

# Approaches to Pre-Enumeration Census and Survey Mapping

Claudia O. Espinosa Villegas and Derek Azar  
Population Division, U.S. Census Bureau  
USCB Workshop on Preparing for CAPI Operations  
September 15, 2022

# Outline

## Introduction

- What does digital mean for statistical mapping?
- What happens during statistical mapping?

## Approaches

- In-Office, Field, and Hybrid Approaches
- Skills, Resources, Advantages, and Challenges

## Downstream Effects

- How does mapping affect subsequent operations?
- What is the effect of changing modes between operations?

## Census and Survey Comparison

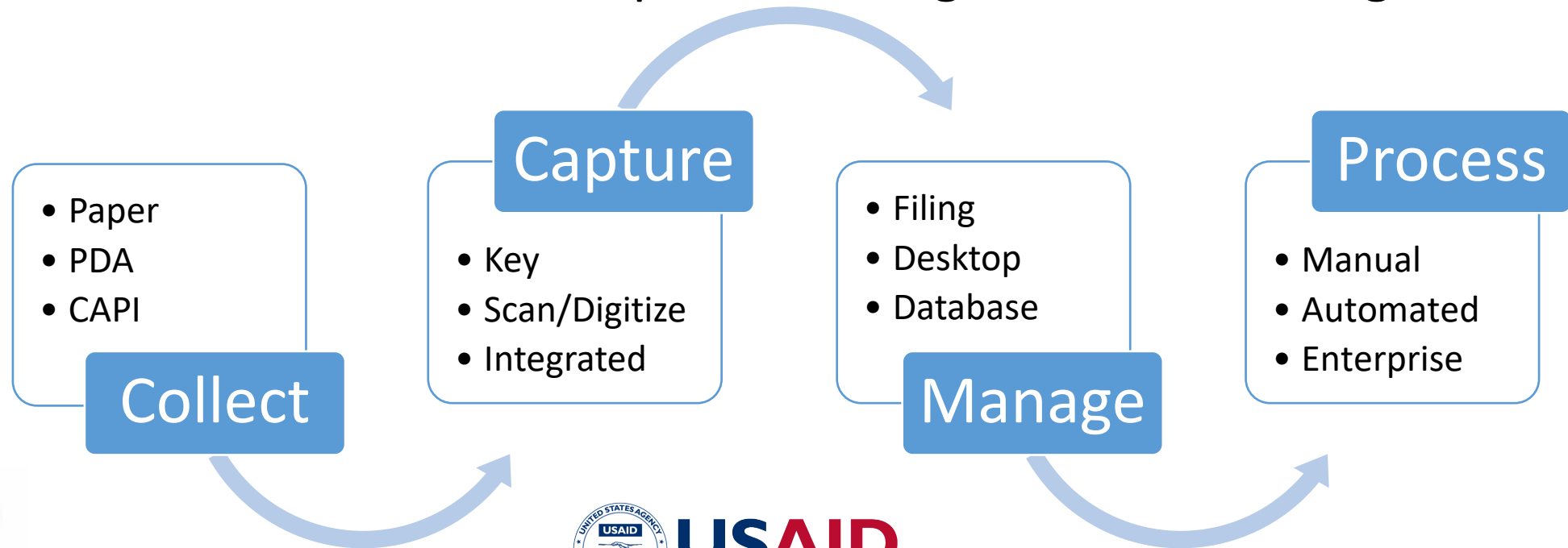
- Why has digital adoption taken longer for censuses?
- What are the effects of scale?

## Technical Assistance

- What are we doing?
- What can we do better?

# Effect of Digital Data Collection on Statistical Mapping

- Strong parallels to how storage and processing of statistical tabular data have progressed
- Steady improvement of processing and storage capabilities
- Gradual and uneven adoption of integrated data management



# Components of Statistical Mapping

- Demarcation
  - Define work and sampling areas for data collection
- Listing
  - Define the in-universe or in-sample units of enumeration
- Operational Control
  - Use of spatial data to track and monitor subsequent operations

# Digital Spatial Data and Management

- We have had digital components to census mapping since the late 1960s. Developed world NSOs adopted digital mapping moving into the 1970s.
- Incremental progress toward digital mapping was made in following 40 years.
- There has been a close partnership between software developers and the statistical geography community.

# Rapid Adoption

leapfrog

verb \ˈlēp-ˌfrɒg, -ˌfræg\

*to move ahead of or beyond (someone or something) in a very quick and sudden way*

*Merriam-Webster.com*

- Developed world NSOs **gradually** incorporated new electronic technologies into the census process over the course of 100+ years.
- Developing world NSOs have the opportunity to **leapfrog** this gradual approach and **rapidly** adopt new technology in just a few years...
  - ... and most are taking advantage of this opportunity!

# Gradual Adoption

- U.S. Census Bureau:
  - **Pre-1950s:** Heavily paper driven, door-to-door enumeration; early computerization (e.g. UNIVAC).
  - **1970s/80s:** Increased mechanization; mail-out/mail-back questionnaires; small spatial databases.
  - **1990s/2000s:** Master Address File/TIGER development and integration (advanced spatial database); laptop questionnaires (some surveys).
  - **2010s:** Internet response; tablet/smartphone questionnaires; optimized field workforce management; more extensive geospatial data management.

# Rapid ~~Gradual~~ Adoption

## ~~Developing World NSOs~~

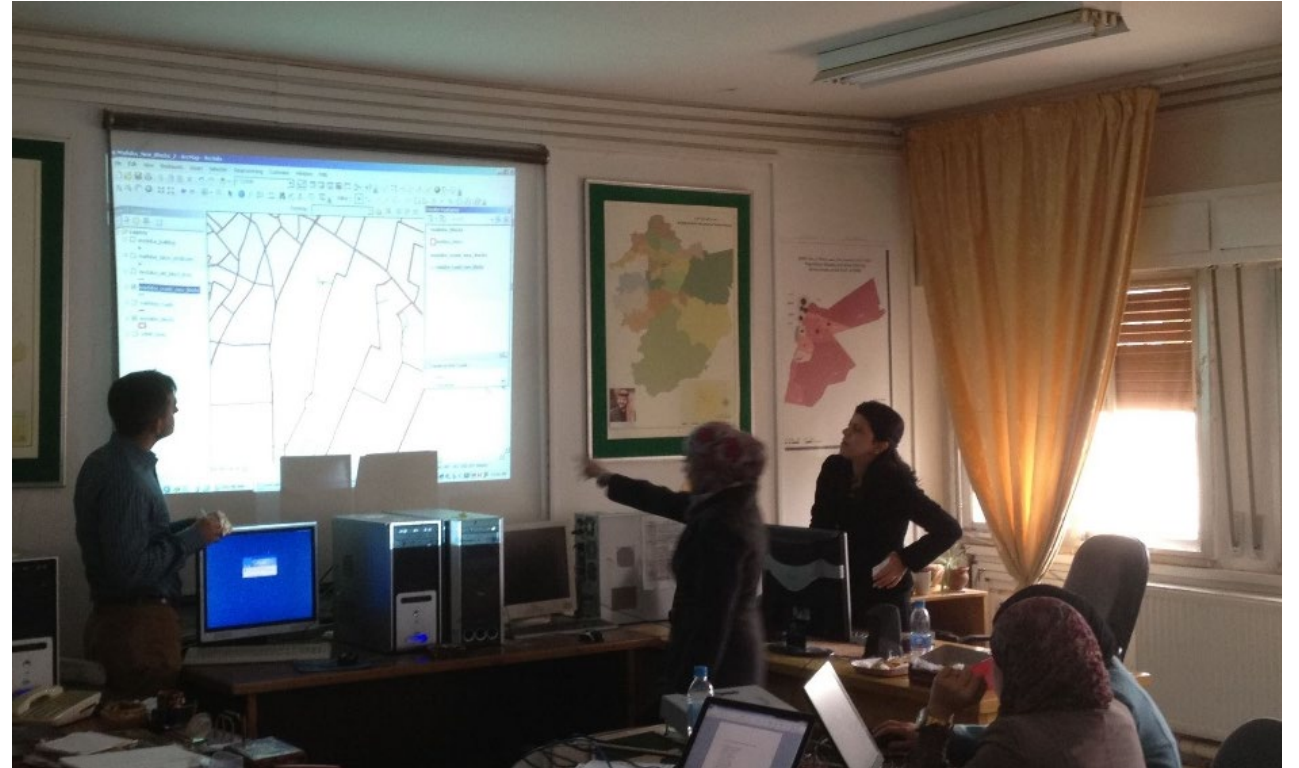
- U.S. Census Bureau **2000s**
  - **Pre-1950s:** Heavily paper driven, door-to-door enumeration; early computerization (e.g. UNIVAC). Desktop + CPro
  - **1970s/80s:** Increased mechanization; mail-out/mail-back questionnaires; small spatial databases.
  - **1990s/2000s:** Master Address File/TIGER development and integration (advanced spatial database); laptop questionnaires (some surveys).
  - **2010s:** Internet response; tablet/smartphone questionnaires; optimized field workforce management; more extensive geospatial data management.

Source: <https://www.census.gov/history/>



# Approaches to mapping operations

In-Office Methods



# In-Office Methods

- Almost certainly requires use of **imagery**...allows for aerial view not available in field without printed imagery (cumbersome) or tablet (technologically advanced)
- **...BUT** NSOs will consider in-office demarcation even when lacking satellite imagery.
  - Not advisable
  - Requires high-quality vector data such as road networks or cadastral data and considerable time for interpretation
  - Ideally vector layers are available even when using imagery



# What Happens During In-Office Demarcation?

- Conflation and interpretation of physical features and pre-existing statistical geography if they exist
- Estimation of the number of housing units
- Digitization of collection geography boundaries
- Enforcement of correspondence rules between statistical and administrative geography

# What Happens During In-Office Listing?

- **Identification** of possible housing units using satellite imagery
  - Multi-family dwellings will be tagged with one point during an in-office listing operation
- **Sorting** between likely residential buildings and non-residential or non-built areas
  - Sometimes difficult to distinguish between features and building-use





# Resource Requirements for In-Office Pre-Census Mapping

High-resolution imagery	High-speed internet	Upgraded workstations	Training
<ul style="list-style-type: none"><li>• At least 10-meter</li><li>• Worldview</li><li>• Sentinel?</li><li>• Quickbird (legacy)</li><li>• Aerial photography</li></ul>	<ul style="list-style-type: none"><li>• Streaming imagery</li><li>• Comparison</li><li>• Data download</li></ul>	<ul style="list-style-type: none"><li>• Processor</li><li>• Graphics card</li><li>• RAM</li><li>• Graphic element overlay</li></ul>	<ul style="list-style-type: none"><li>• Low learning curve for digitizers</li><li>• High learning curve for technical supervisor</li></ul>

# In-Office Benefits

- Reduced travel and time spent on ground
- Reduce resource requirements
  - Map printing
  - Network traffic during in-field operations
- Save on logistical and equipment costs

# Approaches to mapping operations



Field Methods

# Field Objectives





# Update

- Add missing units (housing or enumeration)
- Edit boundaries and attributes
  - Split/Merge
  - May include listing questionnaire
- Delete/flag objects misidentified as structures, vacant and non-residential buildings

# Verification

- Can an enumerator cover this area within the given collection period?
  - Physical features, obstacles
  - Number of housing units (80-150)
- Are all housing units identified and covered by a collection block?
- 
- Does collection geography conform to business rules?

# Capture and Finalization of Field Data

- Data may be transmitted through mobile network or WiFi (e.g. web-based file share, email)
- Boundaries should be checked by headquarters staff
  - Methodology will vary depending on technology used
- There should be a clearly differentiated working and production containers for incoming and verified data

# Approaches to mapping operations

Hybrid Methods

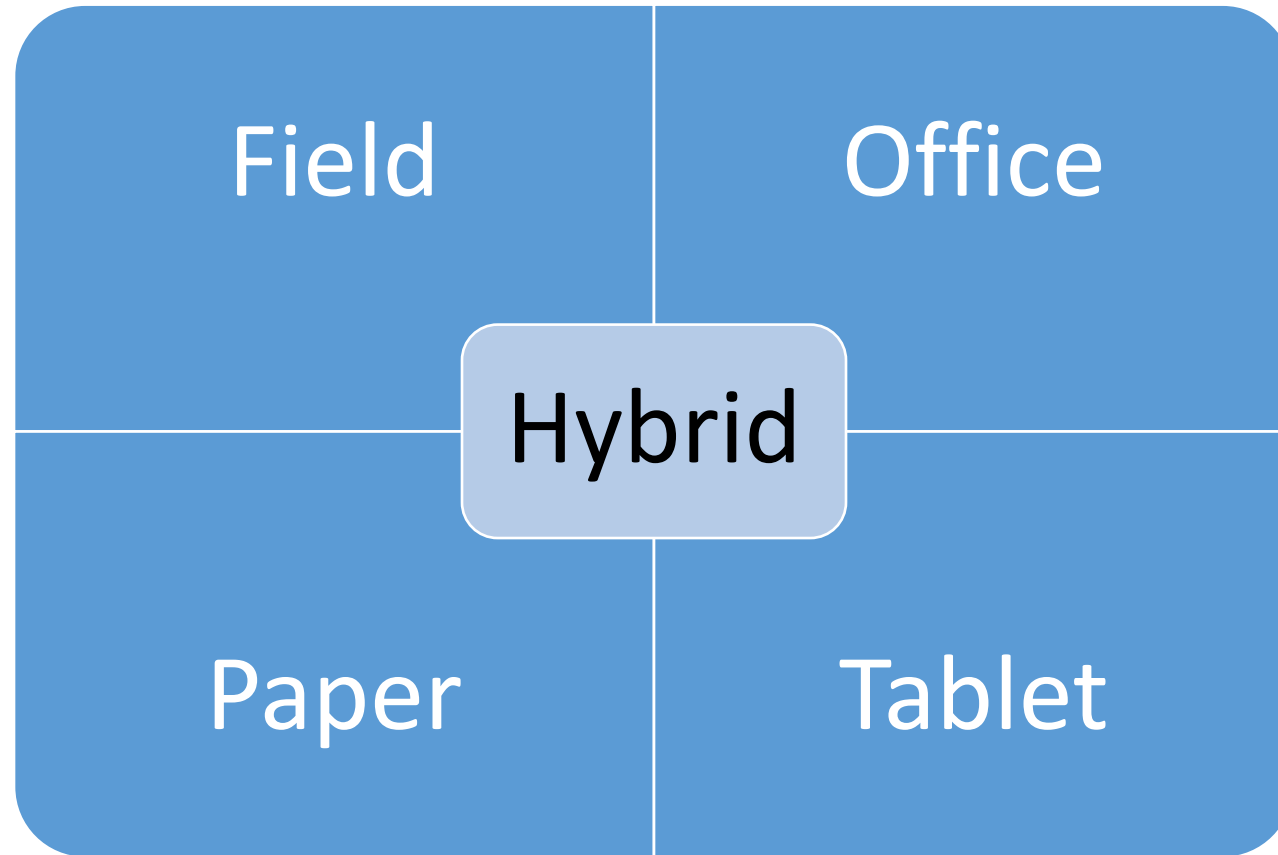
# CAPI or Paper Operation?



# What is a Hybrid Approach?

- Census methodologies may combine approaches
- Can view mapping methodology in a 2x2 matrix (field/office)x(paper based/IT)
- Different combinations appropriate for different skill and resource levels

# Mapping Operations



# Characterizing Approaches

## Field/Paper

- Most traditional
- Visit and interpret each block and HU
- Sketch maps
- Census geography sketched onto physical features
- Relational/Positional data

## Office/Paper

- Collection blocks pre-identified
- Structure points pre-identified
- Possible to clear blocks and HUs without any further work
- Annotation of updates onto positionally accurate GIS data
- Relational/Positional data

## Field/Tablet

- Field worker interpretation of ground-situation with digital tools
- Visit and interpret each block and HU
- Possible GPS guidance
- Census geography digitized onto physical features

## Office/Tablet

- Pre-identification
- Possible to clear blocks and HUs without any further work
- Census geography digitized onto physical features
- Possible GPS guidance
- Positionally correct data generated



# Resources and Training

## Field/Paper

- Printed materials and transportation
- Plotters\* and scanners\*
- Everyone is doing what they have been doing for the past N years
- Field staff may require training if working with GIS maps for first time

## Office/Paper

- Printed materials and transportation
- Imagery
- GIS software
- Plotters and scanners\*
- Training on digitization and photo interpretation
- Field staff may require training if working with GIS maps for first time

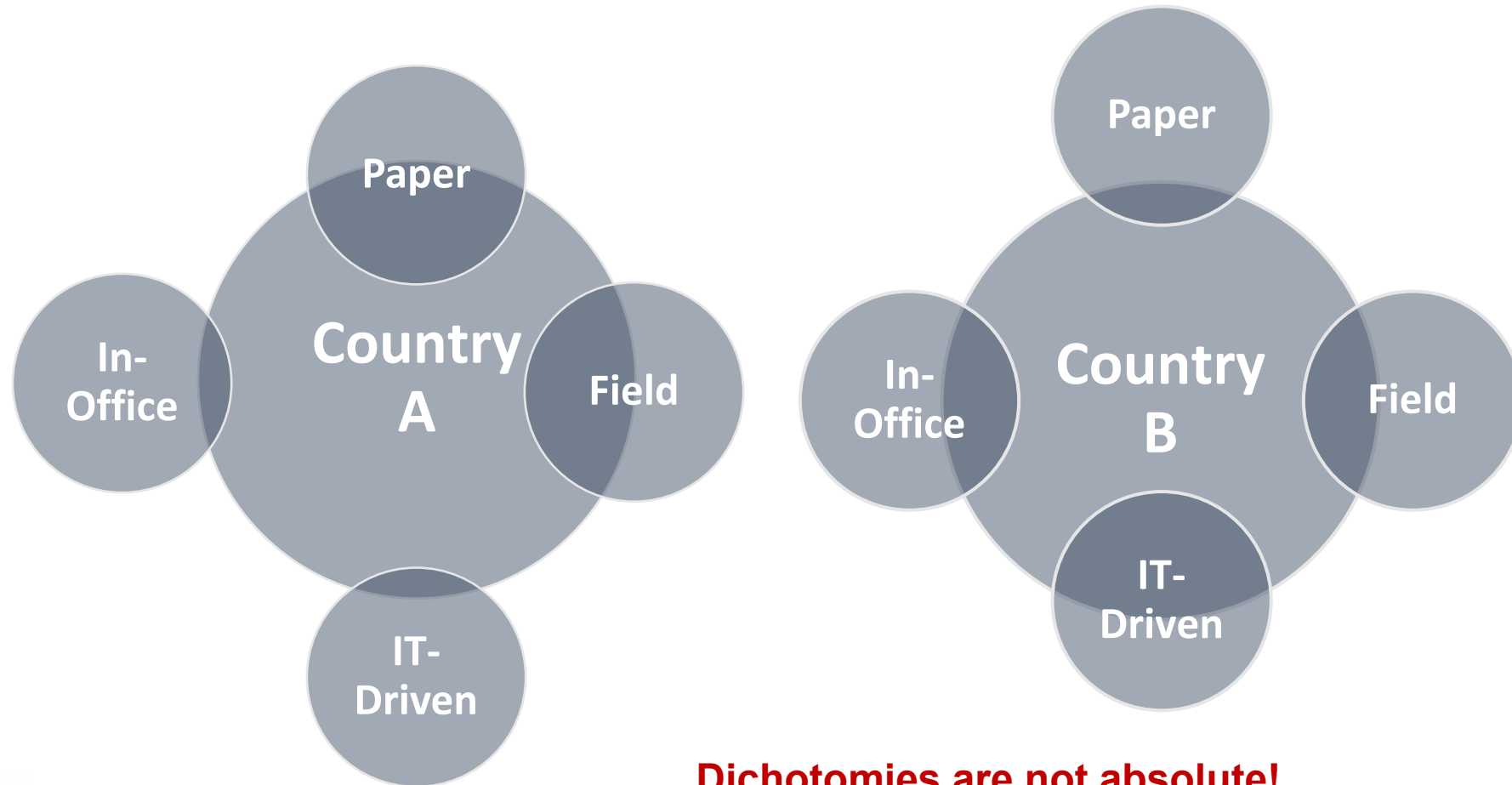
## Field/Tablet

- Devices, transportation, and security
- Mobile enabled spatial data and software
- Operations similar to past practice with different tools
- Tablet-specific instruction for field staff

## Office/Tablet

- Devices, transportation, security
- Imagery and GIS software
- Mobile enabled spatial data and software
- Training on digitization and photo interpretation
- Tablet-specific instruction for field staff

# Characterizing Approaches to Mapping Operations

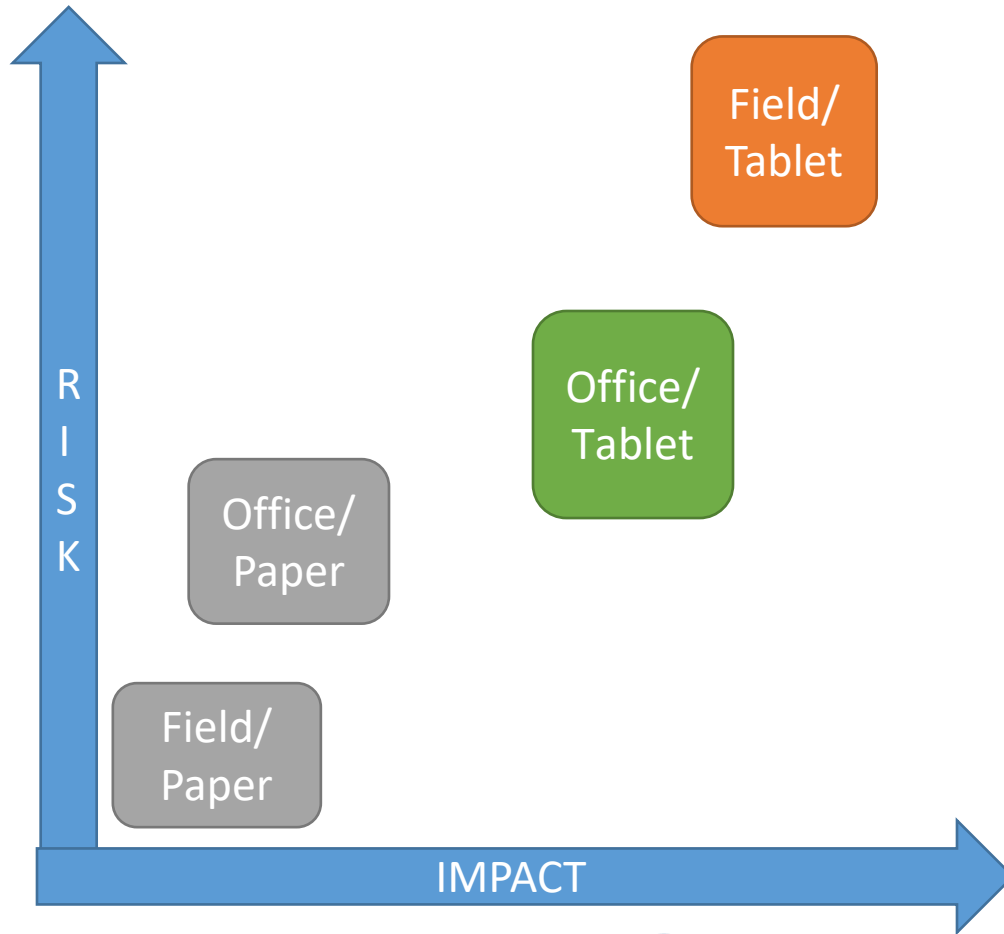


**Dichotomies are not absolute!**

# DOWNSTREAM EFFECTS

Relationship to CAPI

# Mapping Operations Network Infrastructure



- Network infrastructure should be key determinant when developing census methodology
- Security, usability, speed
- Data loss or confidentiality breach not necessarily only way to lose public trust

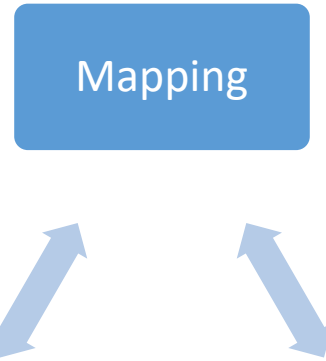
# Operational Control

- Maps used for navigation and tracking data collection progress
- CAPI/network enabled operational control requires fully developed network infrastructure
- Business intelligence dashboards
- Strong role for contractor services

# Changing Modes Between Operations

## Paper Maps <-> CAPI Operation Control

- Boundary interpretation and digitization
- Housing points nearly impossible to capture



## CAPI Maps <-> Paper Enumeration

- Edge matching between work areas
- Transfer and printing

## Network Operational Control <-> Paper Maps

- Manage work at level of geography digitized
- Shared resources can be used to track progress in a non-networked environment

# Censuses vs. Surveys

Size of the dataset is main driver of differences!

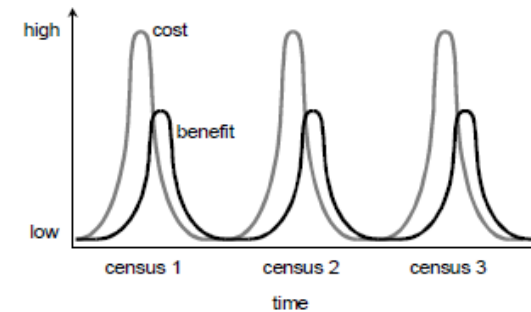
Item	Census	Survey
Records	Millions (respondents) 100s of millions (points)	Thousands
COTS	N/A or Custom	As Is
Accuracy	Positional	Relational/Variable
Integration	Systematic Required	Manual Possible

# Integrated Technology Adoption in Censuses

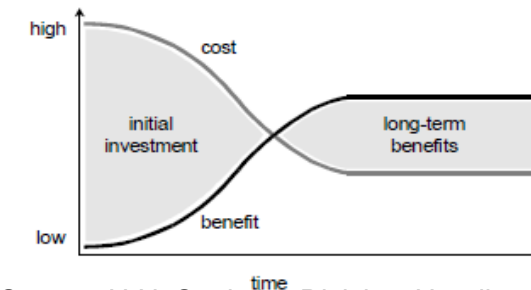
- Strong risk aversion in censuses
- Cost savings harder to quantify
  - High-investment, long return-period model for spatial data digitization
- Does this relationship apply to integrated systems with mapping component?

Figure II.1. Costs and benefits of census mapping options

(a) Traditional mapping approach



(b) Digital mapping approach



Source: U.N. Statistics Division *Handbook on geographic information systems and digital mapping*. New York. 2000



# Role for Technical Assistance

- Outsourcing software development and network infrastructure
  - Not part of NSO core mission
  - Requires surge in capacity and skills
  - Is there still a need for technical assistance in census and survey mapping?

YES!

# Technical Assistance – What We Are Doing

- Introductory Desktop GIS
- Enumeration Area Digitization and Creation
- Pre-Census Spatial Data Management
- Basics of Remote Sensing
- Enterprise Data Editing and Versioning
- Field Map Representation and Interpretation
- Cartographic Dissemination Products

# Mapping Assistance Moving Forward

- Improvement and understanding of day-to-day work
  - New staff and functionality
- Data management and maintenance
  - Schemas and harmonization
  - QA/QC
- Automation of repetitive tasks
- Transition to enterprise data management
- Accurate and useful representation of data in both pre- and post-census maps

# Questions?

# Thank you!

Claudia O Espinosa

[claudia.o.espinosa.villegas@census.gov](mailto:claudia.o.espinosa.villegas@census.gov)