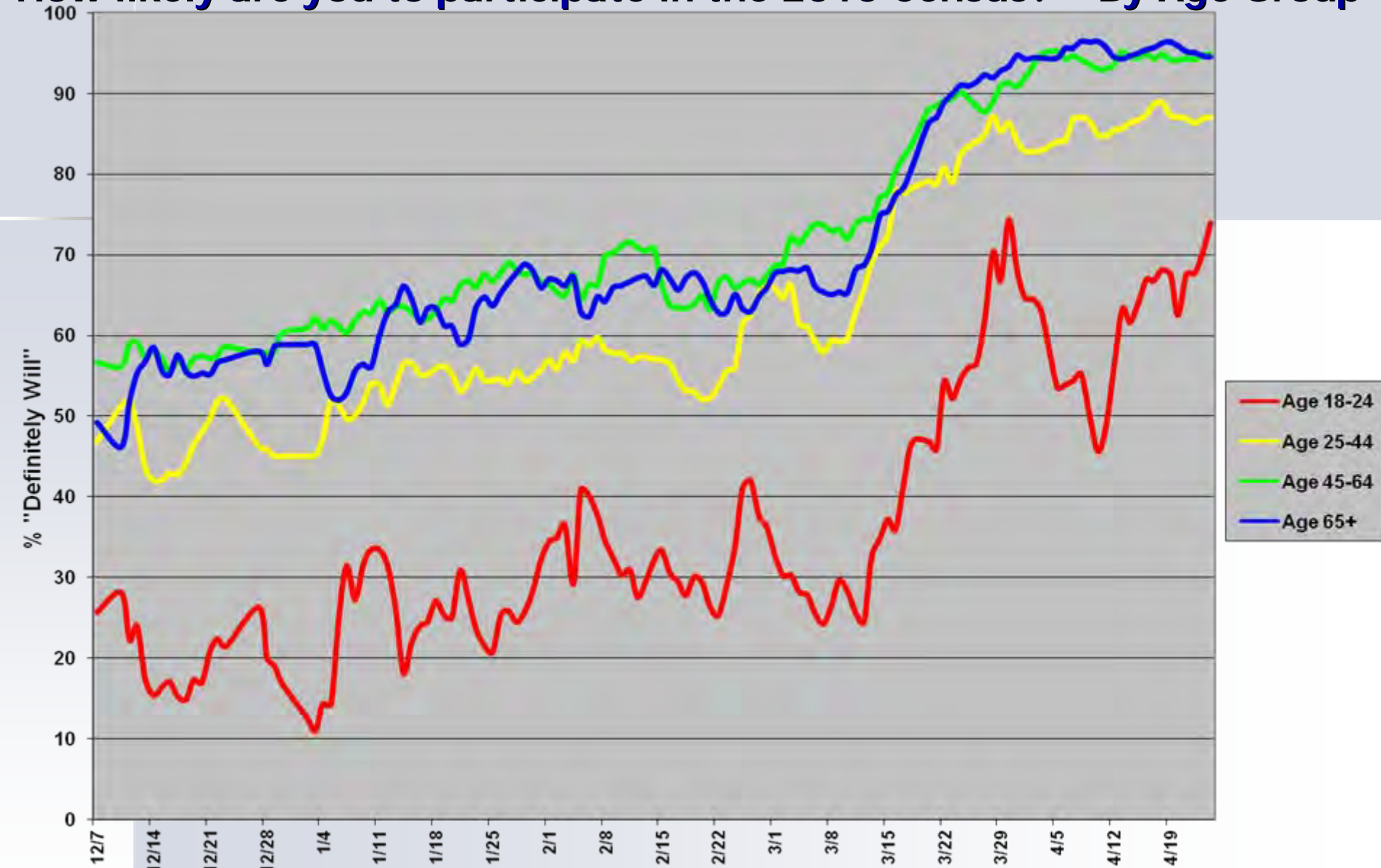
Intent to Participate

How likely are you to participate in the 2010 Census?



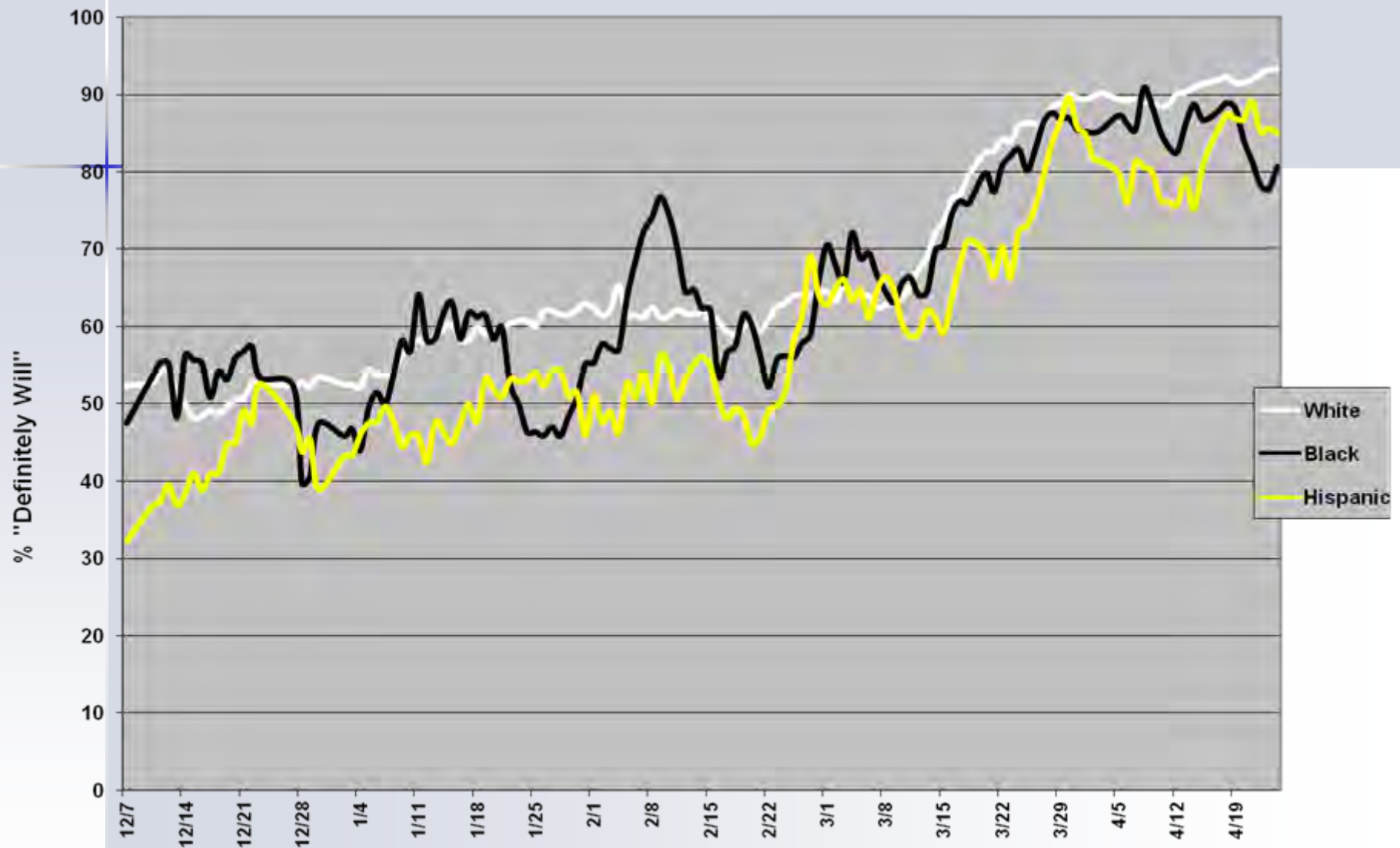


How likely are you to participate in the 2010 Census? – By Age Group





Percentage "Definitely Will" Return by Race/Ethnicity by Day



Intent to Participate, Census 2000 and 2010



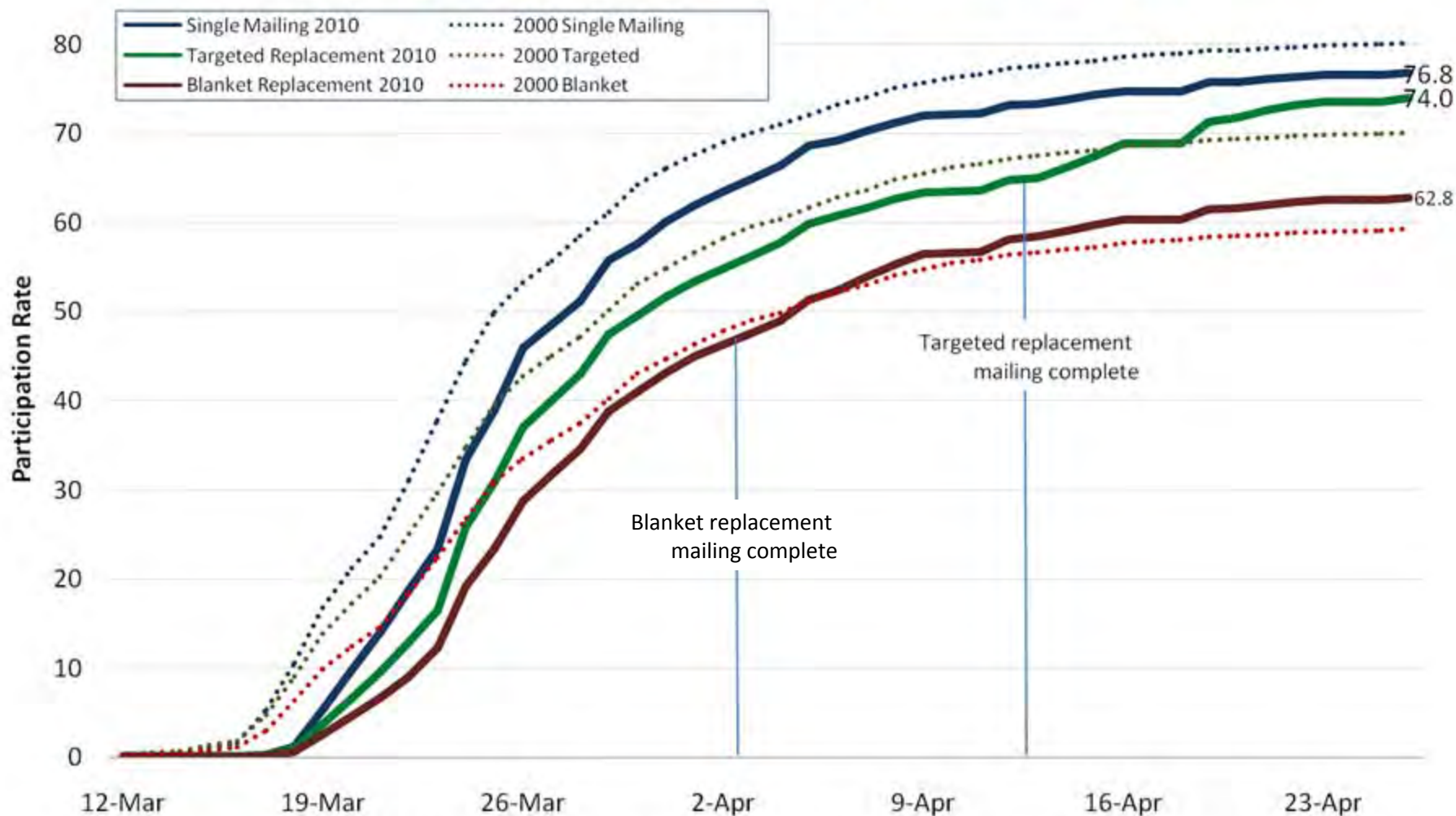
Forms in field,
data on actual behavior
(self report for KN,
returned forms for NORC)

- KN 2000: "How likely are **you or someone in your household** to answer and send back the census? Definitely will, probably will, might or might not, probably will not?"
- NORC 2000: "How likely are **you or someone in your household** to answer and send back the Census when you receive it? Definitely will, Probably will, Might or might not, Probably will not, Definitely will not?"
- Gallup 2010: "How likely **are you** to participate in the 2010 Census? By participate, we mean fill out and mail in a Census form. Would you say you definitely will, probably will, might or might not, probably will not, or definitely will not participate?"

3. Soundness of Process of Data Collection

Feature	Evaluative Indicator Relative to 2000
Short-form only design	Higher participation rates than long form (overall 72% in 2010 vs. combined short-long 69% in 2000)
Bilingual questionnaire to 13 million households with >20% Spanish-speakers	Increased participation rate by about 2 percentage points
Replacement form sent to 40 million households, targeting hard to enumerate	Increased participation in hard to enumerate areas

2000 and 2010 Average Mailback Participation Rates for Census Tracts by Assignment of 2010 Replacement Forms



Source: Take 10 Participation Rates, Decennial Management Division



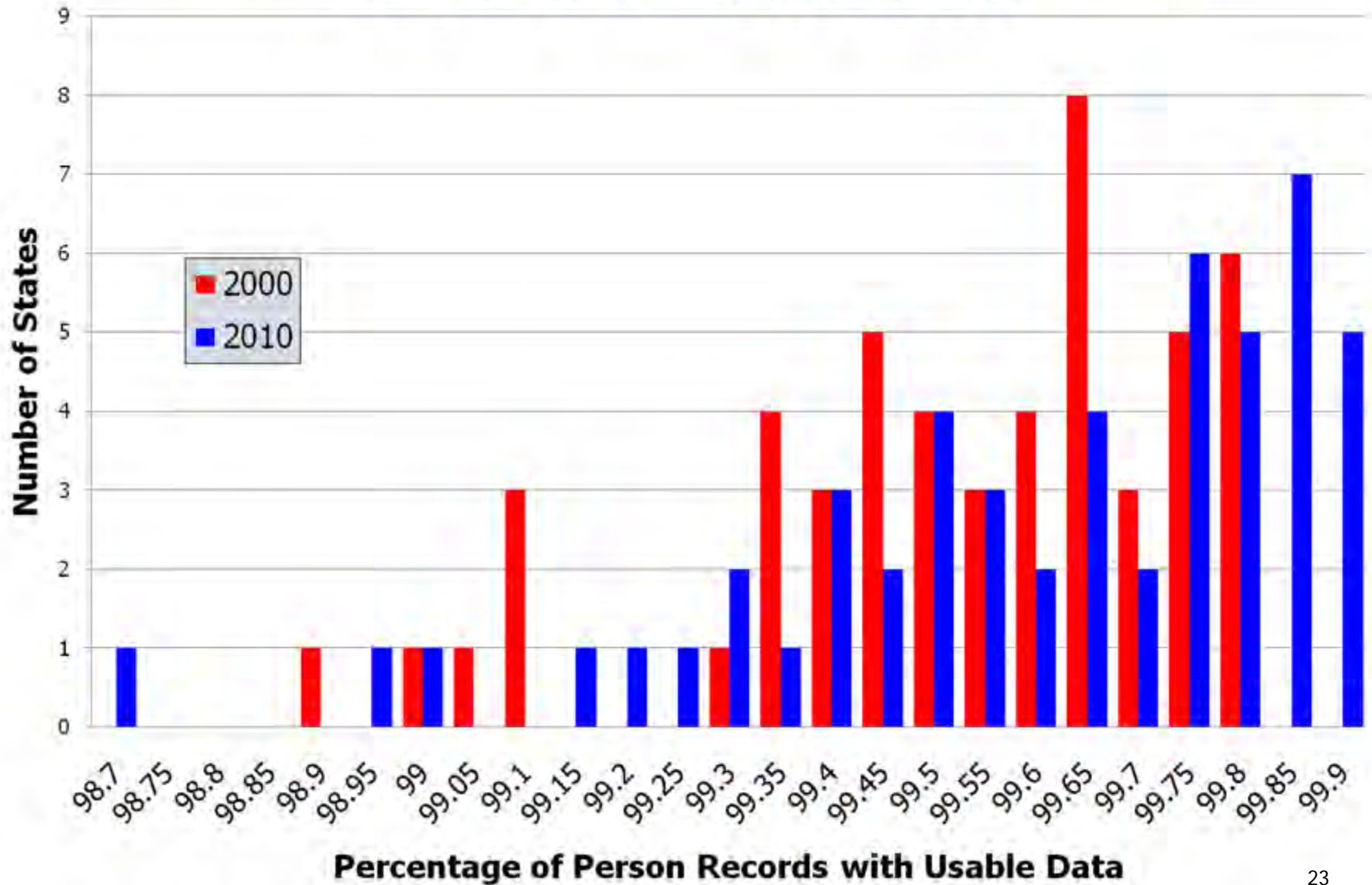
3. Soundness of Process of Data Collection

Feature	Evaluative Indicator Relative to 2000
Verification of composition for households with dynamic membership	Recontact with more households (7.5 million versus 2.5 million)
Assignment to enumerators	Fewer miles per interview
Nonresponse followup visiting rules the same	Higher percentage of population counts in households from proxy respondents

3. Soundness of Process of Data Collection

Feature	Evaluative Indicator Relative to 2000
Unit and Person Imputation Rates	Slightly lower rates of unit and person imputation than 2000
Quality control reinterviews	Essentially all enumerators subject to reinterviews (vs. 75% in 2000); smaller percentage of enumerators failing to meet quality standards
Meeting deadlines on data collection	All 11 operations since mid-2009 on schedule and significantly under budget (\$1.87 billion returned)

Count of States by Percentage of Usable Person Records from Data Collection, 2000 and 2010



Much Evaluation Remains

- Many formal evaluation projects
- Final post-enumeration survey results in 2012
- Of note
 - Post-hoc administrative record census simulation

A New Tool to Help Interpretation of the Census

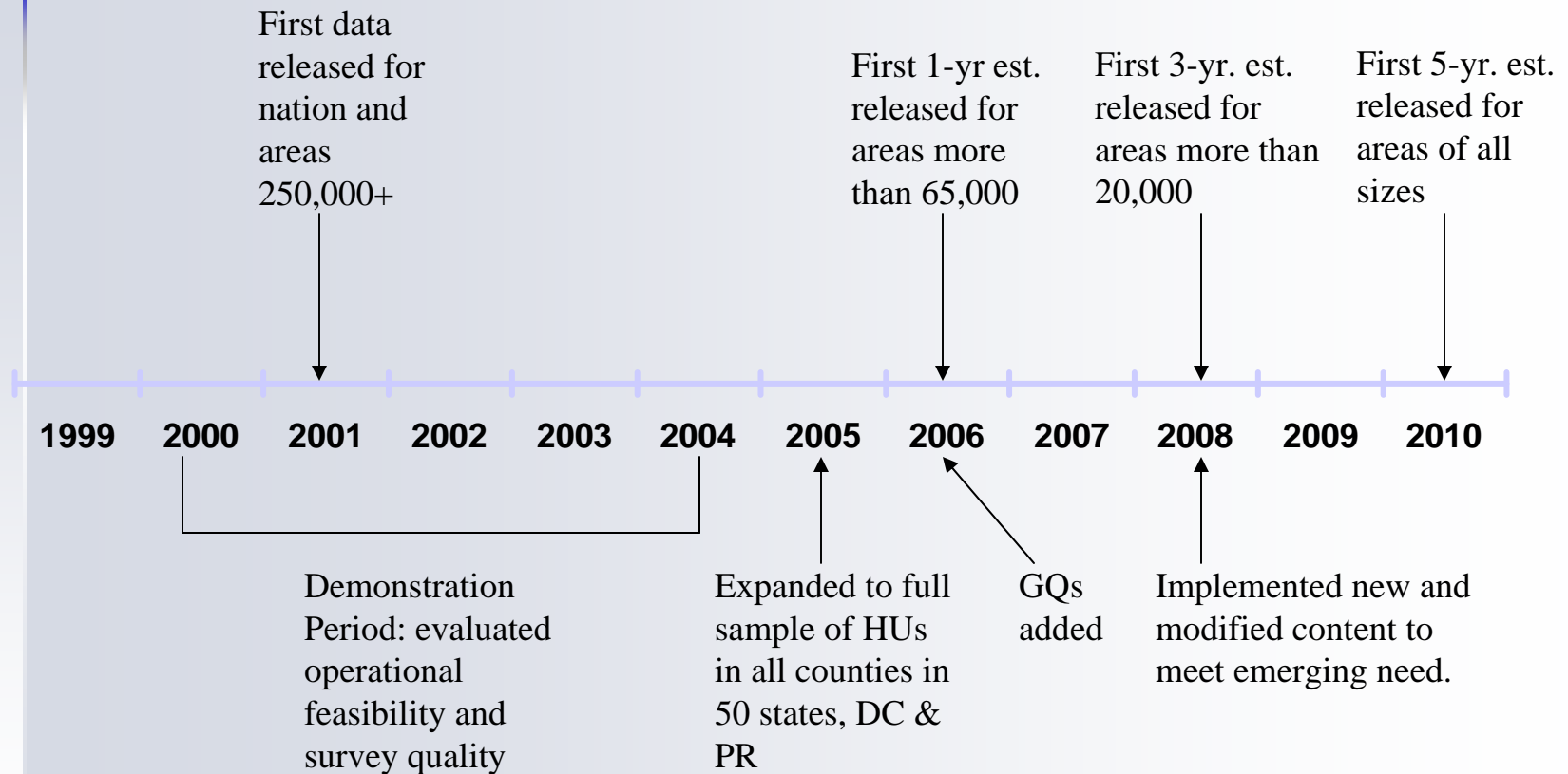
- The 2010 Census is lean in variables
- Comparison with trends of annual change based on the American Community Survey sometimes helps

2020 Planning - Organization

New 2020 Directorate

- It will start out as a small directorate, and then over the years grow into that organization needed to lead the Census Bureau in completing the 2020.
- At its inception it will contain the American Community Survey Office and a small research and planning division.
- Much of its early work will be developing an integrated plan for research and testing , in addition to defining effective integration with the new research directorate, the CFO's office, and the IT enterprise activities.
- A central focus of the 2020 planning will be efforts to reduce the cost of the census. We also want to gain efficiencies in the planning process by using the ACS as a principal test vehicle for the 2020.

History of the ACS



ACS Program Review

- Reassess mission, vision, goals and objectives
- Process Improvement
- Methods (Internet test, multi-lingual brochure, etc.)
- Data Products
- Internal *and* external stakeholders included in review

2020 Census Vision

An efficient and quality census that counts people, once, only once, and in the right place.

2020 Census Mission

The 2020 Census will conduct a census of population and housing and disseminate data to the President, the States, and the American people.

Guiding Principles for the 2020 Census

- Reduce Cost
- Maintain Quality
- Reduce Field Timeline
- Tailor Response Modes
- Leverage ACS
- Support Continual Frame Updating
- Support Agile Decision-making
- Reuse Data
- Leverage Systems and Methods
- Leverage Partnerships
- Support Organizational Solutions
- Build a 21st Century Workforce

2020 Census Goals

**A Complete and
Accurate Census**

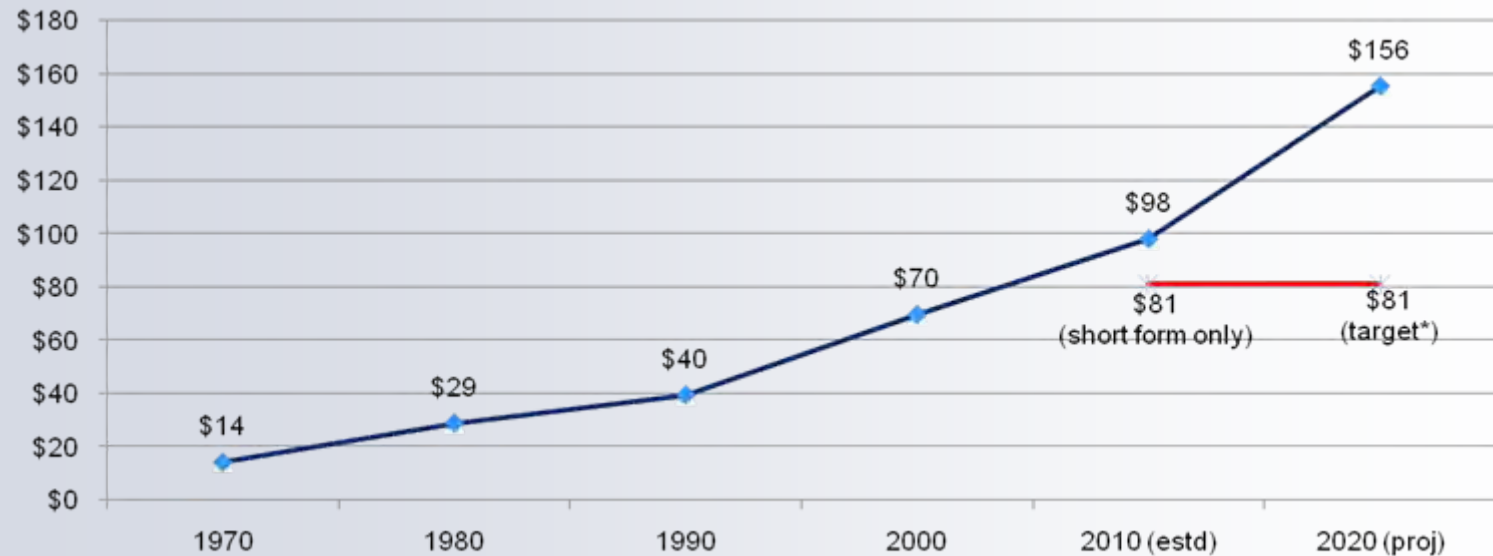
**Embraced and
Valued Results**

**An Efficient
2020 Census**

**A Well-Managed 2020
Census Program**

The Challenge and the Target

Census Cost per Housing Unit (2010\$)



Notes:

Adjusted for projected FY 2010 savings of \$1.6 billion; projected cost per housing unit for 2020 assumes no change in design and real cost growth from 1990-2000 and 2000-2010 averaged (58.3%); includes the costs for the 2002-2012 MAF/TIGER Enhancement Program and 2001-2013 American Community Survey.

(*) One of the goals of the 2020 Census is to conduct the census at a lower cost than the 2010 Census (per housing unit on an inflation-adjusted basis), while maintaining quality.

Four Cost-Drivers

Cost-Driver #1

Increased population diversity and decreased willingness to cooperate with self-response and non-response follow-up

Cost-Driver #2

Limited 2000-2009 updating of the Master Address File and TIGER maps led to a design incorporating “last-minute” (2009) updates

Cost-Driver #3

Failure to link acquisitions, schedule, and budget

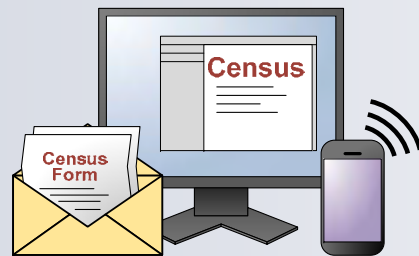
Cost-Driver #4

Demand for absolute accuracy

Potential Solutions To The Four Cost-Drivers

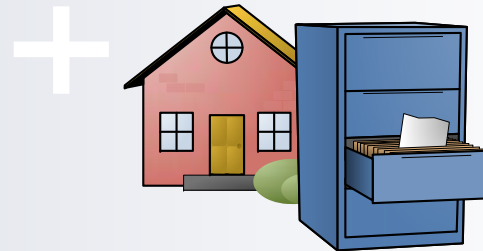
Cost-Driver #1

Increased population diversity and decreased willingness to cooperate with self-response and non-response follow-up



Potential Cost-Reducer

Expanding, automating, and tailoring self-response methods, including an Internet option, and using administrative records for follow-up



Potential Solutions To The Four Cost-Drivers

Cost-Driver #2

Limited 2000-2009 updating of the Master Address File and TIGER maps led to a design incorporating “last-minute” (2009) updates



Potential Cost-Reducer

Continuously updating the address frame and targeting address canvassing and map updates using multiple sources and geographic partnerships, as described in FY 2011 Geographic Support Systems budget initiative

Geographic Support System Initiative

- A program of improved address coverage, spatial feature updates, and enhanced quality assessment and measurement of address and spatial data and systems.
- The objective is to enable continuous update of the, Census Bureau's Master Address File and TIGER database to support programs such as the American Community Survey, other current surveys, and the Population Estimates program. It also supports the goal of a targeted (vs.nationwide) address canvassing for the 2020 Decennial Census.
- The GSS initiative supports ongoing geographic partnership efforts with federal, state, local, and tribal governments as well as commercial entities so that the Census Bureau can acquire the most up to date address and spatial information available.

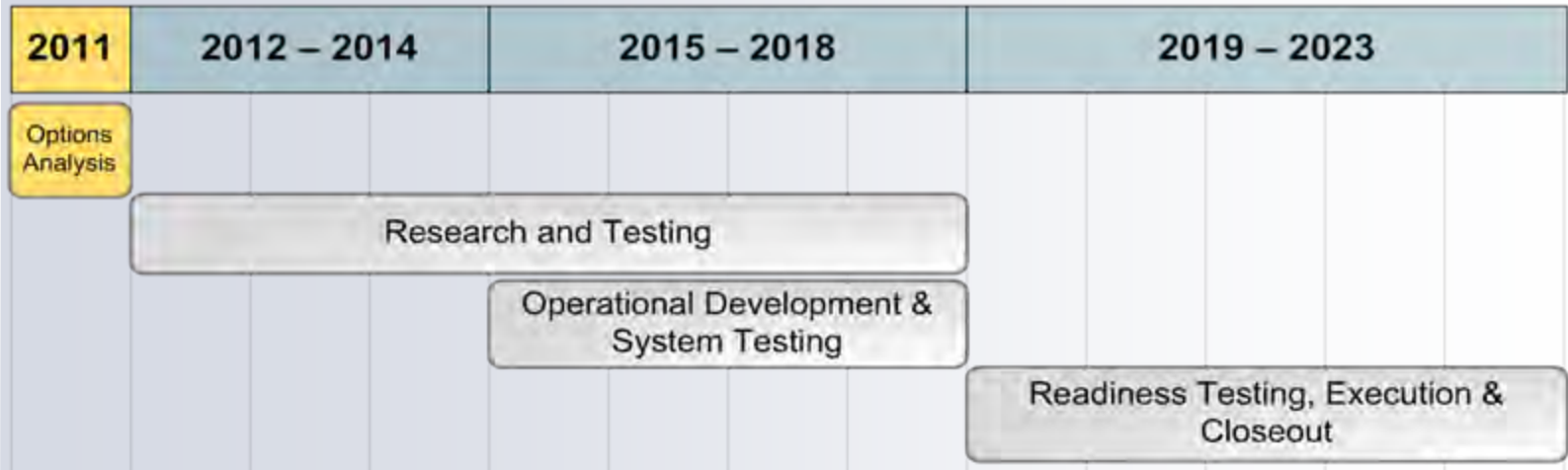
Potential Solutions To The Four Cost-Drivers

Cost-Driver #3

Failure to link acquisitions, schedule, and budget

Potential Cost-Reducer

A rolling approach to planning, budget, and schedule



Potential Solutions To The Four Cost-Drivers

Cost-Driver #3

Failure to link acquisitions, schedule, and budget

Acquisition Management

Identify sourcing and acquisition strategies at beginning of lifecycle

Budget Management

Use a rolling budget approach and successive approximation to continuously refine annual estimate

Schedule Management

Develop a 2020 integrated schedule covering all projects throughout full lifecycle of program

Performance Management

Use performance information from the strategic to the project level to adjust resources and activities

Potential Cost-Reducer

Strong program management strategies to reduce program risk

Risk Management

Build on 2010 experience and establish a risk register and management process from beginning

Human Capital Management

Identify needed skill sets and gaps, as part of broader Census Bureau assessment

Governance and Transition Management

Align governance with major design components

Communications Management

Roll out plans early to oversight entities, the Advisory Committees, and to other stakeholders for feedback

Potential Solutions To The Four Cost-Drivers

Cost-Driver #3

Failure to link acquisitions, schedule, and budget

Potential Cost-Reducer

Full-Life Cycle Work Breakdown Structure

2020 Census WBS

Program Management

Systems Engineering

Frame

Enumeration

Response Processing

Data Products

Evaluative Programs

Infrastructure

Potential Solutions To The Four Cost-Drivers

Cost-Driver #4

Demand for absolute accuracy



Congress



OMB



GAO



DOC IG



National Academy of Sciences



Race and Ethnic Advisory Committees



2020 Census Advisory Committee

Potential Cost-Reducer

Building consensus among stakeholders regarding tradeoff between accuracy and cost



American Public



State, Local, and Tribal Governments



Business/Industry Data Centers



Census Information Centers

Operational Design Alternatives Drive the Research

- Establishing Where to Count
- Enumeration
- IT and Operational Infrastructure

Establishing Where to Count

Three Options

Full Canvassing

Targeted Canvassing

No Canvassing

Continuous Frame Updating

- All options: continuous updating of address and spatial information in the MAF-TIGER database along with expanding the list of updating sources

Supplemented by a full address canvassing in 2019

Supplemented by a targeted address canvassing in 2019

No (special) address canvassing for the 2020 Census

Establishing Where to Count

Research Questions

- How can we build a usable statistical model for Master Address File (MAF) errors, error components, and their magnitude?
- How can the model be used to measure the quality of the Master Address File?
- How can we improve Local Update of Census Addresses (LUCA) procedures for compatibility with continual updating, improved quality measurement, and expanded partnerships with local governments, and a targeted address canvassing operation?

Enumeration

Six Options

Traditional + Internet	Multi-mode Contact	Visit Then Records	Records Then Visit	Near Paperless	Records-based
Initial Contact	First Collect	Reminder	Second Collect	Final Collect	Gap Collect



Enumeration

**Traditional +
Internet**

Multi-mode
Contact

Visit Then
Records

Records Then
Visit

Near Paperless

Records-based

Initial Contact



Direct Mail

First Collect



Mailback

OR



Online

Reminder



Direct Mail

Second Collect



Mailback

OR



Online

Final Collect



Home Visit

OR



Phone Audio

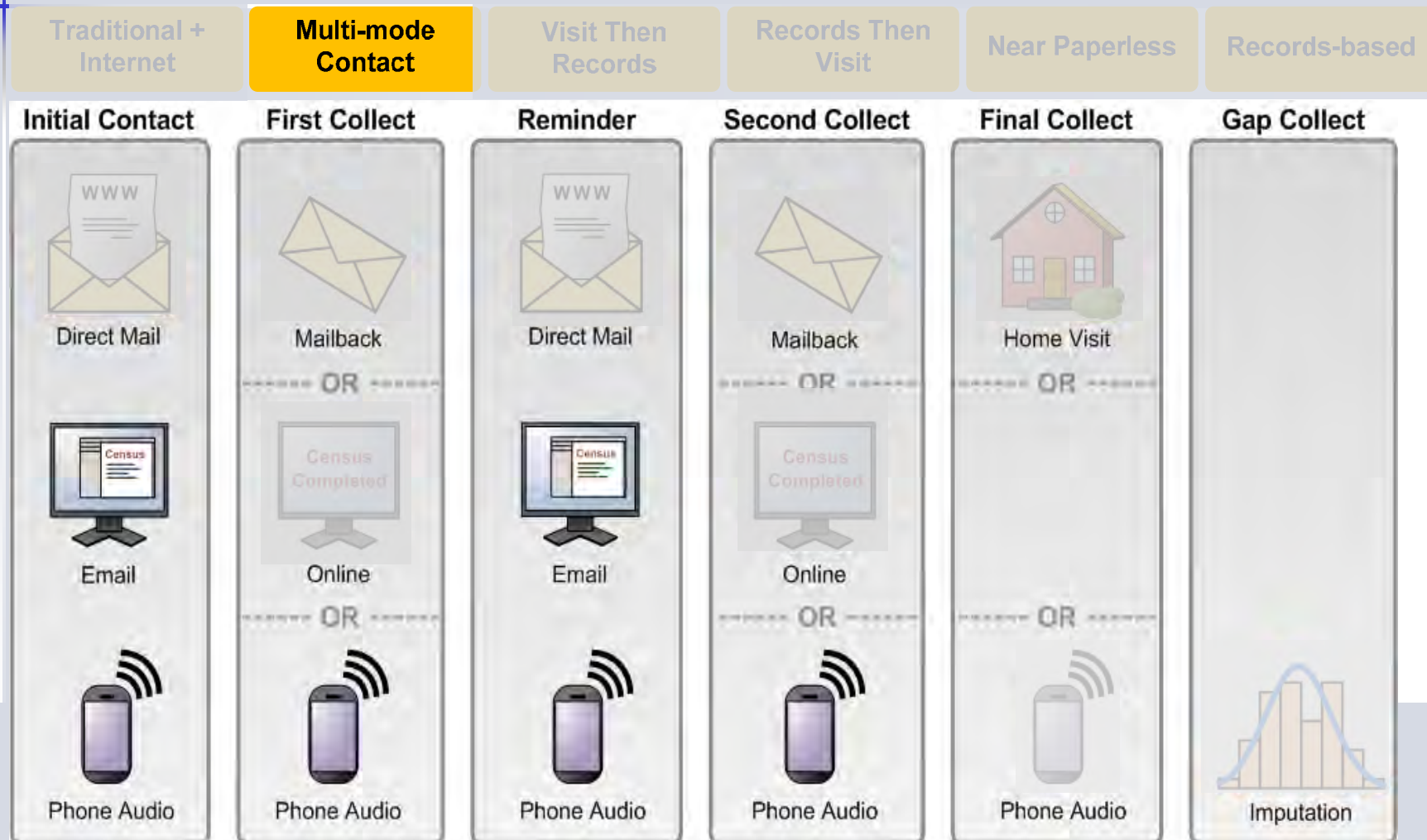
Gap Collect



Imputation

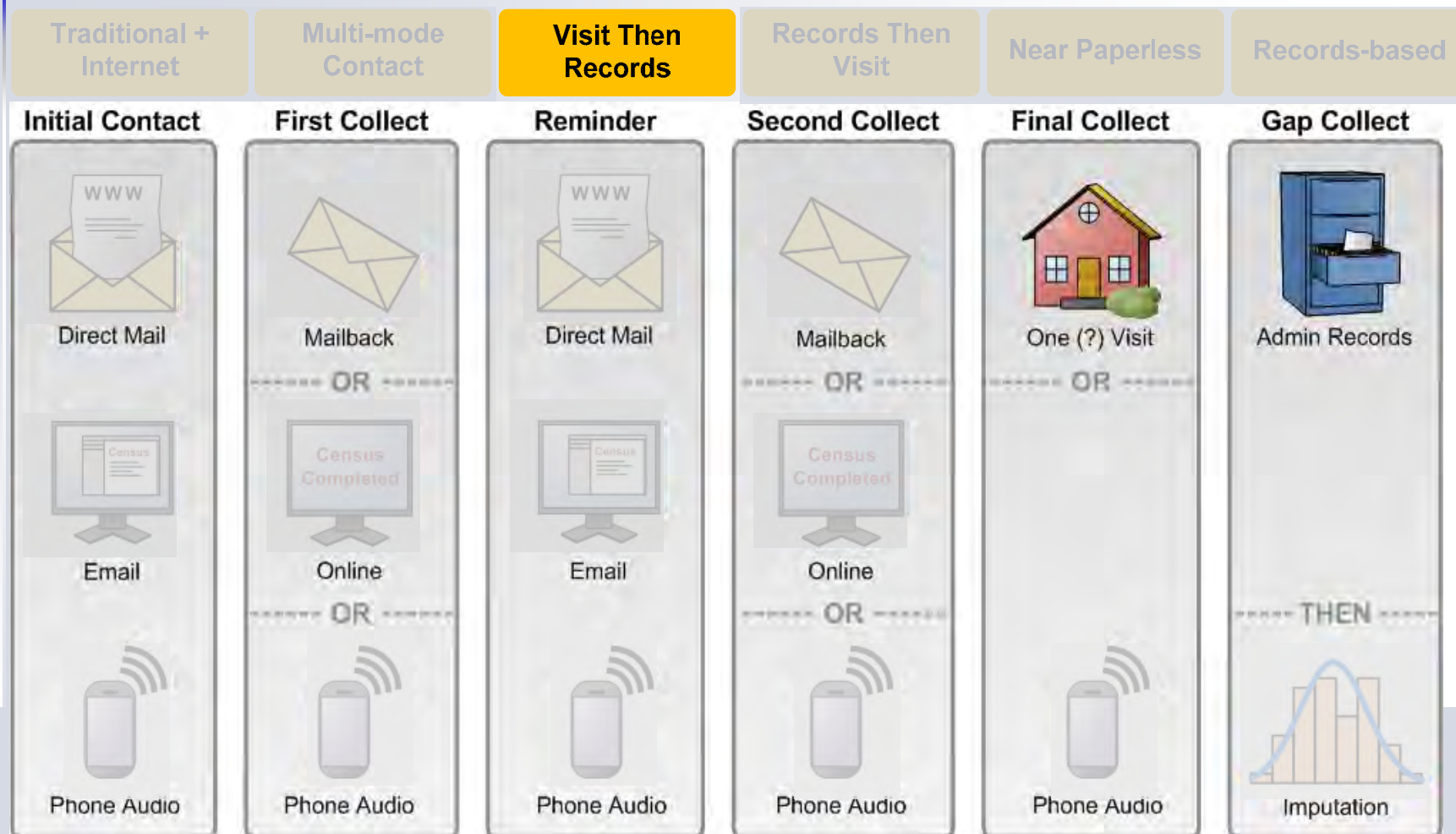


Enumeration

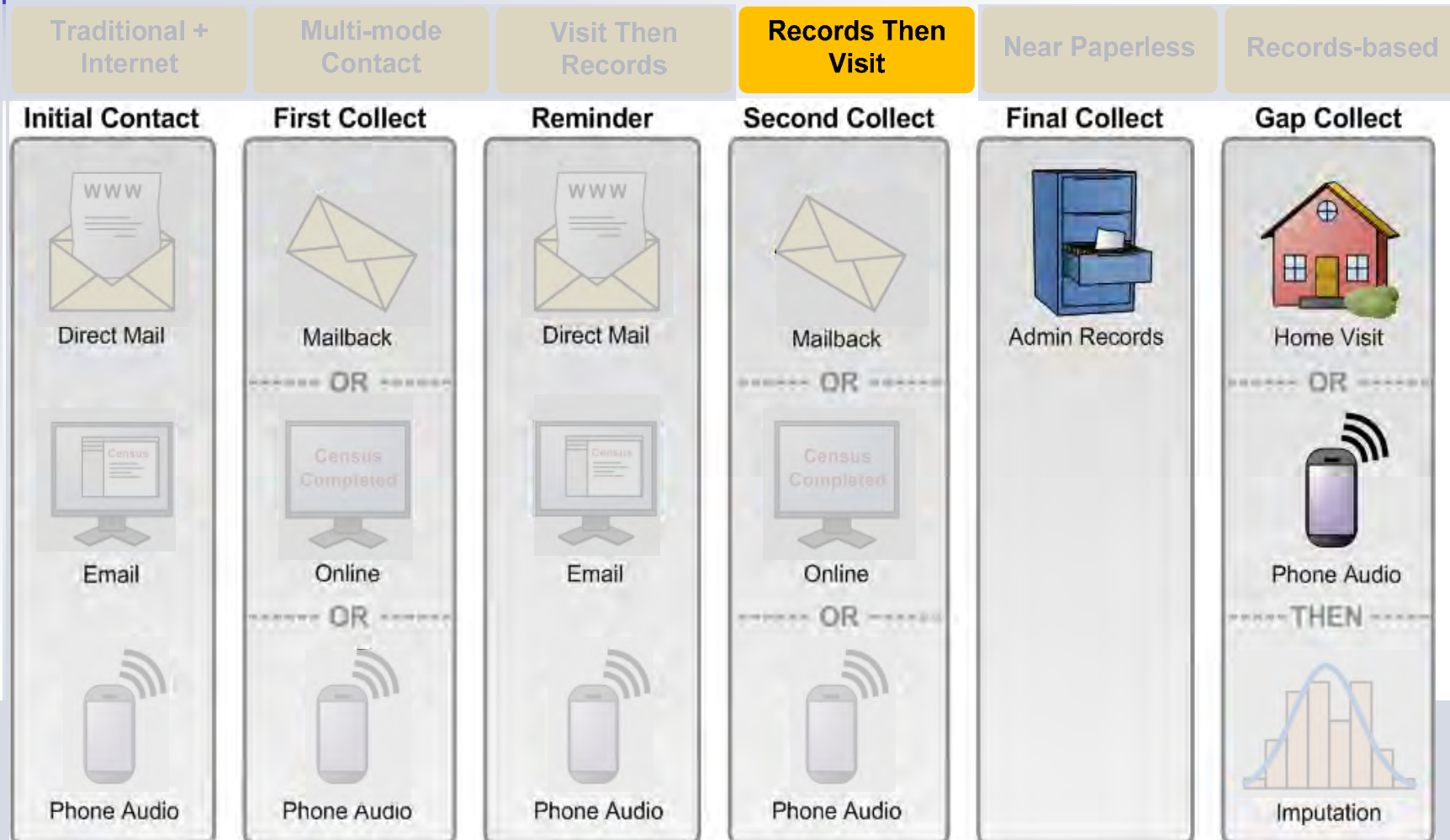




Enumeration



Enumeration



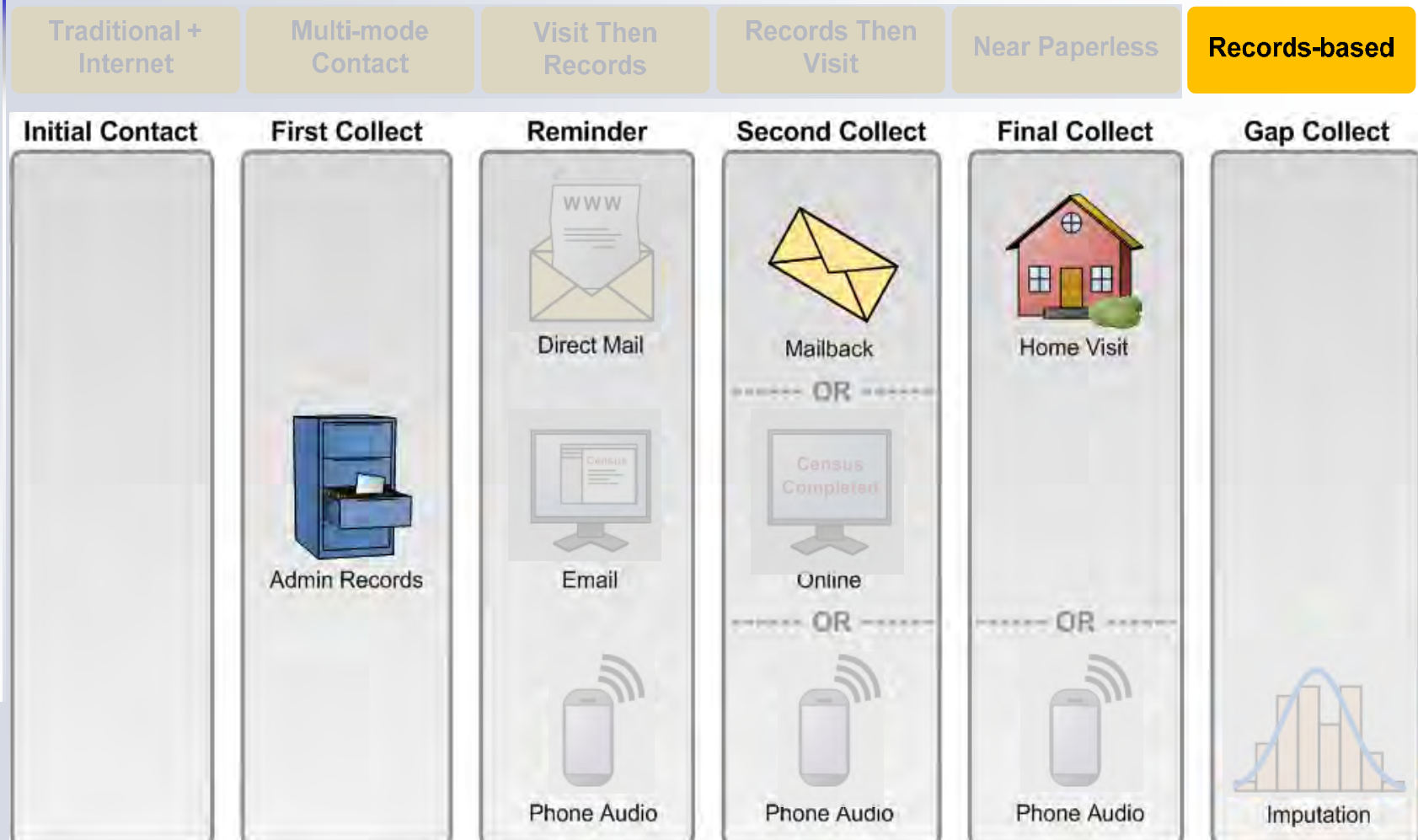


Enumeration





Enumeration



Enumeration

Research Questions

Self-Response Research

- What is the best mix of modes and strategies by demographics and geography?
- What contact frames can be linked to the physical address?
- What technology will be feasible for self-enumeration in 2020?
- How can household follow-up be improved or eliminated?
- What language services and technologies are needed?

Enumeration

Research Questions

Non-Response Research with Administrative Records

- How can we develop and maintain an independent administrative records database?
- How can administrative records be used to minimize census field work?
- What is the quality and the coverage of responses supplied by administrative records?

IT and Operational Infrastructure

Three Options

Decentralized

Hybrid

Centralized

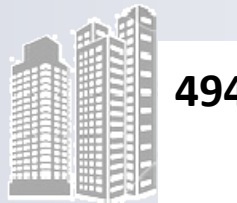
IT and Operational Infrastructure

Decentralized

Hybrid

Centralized

- Modification of current approach



494 Local Census Offices

- Modest HQ IT Improvements

IT and Operational Infrastructure

Decentralized

Hybrid

Centralized

- Centralized approach to managing through automating field work



**Just a small number of depots
(Logistics Support Centers)**

- Shared IT platforms

IT and Operational Infrastructure

Decentralized

Hybrid

Centralized

- Centralized management where feasible due to geographic challenges



Fewer Local Census Offices

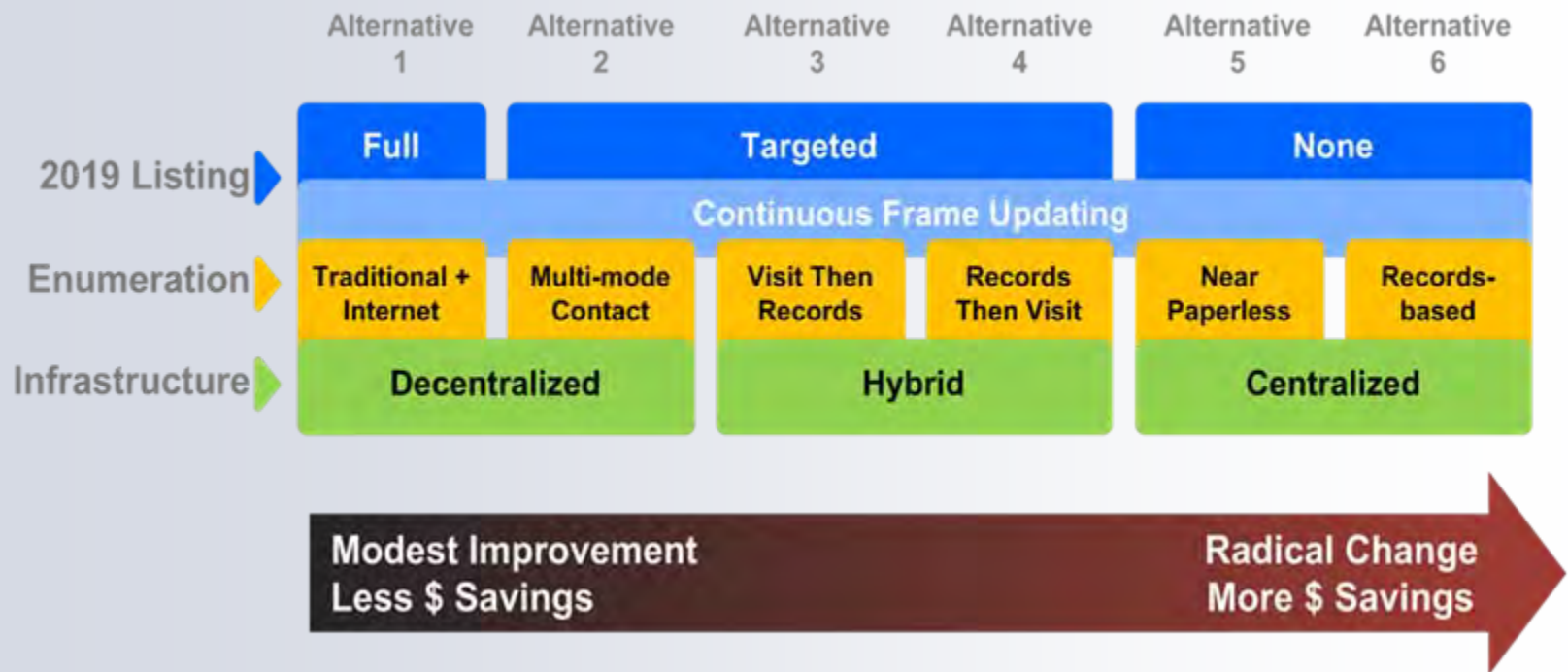
- Shared IT platforms

IT and Operational Infrastructure

Research Questions

- What technologies will be available to support IT and field infrastructure?
- How can we effectively automate and streamline field operations to take advantage of design changes?
- How can we modernize universe control, HQ management, data processing?
- What are our workload management system options and how can we best acquire and implement them?
- How can we use available technology to improve QC?

Six Key Operational Design Alternatives





Policy and Communication Challenges

- ❑ Cost versus Quality Tradeoff
- ❑ Meeting Stakeholder Expectations
- ❑ Public Concern about Privacy and Confidentiality
- ❑ Public Confidence in the Census

2020 Census Early Planning Highlights

□ **Business Plan**

- Provides overall concepts and approach for managing, engineering and designing the 2020 Census.

□ **Design Decision Roadmap**

- An attempt to provide the critical path to get to preliminary designs for the 2020 Census operations during the Early Research and Testing Phase, FY 2012-2014.

□ **Preliminary Concepts of Operations**

- Identified fundamentally different approaches for conducting the 2020 Census, focusing on changes that could potentially reduce cost while maintaining quality.
- Built out preliminary concepts of operations for six (6) possible design alternatives.
- The six options span from modest to extreme changes from the 2010 Census baseline.

2020 Census Early Planning Highlights

□ **Schedule**

- Developed high and mid-level schedules for FY 2012-2014 using milestones identified for key design decision points and deliverables for all 2020 Early Research and Testing Projects.

□ **Program Management Processes**

- Developed strategies for key program management process areas to ensure integration of and consistency, utilizing a standard work breakdown structure (WBS) to integrate cost, schedule, requirements and risk information.
- Key strategy areas include acquisition and sourcing, performance management, risk management, issue management, change management, governance, and schedule management.

2020 Census Early Planning Highlights

□ Skills Assessment

- Completed a skills assessment for what is needed during the 2020 Census Early Research and Testing Phase. This will be used as input for a workforce competency assessment to be conducted agency-wide, and then we will identify strategies for overcoming any skills gaps in conducting early research work for the 2020 Census program. This could identify specific training needed by Bureau staff or help us determine what kinds of skills we need to recruit either as full-time employees or through contracts.

□ Communications

- Kicked off a communications management team to develop communications strategies for external stakeholders and partners, oversight organizations, and Census Bureau and Commerce Department employees.

Conclusions

- ❑ **The purpose of the research and testing phase is to determine how much change is possible; critical design decisions are required by 2015 to avoid past problems**
 - To what extent can the 2020 Census target address canvassing in 2019?
 - How will we collect and integrate respondent information (which response modes will we use and to what extent will we use administrative records)?
 - What will be the extent of field infrastructure (in other words, how far should we go toward centralization)?
- ❑ **Investment is required early in the cycle to reduce cost and risk as the decade progresses (late design changes add risk and cost)**
- ❑ **The more substantial a change to the decennial design that is contemplated, the earlier research must be concluded.**