

2020 Census Program Management Review
Census Bureau Headquarters – Auditorium
Agenda – July 10, 2015
9 a.m. – 12:30 p.m. EDT

9:00 a.m. **Welcome and High–Level Program Updates** – *Shirin Ahmed, Acting Assistant Director for Decennial Census Programs*

- **Decennial Directorate Reorganization Update**
- **New Staff Introductions**

9:15 a.m. **Program Management Updates**

- **General Updates** – *Deirdre Bishop, Chief, Decennial Census Management Division*
 - **Areas of Innovation**
 - **Budget, Schedule and Risk Update**
- **2020 Census Operational Plan Status** – *Ann Wittenauer, Research and Methodology Directorate*

9:45 a.m. **2015 Testing Activities**

- **Address Validation Test Results** – *Evan Moffett, Decennial Census Management Division; Michael Ratcliffe, Geography Division; Pat Cantwell, Decennial Statistical Studies Division*

10:30 a.m. **BREAK**

10:45 a.m. **2015 Testing Activities (Continued)**

- **Optimizing Self-Response Preliminary Findings** – *Jessica Graber, Decennial Census Management Division; Michael Bentley, Decennial Statistical Studies Division; Francis McPhillips, Decennial Census Management Division*
- **2015 Census Test Preliminary Findings** – *Maryann Chapin, Decennial Census Management Division; Tom Mule, Decennial Statistical Studies Division*

12:00 p.m. **2016 Testing Activities** – *Deirdre Bishop, Chief, Decennial Census Management Division*

12:30 p.m. **Wrap–up** – *Deirdre Bishop, Chief, Decennial Census Management Division*

Deirdre Bishop: Good morning everyone. I'd like to ask that everyone take their seats. It's 9 o'clock and we'd like to start on time this morning. Thank you. Well again good morning nice to see everyone. Before we do our welcome I would like to welcome a new member of our team. I'd like to introduce this morning Shirin Ahmed. She's our acting Assistant Director for Decennial Programs.

Many of you may recognize Shirin. She's been with the Census Bureau for many years. She was the Assistant Director previously for Economic Programs providing leadership over the five year Economic Census, and the Census of Governments and more than 70 ongoing surveys that measure and profile the US economy.

We are especially delighted that Shirin is part of our team. She brings a level of experience a wealth of knowledge that we can really benefit moving forward as we progress towards the decennial. So with that I'd like to turn it over to Shirin now.

Shirin Ahmed: Good morning everyone. It's an honor for me to be here today and again to be part of the decennial team and very much a part of this important mission for 2020.

Lisa is on business travel today and is sorry to miss the PMR but she's here in spirit. And so one of the things that she does that I'll do right now is go over a few housekeeping items before we get started.

First of all this is our first half day format. And we're also in the new - in the auditorium to better accommodate all of the attendees. And in our quest to provide greater access to information about our progress on 2020 we are trying to reach a wider audience of stakeholders.

However this does require a few important reminders. The meeting is being broadcast internally on ETV and also externally on the Web. So we need to consider all microphones live at all times when having sidebar conversations.

We also have legal requirements we need to follow when talking with contractors or potential contractors. Mike Palensky, Chief of our Acquisition Division will say to remind staff of the legal obligations when speaking with contractors.

Mike Palensky: Thanks Shirin. So this is a public meeting. We're being broadcast. Everything we talk about here is public information.

So I'd just ask the sidebar conversations at breaks there's contractors here, there's government employees here there's contractors watching us on line they don't have the same advantage being here such just keep your conversations to what's publicly available if people are asking you questions.

So just be careful about what we talk about and not talk outside of that. We're not talking about solutions from them, we're not talking about what about this one about that it's really ask questions.

You can answer questions that are publicly available information. The things that we're talking about today are publicly available. So anything released to the public we can talk about. But anything that's not released to the public then we keep quiet about that. Okay?

Shirin Ahmed: Okay thanks Mike. Okay also remember that when you speak directly on to the microphones to turn on the red button when speaking and then to turn it off when not speaking. Apparently we can only have four mics on at the same time.

Restrooms are out here about 50 feet up on the right before you get to the elevators. And then emergency exits are also right here and then down over there but in an emergency we'll try to direct you out of here.

So with that we'll actually get started with our PMR. Oh, one last thing questions. Q&A is encouraged but due to time limitations we have an email address up there that you can use to send any feedback to us.

All right now we're ready to get started. So today's agenda we're going to give you an update on the decennial staffing.

We're also going to talk a little bit about where we are with the 2020 Census Program. We'll also give you an update on the status of the 2020 Census operational plan.

And then we're going to review the status of the 2015 Census Testing activities for address validation, Optimizing Self-Response test and for the 2015 Census Test. And then we'll end with an update on the 2016 Census Testing activities. So we're very excited that our reorganization did happen on May 31, 2015. Even though it happened we still have quite a bit of work to do.

We're using the next few months to transition the work even though everyone is operating in their new roles and in their new structures we're working to transition the work from the old structure to the new structure.

We're also conducting training below the management level on the roles and responsibilities under the new structure. And then lastly since the last PMR we've launched a team to fill the vacancies that we have quickly and efficiently.

Our goal for the decennial census directorate is very much aligned to the 2020 Census goal to improve the way we do the work so that we do it with high quality and much more effectively.

In terms of additions to our team - next we wanted to welcome Deb Stempowski. I see her in the back. Maybe Deb can just wave her hand. Is she back there yes?

Deb is the new Chief of the American Community Survey Office. And we're lucky to have Deb. Deb previously he was chief of our Economic Management Division in the Economic Programs Directorate.

We're also delighted to have Liz Grieco. She's our new Senior Advisor. She's waving her hand in the back there for administrative records and data linkage. She comes to us from the Demographic Surveys Directorate.

We have a great team of senior leaders and office chief's right here to my left is Deirdre Bishop, Chief of our Decennial Census Management Division.

Tim Trainor in the back there Chief of our Geography Division, Pat Cantwell, Chief of the Decennial Statistical Studies division you'll be hearing from Pat later.

Deidra Hicks, Chief of our Decennial Program Management Office she's also in the back there. Tasha Boone acting Chief of the Decennial Communications and Budget Office and Tasha I know is somewhere around here.

We are interviewing and filling positions. We are close on the Redistricting and Voting Rights Data Office the Decennial IT Division. We're interviewing for the Translation Office Chief and Chief of the Decennial Directorate Support Services Office.

And it would be – I'd be remiss if I didn't mention that we have wonderful staff supporting and doing all the work of the office - of the mission of the offices and divisions. And with that I turn it now to Deirdre Bishop to get us started.

Deidre Bishop: Thank you. One of the nice things you can see about Shirin is that she brings a wonderful sense of calm to the picture. And she really makes everything look so easy. So she's been a great mentor to me since she's been you. Thank you.

I see many familiar faces in the room. And I'm thankful for that. I know it's a busy time. I hope you're all having in a summer. Before I get started I would like to give a little bit of a refresher in terms of where we are with our 2020 Census planning.

Chuck I'm good for the next slide. Thank you. For all but one of our tests that we're conducting this year fiscal year '15 we're winding down our efforts.

We're preparing to baseline our preliminary design decisions related to the 2020 Census by the end of this fiscal year which will include a lifecycle cost estimate for the 2020 Census reflecting our major design decisions.

This is really the point in the decade where we must make the choices that are going to set our major operations for the census in motion.

Our program has spent the past few years doing extensive research, testing and design. We spent the better part of this year narrowing down our efforts to leading us to being able to make those preliminary design decisions.

You can see on the graphic that soon we're going to be moving from our small scale individual tests to refined individual tests. We're going to be moving from proof of concept and prototype IT systems to the actual building of our IT systems and procedures.

Let's talk a little bit about where we are today. I'm not going to go into too many details in the interest of time because we have a lot of speakers and as you heard earlier we've moved to the half day format based on a lot of feedback that we received from our attendees.

But I am happy to report that we have completed the Address Validation Test. You'll recall that this contained two components the Mass Model Validation Test and the Address Validation Test. You'll hear about this from Evan, Pat and Mike coming up soon.

We've nearly completed all of the operations associated with the Optimizing Self-Response Test. That's the test that we conducted in the Savanna, Georgia area.

We have one re-interview operation to go there. And you'll hear from – about that test from Jessica, Mike and Frank.

We've nearly completed our effort related to the 2015 Census Test in Maricopa County, Arizona. And we have one of valuation follow-up effort to go there. You'll hear about her efforts in relation to that test from Maryann and Tom.

I'd just like to take a pause here and thank everyone who supported these tests. We had both thanks to those here in person and to those watching on ETV.

We had many successful partnership and outreach efforts in our Savanna Test. We had many people especially those here in the room go out to Maricopa County, Arizona and observe our field operations and how we're making significant changes to those.

We had members also go out and observe our operations at the Denver Area Operations Support Center in terms of how we're operationalizing the management of those field operations.

You shared your feedback with us and we really appreciate that. We consider those very good lessons learned. And we're applying those to our planning for '16 and the years beyond.

I would be remiss also at this point not to say a huge thanks to the people at this table and to the people in this room and out in the Census Bureau building and in the field. They really committed themselves to the success of these tests. So thank you.

In June we received OMB clearance to move forward with the 2015 National Content Test. That means that we're good to go with the September 1 date for that nationwide test of 1.2 million housing units.

On June 30 we announced the locations of our 2016 Census Test. If you haven't heard yet I'm happy to report that we'll be conducting the test in portions of Los Angeles

County, California and Houston, Texas. And I'll talk about that at the end of our program today.

And finally we've made significant progress in the development of our 2020 Census operational plan. As I mentioned earlier that'll be the preliminary design for the 2020 Census. And (Erin) as she has at previous PMRs will share our progress in those activities.

Now I'd like to focus on our major innovation areas. These were the cost drivers of the last census and areas that we've been focusing most of our research on during our research and testing program this decade.

As I have in previous PMRs I'll just share the goal of each of those major innovation areas with you as a reminder and for those that may be new to the audience into the program.

Our first innovation area is reengineering address canvassing. And the goal here is to reduce the need for a nationwide in field address canvassing walking around to every single census block adding new addresses to our list and making updates to those that have changed.

Instead we plan to use things like aerial imagery, administrative data from our government partners at the state and local level, commercial data from vendors out there who are collecting address and spatial data on a daily basis.

We'll be looking at that information instead conducting a lot of the work that we would have done in the field in the office and doing only a small portion of the work in field address canvassing. So what did we accomplish since the last PMR? We've refined our business process model for in office address canvassing.

We've continued to develop the systems that will be used to support that in office address canvassing. And we've completed the summary of responses for our change detection request for information. We talked a little bit about that at our last PMR.

Looking ahead we will issue the Address Validation Test report. I know many of you are anxious to see that. It is complete. And it's actually with me for a final review and then we'll be pushing it up our management chain so hold tight it'll be really soon.

We also plan to complete in the next few months our plan for address canvassing how we're going to implement that operation in the office, in the field, through our mass coverage study and our quality assurance operations that accompany that.

We'll release the change detection request for proposal. We'll award a contract for national address, geospatial and imagery data sets. And finally we'll design our system requirements for the 2016 address canvassing tests.

The goal of our second innovation area Optimizing Self-Response is to communicate the importance of the 2020 Census to the United States population and generate the largest possible self-response including allowing respondents to respond to the census without a unique census ID. This is what we call non-ID processing.

We're focused on this area because the more we can reduce the Nonresponse Follow-up workload the less expensive the census will be. Nonresponse Follow-up is traditionally our most expensive operation that we conduct in association with the decennial census.

So what have we accomplished here? In terms of our Optimizing Self-Response test in Savannah, Census Day was April 1. We sent the final reminder with paper questionnaire booklets to people between April 7 and April 14. We sent a postcard invitation to previously non-sampled households.

This was a late design change to our program. We wanted to see how people would respond after being exposed to the advertising campaign if they received one postcard invitation. And we did receive a very positive response over 8%. Jessica will talk about that a little later.

We ended official data collection for this part - for this test on May 31. And we closed our data collection system on June 23 that is the Internet self-response application.

As I mentioned we are preparing for our 2015 National Content Test. We've posted the final solicitation for comment in the Federal Register notice on May 22.

We received OMB clearance on June 26. We finalized our paper questionnaire and are other mailing material. All print materials were sent to our vendor for production.

We conducted a forum on ethnic groups from the Middle East and North African communities. We had 27 attendees in over 32 experts commented on our plans. And we completed sample design for this survey so you can see we're busy in that regard.

Looking forward in relation to the Optimizing Self-Response test I mentioned earlier that we have one operation to go.

That re-interview data collection will begin on July 15 and end on August 7. And in relation to the National Content Test we'll begin the self-response data collection on August 24 with census day being officially on September 1 of this year.

The goal of our - using administrative records innovation area is to use administrative data such as federal, state and local government data and third party data that is from commercial sources to reduce the Nonresponse Follow-up workload. The data sources may also be used to enumerate the population in cases of nonresponse.

What have we accomplished over the past few months since our last PMR? In support of the 2015 Census Test we received monthly deliveries from the IRS informing us of the federal tax returns.

We received data from the United States Postal Service about the undeliverable as addressed housing units that were delivered as part of our 2015 Census Test.

This was in support of the identification of vacant housing units. If we identify a housing unit as vacant we can remove that household from the Nonresponse Follow-up workload.

We implemented the administrative records identification of occupied housing units. And we completed data collection for the 2015 Census Test portion nonresponse follow-up.

Looking ahead we'll continue to baseline our report on admin records used in the 2014 Census Test. We'll analyze the data collected in '15 and begin our reporting on that.

We're going to present our findings at an upcoming meeting of the Joint Statistical Meetings in August.

And another addition to the program we're beginning to work very closely with the states of acquisition of SNAP, WIC, and TANF data to help us with the administrative records inventory.

Our fourth innovation area is reengineering field operations. This is one of the most significant innovation areas that we're studying as we look towards 2020.

The goal here is to use technology to more efficiently and effectively manage the 2020 Census fieldwork. How can we replace what humans used to do with technology to help us get our work done better and faster?

We've had a lot going on as you can tell over the past few months. We deployed an enhanced operational control system to help us manage the caseload of Nonresponse Follow-up cases in that Maricopa County test.

We used automated training instead of extensive in person training and verbatim of training to instruct our local supervisors of operations and our enumerators.

We conducted interviews with our non-responding households. Not using paper and pen like in the old days but instead using a handheld smartphone with an application to enumerate and conduct the interviews.

We monitored our operational progress in real time. We created daily optimized assignments through the use of these control systems. And finally looking toward the 2016 Test we released our request for information about how we can possibly use device as a service?

How can we partner with commercial entities in terms of logistics with our smartphones and the technology we use to manage the census?

As I mentioned looking ahead we'll begin our data analysis for the 2015 Census Test. We'll incorporate all the lessons we've learned not only into the functionality of our systems our operational control system and the enumeration instrument but also into the lessons learned how do you manage your staff? How do you conduct the interviews?

We're learning a lot. And again thank you for those of you that went out and provided feedback.

And I would be remiss this bullet is not on a slide but beginning now we're really going to put a lot of effort into planning and implementing those 2016 tests in Los Angeles and Houston.

With that I would now like to turn it over to Ann Wittenauer to talk about the status of our 2020 Census operational plan.

Ann Wittenauer: Passing of the baton early in the morning. Good morning, as Deidra said I'm here to give you a brief update on the status of the 2020 Census operational plan, which I'm happy to say, is on track for release October 1 of this year.

Today you're going to hear how we have refined our work iterating after many rounds of review to provide not only a solid set of documentation but more importantly a program baseline that lays the foundation for a successful 2020 Census. So let's get started.

Next slide please. Many of you have seen this slide so I'll very quickly provide some context for today's update on our work for those who are joining for the first time or who might need a refresher.

This Operational Plan or the first release of it bridges two critical phases in the 2020 Census lifecycle.

The formal research and test phase which ends September 30 of this year and the operational research and design build phases that commenced later this year. And bring a rapid uptick in our March 1, 2020 Census production.

Most of the early research and test work culminates this year with the preliminary design decision, which is the centerpiece of our operational plan.

The goal as most of you know for our preliminary design decision is to conduct the 2020 Census at a lower cost per housing units than the 2010 Census when adjusted for inflation while maintaining high quality results.

The preliminary design thus will focus on the four key areas of innovation that Deirdre just summarized so nicely for us because they have the potential for the biggest cost savings upwards of \$5 billion if they are planned and implemented correctly.

Now with the first version of the 2020 Census Operational Plan coming out three years earlier than it did last decade and five years earlier than it did for the 20 or excuse me the 2000 Census there are things that we are not going to know when we release this first version of the operational plan

So the documentation in October is going to include design issues, additional operational research and testing activities that are still needed in order to mature our design for the census over the next few years.

And thus as you would expect there will be successive releases of the Operational Plan and related documentation as we learn more.

The final version as a result will not come out until 2021. And it will document what we actually did to implement the census.

So besides laying out the preliminary design decision which we're going to refine through further testing the other purpose of our documentation and the operational plans is to lay a foundation for the work needed in the rest of the census lifecycle.

Let's go to the next slide. One last quick refresher you've heard me say the Operational Plan is a set of documents. This slide depicts at a high level what I mean by a set.

The first bucket of work consists of the concept of operations or com ops. It's the core of the operational plan.

Within the com ops we're going to describe how we will execute the census meaning we will include our plans for implementing the preliminary design decision, the four

key areas of innovation and also each of the 34 proposed operational areas that we envision for the 2020 Census. And I'll provide a lot more about those operations in the next slide.

The second bucket consists of something we call supporting documentation. And this set has been expanding over the past few months.

And for example what I mean by that is initially we were focused on exclusively and rightly so on the revised lifecycle cost estimate as a core piece of this documentation and also the IT architecture also very critical.

More recently we have decided to include still more documentation such as a crosswalk for recommendations from many of you in our oversight community.

We're also in planning to include a continuity of operations plan to name just a few.

The third bucket that we think of in our set of documentation consists of a communication tools these things that will be useful for us to convey to our stakeholder community how we're going to implement and conduct the next decennial census.

By this I'm specifically thinking and referring to Infographics. And I'm happy to say that we are working on a draft one in particular called the 2020 Census a New Design for the 21st Century.

And eventually we'll probably have videos all of which are designed to help us convey what it is we're doing with the 2020 Census.

So that's our set. Let's go on to the next slide. Now this slide represents the meat of this morning's update in terms of our progress over the last couple months.

There are three things that are important about this chart. First it shows - it depicts the current set of operations. And I'm going to tell you in a minute how this has changed since our last PMR.

Secondly, this chart depicts the status of each operation. And this is important so that you have an idea of what to expect that is the level of maturity for each operation when our documentation is released in October.

And lastly I'm going to touch upon the alignment of our program and excuse me the enterprise work breakdown structure.

So diving deeper the current list of operations. Well there are still 34 for those of you who remember the last PMR but things have evolved as our planning has matured.

We've been quite busy with internal review sessions. The word extensive comes to mind. And I'm sure Deidre, and Pat and others would agree since we spent the better part of three work weeks in the last couple months holed up with our subject matter experts and leadership to refine and mature the documentation. We call the sessions 'lock ups' so you know it's very serious.

Let me step you through the changes. First there are still nine support areas but we've undergoing quite a bit of refining here. All of this is reflected in the top row. First we have a new operational area systems engineering and integration.

And by this I mean all of the critical work done regarding Requirements Management, Quality Management, Test Management and so forth.

Secondly, you'll see we have merged Strategic Communications and Program Management into one support area called Program Management which as you would expect includes all of the vital functions such as Budget Management, Schedule Management governance and so forth.

We've also decided in the third change that Legal and Policy is better covered within each operational area. It's no longer considered its own support or operation area.

Lastly, in this top row you'll see that we split field infrastructure into two. Now it's Field Infrastructure and Decennial Logistics Management the latter of which will enable us to better plan and oversee the work relative to warehousing, kit assembly, inventory management and so forth.

Now moving on to the other changes again there are 25 operations represented in the bottom two rows but we did make a change in the response data section.

We have split what was called Paper Processing into two operations – now Forms Printing and Distribution and Paper Data Capture.

Another change in the bottom row is reflected by the operation called Evaluations and Experiments. As a result of our extensive reviews we have decided to combine what was called 2030 Research and Planning and Evaluations.

So with this picture you get the latest snapshot of the operations and support areas envisioned for the 2020 Census .

Now for the second point that I want to convey the color coding in each operation depicts the status. Let's talk about that briefly.

Notice a legend in the bottom half of this picture where it says detailed planning has begun, detailed planning is underway detailed planning has not yet started.

We've color coded or operation so that she can get a visual of what types of information or the level of information and maturity you'll be reading about in October.

For 13 of the operations detailed planning has either not yet started or has only recently begun. As a result the Operational Plan is going to have less information about these areas.

The content will be leaner in other words. And by that I mean our documentation is going to include maybe a handful of things such as the purpose, our lessons learned since the last census, key elements of the operational design, any decisions that have been made to date, design issues that still need to be resolved and cost and quality impacts.

On the other hand for the remaining 21 operational and support areas work is underway as evidenced by the dark green shading.

These 21 areas will have significantly more information that we can convey. So besides the purpose and lessons learned we're going to be calling out innovations since the last 2010 - excuse me since the 2010 Census. Cost impacts compared to the 2010 Census, key risks key milestones to name just a few.

Lastly the third point I want to draw to your attention and looking at this slide it is subtle but very important and this is depicted through the background coloring.

And we're using that to show our alignment of the 2020 Census program with the enterprise work breakdown structure.

The enterprise work breakdown structure has eight elements and we have no organized our work so it is consistent with six areas of - with these six areas excuse me that apply to our program.

And the top yellow row you see three of the enterprise work breakdown - tongue twister. Work breakdown structure elements program management, census survey engineering and infrastructure.

Similarly in the middle row there are three more elements of enterprise work breakdown structure in which we have categorized our work frame, response data and published data.

This alignment as you all know is fundamental for good program management in order for us to successfully plan, monitor and control our cost schedule and scope.

So next slide, back in April I introduced a new task to you that our team had undertaken decision analysis and support work.

This effort consists of cataloging and prioritizing the decisions and assumptions that our subject matter experts and our leadership have made to date as well as the decisions that are still needed.

This information then gets documented formally and escalated to the - through the appropriate governance process and managed by change control.

The work had just commenced when we had our last program management review on April 8. And I'm happy to share with you today on the next slide that we have made good progress in this work.

Next please. This slide lists the eighth operations whose decisions we have reviewed and documented.

And you'll see that for two of the eight as I mentioned a minute ago they have been split into two operations so at this point ten of our 34 operations have been through this analysis process.

We've prioritized our work here on the operations that are either early, that are big, or that are forming the requirements needed later this year for example for systems development work or for large acquisitions.

We're going to continue this work through the rest of the summer using this prioritization I mentioned just a minute ago.

Next slide Quality. You know our challenge continue to deliver high quality results for the 2020 Census but contain costs such that they are lower per housing unit when adjusted for inflation than the 2010 Census.

So another critical component of our work these past few months has been zeroing in on the most appropriate quality metrics. We have rightly focused excuse me we have rightly fixed our focus on the key areas of innovation given that they have the greatest potential for cost savings.

And here this morning I'll very briefly mention a few of the quality metrics that we're proposing. For address canvassing we're looking at the number of ads captured and the number of deletes captured.

For Self-Response just a couple examples the projected rates of Self-Response for the United States and by demographic groups. Also we're looking at estimated number of correct enumerations.

And for the biggie Nonresponse Follow-up we're looking at the estimates of total for the US but demographic groups and so forth as compared to the 2010 Census.

We're looking at counts of housing units both occupied and vacant. And the percent of cases resolved by administrative records or by proxy.

Next slide, and lastly let me briefly talk about the very critical work related to our revised lifecycle cost estimate a key deliverable within the Operational Plan documentation.

As you would expect we are in fact conducting what is tantamount to bottom-up estimates for each operation.

But the point of this slide is to reiterate the reasoning behind our work this spring with the cost model.

We have focused our efforts on what I like to call the big rocks in the jar when it comes to cost savings.

The greatest emphasis has been placed on these areas that most likely will account for the majority of our estimated savings the key areas of innovation given their combined potential for savings of just over \$5 billion.

Next slide, in that vein we have been working to encourage excuse me to ensure linkage between the cost estimates for the operations and the systems that will enable them and focusing on defining workload parameters by work breakdown structure to inform our cost estimates.

Next slide, so in conclusion and to wrap up this very brief update on the Operational Plan we are meeting our internal deadlines, our documentation has progressed through extensive reviews.

We're developing quality metrics and workload parameters. And we are aligning our work to the work breakdown structure.

In short, we are on target for the planned release in October 2015 and in just a few months at our next Program Management Review. It's full steam ahead for the operational plan. Are there any questions?

Ty: Am I live? Thanks for all this information. Real quick couple real high level, I don't see scheduling listed in this kind of operational thing with - could you just give us a quick comment on the status of that particularly as it leads towards doing some sort of a scheduled risk analysis at some point and I mean that's resource loaded and all that kind of stuff? Oh go ahead I'll ask a question after.

Deidre Bishop: Sure. In conjunction with the development of the 34 operations and with the lifecycle budget cost estimates we're working very hard to flush out the schedule for the 2020 Census .

We have an integrated master schedule as we've mentioned before. We are moving toward the enterprise tools that we use to manage our resource loaded schedules. So I know you received copies of our schedules on a monthly basis.

Just as I feel the Operational Plan is moving along nicely I'm happy with the progress in terms of their lifecycle cost estimates of our budget and our scheduled as we look first 2020.

Ty: Do you have a timeframe in mind for when these will be resource loaded?

Deidre Bishop: Yes. We've worked with our ORMPE the Office of Risk Management and Program Evaluation. And by the end of this year we will be resource loaded for a subset of the operations related to the 2020 program.

Ty: Thank you. Look for that one when it's coming. A couple more just - I'm trying to see it at a high level.

On the administrative records getting this stuff from the states SNAP, the TANF in particular do you have some milestones in place some targets so that you have a sense that you're on track, you're making - you can manage that?

Deidre Bishop: As you know over the past few years we've been operating with limited resources. We have focused our research related to the use of administrative records more significantly on national level files.

With that said as we approach the Operational Plan milestone by the end of this year we believe we need to know the core set of records that we will use in association with

the identification of vacant housing units and the occupied housing units to help with enumeration.

With that said we're working very closely with CARRA our Center for Administrative Records and Research Application on the acquisition of these state level files. And we do have many in house.

As we look towards the Census Tests in 2016 in Los Angeles and Houston we are trying to acquire state level files for those test areas to see how we can apply that late - that data to our research.

Ty: Okay good.

Trisha: Trisha with the OIG. You mentioned the quality metrics for ad can as being number of ads and number of deletes and is there a goal other than just a straight number?

Deidre Bishop: I think you'll learn a lot more about this during the presentation from Evan, Pat and Mike. So Trisha if you'd be willing to hold that question I think you'll get your answer very shortly.

Trisha: Okay sure okay.

Ty: I have one more if there's no one else. Just really quick at high level you mentioned the governance process the decision analysis.

How would somebody not in the middle of that governance chain know which of those operations have completed what you described having happen for ten of the 34?

Deidre Bishop: So let me see if this helps. For each of the operations we have a team of people who are responsible for carrying out the activities associated with those operations.

They have been involved with the development of the operational plan. They've contributed to the lifecycle cost estimates and to the development of our schedule.

We have had many sessions in which the entire management team associated with the decennial census so not only decennial census management division but geography division decennial statistical studies division.

Sometimes members of the IT directorate and always members of the 2020 Census Operational Plan Team hear what's happening in relation to the development of this plan. We then - once we complete each iteration we share that with Shirin, with Lisa, with Nancy, and with John...

Ann Wittenauer: And Brian.

Deidre Bishop: ...and Brian of course for our IT work.

Ty: Just some kind of a wrap up document or something there's a milestone it's like okay now it's stamped it's done we know it's on this list so that folks outside the bureau if they wanted to take a look you know I mean I'm just trying to find how we would know which ones and what we'd be looking at?

Deidre Bishop: Yes. So as we've talked about in previous PMRs. And as Ann mentioned today we are on track to baseline the Operational Plan by the end of this fiscal year.

We started out developing a narrative associated with the operational plan. And very quickly it became obvious that this narrative was going to be over 300 pages. So how were we going to use that as a tool to communicate information?

We decided that in addition to the narrative we would develop what we call a slide deck library. So for each operation we will have the lessons that we learned from 2010, the key innovations that we're applying for 2020, the key decisions that we've made to date through our work not only on the development of the Operational Plan

but what we've learned from the 2013, 2014, 2015 tests, key decisions that we still have to make, the risks that we're knowledgeable, you know, that we're planning for, the cost and quality impacts and then a schedule a timeline for what's going to occur next. That will all be included in part of this slide deck library for each operation. And I think that will be more easily shared as we move forward.

Ty: Very helpful, thank you.

Man: Real quickly. Once we get the initial design documents and then you move to future iterations how are you going to document changes in those updates to the, you know, baseline decisions, the cost, the schedule is that going to be a major component of sort of the second, third, fourth and so on for people when they're reading these?

Deidre Bishop: We are start - we have been talking about that but we're really starting to focus on that now. So with each iteration - and we expect that there will be probably at least two iterations per fiscal year in terms of the Operational Plan because there are things that we'll be changing and that we'll be making adjustments too.

We've also talked about adding different components to the Operational Plan to be in line with good project management practices as appendices so we'll be feeding those into the Operational Plan next year.

We have a rigorous change control process within the decennial census management – process. And we'll be applying those same rigorous requirements to changes to the operational plan.

Okay well we're right on schedule. So I'd like to ask us now to move to the presentation about the Address Validation Test. Before Evan gets started I would like to announce that we've asked the team to adjust the air conditioning.

They made that adjustment about half an hour ago. So very shortly I hope that you'll see a change in terms of the coolness and the humidity in the room so be patient.
Thank you.

Evan Moffett: Okay good morning. My name is Evan Moffitt, Mike Radcliffe, and Pat Cantwell and I will be sharing with you the results of the Address Validation Test.

So what you're going to hear today is some general background on the Address Validation Test as well as results of the statistical modeling component which is the Math Model Validation Test and results of the partial block canvass portion of this as well.

The purpose of the Address Validation Test is to evaluate our methods for reengineering address canvassing.

In addition we are testing how well are in office procedures can replace in field procedures and to assess our ability to ensure an accurate Master Address File moving forward.

As you've heard in previous PMRs there's a couple of components to the Address Validation Test the first of which is the Math Model Validation Test the MMVT.

We collected data for this portion of AVT from September to December 2014. We use this to assess statistical models which Pat is going to talk about in just a minute.

This is a nationally representative sample and we used our full block canvassing methodology which we have used traditionally in the census. We used the legacy piece of software to collect this data and that was the Automated Listing a Mapping Instrument.

The second component of the AVT or the Partial Block Canvassing Test was conducted between December of '14 and February of '15. The purpose of this was to

test the ability to canvas partial blocks as opposed to full blocks which was done in the Math Model Validation Test.

These blocks were identified through the use of imagery. And we are planning to apply this technology as we move forward with the address canvassing program for 2020.

From a data collection perspective we used our next generation Listing and Mapping Instrument to collect this data. At this point I'll turn it over to pat.

Pat Cantrell: Thank you Evan, good morning. I'm going to give you just a very brief background on the Math Model Validation Test. Many more details are available but we spoke about this at one or two of the prior PMRs.

We took a national sample of 10,100 blocks. Now the 10,000 blocks we took 10,000 blocks a sample from of those with at least one address so in the United States we have about 6.3 million blocks of at least one address. We took a representative sample of 10,000 of those.

We also took a sample of 100 blocks. This was just a convenience sample to do a little further analysis from those with no addresses or models that are for blocks with no addresses are a bit different.

So we're talking about one million addresses in our sample part - in our sample here. And we did a full block dependent canvassing.

So by dependent we mean that the listers started with a list an extract that we took and they would go to all these addresses, verify the addresses as they found them, update anything that they found out there including adding addresses that weren't on the lists that they were given delete addresses that shouldn't be there like that.

So the two main research questions or can we use statistical models to determine specific blocks that need additional action?

Now this could be various types of action but mostly we put these together into two types in office our in field canvassing.

And the second can we use these models to predict the coverage of the math, the quality of the math, otherwise how many ads, how many deletes or other errors are on the math?

But before we give you any numbers let me just summarize briefly the statistical models that we applied in this test were not effective in two major areas first identifying specific blocks with many ads or deletes and second predicting national totals of math coverage errors otherwise the ads or deletes in our frame.

Okay so let's look at some numbers. I want to just give a little back - a little information on this slide so that you understand it clearly.

And let me start with the left - the first column of numbers. So these are just totals. This deals with what's in the universe and what we conducted in the field otherwise it has nothing to do with the models themselves.

So we took our universe and we had about 10,000 - we had 10,000 blocks in the universe. These are the blocks with at least one address.

Weighed it up you can see that they have about – there are about 6.3 million blocks if using the sample weights.

And this gives us about 136.3 million addresses in our frame. As I said we're talking about blocks with at least one address.

As we've mentioned before we also did not go to Alaska, Hawaii and some of the federal lands. We also didn't include Puerto Rico in this test.

And we went to each of these blocks in the samples so we did the fieldwork on all of them so across all of the 10,000 blocks in sample we found weighted about 5.7 million ads so that's the weighted number there and about 7.6 million deletes to our fieldwork.

So let's look at the models themselves or let's look at the numbers that we're giving you for them. We have - I have here models one, two, three and four these are four specific different types of statistical approaches in our models.

And we've talked about these I think in two PMRs before so I won't go into details about them. There's a little information in one of the slides at the back of your package. And again if you're interested we can answer any questions.

But we want to compare the performance of these models and trying to say which blocks need work right now or need some kind of action right now whether that's in office canvassing in field canvassing?

So what they did is it took the list of blocks that we had in the sample and based upon the models they prioritized the blocks from those most needing some kind of action all the way down to those least needing action.

Now to make it some - a fair comparison we did a 20% canvas where the 20% represents the number of addresses.

As I mentioned before we're going to all of these addresses in the field but the idea is what if we could - what if we decided only to go to a certain number of them in our canvas?

So for example looking at model one and went down the list of blocks that are most need of canvassing and when we down to a 47.3% of the blocks it had it's 20% of the housing units in our frame.

In models two, three and four did prioritize the blocks differently according to those models. You'll notice for instance model three only selected 4.4% of the blocks was canvassed even though it obtained 20% of the housing units in a frame.

And of course I'm talking with weighted numbers here. So it actually went - model three selects large blocks in its work.

So let's look at the results down at the bottom two rows. For ads what's the rate of capture of the ads in the universe based upon this 20% canvas?

By the way I mentioned 20% there's nothing special about 20%. We could have done other - we used other numbers. In our report we have results for 20%, 25% and 40% and the results are analogous.

So looking at the rate of capture of ads we can see that models one, two and four did just about performed about equally 47% capture the ads.

For the deletes there was a more variety here diversity here 34% model one, 53% model two which is not surprising model two is actually built to try to capture deletes even though it did just as well as some of the others capturing ads.

So we're doing better than random selection of housing units or addresses at 20% but it's not as good as we had hopes.

Let me look at this two different ways. I'm going to go to the next slide. On the next slide we look instead of at address level results how about block level results otherwise what are the important blocks to capture?

Obviously if a block has no ads, no deletes, no errors in there we would like not to spend money on that block at the time however if it has 50 ads that's certainly an ad certainly a block that deserves some kind of attention some kind of work.

Well where's the threshold? Just for this exercise we looked at blocks with at least five ads. Also we looked at blocks with at least one ad.

So you can see again down – the first column of numbers again the same number of - this is the same row here number of block 6.3 blocks weighted from the sample.

And in the sample we had 188,000 again a weighted number of blocks with five or more ads. So you can see it's a relatively rare event only about 3% of the blocks did have five or more ads.

And how well did our models do capturing these? Well again the models one, two and four captured about 43% to 47% of these blocks. Again much better than random but not what we would like.

Blocks with at least one ad model one actually got about 54% of them and they varied across there. Now let's look at the other side.

The other side is how about the rate of blocks erroneously canvassed? You go out there and you find nothing out there at least not what you were looking for.

So if we consider blocks of five or more ads if you considered an error if the block didn't have five or more ads then our - all of our - three of our models were up around the 95% error rate. Model three did a bit better there.

If you look at blocks with at least one ad we had error rates around 75% or 67% although only 33% for model three so again some slightly positive results but not what we're looking for.

We look at this one more way before we move on to the second research question is let's look at prediction versus observation.

So here's a prediction outcome matrix. On the left-hand column -- and I admit this is difficult to see -- on the left-hand column we have the predictions.

And this is based upon actually model one which happened to be a zero inflated negative binomial model. And we use the average of the predicted distribution which is just one way you could represent a distribution.

And so for instance the first row -- don't go looking down the left-hand column -- the first row says suppose we are predicting between zero and 1/4 of an ad for a block otherwise we're predicting that there's nothing out there in terms of ads.

And what we've highlighted on the far upper right corner so you can see these are the blocks with five or more ads although we've broken them down across areas what did we observe in the field zero, one, two the yellow represents five or more ads.

So if you're looking at this first row which what we're thinking of as least likely to need some action according to our prediction you can see that we're only missing five blocks of five or more ads a very small proportion of the time half a percent for that first row. And for the second row again only about half a percent.

So these numbers are very small. Now this is also not too surprising considering most blocks don't have any ads.

So but what we have in the first four rows which represented 75% of the blocks these numbers are generally between 1/2 a percent and 1-1/2% where you define five or more adds.

So for the low - for the first so many rows we're predicting very little I happen. Very little is happening. However we don't do so well when we're predicting larger amounts.

When we're predicting large numbers ten or more, ten to 20 ads, 20 to 50, or 50 or more sure we do see some results where we have a lot of ads but we also see a lot of results where we have very few ads.

And I also should mention that the standard errors become much larger on the bottom row - we have them in our report not right here.

But the standard errors become much larger because there are very few blocks which we predicted so many ads.

So this - it says that this could be used as a screening procedure perhaps to help say these are the blocks that we don't want to pay attention to right now. And let's concentrate on the other set of blocks the blocks down towards the bottom. But overall the prediction was - it only worked moderately well.

So the question asked about how well can we do predicting coverage of the math and otherwise errors in the math ads and deletes are what we're focusing on here.

So you can see the first row this is just the first row of numbers is just the sample base estimates. This is just based on the fieldwork.

We have 135.9 million addresses. You might if you happen to remember this is a little different from what I had in a prior slide because the prior slide was a sample base number this actually is the actual tally from the frame. They're about the same.

But the next two rows are estimates from the sample. In the field we estimated 5.7 million ads but what would our model have predicted about 8.6 million, much higher.

The deletes did a little better. We found in the field a weighted number 7.6 million ads but our model predicts about 8.7 million still very high.

Now I will say that this is - the model that based upon data from the 2009 address canvassing and we applied it to data that were available at the time of the test.

So we're using old parameters. We wanted to see how well they would work otherwise based on information that we already had available because this would be the case going into address canvassing for further action, or field action, or whatever like this we can only use the information that's already available across the country or across whatever domain we're talking about.

So just to summarize determining the specific blocks that need additional action the rate of capture was too low. The rate of erroneous canvas was too high.

Using the statistical models to predict national totals you can see that we over predicted based upon the information that we used.

The model parameters reflected the condition on the math back in 2009 in which the data were taken the 2009 address canvassing of the 2010 Census.

I also want to say that right the conditions are different in many other ways not to mention like the economy, and state of the math itself we're only halfway through the decade.

Further there have been many enhancements and updates to the math because of our geographic support system initiative though that has improved the quality of the math and made for fewer ads and deletes on the math. Okay I'm going to turn this over to Mike.

Mike Ratcliffe: Thank you Pat. I'll talk about the Partial Block Canvassing Test, and our results, and a little bit about some of the analysis of the imagery review that formed a portion of the

in office work for Partial Block Canvassing and how that related to MMVT listing results.

So we went into Partial Block Canvassing I'll start with Evan had mentioned Partial Block Canvassing was a test of the new methodology going to only a portion of the block as the name applies rather than the full block.

And I'll just reiterate this was a new methodology or is the new methodology for the Census Bureau that we wanted to test. It involved testing our methods for identifying blocks that would be appropriate for Partial Block Canvassing.

But more importantly testing our instructions to listers and other information that we provided to them to help them navigate in the field, find their work location, define their work locations and then carry out the work.

We've talked about full block canvassing already but in that operation the lister - the canvasser receives a block that is their work area.

They navigate to a starting point on the block. And then they traverse the entirety of the block all the roads internal to the block but they move around the block in a very organized fashion.

With Partial Block Canvassing we're sending them obviously to only a portion of the block. And we need to describe to them how to get to that specific portion and do that in a way that they can find that worksite and carry out their work effectively.

So we went in with two assumptions that Partial Block Canvassing would be more efficient in large land area blocks by avoiding the need to traverse the entirety of the block to collect the changes. I've got an example that I'll show of a couple of large relatively large suburban blocks.

And then in blocks with large numbers of addresses that are in the math and can be validated in the office Partial Block Canvassing would save time by focusing our efforts only on the portion of the block in which changes have occurred and that we've not accounted for in our update operations.

So there were a number of research questions. The first three -- in the interest of time I'm not going to read through all of these -- but the first three relate to the methodology and the materials for carrying out Partial Block Canvassing.

And the last one relates to the fourth one relates to how effective are in office methodology specifically our imagery review is at identifying where housing units have changed.

And so the Math Model Validation Test gave us the opportunity to collect information in the field that we could then use to test and assess how well or how effective our in -- our imagery review was. And I'll have more information on that towards - at the end of the presentation.

As part of our analysis we wanted to look at the results from fieldwork in comparison to expectations based on our in office review. Did we collect the information that we expecting? Did we find additional updates in the field?

And in the -- and then in addition for blocks that were in both the partial block and the full block canvassing all of the Partial Block Canvassing blocks were drawn from within the 10,100 sample for the Math Model Validation Test.

So in those blocks compare the results from both tests and assess the reasons for any differences. Did the full block canvassers find additional updates especially any that might not be detectable through imagery review?

This was a real concern for us because as we're looking at the imagery and you can see where change has occurred. And I have a couple of examples. And we'll see that in just a moment.

But there are certain types of changes to housing units particularly within existing structures where you can't see underneath the roof. In that sort of situation the imagery review might not detect the change to housing units but the listers would find those in the field.

So we wanted to gather information about the extent to that and assess that. And then use that information to help hone our techniques for in office work or look for other sources of information for those kinds of changes.

So I've already mentioned we drew our 615 Partial Block Canvassing blocks from within the 10,100 MMVT blocks.

We had 37 assignment areas. We grouped the blocks into assignment areas, 37 different areas, 35 listers went out into the field.

These are all professionals. Given the timing of the development of the program we really didn't have time to hire staff so we used professional staff from headquarters and the regional offices to carry out the work.

This also gave us an opportunity for the professional staff who are working on programs and developing procedures and methods to actually test them in the field and gain that critical experience.

And Evan has mentioned we used the corporate Listing and Mapping Application. This was the first use of the LMA in a production operation in the field.

So let's look at a couple of examples of imagery. Here we can see a typical suburban block. The older image in both of these instances - well in this case the older image is on the left the newer image is on the right.

You can see in the older image where there are - there's undeveloped land. We can confirm - and then in the newer image you can see that there have been houses built on those parcels on that land.

We can confirm in the office comparing to the Master Address File the number of counts. And then we can confirm the addresses. So we confirm that the houses that you see in the older image are accounted for in the Master Address File.

And then we have new houses that we have to - new addresses that we have to collect and then add to the Master Address File. So here we would send the lister only to the portions of those streets where the changes have actually occurred and where we need to collect new information.

In this image and I'll just mention that this image is also on a poster that we have up in the hallway that discusses - Partial Block Canvassing. So you can look at that poster during the break and at your leisure.

But in this image this several blocks in Anne Arundel County. The blue the block boundaries are highlighted in blue.

You've got the full extent of two blocks or several blocks on the left. And then on the right the older image on top and the newer - the more current image on the bottom.

These two blocks together encompass seven road miles. So if we did a - if we conducted a full block canvas in these blocks the listers the canvassers would have to drive a total of seven miles to conduct or walk a total of seven miles to carry out their work.

We can confirm that for most of the block we already have accounted - or most of each of these blocks we've already accounted for the addresses within the Master Address File.

And the area where change has actually occurred is circled in red. In the left we've got one circle and then you can see the two specific cul-de-sacs where changes have occurred circled in the images on the right.

There's 1100 a little over 1100 houses that are already in - that are in these blocks that are already accounted for in the Master Address File. And the change on these two new cul-de-sacs accounts for 20 additional houses.

So again if we canvassed the entirety of these blocks the canvasser would be validating the existence of validating the records for 1100 housing units and then adding 20. In the partial block approach they go directly to the area where change has occurred. We validated the other addresses in the office.

And they go to directly to the location of the change, collect the information for the 20 new addresses, complete their work and then move on to their next assignment.

So getting to some results our analysis of the Partial Block Canvassing results indicates that we can successfully implement this methodology in the field. All listers navigated to their specific work assignment based on the written descriptions from the in office imagery reviewers.

We do need to conduct additional analysis to determine the specific contexts in which Partial Block Canvassing is most effective and most efficient as well as the level of expertise and experience that's necessary for listers.

You can recall that I said we had 35 professionals. Those were mostly geographers that went out to the field.

We need to test this with our typical listing and canvassing staff to understand if there are specific training, and experience, and levels of expertise that are needed to carry out this type of canvassing more effectively or it's carried out effectively.

With regard to the successful navigation we - well all canvassers all listers did reach their work assignments and they were able to carry off work there were some instances in which instructions from that in office reviewers were confusing or were unclear. And those affected their ability to carry out work effectively and accurately.

So we've identified those kinds of specific situations. And we're looking at how we can better describe the work, better identify the specific worksite so that we - so that it's clear to the lister the exact work that they need to carry out.

And the work area polygons that we defined so in each work assignment we gave them written instructions. You'll see some instructions on the next slide.

But we also drew a polygon around the area where they were supposed to work. We gave them a boundary.

And in some instances the work area polygon that was defined based on the imagery review in the office based on the new development that was visible in imagery did not always match the extent of new development on the ground.

So in other words we were looking at the best imagery the most current imagery we had. We drew the boundary around what we thought was the area of change.

And when the canvasser got to that area there was additional development. The development was larger the extended area was larger than we expected based on the imagery.

They were told to work based off of the written instructions and continue on keep - to continue adding - carrying out their work.

But in some cases they stopped work at the boundary of their work polygon. And that accounted for some of the differences that we saw between the full block canvassing and the Partial Block Canvassing.

Here are some - here are the different instruction types. I – you can read these but just a couple in general.

Most of the work assignments were either whole streets or, you know, canvass the entirety of Johnson Court that's fairly clear, canvas Burke Street from 46th to 50 inclusive of those two addresses.

It's the branching roads instruction that was the one that caused more confusion than others. Where canvas starting at the intersection of Freedom Highway and Redwood Drive canvas Redwood Drive and any roads that branch off of Redwood Drive.

This is the specific situation where the - where we saw the difference between the instruction, between the polygon that was defined based on imagery and then what the lister found actually in the field.

And then where you had multiple roads branching off of that trunk line running into the development. In some cases they only listed the initial branch going off of the main road.

In other - although, you know, as opposed to continuing on to every street that continued off of so you can imagine the tree - all of the branches, and the twigs and so on that continue on.

So those are the areas where we need to refine our instructions, refine our information so that we avoid those - that sort of confusion. All right going through specific results well in the Partial Block Canvassing blocks we had a total of 17,627 actions that were taken.

And the vast majority of those or the majority of those were at switches what we were expected - we were going out to focus on change. And so we expected that most of the action taken in the field would be new addresses collected.

Those broke down as 10,189 added - ads broke down into what we call true ads. These are addresses that were completely new to the math.

So Pat mentioned going out into the field with a list just like MMVT the Partial Block Canvassing went out with a dependent list a list of addresses loaded onto the LMA.

So 4300 of the addresses that were added were completely new to the math, 2900, 2,931 of those addresses were in the Master Address File but had not been geocoded.

So we collected information that allowed us to geocode to the block 2900 addresses. So that's good. That's making the math more accurate improving the quality of the Master Address File.

And then 2957 of those ads were ads that matched to the math but there were other kinds of addresses. So there are filtering rules that we applied to decide what types of addresses make it onto the list.

Some of these - most of these addresses were addresses that did not make it onto the list but were already in the Master Address File.

So we need to look at how we I - how we structure how we identify what is a good address to go onto a list for canvassing. So we'll be looking at those 2900 addresses to see what can we learn about their characteristics?

We had a number of deletes, duplicates, changes only two changed to nonresidential and so on. And we verified 7000 addresses. And again those were mostly the addresses that were bounding the work assignments.

Compared to the Math Model Validation Test we see similar roughly similar numbers of action or of ads acquired.

We do have more ads in Partial Block Canvassing than we had in the Math Model Validation Test. We need to look at those – the differences more in more detail but Partial Block Canvassing went to field later then the Math Model Validation Test.

In some cases the partial block canvassers were in the field up to in their block up to three months after the Math Model Validation Test lister had visited that same block.

And there were blocks I was in a number of blocks where there was active construction occurring. So we expect that a number of the - some of the difference between the number of ads in Partial Block Canvassing compared to MMVT is simply because we were in the field later. And we're picking up housing units that were built in that interim period after the MMVT lister was there.

As we expected ad actions accounted for a higher percentage a much higher percentage of total actions in Partial Block Canvassing then in Math Model Validation Test.

But that's because they were - the actions by the MMVT listers included larger numbers of verifications and validations things that we were doing in the office for Partial Block Canvassing.

So some of our key takeaways relating in this slide relating to when Partial Block Canvassing did not find an address that was located by Math Model Validation Test reasons for the omissions tended to be the area was provided to the PBC lister but the instruction was poorly worded or the polygon was poorly defined, mentioned that already leading to lister confusion.

The ad represented a situation that was not detectable by imagery review, changes within existing structures or changes in the use of an existing structure.

So we had instances where there were additional apartments found within an existing apartment building but we also have instances where and we know that this is an issue for imagery review where you have conversion from commercial to residential use.

So you're looking at rooftop. You see the structure in both sets of images but you don't know that the change has occurred underneath the roof.

So in those situations the changes were not provided as a work area. And then the third omission the ad represented a situation that was not detected due to imagery quality or vintage issues and therefore also was not provided to the lister. And we're looking at those.

And as part of our work in geography division we'll be acquiring higher quality higher resolution imagery to deal with some of these issues especially related to areas where you have a dense canopy and the image and it obscures the ability to see images - or structures on the ground.

Additional key takeaways well I've talked about this some of this already polygon instructions resulted in - they tend - they generally resulted in successful navigation but we have learned and are looking at the areas where we had problems and developing refinements to those instructions.

The one recommendation we also and coming back from our listers and our debriefing sessions they all agreed that clearing up and fixing missing and misaligned street features and block boundaries in the office prior to sending out to the field would help improve the process of canvassing and improve the process in the field.

So get the feature - and so improve the quality of the road network, improve the accuracy and the locational information for the roads before going out to the field.

So based on the PBC test we recommend that we test the Partial Block Canvassing in 2016 address canvassing test with traditional listers, that we improve the written - the clarity of written instructions and that we conduct additional analysis on the individual – at the individual address level to fully understand the differences between the Math Model Validation Test, and the PBC listing results and the imagery review results.

I think we're running out of time so I'm going to move quickly into some of our results comparing imagery with the Math Model Validation Test.

So we were interested in - so the Math Model Validation Test gave us the opportunity to and through are in office portion of the PBC gave us the opportunity to compare imagery, and test that and assess that aspect of the work.

Seventy percent so we're looking at the 18,000 ad actions taken in Math Model Validation Test 12,900 of those were detected through imagery review. And we had 1100 or 1168 that were detectable where they were theoretically detectable by imagery but were missed due to imagery quality in the review process issues or reviewer errors.

A little under 1000 were detectable by imagery but was missed due to the vintage. So I can go into more of that later in some of the vintage issues.

And then 2600 were undetectable by imagery and these were again within structure changes. So we have to look for other sources of information about change for those types of units.

Let's see I'm not advancing here. I must've hit the wrong – well so the next slide is while we're waiting for that - oops one more.

Yes. Anyway so the next slide we looked at the - we compared imagery review results against the observed number of MMVT ads, the weighted results for the nation yes that's it were on Slide 28. Go one more okay there we go.

So this is looking at the - what we expected when we reviewed the imagery compared to the observed number of ads in the MMVT.

And the highlighted cells here in 82% of the blocks in which the imagery review expected to find zero ads of MMVT listers also found zero ads.

So that tells us again the imagery review is very good at identifying consistency, stability where there's been no change.

And in an additional 11% of blocks where we expected to find zero ads based on the imagery the MMVT listers found only one so that's encouraging news about the - related to imagery review and the in office work.

At the other end of the scale only in 65% of the blocks in which we expected to find more than 21 ads based on the imagery the MMVT lister also found more than 21. So again that goes back to some of the issues within structure change, apartment buildings and so on.

Looking at this a little bit more so we're - as part of our in office canvassing work we're identifying blocks that we would identify as passive where there's been no change or where we've accounted for everything within the map already. So we would not need to take any action to update the address list.

Eighty four percent of the blocks with at least one address are stable. And 80% - and those blocks contained 85% of the addresses. Again these are weighted results from the MMVT work and the imagery review.

So the passive blocks the blocks that we can review and say these look good. We don't need to take any effective action immediately contained 85% of the housing units of the country.

The blocks where we would need to take action where we're seeing differences between the math and what's on the ground encompassed about 15% of the housing units.

And then the next step would be to identify the appropriate action to take, acquire the update - to acquire updates to the address list. And Evan you want to take the last one?

Evan Moffett: So you've heard a lot of information. And I just want to quickly summarize some key takeaways from this work as it relates to the objectives of the Address Validation Test.

As you can see here the first objective was to evaluate our methods for reengineering address canvassing.

From a statistical modeling perspective we believe that the models were not effective at identifying specific blocks with many ads or many deletes nor at predicting national totals of math coverage errors.

As Mike just discussed we also believe that the Partial Block Canvassing methodology offers potential savings actually not savings potential - a new methodology for conducting in field canvassing

And from a cost perspective we plan to test Partial Block Canvassing and full block canvassing to collect those metrics as part of the 2016 Address Canvassing Test.

The second objective that was presented earlier was how well the in office procedures can replace in field procedures?

We believe that we've demonstrated the utility of imagery review to guide decision making and inform the planning efforts for address canvassing.

We also believe that we demonstrated the value of field work to gather information to assess the effectiveness of in office methods. So we plan to continue this work as part of the Math Coverage Study which will begin next fiscal year.

And then finally one of our objectives was to assess our ability to ensure an accurate Master Address File.

And our key takeaways here is that the math models were ineffective at measuring math coverage error. And we can - we plan to continue to research this and focus on collecting metrics via the Math Coverage Study in the future years ahead. So at this point we want to take some questions.

Trisha: Yes. Thanks for the discussion on the efforts to figure out how to do targeted address canvassing which has been discussed over many of our reports and so on. And so I appreciate these - what you're telling us about what you've been testing and so on in such detail.

Okay I asked the question earlier Deidra about, you know, the metrics, the quality metrics. And I've heard kind of like a lot of numbers. And it wasn't as good as we thought it would be or I'm not sure but I don't hear a lot of like definitive - so if you have definitive like what is many - what is not as much as you had hoped for? What did you hope for?

Pat Cantwell: Actually I don't think we've ever determined some specific number of ads or deletes that missing would be okay.

I mean all this depends upon what's going on in several areas. We have in office canvassing and in field canvassing and both of them are going to contribute towards the work.

So if for instance if statistical model are trying to predict where the ads are or which blocks have the ads but if those are going to be picked up in a different operation and

we haven't even defined all the operations or the sources that we're going to be using then it's let's just say that we haven't decided what would be the appropriate number right now because it just depends too much on the different sources and activities we'll be doing.

Trisha: Right. So there's no target quality metric at this time.

Pat Cantwell: We haven't determined - we have not said this is a specific threshold that's acceptable or not acceptable no.

Trisha: Oh okay.

Deidre Bishop: Trisha, I think the key point that we wanted to get across was that going into the Math Model Validation Test we were hopeful that statistical could help us predict where we should do in field address canvassing.

And what happened was that we found out we should not do that, that we should rely on other methods such as the use of imagery.

That did prove very effective as part of the Partial Block Canvassing Test, that we should continue our research with administrative records from federal, state, local governments and really explore what the commercial sector has to offer in terms of the address and the geospatial data sets. And also really up to date imagery.

Mike Ratcliffe: And I think also it's not really in my mind it's not really a question of what is the expected number of ads and the expected number of deletes but what is the number of housing units and addresses that exist?

And you need to glean that from a variety of different sources that provide that information and then how well do we compare to that?

And then that identifies where we need to collect ads and where we need - where we're too many addresses in the Master Address File.

And where we need to collect updates to the Master Address File s population division is preparing a projected number of housing units for 2020. That will give us one measure to compare against.

As we look at a variety of other sources commercial data files, administrative records, the delivery sequence file from the Postal Service, local government files all of these give us a sense of the number of housing units the number of addresses that exist.

And all can be used to compare against the Master Address File at a very detailed level to tell us how accurate we are in one location and how far off we may be in another location. And then what we need to do to bring those numbers into consistency.

Deidre Bishop: We'll take one more question from Ty and then we'll go to our break.

Ty: Thank you. Ditto, Trisha on how valuable all this information is. I'm not a modeling wonk on the stuff.

But I've always thought models were good in some situations better than other situations. And I'm hearing that they didn't work nationally or they didn't work when you line them up the straight up against the imagery.

But does the bureau know where the models worked well and where they didn't work well? They just seem to be a very relatively inexpensive way to find things if there are certain blocks with a - I mean you have to know in advance where, you know, because I'm thinking like tool kit.

I mean in the bureau has all these different things it's doing and maybe it works in urban areas as opposed to other areas. I mean PBC didn't work in some places but maybe, you know, models worked better there. I mean has that been explored?

Pat Cantwell: Actually Ty, what you're saying is it refers to some of the work that we're doing right now. We're looking at domains that we can break the universe down into because we have the data now following the test.

So we're looking at different areas of breaking them into several as you say inner city balance more rural like that.

And also looking even at what we call ACT codes. I apologize I don't know what it stands for. But it represents the different types of enumeration that we would do address characteristic types. But and again yes we are looking at subdomains to see how well the models work on different subdomains.

Deirdre Bishop: Okay great. Thank you. Thanks for your attention this morning. We're five minutes behind schedule but I would like to offer the full 15 minutes for the break. So let's come back at ten of 11.

And I'd just like to add one thing, Pat had to give his presentation to me twice in order for to fully make sense. So, if you have additional questions for him during the break I'm sure he would entertain those. Thank you.

(((BREAK)))

Deirdre Bishop: Welcome back everyone. We'd like to ask that you take your seats in the next minute or two. Thank you.

Sometimes the most fun part of the PMR is getting to talk with one another at the break so I know how that goes. Let's now move to the preliminary results of our 2015 Census Test.

We'll start with Jessica, and Mike, and Frank providing an overview of the optimizing self-response test in the Savannah, Georgia area. And then we'll move to Maryann and

to Tom to hear about the preliminary findings from the 2015 Census Test in Maricopa County, Arizona.

I would like to stress before we get too far into these presentations these are preliminary results. They will be more anecdotal findings than quantitative findings.

As I mentioned earlier our next push will be analyzing the data that come in from these tests and we'll have more details to share with you during the October PMR. So keep that in mind as you hear the presentations.

And with that I'll turn it over to Jessica.

Jessica Graber: Good morning I'm Jessica Graber and I'm here with Michael Bentley and Frank McPhillips and today we'll give an update on the 2015 Optimizing Self-Response Test including some preliminary results.

As I mentioned in the last PMR our objective for the 2015 OSR test was to increase self-response through the research and testing of contact and communication strategies and this was being done prior to awarding a communications contract.

This test was designed to build and expand upon what we learned in the 2014 census test. To refresh your memory about the 2014 census test at that time we tested a pre-registration option called notify me and that's where respondents could sign up to receive notification when the Internet instrument was available.

We also tested the substitution of an email invitation for the initial mailed letter as well as offered the option to respond without a census ID. And all non-ID cases in 2014 were processed in batch mode.

The 2015 census, 2015 OSR test took place in Savannah, Georgia, Savannah, Georgia Media Market and in comparison it included partnership in outreach activities, a

contact approach designed to promote Internet self-response, an early enough an offer notify me but this time paired with advertising.

And the option to self-respond without a unique census ID and in 2015 these cases were processed in real time. The 2015 Optimizing Self-Response Test included our most successful contact approach from 2014. This was called the Internet push strategy.

And it's a series of contacts beginning with a mailed letter of invitation followed by two reminder postcards and finally the mailing of a hard copy questionnaire.

Next I'll present our mail panel design and how two important design changes that Deidre spoke of earlier. These changes allowed us to fully optimize what we can learn from this test.

This slide shows our initial mail panel design with three panels each made up of 30,000 housing units. The first was our pre-registration our notify me panel where we mailed a postcard invitation inviting them to participate in this notify me option.

While we didn't see high levels of response to the notify me option in 2014 we wanted to evaluate this in the context of outreach and partnership efforts. All non-responding housing units who did not take advantage of this option received our standard Internet push mailings.

The second panel labeled non-ID also received the Internet push contacts but they were not provided a census ID with which to respond. We had a similar panel in 2014 but as I noted earlier in 2015 we used real time non-ID processing instead of the batch processing.

And the third panel is considered our control panel and received the identical number and schedule of mailings but were in fact assigned to census ID.

So this next slide describes the first design change unit that we made. While our initial mail panel design included 90,000 housing units the entire Savannah Media Market of approximately 400,000 households was in fact eligible to participate with those not sampled eligible to use the non-ID option.

We anticipated that we'd identify any response from the remaining households that they would have been the results of our outreach and promotion efforts.

So early on in the data collection period we realized we had an opportunity to test an additional panel in our design when that took advantage of this test environment of this outreach environment and would help us understand the impact of providing just a single mailing but providing some tangible that would remind people and prompt their response.

So we selected a new postcard panel of 30,000 housing units that included only units not previously sampled. Each of these 30,000 housing units received a single postcard with a message letting them know that it's not too late to respond.

While not explicitly stated the mailing assumed they would have been exposed to the ongoing promotion and outreach in the area. None of these cases in this new panel were provided a census ID.

The second design change also noted this morning was the addition of a re-interview component. And the objectives of the re-interview is to validate responses that we've collected over the Internet.

Cases selected for validation include those who responded both with and without a census ID and our plan is to re-interview a total of 5000 cases either in person or by phone.

Here you can see a timeline of test operations. We've completed primary data collection in Savannah and re-interview activities are scheduled to begin on July 15.

And now I'll pass this to Mike and Frank who will provide some preliminary findings from the test and then Deidre will speak later on today about how these results inform our plans for the 2016 test.

Mike Bentley: Thanks Jessica. Hi everyone I'm pleased to present some preliminary findings on response rates from the test in Savannah. We are still processing late mail responses for another few days or so and the post processing is still being finalized but nearly half of the sample in the three initial mail out panels has responded.

The total number is highest in the control panel where we provided an ID. Weighted to reflect the sample design, 47.9% responses of June 29 with 33.6% of the sample responding by Internet.

Those rates are statistically higher than both in the panel without an ID and with a notify me postcard panel which have response of about 44 and 47% respectively.

You may recall as Jessica mentioned that we also tested notify me and non-ID responses as part of the 2014 census test last year in parts of DC and Montgomery County but we also wanted to study these in an environment with advertising and promotions.

The results we saw in 2015 in Savannah are consistent with what we learned last year. Not providing an ID to the address will lead to a lower response rate overall.

Primarily because some of the non-ID cases won't be matched to our address frame. Similarly we learned that the added burden of the notify me postcard in which they are asked to register for their preference and then to come back and respond may have inhibited overall response for that panel.

I'll talk a little bit more about the notify me results in just a few moments. As (Jessica) previously mentioned during the test as we began to assess progress early on we

decided to include an additional postcard mail out which was sent to a sample of addresses outside of the three initial mail out panels to households who had not yet responded to advertising or promotional efforts.

This postcard resulted in about 8% of the 30,000 sample responding without an ID mostly by Internet with the remainder providing a response by phone. Further the promotion and outreach efforts were very successful in bringing in more than 35,000 respondents from outside the mail panels in Savannah to respond.

Recall that these respondents did not receive any mail materials at all but instead were motivated to respond to be counted solely on an ad they saw or heard, word of mouth or perhaps by attending one of various partnership events throughout one of the 20 counties in Savannah.

Slide 9 please. Finally the last preliminary response result we want to discuss today is the notify me census of encouraging people to pre-register to receive an email or text invitation as their contact method.

During this sign up period a total of 1,925 participants pre-registered. Of these 1,341 signed up before the cutoff date and were matched and of those 1,203 were in one of the 20 counties comprising the Savannah area.

The majority over 80% of these selected email as their preferred contact method with the others opting to receive a text message. About 93% of these participants ultimately came back and responded most by Internet.

Overall these results are very similar to what we learned in last year's 2014 census test where participation with pre-registration was also fairly low. Together the results of the notify me testing both last year and in 2015 our upcoming 2020 design decision later this year.

And now I'm going to turn it over to Frank for an early look at some of the interesting results on the non-ID processing in the test.

Frank McPhillips: Thanks Mike. Good morning everyone. For today given the brief time we had I just wanted to cover a couple of things in our early results analysis. So we're going to look at when we got the non-ID responses to kind of give you an idea of sort of the distribution of workload because I know that was of interest going into the test in terms of the system performance and so forth.

And then also I'm going to show you a little bit about what happened during real time matching and geo-coding of the non-ID responses and then the influence of these administrative records data following that real time processing.

And at the tail end I'll just tell you a little bit about other stuff that we're tabulating. So, (Mike) - got it.

So slide 11 talks a little bit about the distribution of the response over the course of the test. So this is basically the last week of March through end of May.

And as you can see the bulk of the response is up front around when we were doing most of the promotional efforts and there were mailings going on and so forth that would then give us another bump.

Essentially though, you know, things went comparable to the way that we would have expected and you see some small spikes at the tail end of the test again when we did additional mailings and tried to promote response.

So this was covered in some of the other material. And we did see in terms of days of the week definitely weekends saw a lower response and Monday through Wednesday was typically the greatest amount of response and actually later in the test that shifted a little bit to Wednesday, Thursday.

So the next slide breaks it down by day of the week and obviously Wednesday had the largest percentage at almost 20% and then again Saturday and Sunday the smallest and then you can see the progression very comparable for Monday, Tuesday and then Thursday catches up a little bit.

Before I talk about the results of real time versus administrative matching I just wanted to review a little bit about the process. So real time matching occurs and geocoding I should say occurs during the self-response.

So the respondent enters their address, you know, next screen they see the standardized version of their address so we can confirm this is a good address to process.

Then the next step is before they continue to interview we hit the matching and geocoding service and we get that result. And so after that though for cases that do not match we turn those all over to our center for administrative records and research applications area who use a process to apply a composite of commercial and federal administrative records data to try to pick up missing data elements or correct misspellings things like that to see if we can get another shot at matching.

And so that's just a reminder of what's in that composite and what we used before we go to the actual results.

So this shows at a high level the matching results and what we're showing here is in the first row matched and geocoded is really representative of a geocode, a match to a record that has a census block assigned already and where we are very strict about the rules in which we permitted the record to be considered resolved if you will.

And we could consider that case as a response and therefore remove it from non-response followup. And so for this test only we just used very strict rules that allowed us to identify operations from the 2010 Census which would have physically verified the existence and location of the address.

And then the next row shows actually a more broad category where we also got matches and either the record and the master address file was un-geocoded or it was another kind of geocode.

And this is what I was alluding to in terms of the strictness of the rules. So in this category we didn't again just for this test consider some of the more recent updates that we've gotten to the math.

And what that allows us to do is look within that category and kind of reevaluate those rules and see, you know, what areas where we might want to reconsider, you know, accepting that geocode that goes with that math record.

The bottom line as you can see that we had an extremely high match rate actually during real time, you know, just under 97% and so if you include both rows.

And then when you add in what the administrative records composite does for us that rate gets up to like 98.5%. So very encouraging matching results and this is of course inclusive of both the cases that (Michael) was talking about in terms of the sample to non-ID cases that were in panel as well as the rest of the site, those that responded without us getting or sending them any direct mailings.

So this next slide just breaks it down a little bit further in terms of those matching results. And I just want to qualify that unacceptable geocode in the second row that was just our terminology that it didn't meet our rules for this test.

And so that's the category we can really look at in terms of refining those rules, looking at those sources of the math input and the geocodes that were associated with it and then see if, you know, what part of the rule we want to revisit.

And we already have a good idea of that from our preliminary analysis. And then the match but un-geocoded is really another opportunity. This was typically cases from the

delivery sequence file I should say addresses that we received from the delivery sequence file that perhaps we just don't have a TIGER feature yet or an address range for that TIGER feature.

So we haven't assigned the geocode yet, you know, using our typical methodologies but certainly the work that's going on through the geographic support system work and the typical work we do to maintain the math and TIGER during the decade will put us in a better position in 2020 to have a lot of those DSF records already geocoded so that when we make the match we can call a case done.

So bottom line though you're really seeing an extremely high match rate and again very encouraging.

So just a quick look ahead through what you'll be able to expect to hear about in the next PMR or at least in the analysis report that will come out for our team is we're also going to be looking at the demographics of non-ID respondents.

You know, what are the characteristics of the folks that decided not to use their ID and in particular those that were not in our sample group but particularly interested in the ones that just opted to go that route.

And that might help us in terms of planning and, you know, targeting, you know, certain groups for this kind of response and seeing if there are those that just have that propensity.

We're also looking at comparing our results from the other site tests, so this is Maricopa County and we did self-response that included a non-ID component. We did not do real time processing there but we used the same business processes to do the matching and geocoding so we have a basis for comparison between the sites.

And, you know, just the little foreshadowing one of the things we saw was that the real time processing seemed to really help us boost those (metrics) we could see that difference between the two sites.

So looking forward to letting you know a little bit more about that the next time we talk.

And then finally we're going to be looking at the workload that in a traditional non-ID processing operation for decennial that we would have processed in a manual fashion.

So this is things that we couldn't resolve during automated processing. So we'll be able to give you a sense of those cases that fell out of automated processing, you know, for example we weren't able to obtain a geocode or we could not match our geocode.

How many of those we could have subsequently resolved. So looking forward to getting to finish digging through the data and getting all the good results out of that.

Deirdre Bishop: Questions for the team?

Ty: Since it's such a large group could you just give one example of what you mean by a geocode that's unacceptable? I mean what are we talking about there?

Frank McPhillips: So we were very conservative in the rules for this test and again some of the things that we would have accepted as far as a source for our geocode for our math record in this case would have been a 2010 operation.

Where we had, you know, such as address canvassing where somebody physically, you know, observes this address on the ground. Now in the category that you're asking about a big group is the records that have come in from our partners through the geographic support system efforts.

Certainly, you know, we can know those to be good quality addresses and we have, you know, subsequent work that's happened and the geo has been doing for the last few years as well to validate those inputs from local sources.

So that category is going to move into the acceptable geocode, you know, that's probably the most classic example.

Deirdre Bishop: Okay before we move onto the preliminary findings from the 2015 Census Test I just want to take a brief opportunity to make this real and sum it up a little bit as we move into the design decisions associated with the operational plan.

So as you heard from the team findings from both the 2014 and the 2015 test showed us that asking people to pre-register or sign up for notify me hasn't really been successful.

You heard that people don't really mind entering in that 14 digit ID that we send to them. However with that said because non-ID processing has been so successful and we're receiving such a good metric it is a very good mechanism for helping to make the census more mobile as we move forward giving people an opportunity to respond anytime, anywhere.

So with that I'd like to move to the results from the 15 test and Maryann will kick that off.

Maryann Chapin: Okay, good morning as Deidre said my name is Maryann Chapin. This morning Tom Mule and I are very happy to be here with you to present some of our preliminary findings from the 2015 Census Test.

The 2015 Census Test was really a tremendous experience for us and overall we deem it a huge success. We've learned a great deal from our experiences that will inform the design decisions we'll make later this year.

It will shape the enhancements that we plan for the 2016 Census Test, other future tests and the 2020 Census itself. The 2015 Census Test was not intended to be the full or final solution that we use for the 2020 Census.

So we fully acknowledged that some aspects of what we implemented in Maricopa County were more mature than other aspects.

So over the next - I'm sorry, over the next 45 minutes Tom and I will share with you information on our administrative records processing and we'll present a representation of the more prevalent qualitative findings reported by participants and/or observers of the test.

What is presented is not going to be an exhaustive list, additional observations, the analysis that will be conducted on the data that was collected and the remaining activities for this test will continue to inform our understanding of what happened during the test.

In the presentation we'll touch on topics related to the United States Postal Service support of this test, our approach to training, the COMPASS application that was used for data collection, some procedural situations that we encountered and field reengineering.

We'll close with an update on the bring your own device testing, the evaluation followup and talk some about our next steps. Before getting into the specifics of the test the next couple of slides are provided to give you sort of a refresher about some of the specifics of the 2015 Census Test.

Our focus for the 2015 Census Test was on testing innovations and collecting data that will inform the preliminary design decisions that we'll make later this year.

In the 2015 Census Test we tested the reengineering of roles, responsibilities and infrastructure for conducting field data collection. We tested the feasibility of fully

utilizing the advantages of planned automation and available real time data to transform the efficiency and effectiveness of our data collection operations.

We further explored the use of data that households have already provided to the government to reduce the non-response followup workload and increase the overall non-response followup productivity through the use of administrative records, field re-engineering and adaptive design.

We also tested the operational implementation of a bring your own device option for our enumerators and we will conduct focus groups to explore the reactions of the respondents from this census test on contact methods, administrative records usage, privacy and confidentiality concerns and how we as the Census Bureau may be able to address those concerns.

As stated earlier the 2015 Census Test was conducted in Maricopa County, Arizona. Block groups for the test were identified based on their diversity of socioeconomic characteristics.

The block groups in Central Maricopa that are shown in blue were selected for their high concentrations of both vacant and Hispanic populations. So block groups located in the cities of Chandler and Mesa shown in yellow and pink respectively were identified as areas with higher 2010 Census return rates and lower mobility.

The block groups in the outer ring on the northern border of Maricopa County that are shown in green were selected to allow us to test in more remote locations.

Seen on this slide are some of the key milestones for the 2015 Census Test. The items shown in black with a checkmark are those activities that were completed as of our last PMR in April.

Those shown in blue with a checkmark are the activities that we have completed since that PMR and those that are shown in blue with an open circle are the items that are still to be completed.

At this point I'd like to turn the presentation over to Tom Mule. Tom will share with you some information on our administrative records identification and processing.

Tom Mule: Thank you Maryann. Today I'm going to share some of the results that we had from the administrative records identification that we did while production was happening for the 2015 Census Test.

I'll go over the data sources that we used and I will go over the results that we had from the identification, the phase one that we identified on May 12, this was two days before the start of the non-response followup operation.

And then also document we were able to identify and where we are in phase two. This was something based on how the data was coming in from the Internal Revenue Service.

This was something that we were able to add on as the test was going on and additional identification of administrative records while the test was going on. So we would document both of those phases and then also share some of the characteristics of these identified units that we have.

So going first with the administrative records which file did we end up using in our identification? So the first part the first topic the federal sources that we did end up using we started by using the Internal Revenue Service individual income tax returns.

And we owe a big thank you to the IRS because they were able for the processing for this test to be able to start delivering the 1040 individual tax returns, they were able to start delivering that to us on a monthly basis starting in March.

So for our processing especially for our phase one delivery we had their deliveries from March, April and the delivery that they were able to do the beginning of May. So we had that 1040 information for people who started filing in February available for out use in this test.

We also utilized the Internal Revenue Service the information returns your 1099 information about interest and dividends. That was another source that we were able to use to be able to develop our rosters and use our information of determining for our predictive models.

We also continued to use the center for Medicare and Medicaid services the Medicare enrollment database. This is another major source that we've used in the past which gives us coverage of the elderly population and also sources picked up from the two Internal Revenue service sources that I just mentioned earlier.

For this test a new source that we did end up using was the Indian Health Service this patient database. Since we were conducting this test in Maricopa County which has a concentration of Indian populations this was another source that we ended up using when building our rosters.

We also utilized the social security number identification file or the numident. This was used as part of CARRA their processing of their being able to assign the protective identification keys to the record.

And plus this also was allowed for the persons - it was a major source of age and sex information for those people. So those are some examples that we used for persons to build rosters for our administrative record household.

We also continued to use the United States Postal Service for the 2015 Census Test. We did conduct four mailings. With those mailings we requested from the United States Postal Service if they could - unable to deliver as addressed they provided that information and also those reasons.

So we were able to use that information from the second mailing which had an in home delivery date targeted for April 1 as part of our information as well.

We also as part of this test we did use a commercial third-party file. We used the targets federal consumer file. One thing with this is we did not use this file to build rosters, we did not take persons from this file and put them into the household that we were building.

But we did use this as another piece of information for persons that we had rostered from the federal sources I listed earlier we used the targets as another source to say was that person associated with our sample address or is that person associated with another address.

And we also used partnering with the fitness for youth team. We were able to use best race and best Hispanic and origin assignment that (Sonya Resto) and her team was able to do.

They were able to use 2010 Census and other federal, state and commercial sources to be able to assign for the administrative record what their best race and Hispanic origin information was and this was information that was based on 2010 processing.

The one thing Sonya did want me to point out is that additional work and taking into account additional census responses or also additional federal or state or commercial responses could have additional improvements in the future.

So let's go over some of the administrative records identification results. So two days before the start of the non-response followup operation we conducted our first phase one of our administrative records processing.

We identified all of the cases that were eligible for the non-response followup operation that hadn't received a response yet. We proceeded to implement our

production approach which for this test we ended up using predictive models as compared to the rule based approaches that we did in the past.

Applying our predictive model of methodologies we were able to identify for the records that we had 11.6% of those we identified as being administrative record vacant.

Then we proceeded to also identify for 18.1% we proceeded to identify that they were administrative record occupied based on the roster of people that we're able to put together for those housing units.

So that did result it was 70.2% which at this point in time we did not make a determination. For these cases they proceeded to go out to the field for the enumeration especially for the two experimental panels where the administrative records information was utilized.

The main things in terms of this is either we did not have administrative records either person or from the United States Postal Service either undeliverable as addressed information to make a decision or we had information but they were below our cutoffs.

So I did mention that we did implement a phase two of this in early June. One of the things that we were able to see with working with the center for economic statistics and also with CARRA.

That with the Internal Revenue Service was delivering these monthly files as we had a very quick turnaround in being able to have these files available to us and IRS was delivering these files to us in a very timely manner.

So in early June we got the next delivery of the Internal Revenue Service 1040 tax returns. So based on that we implemented an additional processing of identifying administrative records.

So at this point we took the cases before that we could not determine and we took the new information that we had and we proceeded to run our predictive model approaches.

And so the result of that is out of that 70.2% that we had before there was another 1.7% that we were able to identify as administrative records occupied. We were able to process this information, have this available on June 5 so when June 6 came around it became time to start doing the field work for our full removal panel the experimental panel, these cases were removed from the field work on June 6. They got no more visits.

For our hybrid removal in which those cases were getting one visit it's those cases that have already been visited, they were turned off and then we provided the information that for these cases any point after that once they got those visits the work for those can stop.

So one of the things we try to do as part of doing this research is what things have we learned in previous tests and previous, you know, how are we carrying those forward.

So this slide here is showing results that we showed back in January. What was one of the things that we were seeing in January related to the performance of administrative records especially when we were identifying occupied units.

So one of the things that we ended up seeing in the 2014 test is when we were looking at the results of how the account comparison how the account that we had based on our administrative record household how did that compare to the census determination based on field work.

We ended up seeing that there was three household composition categories that performed better than others. So we were able to classify cases based on their administrative record composition based on the number of adults that were there, based on 18 or more and whether kids 17 and under were present or not.

We ended up seeing in the 2014 test that the one adult not children, two adults with no children or two adults with children those three categories were performing better in the 2014 test.

They had count agreements between 57 and 65%. When we were seeing those and we were implementing predictive modeling approaches for the 2015 we were searching ways we might be able to account for this and have this in our predict models.

So one of the things I want to show from our 2015 results is we ran our 2015 processing with our predictive models. We were able to identify the administrative records occupied cases that's 18.1% that I described in phase one and the 1.7 that I described in phase two.

So how did those go but how were those distributed across these adult and these household composition categories? And one of the things we were able to see as part of our predictive approaches that we're using for the 2015 test is that the cases that we are identifying were primarily coming from the three categories that performed the best in the 2014 test.

So with the one adult zero children they had - we were focusing on phase one, 34% of the cases we identified fell into that category, 30.1% falls in the two adult no children category and 25.2% fell in the two adult with one plus category.

So we were able to see with our approaches that we are identifying cases that were in categories that performed better in the 2014 test.

One thing which we did not go into detail in the January 2014 PRM when we were discussing 2014 results is what we were seeing related to race and Hispanic origin.

So how often was race and Hispanic origin information available from the administrative record sources? So as part of the 2015 Census Test there was a combined question of race and Hispanic origin in the same question.

Some of it presents some results based on a combined race in Hispanic origin categories. There we some results based on eight categories based on Hispanic and then there will be six categories of non-Hispanic with each one of the races alone and then non-Hispanic or two or more races.

So we're going to be able to look and be able to say for these administrative record cases that we have identified, looking at our past census sources and as they work with the fitness for youth team was able to provide based on best race and best Hispanic origin how often do we have race and Hispanic origin available to us.

So in the 2015 Census Test we are able to see that overall for phase one we had a race and Hispanic origin available 82% of the time and for the phase two cases identified it with 77% of the time.

Also this figure showed results for the four areas that Maryann highlighted earlier. These were the four different areas of the test so we were able to see the central area which in addition to the characteristics that Maryann mentioned also had a very heavily Hispanic population.

We've been able to see somewhere performing results across all the four different areas. Chandler does have a 72% for phase two but that only had 1.7% of cases identified overall so there's definitely a small sample size there.

So the other thing which we did want to look at was being able to see four, these four different areas breaking down where we had race and Hispanic origin available, how much did we have for the eight categories that I described earlier.

So some of the things that we were able to see is that in the Central Maricopa area 55.3% of the time we had a Hispanic response on our best race and Hispanic origin results.

We were generally able to see for the three other areas that the non-Hispanic white alone that those results were between 51 and 66%. We were also able to see that in the Mesa area that the American Indian and Alaska Native the non-Hispanic alone group there was a 3.6%.

We generally saw for the non-Hispanic black alone that the results ranged between 2.1 and 8.3% and then for the Asian alone we did see in the Chandler area there was a 5.2%. So those are some of the results that we saw related to the race and Hispanic origin.

The one thing which we are going to do is our further analysis is going to be definitely digging deeper. We are working with CARRA to have the 2015 census responses, have the protective identification keys assigned to those.

So with our control panel where we did do identification but those cases did go out to the field we are able to compare on an individual record basis the race and Hispanic origin information that we have versus what was collected in the field.

That's definitely future work that's going to be done related to this and they also showed some results related to the count comparisons earlier and will continue to do those again for the 2015 test.

But one thing this test does have it does have an evaluation followup component where additional interviewing and results will be available for those comparisons as well.

At this point I will turn it back to Maryann.

Maryann Chapin: Thanks Tom. I'd now like to focus the balance of our presentation on our operational experiences with the 2015 Census Test. In past PMR's we've talked about our target sample sizes for the test.

On this slide we'd like to share with you some information about the cases that were included in the non-response followup field data collection workload.

Before speaking to the DTL's of this table although not a main objective for this test we did achieve a self-response rate of approximately 56% in Maricopa County.

So now as a reminder there were three panels in the 2015 Census Test, a controlled panel and two experimental panels. The two experimental panels being the full removal panel and the hybrid removal panel.

The controlled panel employed similar procedures to those used in the 2010 Census. The full removal panel implemented an adaptive design context strategy and reduced the initial non-response followup workload by excluding any addresses identified as vacant or occupied based on administrative records prior to any contact attempts being made in the field.

Remaining non-response followup cases were visited at least once. In the hybrid removal panel the initial non-response followup workload was reduced to exclude any addresses identified as vacant using administrative records prior to any contact attempts being made.

For all remaining addresses enumerators made one personal visit. After that initial attempt for those cases that were not resolved the non-response followup workload was further reduced to remove any of the addresses that could be enumerated using administrative records.

The cases that remained after the additional administrative records removal had at least one additional contact attempt made.

So now looking at the table I'd like to draw your attention to the column labeled final workload it's the second from the right on the top part of the table.

The numbers in this column reflect the non-response followup workload or the field data collection workload for each panel. The workloads factor in the initial self-response workload for each panel, sub-sampling that was done to arrive at our target sample sizes.

It reflects the administrative records removal of vacants from the hybrid panel. The administrative records removal of the vacant and occupied from the full removal panel and the removal of late self-responses prior to the first day of our field work.

The lower portion of the table represents how the cases included in the non-response followup workload were resolved. I'd like to draw your attention to the line labeled discontinued with administrative records.

The discontinued with administrative records reflects the number of cases removed from the non-response followup workload in the hybrid panel for those cases unresolved after that first attempt and for which we had administrative records that could be used for enumeration.

Finally, I'd like to point out the row labeled late returns after the NRFU start. These cases reflect or could reflect what really were genuine and legitimate late self-responses or they may be self-responses that were generated as a result of the notice of visits that we left on the door at each attempt that directed the respondents for the experimental cases to either the Internet, to questionnaire assistance or to return their paper form.

I also want to note for the one case shown for the control panel that was not resolved we did determine that this case could not be closed out because the information that

we had for this case did not meet the COMPASS requirements that and address had to include either a house number and street name or a location description.

For this case we only had a street name, P.O. Box number and map spot. So this situation is going to factor into the lessons learned from this 2015 Census Test and how we move forward for 2016.

For the 2015 Census Test the Census Bureau partnered with the United States Postal Service to assist in the steps necessary for our job applicants to complete their enrollment steps which included fingerprinting, swearing of the oath of office and having a photo identification taken.

Twelve United States Postal Service locations across Maricopa County were identified to support the 2015 Census Test. Early in the test we did experience some challenges in the scheduling of the appointments for our field staff to complete their application process.

To work with the Census Bureau the USPS staff were required to attain a special sworn status. We experienced some delays in some of these offices in having their clerks sworn.

This resulted in not having the clerks trained on the necessary procedures and equipment that would allow them to complete the application process for our enumerators.

So on April 3 which is when we expected to have all 12 of the locations functional we had 10 of the 12 fully functional. Also in an effort to be proactive the Census Bureau began steps to select our field staff earlier than we had originally communicated to the Postal Service.

So while they were expecting only a very few number of staff to be scheduling appointments during that first week we overwhelmed the Postal Service offices by having approximately 300 people trying to schedule appointments.

In addition we experienced some challenges with the actual scheduler itself. Applicants were having difficulty getting through to the United States Postal Service.

This was a concern since our applicants have only 10 days to visit the USPS for their fingerprinting and their paperwork or their offer of employment could be rescinded.

To rectify the situation the decennial service center was able to receive calls from the applicants, schedule the appointments and send to the United States Postal Service offices a secure file that contained the information they needed on the appointments.

Then the USPS could confirm the appointments with the applicants and this alleviated the challenges that we had with getting the appointments scheduled.

We also experienced a gap in service on April 15 for tax day where we had limited sites and appointments available. This was not something that we the Census Bureau had expected and really was primarily just a communication breakdown in something that we should have anticipated.

In addition we did experience some other minor situations that really could be expected the first time you try to test something like this but all of these experiences are going to better position us for the future endeavors when we partner with the United States Postal Service.

I do also want to mention that the Postal Service played an essential role in a couple other aspects of the 2015 Census Test. First, as Tom mentioned earlier they provided the detailed reasons why mail was undeliverable as addressed which was integral in our ability to identify and remove vacant from the non-response followup workload.

But then in addition we were able to take advantage of a service that's offered by the United States Postal Service that assisted in our early recruiting efforts in areas where we were having a challenge getting the desired number of applicants.

We were able to take advantage of the every door direct mailing that allowed us to have a postcard generated and delivered to target areas that communicated to the people in those areas about the census test and the job opportunities available.

Over the next several slides our focus will be on the experimental panels and the enhancements and innovations used for those panels. We've learned much more than we're going to be able to share in a few slides.

But the next few slides will attempt to hit some of the key points but is not intended to be a comprehensive listing of our successes, our challenges or the opportunities that the 2015 Census Test did afford us.

First we'll discuss our new approach to training for the experimental panel enumerators. As a quick reminder the 2010 Census and for the control panel the training used was a verbatim approach.

For the experimental panels the training used a blended methodology that incorporated best practices of adult education. It involved nine hours of online pre-classroom enumerate independent study, 7 hours of in classroom training and then an additional 2 hours of post classroom online training.

Overall with this training approach we were able to reduce the number of training hours as compared to the 2010 Census from 32 hours to 18 hours.

This training approach was successful in providing standardization of the information and experiences for all the enumerators. It provided for tracking that enumerators had completed their online training and offered various learning methods.

A key success for the 2015 Census Test was that after completion of the online training and the classroom training we did identify several topics where the enumerators required some additional information to clarify various aspects of their job.

Very quickly the rocket team was able to develop training videos that provided the needed information. A link to the video was texted out to the enumerators allowing them access to the information delivered in a timely, efficient and consistent manner.

Another aspect of the 2015 Census Test was the employment of a concept of a training specialist. This training specialist was responsible for training the experimental panel enumerators who were to perform the field data collection.

These positions for this test were filled by existing Census Bureau employees relied on their knowledge and experience and some minimal training specific to the test.

We heard from some enumerators who participated in the training that they felt the training specialists were not knowledgeable on the subject matter, the COMPASS application or the experimental panel procedures and were not necessarily invested in the success of the test.

Our long-term vision really is to establish a position where the primary responsibility is the training of our enumerators and that the people that fill these positions will be subject to a rigorous train the trainer program.

So for the 2016 test we'll take another step toward that ultimate vision with establishing the training specialist position and putting those training specialists through a robust training program.

From our enumerator debriefings we did hear that in general the online and classroom training was well received but we did hear about some specific areas where we need to consider change and these included including more practice on handling difficult cases and specifically on the harder to refusal type situations.

Clearer instructions on who to call for assistance and more instruction on how to text and set up their voicemails.

We'll now move onto the COMPASS application. As a reminder the data collection for the 2015 Census Test used the COMPASS application on an Android device.

In general enumerators found that the Android devices and the COMPASS application were easy to use. However, enumerators and observers reported crashing or freezing of the COMPASS application.

When this occurred the smart phone screen would go blank or the screen would freeze causing the enumerator to have to reopen the case before being able to proceed with the interview.

The crashing or freezing of the application was attributed to several causes including excessive heat of the device and memory demands while we had maps open in the background.

Further research is needed on these situations in terms of the crashing and freezing but it is subject to the reports that were called into our (desential) service center when the enumerators experienced that and we do think that that was severely under reported to the decennial service center.

In addition, enumerators were concerned and/or frustrated when cases that they believed they had resolved were not disappearing from their case list. It seems that the root cause of this may have been a training issue meaning that enumerators may not have been trained well enough on what happens when a case becomes proxy eligible and that it remains on their active cast list.

So after completing an attempt and coding it appropriately if that case then becomes proxy eligible because it's exhausted all of the attempts available the case will remain

on the active case list for 90 minutes and this was to allow the enumerator's time to locate someone who could provide an acceptable proxy response.

If the unit was also proxy eligible and was also a multi-unit the case would remain on their active case list until 11 o'clock at night. This allowed the enumerators to work any additional units within that multi-unit structure and to have time to go to the management office and obtain information specific to the proxy responses for those cases.

Finally, observers also reported respondent frustration with our coverage questions. As an enumerator approaches the completion of the interview he or she must ask a series of over count questions to ensure that we are not including anyone in the household enumeration that may have or should have been counted elsewhere.

It was reported that these questions created some frustration and undue burden on the respondents and modifications to these questions are in the works for the 2016 Census Test.

And as a last comment over the course of the test we were able to release several new versions of the COMPASS application that addressed various situations encountered by the enumerators.

The new releases fixed things like fixes to reduce the impact on the map usage that I mentioned earlier that was contributing to the crashes and freezes. We were very pleased with how these software releases occurred.

This wasn't something we were able to test in the 2014 census test and so we see the deployment of these releases as a significant accomplishment for the 2015 test.

We'll now cover various components of the 2015 Census Test specific to some procedures and the interaction of the COMPASS application and the management of assignments.

In the test enumerators encountered situations such as proxy interviews, vacant housing units, addresses that did not exist and other types of non-interviews.

And the COMPASS application did have the functionality needed to handle the outcomes of these non-standard situations. However, we learned that the procedural and operational implementation requires improved alignment of the functionality that exists within the COMPASS application.

Learning from our 2015 Census Test experiences are needed on the various non-standard scenarios to more fully understand and define the expected enumerator experience.

Such improvements to the COMPASS functionality and the pathing through the application can enable the enumerator to more successfully perform their job.

Similarly greater coordination of the assignment of non-response followup cases within multi-unit structures is needed. Observers reported multiple enumerators visiting the same multi-unit complexes on a given day or from one day to the next.

This resulted in leasing managers being contacted by multiple enumerators requesting similar information on the occupancy status of units. Additional thought and planning is needed with regard to how to handle assignments within multi-unit structures.

An approach to the multi-unit structures it may also have applicability as we think about how to handle assignments within gated communities.

We'll now cover various components of the 2015 Census Test specific to the reengineering of the management of the field work. The area operation support center or AOSC was observed to be very well run and much quieter than a typical local census office.

Very noticeable was the lack of paper due to automation of data collection and the electronic payroll. The area manager of operations or AMO and the field manager of operations or FMO's were well equipped to perform the functions that we asked of them.

It should be noted however that the AMO and one of the two FMO's were experienced Census Bureau staff and were able to master the requirements of their job without problem.

The second of the two FMO's was hired off the street and experienced a steeper learning curve that's more likely to be reflective of what we will see as we staff up for the 2020 Census .

The alerts that were generated by the operational control system or mojo were very successful in identifying situations that required in some instances us to contact our enumerators and take some corrective action.

Another very successful aspect of the test was the entry of enumerator work availability. We found that when the enumerators indicated that they were available for work and that when work was available and assigned to an enumerator they did in fact work the hours that they had indicated. This even includes the Memorial Day holiday that was in the midst of this test.

The workload optimization that considered an enumerators home location, the location of the non-response followup workload and the use of a response propensity model to predict when respondents are more likely to be home was observed to be very effective.

In fact when we had observers from headquarters out in the field we did have someone from our national processing center there and they were interested in how they might be able to use this in informing when to call respondents for the cases that go to the call centers.

And then finally overall while we consider the routing of cases in the 2015 Census Test of success there were observations reported that the order in which enumerators were instructed to work their cases at times was inefficient.

Observers reported instances where enumerators would have to backtrack returning to addresses that they had passed on their way to another address. To address this situation during the test we were able to make some slight improvements that addressed some of the inefficiencies.

Since that test we've had much more detailed discussions between the staff that are designing the operational control system and the routing algorithm with addition from our geography staff to identify how we can improve the routing algorithm that we'll use for the 2016 Census Test.

Our efforts to test the technical implementation of bring your own device have concluded. The enumerators who participated in the bring your own device testing were recruited from the enumerators who conducted the non-response followup control panel data collection and who had a mobile device that met established operating system requirements for IOS or Android platforms.

Selected enumerators used their own device for this data collection to contact addresses. We had the potential to contact up to 4500 addresses in the test, ultimately over the 12 day period for the data collection we completed work on approximately 2300 cases.

Approximately 80 enumerators participated in the bring your own device testing. Prior to conducting the data collection the enumerators participated in one of six half day sessions where they validated that their devices did in fact meet the minimum operating system requirements.

They downloaded and installed the COMPASS application, they reviewed and signed the acceptable use policy and reviewed and signed the reimbursement policy that was specific to this test.

Some initial observations associated with the BYO testing include that the training was fairly labor intensive, validating the minimum operating system requirements and ensuring that the COMPASS application was downloaded and installed correctly required a heavy involvement from census headquarter staff that were present to assist enumerators.

Based on that thought is being given to the development of a pre-training device preparation instructions and perhaps conducting separate training sessions based on the device operating system whether it was IOS or Android.

From July 20 through August 14 enumerators who are specially trained on re-interviewing techniques will conduct the evaluation followup operation. There are two major objectives to the evaluation followup.

Those being to obtain the most accurate status of the housing unit on census day and to identify people associated with the occupied housing unit during the calendar year as well as the timing of their association with the housing unit.

This will ultimately help determine the most accurate household status and roster for census day. The evaluation followup will consist of approximately 4100 cases, sub-sampled from the non-response followup control panel cases.

Where housing information collected in the non-response followup differs or conflicts with information we have from administrative records for that housing unit.

Training for the evaluation followup enumerators will occur next week and will begin the data collection conducting it from mid-July through mid-August.

With so much accomplished we still have so much more to do. Next week we'll conduct the focus groups aimed at the specific aspects of the respondent experience.

We will still have to complete the data collection for the evaluation followup and with the majority of the field data collection complete staff are now turning their attention to data analysis and understanding what that tell us about the approaches to administrative records for the identification of vacant addresses and occupied units.

As well as what the data tell us about the reengineered field management structure, automated training, adaptive design, optimized assignment management and route optimization.

The results of the analysis will inform the preliminary design decisions that we'll make in the upcoming months. And finally the results from the field observations the debriefings our lessons learned and our data analysis will inform our planning for the 2016 Census Test and beyond.

And with that Tom and I will be happy to take your questions.

Trisha: Thanks Maryann and yes the 2015 test was really - I enjoyed going out in the field and seeing ROCKeT and COMPASS, very interesting, very different. I wanted to go back to the table and I apologize if you already addressed this on page 11, no 15 I'm sorry 15 yes on page 15.

The row that says discontinued without address did you discuss like those numbers which seem rather large?

Maryann Chapin: I didn't discuss those but those are primarily for the cases where they exhausted their maximum attempts and were then the work stopped in the field after they'd had the opportunity to conduct a proxy interview.

Trisha: But they were not completed?

Maryann Chapin: Right they didn't end with a respondent provided response.

Trisha: Okay, so you'll be analyzing more about that?

Maryann Chapin: Yes.

Trisha: Thank you.

Ty: Thanks again for all this information. A couple quick observations and then I think a question related to them. You've got a lot of lessons learned and feedback right.

And I mean I'm just listening to what you've said and so it sounds like they're going to directly inform the design, somebody directly inform some tools, others are going to inform future tests, how you do the tests and some it sounds like might be informing how you might be trying to interpret the data from the test as well.

The other observation is I know you have a governance framework for a recommendation followup we've heard about previously and I guess the question is how much of what you've talked about learned and I know we've had lots of conversations and you've heard from lots of other observers as well.

How much of that is being captured kind of formally by your governance framework for ensuring that these many, many, many, many different lessons get tracked and followed through on?

Deirdre Bishop: First we have the responses from our focus groups and our debriefings, formal results of those activities are taking place. Next, any person who went out into the field to conduct an observation on behalf of headquarters or the regional offices was asked to submit a trip report.

Those trip reports we're tabulating and tallying lessons learned from those reports. Everything that we heard from the managers that were working out of the AOSC the FMO's, the LSO's and then the enumerators in the field we're documenting as well.

And as you'll hear when I give the 2016 Census Test update we're applying what we've learned through our planning for the future.

Ty: I guess is there - could you say more about tracking I mean I guess there were a couple things for example like in the multi-units that, you know, you struggled with in 14 but, you know, still struggling with them in 15.

And just could you say more about how you're ensuring that I mean some of these things you can't allocate resources to deal with and won't doing so but there's kind of a control mechanism in place over making sure things aren't falling through the cracks.

Deirdre Bishop: I should have mentioned that we just last year developed a formal knowledge management database and that is one way. And then another way Ann talked about the decision analysis that we're doing as part of the development of the operational plan.

That is another way where we're bringing all the key leaders for each operation into the room and documenting the key questions and the answers moving forward.

Man: Maryann regarding the mojo alerts you classified those as very successful during the 2015 test. Could you talk a little bit more about how you came to that conclusion?

Maryann Chapin: I think primarily so we had several types of alerts that were thrown. There were things like the long-distance flag which gave an indication of whether enumerators were exceeding an acceptable or established distance from where we expected them to be conducting the interview.

We had short interview flags which we looked for enumerators who may have had cases being conducted very quickly. We also looked at or there was an alert associated with if an enumerator had I'm going to say the number of sufficient parcels they had was out of line from what their peers were experiencing.

So having these alerts thrown we were able to when they occurred take some immediate actions to contact the enumerators and talk to them about what was going on.

And sometimes there are legitimate explanations for what they were seeing but it allowed us to monitor this in real time and that was very successful in terms of reaching the enumerators and talking through what we were seeing and ensuring that we weren't seeing something egregious occurring in the field.

Man: Okay I think you just explained how they were supposed to work but have you done any work to analyze whether or not they did work when they were supposed to work?

Maryann Chapin: I think some of that analysis is going on now.

Deirdre Bishop: Dan.

Dan: Just a couple of clarifying questions. On the bring your own device panel I just wanted to clear up a my own mind. You can use either an IOS or an Android device but what is the minimum specification where you - any idea how it broke down between IOS and Android or were they required to have a table class as opposed to working on the smaller iPhone size?

Maryann Chapin: So let me see if I've got your question. Are you asking me do I know the break down between the IOS and Android in terms of the participants? I don't have those numbers readily available with me and did we allow for a tablet?

Dan: A minimum standard for device...

Maryann Chapin: Right, I don't know those off the top of my head. I believe they went back - I'm trying to see if I can see a face in the audience that might be able to give me that information.

It may have been like a couple iterations older than maybe the most current operating system available but again I would need to confirm that. Brian may know he looks like he's going to...

Brian McGrath: So Dan, I don't have the specifics but we did not require the latest greatest device. There was several iterations of, you know, N minus 4 I think comes to mind for me of what was available.

So it was a pretty wide inventory of commercially available devices.

Dan: Okay. And then the second one on the evaluation followup component of this which will be starting up. What's the introductory script? What do you tell the people that you're going back to who have been contacted a while back of why people are back knocking on the door again?

And then is the exact content of the interview after that the same or is there any variation of the coverage probe questions in particular that you go through in the evaluation followup interview as opposed to the initial?

Tom Mule: Yes, I don't know the introductory script so we can look into that and provide it to you. But one thing I can clarify we do have some more additional questions relating to their residence that were not asked as part of the coverage followup questions so we do do that.

Since we are interested in whether the unit was vacant on census day or not we do have some additional questions related to vacancy that were not asked. And one of our followup groups is where we had an administrative records response but the non-response followup response was provided by a proxy respondent.

So those instances we're trying to go back to a household member and try to get the interview with the householder so that makes one group of cases where we're not going back a repetitive time.

But I mean you do bring up a good point we try to think of those since we are going out again for another visit we try to take those into account while designing the survey.

Deirdre Bishop: How many people had an opportunity to go out and observe in Maricopa County in the heat? Many people in this room and how many people have a much finer appreciation for what those people in the field have to go to get those interviews done?

Yes thank you to the team. I'm going to try to keep us on schedule with a presentation about our upcoming 2016 Census Testing activities. And Chuck I'm going to move right into slide 2.

We have two significant tests planned for fiscal year 2016. First the 2016 Census Test and second the address canvassing test. Our 2016 Census Test as I mentioned earlier we're going to begin moving from small scale individual tests that we've conducted over the past few years where we were looking at proof of concepts and prototypes to much more refined tests as we move forward and the building of the systems that are going to help us support the 2020 Census .

This test again will have an April 1 census day to mirror what will occur in 2020. We've selected two urban areas for this test. First within - we're going to conduct the test within contiguous areas within Los Angeles County, California and Harris County, Texas specifically in Houston.

I was really pleased to see those two maps outside the door that are blown up showing the test areas so if you haven't taken a look please do so on your way out.

Why were these two sites selected? Well we had a lot of objectives that we wanted to meet. These two urban locations afforded us the opportunity to test our objectives.

Los Angeles is the second largest city in the United States with a population of 3.8 million according to our ACS data and Houston is the fourth largest city with a population of 2.1 million people.

These areas contain large degrees of language diversity. During the 2015 Census Test we had to de-scope our work related to the language program but we really need to ramp that up during our next few tests.

Both of these areas contain language diverse populations with strong concentrations of people who are limited English proficient. In Los Angeles 16.3% of the population is limited English proficient and in Houston it's 13.59%.

These areas have strong degrees of demographic diversity, a good mix of Hispanic, Asian, white and black populations for us to do our testing. In Houston we have an extremely high vacancy rate of 12.3% as compared to the national average of 7%.

Los Angeles is very close to the national average of 6.8% but the high vacancy rate is good for us in terms of testing the use of administrative records and removing those vacant housing units from the non-response followup workload.

As we've studied in other tests we are extremely interested in knowing the degree of Internet usage within the area. In Los Angeles and Houston we have about 79% of the population that has access to high speed Internet. This data too came from the ACS.

Okay so the next slide Chuck. You'll recall that during our 2015 Census Test we focused on self-response during the optimizing self-response test in Savannah.

We focused on non-response in our test in Maricopa County. This 2016 Census Test is going to bring those components together. Our self-responsive objectives will include

providing support for respondents with limited English proficiency and how do we plan to do that?

Both in terms of the way we contact people so via the invitation letters and the postcards we'll do that in multiple languages. And also we'll do that via the response options that we provide, so on the Internet, on paper, during non-response followup we'll offer the option to respond in different languages.

Right now we're planning to support the English, Spanish, Chinese and Korean languages. In addition we're going to use partnerships to reach the demographically diverse populations.

We know from the 2010 Census and from our preliminary work in Savannah that these partnerships are critical to reaching those harder to count populations.

I'm happy to report that for the first time we're going to use text messaging to reach out and help motivate self-response. We're going to continue to refine our non-ID processing methodology.

We've seen very promising results thus far and we want to continue to see what we can do in that regard. And finally we'll test the use of a cloud based infrastructure for the first time for both our self-response and our non-ID processing options.

The 2016 test will be the one that allows us to operationalize our new methods and new technology across multiple locations and time zones during the non-response followup operations.

Our objectives to non-response followup include determining our strategy for the 2020 Census . How are we going to use administrative records to reduce the non-response followup universe and to determine the number of contacts that we're going to make with each non-responding household?

We'll continue to refine our field management structure. What is the appropriate ratio of enumerators to local supervisors of operations and local supervisors of operations to our field managers?

I think it's pretty well known that in the 2010 Census we had a ratio of 8 enumerators to 1 crew leader. During the 15 test we had a ratio of about 16 enumerators to 1 local supervisor of operation. We're going to see what is the best ratio.

What else? We're going to enhance the sophisticated operational control system that's used to manage our caseload and we're going to make refinements to our COMPASS application, the application that we use to collect the interviews on the handheld device.

(Ty) alluded to the problems that we experienced in multi-unit structures. We talked about this the other day. Many of us think of multi-unit structures especially those of us who kind of grew up in the New York regional office as those high rise buildings.

But in Phoenix and especially in the Mesa area we saw many gated communities. And so when we send our enumerators there what is the best way to do that?

Should we send one enumerator, should we send two? We've learned that we should definitely send too many. It's not good in terms of making relationships with the community management and it's not good in terms of ensuring stability within the census framework.

During this test we're also going to really study quality control to help identify possible falsification that could be occurring or errors that perhaps enumerators are making.

We'd like to do that in real time and fix the problems sooner rather than later. More specifically we'll use data that's collected during the interview. First global

positioning system points, GPS points to help us determine where the enumerator actually is during the interview.

And second, the time of the interview and the length of the interview. All of those things can help us determine problems that may be occurring in the field.

We have not had re-interview functionality built into our application in conjunction with the actual interviewing capability. We'll do that for the 2016 test.

And finally we'll test the use of administrative records to help us with in house quality control as part of this test.

Now let's turn to the address canvassing test. **Evan** and (Mike) and (Pat) talked about the research that we did as part of the address validation test and all of that is helping inform this address canvassing test that's planned for the fall of 2016.

I think it's important to mention that this test will stand both fiscal year 16 and 17. We'll start in September of 16 and then move into October. Our objectives here include implementing our new methodologies for both in office and in field canvassing.

How do these two operations work together? We'll use the listing and mapping instrument, one of our enterprise systems for collecting addresses when we're in the field.

We'll test the use of a new geography to collect our data. The use of a basic collection unit as opposed to our traditional census geography. Finally we'll use the data that we collect during this test to update our math TIGER system.

In terms of the location this test will be conducted nationwide. We'll do infield address canvassing across the nation, we'll also do in office canvassing across the nation.

With that said we're interested in building the foundation for our 2017 test. So we're going to focus on specific areas to help us do that. Build a good frame in these areas, use those same areas when we go out in the field in 2017 for that test.

So we'd like to focus on an urban area ideally one that's been a participant in our geographic support system initiative to see how data that they've been providing over the decade informs the building of the address list at this point.

We'd like to again include representation of our limited English proficient populations and we'd like to consider the degree of connectivity in terms of Internet connection that we have there.

We will focus on two Indian Reservations to help us with the question of tribal enrollment and we'll also focus on Puerto Rico in this test.

Finally as we refine our operations and we move from research and testing to implementation we're beginning to do the same with our system. As I mentioned earlier we're now moving from our proof of concept, our prototype systems, we're adding functionality as we go along and we're building the systems that will be used for 2020.

This slide does a good job of making the connection between our operation and the system we'll use. Just to give you a few examples of the enterprise systems that we're using to help us with address listing we'll using the listing and mapping instrument.

To help us with Internet response we'll use a Primus. Many of you had heard that we've been working with Centurion. We're considering the use of a new application to help with Internet response.

We'll use ICADE our integrated and capture and data entry system to help with paper response. We'll use a multi-mode operational control system to help with non-response followup.

And then we're using some systems that were developed specifically for 2020. Our non-ID processing systems as Frank talked about earlier and then COMPASS to help with conducting the interviews.

So that in a nutshell is an overview of what we have planned for 2016. On June 30 a press release went out. We have our Internet site updated with frequently asked questions about the two tests and then again we have the maps outside the door so please take a look.

Any questions about 16? Brad.

Brad: I didn't hear you mention training for the 2016 Census Test. Does that mean you're going to replicate the training approach you used in 2015?

Deirdre Bishop: We will continue with the use of automated training. We do have lessons learned and we are going to apply those to the training development and the actual implementation of the training.

Brad: Okay that was my next question. You're going to use the knowledge six and the focus groups from BYOD and the debriefing session?

Deirdre Bishop: Definitely yes.

Brad: Okay.

Deirdre Bishop: Other questions? Sharon.

Sharon: I'm just curious because you're now wrapping up so you're going to be needing greater funding in 16 and 17 than you've been using or existing in the past.

How are those years funding looking to you with the past back information that you have as well as what Congress is debating about giving you for 17?

Deirdre Bishop: So I'll start by saying over the past few years we've done a very good job of sharing what we've done very well. With limited funding people have worked very hard focusing on our key innovation areas.

The areas that we've talked about today, reengineering, address canvassing, optimizing self-response, using administrative records and reengineering field operations.

I think that our progress over the past few years has been very good and it's really due to the people that are sitting in this room. People have been working very hard, they're very committed to the job and they've ensured success for these census tests.

With that said we've put things to the side. We haven't focused on some of the operations that are very important to the census. It's time now that we need to focus on those operations.

I talked about how we de-scoped the language program in 15. We have to get back to that in 16. We have not worked on group quarters operations those populations that are related to the military that live in nursing homes, college dormitories. We have to focus on that now.

We really have to begin working on our legally mandated programs so the redistricting data program, the local update of census addresses. We have to make those happen in FY 16 otherwise we're going to fall behind.

And as I mentioned it's critical that we begin to really move from our proof of concept, our prototype systems to the actual systems that we'll use for 2020. Many people say, you know, what is your vision what do you think?

I have to be ready for 2018 and to end test. That is almost my census right. We have to be good to go, we have to be able to know the systems we use in 2018 are those that we're going to use in 2020.

We only have time to make minor adjustments to those systems to be prepared. With that said we could potentially take significant cuts to our budget.

We are right now working with our budget shop in house, we're working with the department with the Office of Management and Budget and the impacts on the risks and I'll leave it at that. Dan.

Dan: Just one quick thing. When you said - it may be a grammar question more than anything else but when you say refinement to real time non-ID processing methods including respondent validation does real time apply to respondent validation?

Is that trying to implement that person check in line when they're trying to reply or is respondent validation still being conceived of or at least for the 16 test as being a post hoc check?

Deirdre Bishop: I'm going to give you the 10,000 foot answer and then I'm going to look to Evan to provide more details. So yes we are looking at how to validate respondents in real time as they're responding to the census. And now I'll turn it over.

Evan Moffett: So from a real time respondent validation perspective we have put a request for information out on the street and we've done an analysis of the results of that RFI.

I think in general the recommendation is that respondent validation be a combination of real time and post process. In addition to that we have begun an engagement with a group that will provide us with some recommendations on the most efficient and effective way to do this on October 1.

So at that point we will take all of those recommendations and make some decisions about the methodologies that we will employ in the 2016 Census Test to do this in the most efficient and effective way possible.

Deirdre Bishop: Okay we're right on schedule so I'm going to keep us moving. I'd just now like to wrap up what we heard today to summarize what the panel has shared with you.

First, I think it's very important to stress that we are on track to deliver our initial 2020 Census Operational Plan the baseline by the end of this fiscal year. Ann mentioned that in the past we didn't do this this early in the decade.

So in relation to the 2010 Census we did not release our Operational Plan until 2008. In relation to the 2000 census we released the plan in 2000. So we are well ahead of the game and I'm really pleased about that.

In answer to Ty's question I mentioned that we will not only be sharing the narrative what we're calling our working papers as we announce our Operational Plan but we will have a slide deck library to make it a little bit easier to get through the meat, the substance of the documentation.

In terms of the address validation test we learned that our models were not as accurate in predicting areas of change as we thought they may be that we really need to turn and look more closely at those in office methods such as the review of imagery, the use of commercial and administrative data sources and that in certain areas in field address canvassing is still necessary.

You heard that we will conduct another test of partial block canvassing as part of our 2016 address canvassing test. We learned from the optimizing self-response test in Savannah that the Internet continues to be a very effective mode of response.

Our Internet push strategy where we invited people to respond to the Internet before doing anything else is working. We've seen a higher response from those people that we shared a census ID with, very interesting finding I think.

Notify me is not showing real promise and finally we are very effective in doing real time matching for those non-ID cases coming in. In relation to the 2015 Census Test we've learned that administrative records are proving very effective in helping us identify those vacant housing units and in reducing the non-response followup workload as well as identifying race and Hispanic origin data.

I'd like to stress here and Tom mentioned this, we have a very strong relationship at this point with the Internal Revenue Service and with the United States Postal Service and we're very appreciate for that and for their help.

Automated training is proving really effective as is using technology to manage the caseload. We know that everything wasn't perfect. We did have as Ty mentioned a lot of lessons learned, a lot of findings, you saw that when you were out in the field.

But the work that we've done has shown great progress. We know we have a lot more to do but we are at a very good point.

Finally in terms of the 2016 tests we have these two tests gearing up now. Not only are we going to analyze what we learned in 15 we're going to apply it and we're going to plan it and implement 16.

With that I'd just like to say one more thing. We're really interested in knowing what you think about this new format. So do you like the setup in the auditorium versus the conference center?

Please let me know after the meeting or at a later time. The conference room afforded us with a little bit more intimate or conversational setting. This seemed a little bit more

formal but it also felt a lot cooler, a lot more - we had a little bit more space to work with. So let me know about that.

And also what do you think about the half day meeting as opposed to the full day? Somehow I always seem to get the first presentation or the last presentation of the day and the past two PMR's at 3 o'clock when I looked out at the audience everybody looked like they'd had enough.

And I think the panel felt like they had enough but today people seem to look happy and fresh. So let me know what you think about the half day setting as well.

And thank you again we know it's the week after the Fourth of July I was afraid we wouldn't have a good turnout but we did. We had a great audience. Thank for your questions and your continued support.

END