

First Meeting of the Commission on Evidence-Based Policymaking

National Academy of Sciences Building
2101 Constitution Avenue, NW • Washington, DC 20418
July 22, 2016

Katharine G. Abraham: We've got a pretty crowded agenda. But we should go ahead and get started. It's my honor and privilege to open this first meeting on the Commission on Evidence-Based Policymaking.

And I'd like to thank the commission members who are here, our distinguished and the speakers that we'll be hearing from this morning and members of the public who are joining us in person and on the phone.

Maybe I can begin. One of the first things we want to do is go around the room and have the members of the commission introduce themselves. Maybe I can start by introducing myself.

I am Katharine Abraham the chair of the commission. I'm a labor economist by training and background, but actually spent most of my career in academia, mainly at the University of Maryland. But I've also been able to enjoy spending a fair amount of time in government.

I spent eight years at the Bureau of Labor Statistics, was the commissioner of Labor Statistics. I know most of you are familiar with the BLS the agency that's responsible for the production of data on the U.S. labor market and other statistics that allow policy makers to monitor the state of the economy.

More recently I spent two years at the Council of Economic Advisers but one of the things I most enjoyed was the chance to become involved in the federal government's ongoing efforts to strengthen the use of evidence to policy.

So I'm very happy to see you all here. This group of commissioners that has now as of Wednesday been fully assembled is an exceptionally accomplished and distinguished group that brings a diversity of valuable perspectives and experience to our task at hand. I'm really looking forward to working with all of you.

I've spoken with many, not all of you, but many of you in advance of the meeting and I know that all of you share the belief that sound policy really has to be grounded facts and in evidence.

Facts about the social and economic conditions prevailing in our society and evidence about those consequences of the various policies that might be adopted in an effort to improve those social and economic conditions.

The creation of this Commission on Evidence-Based Policymaking really creates a unique opportunity for all of us to focus contributions to improving the efficiency and the effectiveness of our government.

If we do our work well and I'm confident we will. I think that we can provide a practical blueprint of creating a data infrastructure that enhances the information about social and economic conditions available for policy makers and allows ongoing evaluation to become a more routine part of the management of federal programs.

We have, as you know, a lot to do in a short period of time but I - I think we're up to the task and I really look forward to working with all of your to carry out the charge we've been given. Ron.

Ron Haskins: I'm Ron Haskins. I'm at the Brookings Institution. I've been here for 15 years now. I run a center called The Center of Children and Families. My discipline is developmental psychology.

I was talking to someone this morning. I once (unintelligible) which is the primary term in the field for something like 8 years of studies and added up the ends.

And the average number of subjects for a study was on the order of 550-something. It's not like 45. And now we have studies (unintelligible). So things have changed quite a bit.

Previously I was with the Ways and Means after leaving North Carolina, the Ways and Means Committee group, 15 years. I had many adventures working on legislation especially once Republicans began - Republican staff there and once Republican became the majority life changed quite a bit.

And we passed lots of legislation. So I have a fair amount of legislative experience. I think I know a fair amount about how Washington works. At some point hope that will be useful to discussion...

Operator: The host has a recording feature on.

Ron Haskins: I said something that shouldn't be recorded. I don't know what it was.

((Crosstalk))

Ron Haskins: So and then I was in the White House for a year and then at Brookings since then. I came to Washington in 1985 from UNC for one year on a fellowship. That was 30 years ago. And my intention was to study how social science evidence can be used to improve public policy for children. That was my goal.

And I had been here maybe 10 minutes or so and figured out it would be the shortest book ever written because the evidence was not exactly a major issue in Washington. And it still is not - well it should be.

But the things have changed dramatically. And the one substantive thing - the impression that I want to leave you with is that we are operating in fertile soil for sure. Things have changed dramatically. There are many, many branches of the evidenced based movement now.

Recently the Association for Public Policy Analysis and Management titled their conference this last year. The golden age of evidence-based policy. And I think it's true. We have great things going on.

Many of you may know about the results first initiative by Pew-MacArthur where they're actually trying to get states to do - which is the most fundamental thing I can think of in evidenced based policy which is stop spending money on stuff that doesn't work and spend it on programs that do.

And they are known to work because of good usually random assignment studies. And so we need to - they actually have more than 20 states with senior staffers and agencies that go through their programs and find out all the evidence on the programs that they're using.

And then compare that with other programs that they're not using and what the evidence on those programs is. And as a result many states are actually

changing the allocation of their dollars which is one of the most fundamental goals I think of evidence-based policymaking.

So and there are many other examples of that. I just want to mention to you one - we're very fortunate that OMB is a close associate as well as the Census Bureau. But OMB especially in the Obama administration has been a spectacular leader in evidence-based policy.

And they have their fingers in everything. They're great managers, tremendous background on ... I don't know if you've had a chance to read the background papers yet. But I give you an example of how great OMB and the Census Bureau are.

So, OMB is really in many ways the quarterback in the evidence-based movement. And I was fortunate to be a member of a group of people here in Washington behind the (unintelligible) foundation that do what we could to make sure that the next administration will take over where Obama leaves off.

And expand it even more so we wouldn't lose any of these great evidence-based activities that are now ongoing. So I guess I think it's a good thing for you to know that there really is, there's a big demand for what we're going to do.

And if we do our job properly which I, given the - I have no doubt that we will, given the perfect composition of this group - that will really be - it will enhance the evidenced based movement greatly.

And probably the major thing that we need to do or certainly one of the major things is make it easier for researchers to get access to massive sets of data from the federal government.

Katharine G. Abraham: Great. What I'd like to do next is were are all getting to know each other. So if you could go around and have the commission members say a few words about their background. And if you have any particular thoughts about things that you think the panel should be focused - the commission should be focused on this is a good time to mention that as well.

So if you don't mind Bruce we'll go ahead and start with you.

Bruce Meyer: Sure. I'm Bruce. I'm Bruce Meyer. I'm a professor at the University of Chicago where I've been for a dozen years. Before that I was at North Western University for 17 years.

I'm a public finance and labor economy by trade. I've spent most of my career, well even before I was professor, using administrative and survey data to evaluate government programs to try and figure out what their effects on people and firms are.

And with the goal of trying to figure out what works and what doesn't and how we can improve the function with government programs. And I'm expecting that this commission will do all kinds of good things to push the evaluation of government programs and in addition more fundamental policy research forward.

Katharine G. Abraham: Bob.

Robert Hahn: Good morning. Bob Hahn. I'm a professor at Oxford University in the UK. And my background is in microeconomics studying the field of regulation, focusing lately on some environmental issues.

I became very interested in what I thought was evidenced-based policy about 15 years ago when I helped start something called the AEI-Brookings Joint Center for Regulatory Studies with Bob Litan.

And now I work on doing a lot of field experiments with people at Oxford and people at Chicago. And I'm very interested, as I think Bruce is, in seeing how we can - and Ron based on his remarks. Seeing how we can promote this within government and in the private sector and at least at a minimum get a better information base upon which decision makers can make decisions.

I personally would like to move decision makers more towards weighing cost benefit analysis broadly construed sort of in a Cass Sunstein framework to consider that as a major criterion in their decision making for public policy but I think that's a separate issue. Thank you.

Sherry Glied: Hi. I'm Sherry Glied. I'm currently the Dean of the Robert F. Wagner Graduate School of Public Service at NYU and an economist by training. So you can hear a theme coming out here. My research has been in healthcare economics.

I had two stints in government. The first one way long ago back in 1992 and 1993 I was a senior economist in the Council of Economic Advisors in the waning days of the George Bush administration and the beginning of the Clinton Administration.

I just served two years in 2010 to 2012 as Assistant Secretary for Planning, Evaluation at HHS where we did a lot of work around evidence-based policy in a couple of important areas.

I'm actually particularly interested in how we can make routine use of evidence not only in evaluating policy but also in the way the regulatory process and the

process of implementation and quality improvement within government continuously - on a continuing basis rather than on a one off basis as well as of course the use of data research.

Allison Orris: Hi. I'm Allison Orris. I'm currently the Associate Administrator at the Office of Information and Regulatory Affairs, OIRA, within the Office of Management and Budget.

I am not an economist by training. I'm a lawyer by training and practiced law for several years focusing on healthcare before moving into healthcare policy. I spent about five years at the Department of Health and Human Services working primarily on Medicaid legislation, policy, and implementation.

And then since coming to OMB I've worked primarily on regulations but one of the main roles that OIRA supports is a team that works very closely with statistical and oversees and coordinates with the U.S. statistical systems.

So I've learned a lot about that field and we also - and I appreciate Ron's remarks about the evidence building work that we at the administration have done. We have done a lot of work in that area.

And I am excited to be part of this commission to continue to support that work and to think about new ways that we can build linkages and use evidence when policymaking.

Paul Ohm: Hi. I'm Paul Ohm. I'm a professor of law at the Georgetown University Law Center. I just arrived a year ago having spent a decade in Boulder, Colorado, on the faculty of the University of Colorado.

The weather's different back there. And I've had two stints in government myself. I was a computer crimes prosecutor at the Justice Department about 15 years ago where I thought about evidence in a different way and we would read email and access people's credit files.

And then more recently I spent one year as a senior policy adviser at the Federal Trade Commission. An agency that really does think a lot about economics and evidence in other ways.

I am a lawyer by training not an economist but I do have under graduate in computer science. I have spent time building and defending networks and also helping to do some database management.

And I think some of those admittedly amateurish scales to bear. My primary focus is on privacy scholarship and Ron started by talking about the golden age of evidence-based policymaking. In many ways we are also in a golden age of information privacy law scholarship and thinking.

And one of the things I hope to do for this commission which I'm quite excited to be part of is to really bring to bear the latest and most sophisticated thinking by other people not by myself about what we mean when we use this very complex term.

Privacy, what are some of the tradeoffs with utility because unfortunately there is an intrinsic tradeoff. And, you know, how you can try and have the most of all worlds but never the best of all worlds. How you could have utility and some policymaking and evidence but at the same time protect the rights and respect the people involved in the data.

So once again thank you. It's very nice to meet all of you. And I look forward to the work of this commission.

Kathleen Rice: Good morning. Such a pleasure to be here. My name's Kathleen Rice. I'm a counsel at Faegre Baker Daniels in Indiana. I'm also a lawyer. Starting a new trend over here.

I help lead the firm's privacy and data security practice. And before that I served for eight years as the counsel on the Senate Intelligence Committee where I was involved in a range of intelligence and legislative issues, including on data collection.

Prior to that I was an assistant general counsel at the Federal Bureau of Investigation. And, like Paul, I was a federal prosecutor where I also worked on computer crimes and child pornography and other issues all dealing with evidence and reasonable expectations of privacy.

So I share Paul's comments. And it's a pleasure to be here. Thank you.

Latanya Sweeney: Hi. I'm Latanya Sweeney. I'm a computer scientist, starting a third trend. I'm a professor at Harvard. And a lot of my work has been in the demonstrating privacy risk of current practices and why they're vulnerable and why the explosion in data in today's society makes us have to think about privacy in different ways.

But also the computer scientist developing innovative solutions to - so that we can share the data widely but with privacy guarantees. I too served in government. I was the Chief Technology Officer at the Federal Trade Commission where we'd address all kinds of issues and all kinds of evidence.

This was pointed out as well a moment ago.

Kenneth Troske: My name is Ken Troske. I'm a Professor of Economics at the University of Kentucky which is what we think of in Kentucky as UK. So this is - I have had a couple of stints in government as well.

I will say the last time I was on a federal commission I was appointed to the Congressional Oversight Panel which oversaw the Treasury spending on the Relief Program Fund.

And I was - there were five of us. And I was - there were four lawyers and me. And so I was the lone voice in the wilderness. I'm glad to see there are other economists on the commission.

I started off - I've used administrative data for most of my career as an economist. I started off using EI wage record data from Wisconsin on my dissertation. And I can actually literally recall going up with my dissertation advisors to the folks in Wisconsin and them given us their reel to reel tapes and us bringing them back and using them.

And I had them for a long period of time. I don't know that they knew that. I then - once I finished my dissertation went to work at the center of economic studies at the Census Bureau.

And I spent a number of years there. And I'm familiar with many of the - I still remember a few of the people that are at the census. And started there when we were using - starting to link data together.

The reason I went to the census was because my goal, and I accomplished it, was to link decennial census data to the economic census data. So matching

workers in the firms. I created one of the first cross-sectional databases that did that.

And I left the Census Bureau and went to University of Missouri and then now at the University of Kentucky we continued using a variety of state level administrative data.

And I'm well aware of the challenges. So I had access to UI Wage Record Data from Wisconsin on a case and then I've had my access from Kentucky pulled.

I've been getting the data for years. And then some administrator somewhere decided that well maybe we don't want you to have access to it anymore for no apparent reason.

And so I am well aware of the challenges of accessing this data and security concerns. And one of my goals hopefully is to figure a much more systematic fashion where we can use - I was involved in the study of WIA on which we were supposed to get data from WIA data from all 50 states.

And the letters from the Department of Labor went out and said please help these people. Because they're doing an evaluation of a very important program. And we got data from a grand total of 12 states despite the Department of Labor's best efforts to convince the states that this was an important thing to do.

And so one of my goals is to figure out how we can ensure that there is some uniform rules regarding the use of this data. Because, you know, we wrote a fairly influential paper about the performance of WIA.

It sure would have been nice to have data from 50 states as opposed to 12. So this is something that I believe very passionately about. I recognize the importance of protecting the privacy of individuals but these data are key if we're going to understand how these programs function.

And we are going to continue to design. And it's not about killing - necessarily about killing inefficient programs. It's about making programs work better to serve the people that they're supposed to be serving.

If you really care - if you really think these programs are important that should be what you're interested in. Unfortunately my own experience has been many people who have been administering these programs are not interested.

Are more concerned about simply getting the programs running on day to day funds as opposed to thinking about ways to improve it. I'm hoping we can make a little progress.

Kim Wallin: Good morning. I'm Kim Wallin. I currently am the CPA with my own practice in the state of Nevada. I say that because I have an office in Las Vegas and Reno. I'm certain Reno and Las Vegas were a lot cooler. I'm dying here right now.

I think I'm probably the only accountant on the commission. So I'm very alone. It's not to know that you're not alone anymore.

Kenneth Troske: Yes it is.

Kim Wallin: And I think you're probably thinking, why is she on this commission? What does she have to do with that? Well I spent two terms as Controller for the state

of Nevada. And as Controller for the state of Nevada I happened to come in 2007. So right before we had the Stimulus Act and all that.

And when I first came into office everything was great and I was finding out that our agencies in the state would come to the legislature and ask for money. And the legislature would say well what have been the results of this program when we need more money? Well I don't know. We have to go to the contractor?

And I thought that should be the first thing that they're coming forward with is, all right, what are the results with (packing) and stuff. So I became a big proponent on let's see what the evidence is and what is actually happening.

And then we started into the Stimulus Act. I was very active in coming back as to how should we bring the data forward from the states of the federal government. And it all happened pretty quickly then. The proponent of what is called XBRL which is Extensible Business Reporting Language.

The FDIC currently uses it. Treasury is starting a program and using it in the financial area. And the FCC of course is using it as well. And it allows the data to become machine readable. But that's something that I think is important here is how can we gather this data and make it available for everyone to use and that might be a format to go and do. Some of the things that I'm very passionate about here is the data quality to make the better decisions. Standardization of that data so that way we can have better comparisons of apples to apples.

And that was something else I found at the state is that when an agency department we call fuel, fuel and everyone would call it supplies. And, you know, that's a pretty deep dive down there. But this is evident throughout.

And then also I - I'm also a big proponent of the Data Act. I came back and testified on the Data Act from the very get go. And I know that it's been passed. So I think that there's some opportunities for the commission to maybe look at that. How we can kind of interface with that as well as we go forward.

I know that the Data Act is your financial data and your administrative data but that's part of it. And I think that Robert you were talking about the cost benefit analysis and I think that that's something that is very important for us to do.

And one of the things that the cost benefit analysis it's great to say this but we also - we have a social side of it. We have to bring in the thought skills there. And I am truly honored to serve on this commission with all of the economist and attorneys here. So, Thank you.

Katharine G. Abraham: Jeff.

Jeffrey Liebman: Hi. Can you hear me?

Katharine G. Abraham: Yes. We can hear you.

Jeffrey Liebman: Excellent. Excellent. Good news. It's great to be here at least remotely. And I look forward to being in person next time we convene. I think there are three things in my background that are worth mentioning.

So I'm a professor of public policy at the Harvard Kennedy School. My research, a very low share of it, involves administrative data. Indeed I'd like to acknowledge that someone else went all the way back to their dissertation. I'll do the same thing.

When I was working on my dissertation I spent a summer at the Census Bureau using data that matched the CPS to extracts from tax returns to figure out what fraction of the earned income tax credit dollars were going to poor families with children and what fraction because of errors or fraud were being siphoned off to families that were not the target of the program.

I was using the census data. I worked in the Census Data Centers, the remote ones with the SIPP matched to the Social Security earnings histories and benefit records. And also I've done several randomized experiments in which we matched UI records or other records to figure out the measure of some of the key outcomes.

Second thing about my background is that which is potentially relevant is that I served in government twice. During the Clinton administration, I was at the National Economic Council coordinating the Social Security Reform Working Group.

And at the beginning of this administration I spent two years at OMB in various leadership roles and played a small role in the work that Ron so generously described earlier.

And then for the last five years I've been doing something somewhat different. I've built a team here at the Harvard Kennedy School that I call the government performance lab. It's now 40 people, mostly recent graduates of public policy schools who are employed full time.

And put them on the ground in state and local governments around the country to help social service agencies in those state and local governments use data to better administrative programs in real time.

And so rather than solely retrospectively evaluating problems we're helping government use their administrative data systems to meet with service providers weekly or monthly around data to figure out how to re-engineer the system and get better results for multiple populations.

And so I'm very focused in how we can make federal data available to state and local governments to better run their programs. Not just to evaluate it after the fact which is very important.

And certainly I'm in favor of that as much of that as possible but also helps to in real time figure out which providers are doing better than other providers, figure out when systems can be improved.

Katharine G. Abraham: Great. Thank you. There are a couple of other commission members who are planning to join us for at least part of the day. But I don't know if they're on the phones yet.

Eileen LeFurgy: They are logged in to the WebEx but they haven't set up their camera. So I'm just - I'm waiting for them to - is the only one that's really - they're logged in. So they're probably listening.

Katharine G. Abraham: Hilary, are you on the phone?

Hilary Hoynes: Good - I am. Good morning.

Katharine G. Abraham: Good morning.

Hilary Hoynes: Calling in from the west coast. Bright and it's early.

Katharine G. Abraham: Right. I am impressed. Very bright and early.

Hilary Hoynes: Good morning. So thanks. It's just such a pleasure to be serving on this committee and I look forward to meeting you all in person at the next meeting. I am also an economist. And my appointment is at the University of California at Berkley.

And I have appointments both in the economics department as well as in the Goldman School of Public Policy. And throughout my career much like Jeff and Bruce and Ken have worked a great deal with administrative and survey data.

My experience with administrative data really spans those federal and state and local sources. I also have been working with the - had experience with the remote RDCs for accessing sensitive data.

And my research generally focuses on tax policy and government programs and their effects on poverty and enequality and quality with a particular interest in expertise for the disadvantaged population.

And then I come in to this commission with a great deal of enthusiasm and optimism about opportunities for - just to reiterate what others have said, for, you know, improving what we learned for the functioning and improvement and implementation of policy at all levels of government.

But I also come into it with a great amount of interest and optimism in moving us forward towards harmonizing linking and providing secure access to the incredible resource that we have at our disposal.

Through these administrative data sources that are sitting and in existence. And I have a kind of particularly interesting sort of spin on this I think. For the last

six years I have been a co-editor at the American Economic Review which is sort of the premier academic journal run by our association the American Economic Review.

And as such and as a co-editor that handles a great deal of the work that is submitted in the area of sort of applied microeconomics and policy. In that short period of time in the six years there has been an incredible shift so that essentially almost all of the work at the frontier in our science is being done in administrative data.

And it's quite interesting to see that the lack of access and availability of these linked administrative data sources in the U.S. really puts us at a disadvantage at that frontier.

So the vast majority of papers that I handle now are using from Germany, Denmark, Sweden, Norway, Spain, Portugal where Chile other countries around the world they're really, you know, have these clearinghouses and data repositories that have furthered their - what they do.

Besides governments that also very notably at the scientific frontier. So I'm very excited about taking the opportunity for moving us forward through the functioning of these programs and also to further science.

So thanks very much.

Katharine G. Abraham: Great. Thank you Hilary. This meeting has obviously attracted a good deal of interest from a number of people. Maybe we have a few minutes. Maybe we can just go around the outside of the room and to the back and people just say who they are and where they're from. That would be great.

I think we should start with the people who are here with us at the table who are not commission members and then start and go from you, Connie.

Jeri Mulrow: Good morning. I am Jeri Mulrow. I'm the acting director for the Bureau of Justice Statistics which is the federal statistical agency within the Department of Justice. I'll be speaking later today on a project that I think you might interesting.

Nancy Potok: Good morning. I'm Nancy Potok. I'm the deputy director of the U.S. Census Bureau.

John Righter: Good morning. My name is John Righter. I'm the deputy staff director of the Health, Education, Labor and Pensions Committee here for Senator Patty Murray.

Devin O'Connor: I'm Devin O'Connor from the Office of Management and Budget.

Kate McNamara: My name is Kate McNamara from the Census Bureau.

Mary Bohman: I'm Mary Bohman. I'm the administrator of the Economic Research Service which is a federal statistical agency in the U.S. Department of Agriculture.

Connie Citro: I am Connie Citro. I'm director of the Committee on National Statistics which is a standing unit here at the National Academy of Sciences.

((Crosstalk))

Katharine G. Abraham: Thank you all. I should be maybe add to these introductions that Shelly will be the Director for the Commission. So thank you. I'm hopeful to know at least some of the groups that are interested in the work of this commission moving forward.

So I see that Ted is not here yet. So maybe we can just - if you don't Devin, take these remarks out of order and start with you. Devin is the Associate Director for Economic Policy at OMB. If you'll give some opening remarks about the works of the commission.

Devin O'Connor: Sure. I'd like to start by thanking the commission for taking on this important work. And I should know it's one thing to have a really impressive commission which you do. It's another thing to have a really impressive audience. Which you also do.

I also want to thank everyone who worked to make the commission legislation possible and especially Ted and John and also to thank Census for setting up this first meeting particularly and all their preparing for the commission.

I thought I'd use my time briefly just to touch on a little bit about how the administration views evidenced based policy. And then also just a couple reasons that we have particular excitement at this time around the commission.

Evidence driven policy is an important part of the President's vision for modern government one that taps emerging innovation and technology to deliver the best results for the American people.

The President made clear that policy should be influenced by evidence. The evidence of what works and what doesn't. And also the evidence or where we should be applying resources, where our greatest challenges are and where our greatest opportunities lay.

We use three simple guiding principles to try to determine how to apply evidence-based policymaking in the government. First where we have strong evidence we act on it. Where we need stronger evidence we try to build it.

And when we see the opportunity to build capacity to make this work easier in the future we invest in it. You can see how we executed on those principles throughout this administration.

I think one really high profile example of that is the administration's embrace of innovating funding models like secured evidence grants. These grants both prioritize funding for scaling up the things where we have the strongest evidence but also continue to do the work of pushing evidence forward by investing in promising practices.

Testing, retesting new things, updating what we know and just keeping everything current so that it can continue to move forward. I want to call attention to the third principle though too.

I think that one's really important. Which is support for an evidence-based approach is going to require and continue to require investments in capacity building. Whether that's improving data quality and data systems, sometimes going out and getting more data, and often if you want to unlock the great value of that data it means making the data easier for use.

A good example of that is HHS and its Medicare and Medicaid services and their investments in their virtual data research center. This is the hub they use to get out information for Medicare and Medicaid on the claims beneficiaries.

A really important data set. And they've significant investments in making that warehouse both easier to access, cheaper to access - which is an important point. And also making the data that's available more recent data.

And I think you can see some of the motivation behind that is it's taken on a really important mission in trying to reform service delivery. And that's going to involve both a lot of research and a lot of work to try to drive towards what are the best ways of moving this forward.

And so kind of advancing data transparency is an incredibly important part of that and tapping outside resources and researchers and getting the information they need in order to drive that mission forward.

The administration's embrace of evidence driven policy had led to significant results all across government whether that's reducing the time it takes for students to apply for federal financial aid, increasing retirement savings for service members, reducing veteran's homelessness or making international development aid more effective.

And you can see it's not just in one area or one set of programs but across government that we're making the effort. And the President's budget includes a number of proposals to further that effort.

If the commission takes up this charge I just want to highlight two places where we have particular excitement. First is I think many folks and a lot of the folks in this room would probably date their time in pushing the evidence agenda forward in government in decades not years.

But if you take any backwards sort of looking review of where we've been and where we are now I think there's just a lot of reason to say we've made significant progress.

And I think one of the most striking aspects of that progress, one of the most notable aspects of that progress is just advancing sophistication with which the evidence-based community and the government discussed the use of evidence.

Which is moving beyond sort of the conversation that takes a single federal program and sort of passes or fails it and says this is effective, this is ineffective. And instead looks within programs and says what are the most effect ways we can advance what we're doing?

How can we cross apply the lessons that we're learning here and how can we take more cross-cutting and comprehensive action on the basis of that evidence?

Second, it does feel like we are a bit of an opportune time. And you hear a lot of people discussing this in different ways. But it seems like the ease of access to data or the cost of data has started to go down.

There have been technological advances. There have been other advances that sort of create this opportunity that combined with modern methods really does seem like you can lead to a new era in research and valuation.

Those things could get through. There could be a virtuous cycle going on to make things easier and make the leap forward really possible. Looking for additional ways to make better use of the data the government's already collecting regardless of the source of the data or what that data is, is always going to sort of have great potential for results.

I think administrative data though among that seems to have vast and untapped potential for doing that task. There's a whole broad range of analytical and management tools to provide the evidence infrastructure.

But with the administrative data in particular but I think what you see as a chance to get more data on more people, reducing the sort of - making it so that with a larger sample size you may be able to detect things that you wouldn't have been able to detect before and avoiding some of the costs that it would require to do kind of that same level of data collection through a survey.

So with that in mind it does seem like there's just a chance to really reduce the cost with what we're doing. And once you reduce the cost of what we're doing we can increase the frequency. We can increase the size. We can really start to sort of ride that wave.

And I don't want to sound too Polly Ann-ish because I want to acknowledge that both the progress we've made and often times either long or tough or piecemeal. And also that administrative data in particular comes with a lot of challenges.

Whether those are challenges of data quality or a need for data redesign or thinking through how we collect those data. In particular privacy protections and confidential protections that are going to go along with it.

But having said that unlocking that data has immense potential. So the one other thing that it's important to note, especially on the administrative data front though I conclude that oftentimes unlocking that potential is going to mean additional upfront costs.

And figuring our way forward on that is really important as well. So those are just two of the reasons though that the administration's embraced Senator Murray and Speaker Ryan's proposal. And see a lot - there's a lot of potential there.

So we're very excited about the commission's work and I wanted to end by thanking you all again for taking on this great task.

Katharine G. Abraham: Thank you. If people have questions are you willing to respond...

Devin O'Connor: Of course.

((Crosstalk))

Devin O'Connor: Are we doing them now or are we going to at the end?

Katharine G. Abraham: Why don't we do you ...then John and Ted.

Devin O'Connor: Sure.

Katharine G. Abraham: Does anyone have any questions for Devin?

Devin O'Connor: Oh, I'm getting off easy.

Katharine G. Abraham: You're getting off.

((Crosstalk))

Devin O'Connor: They saw Ted enter and they're holding on to them.

Ken Troske: Let me ask you something. Has the administration - does the administration have plans to work with the next administration? Especially on this issue of evidence-based policy?

You mentioned a few things this administration has done (unintelligible) I know over initiating all kinds of new things and it would be a real shame on the one administration to the next that they step backwards.

So have you thought about that? Is there group it's going to focus on? Do you need documents? You're going to meet with people and so forth?

Devin O'Connor: Yes. That's a great question. I would say - two things I would say to that. One is yes the administration is thinking very carefully about what we can do to sort of advance this work going forward. What we can leave behind. How we can help that transition be smooth.

And we will take those steps and do those things to try to make that happen. I think this also though is a good example of a way in which this commission will be incredibly important and the timeframe doesn't line up exactly right.

But oftentimes I think there's a risk that some of the things that were done during this administration - you don't want.

Depending on what the next administration comes in as and what their sort of relationship to the previous administration was often times things can get a little bit either pigeon holed or sort of treated as oh that was the thing they did. And we want to do our new thing.

I think the commission because of the stature of the commission when you come out with your recommendation, in a lot of ways what you can do is sort

of take the state of affairs that we have and for Congressman in the administration today look here are some places where we've made really significant progress.

We should continue to do those things. We should do those things more widely. We should build on those things. So I think it's both something we'll do as a matter of transition and one that we were given a lot thought and starting to spend effort and energy on. But also on behalf to this commission probably could be really helpful in further...

Ken Troske: I have one question about most of the work that most of the people have talked about here in the view side focuses on programs that focus on individual people. We have a lot of government programs that focus on each other.

I think businesses, economic development activity, my casual empiricism is that there's no effort to evaluate the - very little effort to evaluate the effectiveness of many of those program and don't seem to take in a lot of the ongoing research that is occurring in economics like (unintelligible) and others that suggest ways to improve those programs.

Could you tell me some of the efforts the administration - is this the focus - has this been a focus of the administration and maybe the other folks that have been in talk.

If you could comment on that as well because one of the things the Census Bureau did was push a lot of business data they have that would allow any type of research.

But it doesn't seem to be going on very much.

Devin O'Connor: Yes. I think it's fair to say that there are places where we've - where the evidence agenda and these evidence in government has advanced further than others.

I would say that the administration's efforts and the push for this has been to - has been across government and all programs and not just individual serving programs about how we're - in every issue and every way in which we're sort of trying to advance from where we are forward.

I do think there are probably others who will speak better to the efforts we've done. I think you've sort of pointed out the energy has often been around the data so far and not necessarily around the design of the programs yet.

But I think particularly within the last few years of the administration what you've seen is sort of things that have been developed and mature in some areas are percolating in others. And they're starting to come through.

And you're starting to solicitations that make better use both of the available evidence but it's certainly the case that we're - some people are further along than others. And some areas are further along than others.

Katharine G. Abraham: Thank you very much Devin. So I would like to turn now to Ted McCann and John Righter to talk a bit about the motivation behind this legislation and what you and your bosses hope will come out of the commission.

((Crosstalk))

John Righter: I'll give Ted a few more minutes here and I'll go ahead and start. Good to be sitting next to him, good to be here to talk to you. Thank you Co-Chairs

Abraham and Haskins and members of the commission for inviting me to appear here today on behalf of Senator Murray to talk at the inaugural meeting of the Commission of Evidence-based Policymaking.

The origins of the commission date back to the fall of 2013 when now Speaker Ryan and Senator Murray were involved in another important effort to try to show the American people that the government can actually work well for them.

They announced a bipartisan Budget Act of 2013 but those negotiations invariably would hit a rough patch. And there were many of them. They would take a pause and talk about other things just to try to reset the dynamic in the room. A lot of that would involve conversations about whether Senator Murray calls the three Fs, family, football, and fishing.

But it also spent a lot of time talking about, you know, what are the things we could do to sort of change the dynamic in conversations in other rooms around here in ways that makes the process work better for families and communities. You know, particularly in the time of deeply divided Congress.

As both of them are very institutional and want to work better. And from - born out of those conversations was the idea to create the commission here now. So as Senator Murray wrote in a rich statement that she's provided for today's hearing - and I'll read from it.

Ensuring policymakers have the tools to make the best possible decisions shouldn't be a partisan issue. It should be about doing the right things with families and communities we represent whether you think we need more government or less government.

To be clear there's a very healthy disagreement between the two principles on that very question. We should agree that we should at least have better government. And several of you made that exact point.

Ken just did and Devin, also. So, you know, just as a - I felt like, you know, Patty Murray and Paul Ryan were the right two people to get us out of the gridlock we were in in 2013.

They were clearly the right two people to advance the cause of evidence and the concept of the commission and get it through Congress. And get it on to the President's desk so that we can be here today.

And the timing is certainly good. Congress today just passed a major bipartisan deal, the budget. But they were also were sending two committees that do a lot of work in this area.

The Senate Health Committee that I sit on and the Ways and Means committee that Ted came from and Donald Schneider. Paul Ryan was on before he got yet another good promoter now Speaker.

And, you know, they've both had a track record in this area. Paul Ryan has made evidence a big part of this expanding opportunity agenda. And Senator Patty Murray has a long history of working to make real changes in major bipartisan legislation.

She was an early backer along with Senator Capon on the Maternal, Infant and Early Childhood Home Visiting Program which is one of the gold standards in this area.

She also wrote the Workforce Innovation and Opportunity Act with Senator Johnny Isakson and this Congress. She did the another Act with Chairman Lamar Alexander.

But she knows the possibilities here. Obviously realizes we can do a lot more. The commission offers a unique opportunity to take that next step forward and the use of evidence-based analysis and federal decision making.

I'm not going to go into too much detail regarding the specifics of the charge which you already know. But I will say that, you know, there's sort of three board pockets they're interested in.

First was the idea of how to build the right architecture today and also it's going to be sustainable going forward to maximize, including through integration of the quality and the utility of data as collected by the federal government.

I should say collected by the government because I think someone might have also pointed out is also involved in state and local governments as much as federal.

And piecing together that puzzle. They've asked you to think about the pros and cons of doing a centralized clearinghouse. And if you think that's a good idea. And it's a clearinghouse or something for them to get back to the idea of accessibility of the data and the way it's possible to maximize that for researchers and policymakers.

And if that idea has merit how to go about collecting and organizing documenting, sharing, securing and keeping private. And that's really important keeping private and confidential the data.

And finally they've asked for your guidance on how to better incorporate outcomes management into their upfront design programs and how to improve that and deliver better results. Which is, you know, fundamentally what they were most interested in today.

But of course all of this needs to be done with those air tight procedures around confidentiality. I wanted to make a couple general observations for including and one of them - and Ron's already mentioned this a couple times.

And I couldn't agree more. In a perfect world we would have the commission complete its work under this administration. If for nothing else just for continuity's sake.

But I think Ron said that as well. I think this is actually an excellent opportunity for the commission to be very impactful with the next administration.

The fact that your work will take you well into 2017, an ability to be - think creatively about how to be a resource to make sure that the excellent work under the Obama administration is carried on whether that's in the form of interim report or targeted hearings, meetings to transition the team, sharing tools. I think there's lots of ideas that you guys can be helpful there.

And I think it's an exciting possibility. And I encourage you to think about how to partner up and be helpful to keep up the work. Also there's been something Ron mentioned, you know, and to use the word branches is there is a lot other really exciting work going on, on a lot of other entities that are working in the space and doing good work.

Good examples, the evidence-based policymaking collaborative that exists at Ron knows well. Every one of them kind of representing one of the organizations on it. There are a lot of experts that have a lot of - that are well versed in the uses of evidence.

I really encourage you to also think creatively about how to draw in those entities that are clearly very friendly and interested in what you're doing to help complete your task, to be a kitchen cabinet of sorts and just to help you work. We gave you an incredibly complex time and a very generous amount of 12 to 15 months to complete it in.

So I think - again I think there's folks that are willing and able to be useful to you. Along those lines and getting back to themes of being, you know, making sure you're air tight and as much as possible. I think Paul made the right point.. You know, you do the most you can. You are not going to be perfect.

But I think with privacy and confidentiality it would nice if it could be a similar collaborative of similar entity on the privacy. But also sort of help inform and reality check and have partners and to help you and me to get to the second phase of this.

Where Ted and I will be much more involved and making sure that Congress in the next administration follows through on the implementation and the adoption of these things.

And along those lines I think I would encourage you to extent that you can the really in clear and compelling ways real test case examples things that really show the federal, state, and local policy officials the potential of this new architecture and data.

You're going to hear later from Dr. Chetty some great examples of, you know, really excites of both sides of the aisle where that is getting done and how in the neighborhoods. You know, stories like that are very powerful to help us when we're, you know, making sure to get in and follow through.

I think there's also even lessons in Congress and things that worked really well in this case. We're behind the administration in a lot of ways. I think we're clearly catching up when you look back the House is something from Ted's old committee what they're doing with the Family First Prevention Services Act and other areas really it's something that's made five to ten years ago. It's hard to, you know, to imagine. So I think there's good work there to talk about, you know, where the success stories are in Congress and doing that program design that's the third element.

I mentioned also things that on the research side are really compelling, a couple reasons that you guys are here. And then I'll second that point that Devin made about making this, you know, your work as broad as possible. And looking at all areas and I've also mentioned this with, you know, on the business side and the progress and purpose to be as broad as possible.

To try to think about it in all areas of evidence. So I want to, on behalf of Senator Murray, in wrapping up to thank each of you for agreeing to serve on the commission. You know, you're taking time away from the family and your friends and your important day jobs.

We do appreciate it. I can assure you that Team Murray and Team Ryan will be, you know, ready to help along the way for the next, you know, 12, 15 month but also at the end, you know, take back to the baton and do what we can to make sure this is implemented.

You know, as a side I learned that it's a good thing to have the Speaker on your bill. It took us about eight months to get it out of the Senate and back over to the House. And it literally took them about eight hours to pass it and send it over to the President.. So but it was good.

So before I close I want to acknowledge and thank Ted McCann who's a great part - and Donald Schneider who also was out back on the old days on budget doing this as well.

They've both been great to work with. As has the staff of OMB and Census have really been a huge help. Great supporters with a lot of great resources on helping organize and get it set up and we really appreciate it. Thank you very much.

Katharine G. Abraham: Ted.

Ted McCann: I think number one we'd like you guys to look at the major transit patterns in metropolitan areas. He - John mentioned Professor Chetty's work and the work that he did. He also mentioned the Speaker's expanding opportunity in America Project.

That was really for the Speaker where the interest in this was originally peaked and we had a conversation with Professor Sullivan who's Bruce's co-writer. And he mentioned basically what the work that Professor Chetty was doing. We had already read it. We enjoyed it.

And he just said look, most of us can't get access to a lot of this data. And, you know, I think the Speaker thought, well that doesn't make a whole lot of sense. And so that was sort of the beginning of our thought process and how we want to move forward.

And then we had again this wonderful working relationship with Senator Murray, you know, through the BBA. We went back and forth on a whole number of issues. I think they think they won. We think we won. It was a good process. And between the two of them they gave me a lot of trust in each other's motives and all of that.

And so at the end of that process the Speaker basically asked Senator Murray and Senator Murray said yes let's work on this evidence-based policy commission together.

When we were looking at the various social assistance programs generally and he mentioned business programs as well. We did find that there was - there is a lot of work done through the Bush administration and then through the Obama administration which has really put emphasis on evidence-based policy.

But it seemed to be sort of lacking generally in terms of the data access that researchers had. And then at the same time there was a lot of concern at the time as well about privacy issues.

And so I think our thought process was well let's figure out what's working here, what's not working. And also at the exact same time figure out how we can best protect American's privacy through an agency that collects a ton of this data.

So I think from a Congressional standpoint John mentioned a lot of the areas. We made a lot of progress on this as we owe - we just passed to the House I told you that Financing Bill that we think is also very useful.

When I was at Ways and Means we tried very, very hard to get some of this more evidence-based policy involved in the DI program. But there is I think wide bipartisan support for promoting evidence-based policies.

I think what we lack to a certain extent is in some ways the expertise that we need to know exactly how best to open up some of the data - administrative data, researchers more broadly so that we can figure out again in social policies here what works and what doesn't work.

I think - I think John sort of mentioned, you know, most of what we would like to do. It is very much, you know, what can be done administratively. What needs to be done statutorily? What is the benefit of a clearinghouse of some sort or the drawbacks? How can we, in Congress, support what the administration is doing right now?

How can we support that going forward? I do think this is an area again where there's enough bipartisan support where there should be a consistent push coming from Speaker Ryan in the future.

Senator Murray - at least at that level - I'm not sure about anybody else. It's interesting in this time year you've got pushback on nearly everything. This is not one of the things that you hear pushback on. It's very much - there's concerns.

You know, there's privacy concerns. I think those are accurate. But everybody wants to know how can we better improve the data that we're getting on these programs? How can we better target outcomes?

And how can Congress incorporate that into the policymaking process at large. Which I think is important, one of the more important things. So with that I mean I know five minutes is sort of the general - sort of the keep it short.

As John said I want to thank all of you guys for participating in this. I do think this is probably one of the more important things that Congress will do and act upon over the next couple years.

And so this is going to be a high priority for the Speaker going forward. And we will maintain a high priority going forward. So thank you guys so much for doing this.

Katharine G. Abraham: I thank you both for your comments. I'm particularly happy hearing the emphasis that both of you placed on once the commission's recommendations come in wanting to work to make sure that something happens to them.

And I know that we'll want to be working closely with you as we move ahead. And make sure that you understand what we're thinking and get your feedback as we move along.

Do people have questions for either Ted or John? Paul.

Paul Ohm: So thank you very much. That was very useful in figuring out the kind of genesis and the framing and what we should think of as our charge. I don't recall a statute that - and I'm probably mistaken.

That, you know, specifically designates that fully one-third of the members need to be versed in database management, confidentiality and privacy.

I wonder if you feel at liberty to express at all kind of the genesis of where that requirement came from. And I think what you're saying is do whatever you want as a commission.

But what do the members you work for or maybe the broader body think you can tell us about what we do when we face really difficult questions about, again, the tradeoff between trying to get the data and also real concerns about privacy?

Right? Does the two-third - one-third split with the commission suggest that tie goes to utility or on the contrary I mean is there any general wisdom you can give us as we run into that recurring issues?

Ted McCann: I - from the Speaker's perspective privacy is very, very important. I mean we came from the Ways and Means Committee is sort of one of the most important things out there.

And so we want to make sure that perspective was heard and understood. You know, I think the other side too is we collect a ton of data. Sometimes that data is not extraordinarily safe. So it's important to know what are we collecting. What makes sense? What doesn't make sense?

And so that's where the speaker is. And I think that database management was more of a technical type of thing so.

John Righter: I think in our mind we want to make sure there's sort of three areas. And I think there's a reason there. It will still take 12 to 15 of you to be able to formally make recommendations back to us. Which in itself is a high hurdle.

I mean it just needs to work so that researchers obviously feel like this is an improvement over the current system and it stresses that it's the issues that they've identified by our program administrators who actually live and do this on a daily basis.

Also to a degree it works giving - there's an excellent paper that I think you guys will see from the staff on the barriers that are very well. Does it work with the real world they live in? And then does it also work with the privacy community and confidentiality?

Those guys obviously a good thing to get as much data out there as possible so that we can start talking about the real complex problems. But it's got to be done in a way that protects our confidentiality, privacy, and the trust of the American people put in the government when they handed over that data.

So all three of those. And it's really not just that, you know, 12 of 15. I think it's also, you know, I think getting a majority of all three of those buckets, the programs, the administrators the researchers and the privacy experts.

Katharine G. Abraham: I have a question I guess. In terms of thinking about the data that are potentially useful for evaluating policy, there's data that are currently collected by federal agencies.

There are data that are currently collected by state governments in the course of administering programs that are wholly or partially funded by the federal government.

Just thinking about those categories of data as we contemplate a data infrastructure, can you give us any insights about whether there's a place where

you would draw the boundary in terms of what you think we should be looking at?

John Righter: We purposely made this as broad as possible. I mean there's obviously three specific charges. But we really did try to open it up because I don't know a tenth of what any of you know in your respective area. I know what my - what Senator Murray wants which is a way to make this better.

Because, I mean, her background is a preschool teacher. She sees the impact of kids over time. She's really excited about a lot research by Dr. Chetty and others are doing as well.

And I think she also sees it as not just a tool to learn more and make an impact in those communities but it's also going to change conversations within Congress. So we were trying to leave that to you to figure out on the specifics.

I know that we hear a lot about they can't just be on the federal side. You've got to figure out about how to involve state and local who are not doing this the right way. So we're very sensitive to that.

And then, you know, to get back to Paul's question. I think we go along Ted and I may not be sitting here with you here next time but we'll be engaged and we can have those conversations about what from the outside of what makes sense using - we still have a very important role in making sure that it gets done at the end.

But I'd love to have that conversation. I think that's - at this point starting today our main directive was to try to make sure we weren't overly like prescribing you how to do it. But giving you as much running room so that we get the smartest possible minds in those three areas.

And let you, you know, decide that. And we can have conversations as we go along.

Ted McCann: The only thing I add is to earlier what John said, from a Congressional perspective things are slightly easier when there is a federal nexus as a general rule. But I wouldn't want to restrict it as that. So that's the only thing.

Ron Haskins: We hear we have 14 months left. A month has already expired. And I think in 13 months, three weeks, and six days we're going to decide even when it usually happens to commission reports. They sink and no one pays any attention.

So we're going to devote some thought, I'm pretty sure, about this. How do we make sure that our report gets attention both publically and with the right people in Congress?

This is a role that both of you could help us with a lot. And one thing that I'd just like to mention that might be a possibility if you would do a joint hearing that we could work with you to bring the right people and bring a lot of attention as much as possible to the recommendations and to what will hope Congress will do.

John Righter: That's a great idea. I mean I can't speak for Senator Murray. But that's a wonderful idea. We will find - you have two members who are good at getting things done. They have a very good track record. And they're very both very committed to this.

So we look forward to getting that report in, what did you say, 13 months and three weeks and six days and twenty-three hours.

((Crosstalk))

Katharine G. Abraham: By my reckoning, the date is September 6, 2017.

Ted McCann: And I would just say when we put out the most recent product on opportunity mobility for the better way agenda. I know one of the big points in that was improving evidence policy and incorporating evidence technology into our social welfare programs generally.

And not just those welfare programs but all of our programs. So this is at the top of the Speaker's agenda. I think this is something that he is eagerly awaiting to see. And I think our press folks and all that understand that too.

Woman 4: I have a question going back to scope and primarily the question of commercial data. You know business analytic companies are the fastest, one of the fastest growing sectors in the data space. And they make their products primarily by personal data that's provided and sometimes government data.

And we're seeing them increasingly then turn back around and sell those data products to government to predict about analytics and so forth. What - and also we're seeing government also turning some of this function questions over to big data through those commercial entities as well.

Was there thinking about the scope of the relationship between the commercial data and government data or is it only from the perspective of data first collected in the government?

Ted McCann: I think the scope again. Probably broader. I know we spent a lot of time thinking about much - like which data type of thing.

We use it a lot from an anti-fraud perspective which our subcommittee does a decent amount of that as well. Currently there's a do not pay portal where they try to use some commercial wage data.

I think this probably goes a bit into the cause of the conversation, all of that. I think you guys just - you know a lot more than us. And I think we're trying to figure out what makes sense with them. Make sense - so okay.

John Righter: I know one of the commissioners who's not here today Bob Groves has spent a lot of time thinking about that question and how you sort of bring in outside of the federal government. I've had some interesting conversations with him about that.

Certainly nothing in there would preclude it. I can't say that I would put it as the primary thing that you were thinking about when we drafted it. But I mean you've convinced me that, you know, that it's in federal government obviously but does a ton in this area.

But there's a lot going on in Silicon Valley and elsewhere to make enough for this commission maybe thinking about the next steps (about) how you start leveraging that in as well.

But again a scope that by design was meant to be - to give you guys as much running room to figure out what makes the most sense once we got the 15 of you in room not to be overly restrictive.

Katharine G. Abraham: Any other questions? Thank you both very much. And then our next speaker is Nancy Potok. She already indicated that she's the Deputy Director at the US Census Bureau.

Nancy Potok: Good morning again. And thank you for inviting me to speak at your inaugural meeting. I'm from the Census Bureau but I'm speaking on behalf of the Federal Statistical System this morning. Many of my colleagues from the Federal Statistical System are here in the room.

So I hope I'm representing you well. I'm really excited to see the commission up and running. And quite honestly the timing couldn't be better for the commission to be up and running. Because the Federal Statistical System right now is at a critical juncture.

We're facing what I would call a perfect storm of factors. And this perfect storm requires us to move from the message that we have used to produce statistical data in the mid-20th Century to the methods that will work for the 21st Century.

So what do I mean by that? Well, I'm sure it's no secret to all the members of the commission most of the people in the room that one of the primary methods that the Federal Statistical System has used to collect data has been serving censuses.

But this extensive reliance on pretty much voluntary surveys for the most part truly is not sustainable. And the trends are clear. Public cooperation including household and businesses has been declining. Costs are rising are rising even in budgets are holding steady or budgets that are declining.

Data users, including researchers, policymakers and other members of the public want high quality data faster with more geographic granularity and with the ability to link data sets from multiple sources.

We talked a little bit this morning. I heard about commercial sources. Well commercial are putting out timely data. It's granular – but there are often no measures of quality, no guarantees of continuity, no transparency, varying levels of coverage and potentially from those companies who have a process modus that could introduce biases into the data.

So our challenge is in the Federal Statistical System that I'm putting for you are we really have to contain costs on data collection. We have to release data sooner. We need to reduce the respondent burden.

We have to offer data at lower levels of geography. And yet we still need to maintain very high quality that can be measured and that we can describe in terms of its consistency, its usability and its reliability.

And in addition while the Federal Statistical Agency has been very focused historically on data collection we really need to think more broadly about data dissemination.

That's only in traditional terms. But in terms of open governments. And that means enhancing the ability of researchers and practitioners to take on data and combine it with other data to better understand our society and economy.

I think that is a theme that we are sure is repeatedly over and over this morning because we need to learn how our government programs are meeting their intended goals.

And also we need to assist communities. This has not come up this morning but we really need to assist communities in improving their service delivery in the areas where they're looking at traffic. Right? Transportation, location of services, other activities that are happening in communities in real time.

I think there's a lot of examples of these types of applications they are seeing. Usually we talk about smart cities. And we see an increasing in role in multiple communities for chief data officers who are taking federal data, program data and combining it with the local data and really improving life for the people who live in those communities.

And that's - the Federal Statistical System is an avid participant in that. That's our mission is to put the data out there to really make these kinds of improvements that improve programs.

So one common thread moving forward is really the ability to link data but in a controlled and curated and in a secure way that protects against cyber security attacks. And also protects and respects the privacy and the confidentiality of the data provider, whether they're individuals or businesses.

And if I may I'd just like to illustrate this by talking about the Census Bureau for a minute. So in the Census Bureau we've been working with the administrative data from other federal agencies, from states, from local governments and from commercial third party providers for decades.

We have gained considerable experience in some of the issues surrounding record linkages, data disclosure, data stewardship in curation and combining survey data with administrative and third party data to create rich often longitudinal databases that can really inform research and public policy.

We are currently working on several joint projects with other statistical agencies and partners. And I think you'll hear about a few of them later in the morning.

And there are problems. We have a lot of obstacles and barriers. And I think the more that you do this deep dive and look at these records from various sources is, you know, you asked questions just a few minutes ago about the scope. How far to go looking at state and local records.

These are real issues. I mean some of these databases that we put together, their researchers are depending on that. It took years to negotiate different agreements with states to get some of that data.

We don't have really years to negotiate that individually agency by agency. And then every time you want a different use there is various amendments it can take forever. And it's not a good process I can tell you.

But I don't want to delve into all of the issues right now. I just would like to tell the commission that we strongly believe we have to be looking at the future. Not just to the present and not to what we've done in the past.

So we're researching now new data sources. We're researching collection methods, usage of technology. And that really allows us to save money to reduce responder burden and to create the type of data product in a timely way that the users are asking for.

You know, one example that I throw out there because I know you're going to hear about it from other projects. But the biggest example at the Census Bureau right now is the use of administrative records for a variety of uses to save billions of dollars in the 2020 Census.

So it's just one example of the kind of return on investment that you can get by looking at the infrastructure, by looking at the technology. Yes, it's true there is some upfront investment. But it really pays off in the end.

So that type of return on investment is why at the Census Bureau we have undertaken research in a number of related areas. And we are working out collaboration with academics, with other statistical and federal agencies and the private sector data analytic and the aggregation companies.

So in our minds they're all in scope because we're doing their research now. And we're fortunate at the Census Bureau because we're authorized in legislation to acquire and use a variety of administrative records.

Not everybody has legislation on the corresponding side that says they have to give it to us when we ask for it. So we are authorized to ask for it and receive it.

But in speaking on behalf of all the federal statistical systems, progress can be very slow right now. And in part it is due to the legal barriers toward data sharing. Even between physical agencies for statistical purposes.

So we're hoping that the commission's work can really look at this and help eliminate some of these barriers and move us forward where we need to be. You know, the research we're conducting right now I would call incremental.

But our vision for the future is really transformational. We really need to eliminate these unnecessary barriers. We have to build the infrastructure to support data sharing in a safeguarded environment. And safeguard is bump up the cyber security and the privacy and confidentiality in the sampling.

We have to aggressively create new partnerships with the research community and other data users and data providers and take advantage of the technology that's out there and the continuing advances that are being made.

And as equally important we have to develop ways to understand how the quality of combined data can both be maintained so we're really putting out high quality data but also measured.

It's very important to develop those kinds of measurements. So if you noted quite a bit the commission has a lot to do in a very short time. But I will say in the Federal Statistical System we are optimistic that the commission really can help us move forward so that we're ready and able to meet the data needs of our nation's policymakers in the research and evaluations community.

And provide the data that people need to do evidence-based policymaking. So I look forward to working with you. On behalf of the Federal Statistical System we all look forward to working with you.

And on behalf of the Census Bureau we're really delighted to be able to support and assist the Evidence Commission in achieving its mission in any way that we can.

Katharine G. Abraham: Questions for Nancy? Ken?

Ken Troske: So I know that Katharine recalls the lunch we had when I used to work at the Census Bureau maybe 20 years ago. There was a number of us sitting around and talking about how we could improve the sharing of data across agencies.

And I think there's been improvements made since I've been there. But you're right. It is incredible to me in looking around at how challenging it is for agencies in the same governmental - the federal government, much less state government to share data.

And, you know, Hilary mentioned these other countries in which they have developed these fantastic data resources. I would argue that one of the main reasons they have it is because they have one federal statistical agency not 13 or however many we have.

So I view that as - when we start talking about making these data available through a single portal and I know there have been efforts, federal law to increase the ability of statistical agencies to share data and those efforts seem not to have succeeded for a variety of reasons.

I guess help me understand why you think - you still seem to think that this is an issue that needs to be addressed. You know, we've been talking about it for 20 or 30 years. I think the ways you think we're going to be successful when other people haven't?

And are there things that you think we need to be focusing on that are going to help us crack this nut that has been almost impenetrable for a long period of time?

Nancy Potok: I think this is an opportunity to crack the nut. I hope it's successful but again a lot of this depends on one - what the commission wants to recommend as a way of moving forward evidence-based policy.

So that's up to you in terms of your recommendations. But I also think that some of the strategies that you've been talking about in terms of how we roll out the recommendations and get attention can be very helpful in moving some of these things forward.

Part of the problem up to now is that a some of this is either addressed in a broad sweeping way like let's reorganize the entire Federal Statistical System and put it in one agency, which is a nonstarter. Right? You all know that.

It's talked about on and off. And I know there's multiple people in this room over the course of their careers who have worked on that...

((Crosstalk))

Ken Troske: That's what I was advocating when I was having lunch with Katharine. And they all told me it would never work.

Nancy Potok: So - but...decades ago...

Ken Troske: I'm aware of that.

Nancy Potok: But in this day and age - this is my personal view okay? You don't have to necessarily reorganize physically the federal government. You can create a virtual statistical system through more data sharing.

If some of those barriers that - you know, we have CIPSE. It's not like all those statistical agencies don't protect confidentiality, privacy the same way.

So we should be able to freely move data back and forth within the statistical system if there are criteria there that released these are the conditions under which this can and should occur.

But yes. I can think of - you know, I see people in the room now who have spent since CIPSE was passed, you know...

Katharine G. Abraham: Most people will know what that is. But maybe you can...

Nancy Potok: Well CIPSE and Shelly probably knows...

((Crosstalk))

Katharine G. Abraham: Confidential Information Protection and Statistical Efficiency Act of 2002.

((Crosstalk))

Nancy Potok: Thank you.

Ken Troske: Very impressive.

Nancy Potok: That is impressive.

Ken Troske: Were you involved in that Katharine? Is that why you...

Katharine G. Abraham: Yes.

Ken Troske: Maybe a little?

Nancy Potok: Well so what - that was really the Act that gave all the federal statistical agencies very similar confidentiality responsibilities and protections that could be passed along to like survey respondents and others who are providing data.

It's very similar to the Title 13 protections for Census data. You know, penalties for violating confidentiality. And it was a big step forward. In addition I know Katharine knows this part.

There was a piece in there that allowed the Bureau of Labor Statistics, the Census Bureau and the Bureau of Economic Analysis to start sharing business data. But there was a companion piece of this which sort of allowed the tax data that underlines some of the key business data in terms of names and addresses of businesses and address register.

That BLS keeps one. The Census keeps one. That companion piece of legislation has never been enacted to this day. So to this day, BLS is collecting a business register on an address list for business and coding it according to what the businesses are.

Census Bureau's collecting a different one. BEA needs to get data from the two agencies and try to harmonize it in some way. And we can't just share our addresses really and say let's just get one list for the federal government and agree what the codes for these businesses ought to look like.

This is such a simple, kind of a no brainer, kind of thing. But when you get into all of these issues of whose data is it and how can we share it and, you know, there is like multiple people with jurisdiction over this data legally.

It just gets complicated. And when you multiply that across all the different data sets that statistical agencies have that we're trying to reconcile, get rid of redundancies, standardize the data, link data to put longitudinal data sets together that really are rich and can provide just multiple opportunities for researchers to do the deep dive.

Every single one of those, you know, has a story like this. And it - our hands are tied. And we need help.

Katharine G. Abraham: Kim?

Kim Wallin: I just, you know, and this is something I saw in state government too. I'll call in the silos. Right? I mean Transportation Department was supposed to be preventing deaths on the highways but then their performance measure was to prevent death.

And then Public Safety didn't have that. But they kind of work together, right? And you threw out there because Ken kind of brought it up about having 13 different statistical agencies and being able to consolidate them, reorganize them is a nonstarter.

That was 30 years ago. I think we need to stop, you know - we need to talk about where, you know, the silos and working together here. And are there any other sacred cows, Ted and John, oh these are nonstarters?

I mean...

Man 1: We have a different list for both of them.

Kim Wallin: But I just have...

((Crosstalk))

Kim Wallin: I just have concerns there. You know it's like we should be able to - you know, we'd like to approach this as something that we're in the 21st Century now and we want to do something different.

And I hope that as we go forward as a commission that people will be open to the changes. Because that will probably be the biggest hindrance that we have is the resistance to the change. And I think we need to address that.

But are there any other sacred cows we kind of need to know about that might hinder this commission's ability to their jobs?

Man 2: Go where like the evidence takes you.

((Crosstalk)) Laughter

Ron Haskins: It would be very helpful I think for the commission - okay. I'll talk just for myself. It would be very helpful for me to have something like the top 10 hit list of barriers that are preventing the statistical agencies from operating at maximum performance.

Some of these are really - we do have one paper. But it doesn't have - I don't think it has - I can't remember. But we need to know what these - the specific barriers.

Because this is going to be one of the few chances where statistical agencies will have the opportunity to make recommendations that we could actually - you sat here and listened to the gentleman next to you say help us turn it into a legislation.

That opportunity does not going to come along very often. So it's a real opportunity to teach us what are these barriers and how could we do the most possible to eliminate them including if it requires statutory changes.

Nancy Potok: We'll be delighted to get you our top 10 hit list.

((Crosstalk))

Latanya Sweeney: Twenty years ago when I was hanging out with people in the Census Bureau one of the - a really interesting problem was - in this particular example was census data wanting to be linked with data from IRS.

And the linking was the problem. But the result after the linking would be de-identified to the standards that were set back then. And so there's sort of this interesting issue that came out of that is if two sensitive data they're only sensitive because the sensitivity is needed to link up individuals.

But then after the linking is done the redaction is done. It renders the result as somewhat de-identified. So this became a really interested problem at MIT where I was a student at the time took that back to MIT.

And we now have technology that could really help that. And so I think this idea of whenever the problems are linkages it would be - it's kind of interesting to revisit that because there might be some ways of softening the issues.

So if we could know...

((Crosstalk))

Nancy Potok: I agree. I think there's a technology has made great advances in a lot of it. But we also are thinking a lot about, you know, what's appropriate for linkages of micro data that's available to researchers in a, you know, controlled environment, approved a project versus what are things that are available to the public.

And when there are linked data sets that are made. For example like the Census Bureau has - it's another acronym, LEHD. But it's Longitudinal...

((Crosstalk))

Man 4: Employer-Household.

Nancy Potok: Employer-Household Dynamic. Thank you. So one of the inventors of that - they are looking at me. But that's a public data set. And yet - so, you know, there are other things too that we're looking at in terms of synthetic data, modeling, what about disclosure review on publically available data.

In addition to what should we be making available through the next generation and the next version of a Research Data Center or clearinghouse for researchers in a curated environment?

So those are - they're really two related but different kinds of issues that I think also, you know, the commission might be thinking about.

Latanya Sweeney: But my question also - there were two parts of it. But the other part is this question around what are the key data sets that have this characteristic that if they're linked they would satisfy lots of research requests.

Do we have a way of figuring that out? Because it could be the case that, you know, if there are 20 such data sets in various combinations linked together. And that was done in a secured way for some definition of secure.

And that result came out - the result of that linkage then de-identified could maybe soften the blow as opposed to everybody trying individually to go get these two data sets and link them.

Nancy Potok: I think that an answerable question. I don't think I can give you an answer off the top of my head. I would probably consult with a lot of people before answering. But that's something that is - that the commission is interested in. We can certainly be looking at that and provide our views on it.

Jeffrey Liebman: Hi. This is Jeff - can I jump in for a second here?

Katharine G. Abraham: Yes.

Jeffrey Liebman: Okay. I think one of the important tasks we have is to reach out to the various communities that we hope will make better use of the data that will come out of this enterprise and get each of the different communities with those barriers.

And I know I've talked to some of you and Hilary and I are going to make sure next week at the MBER Summer Institute that we collect data from economists who are there about barriers that they've run into with accessing administrative data.

But we should think of systematically how we go out to other parts of the academic community not just economists but sociologists and political scientists and other people who use administrative data and get their lists of barriers.

And then separately or similarly we should be getting those within government both at the federal level and the state level of barriers so that we can come up with a list of the top 10 barriers to remove.

And so I think we need to pretty systematically think about what the communities are that are going to be users and find ways to get that to report to us on their barriers so that we don't miss some high priority ones.

Katharine G. Abraham: Thank you. That's a really good point. Sherry?

Sherry Glied: Nancy I just want to pick up on something you said at the beginning of your remarks and just so that we keep it in the back of our heads. Or how we should keep it in the back of our heads moving forward which is the growing problems that the statistical agencies are having with their existing survey methodology.

So the notion of linking the surveys to the administrative data or the surveys to one another is complicated by the fact that we're getting low response rates. It's harder and harder to get, you know, representative sample and so on.

How do - how should we think about that problem in conjunction with this problem?

Nancy Potok: I think there's multiple ways to think about it. In a way it's the idea of, you know, let's - if we - we kind of know why which surveys we think are really important to continue and that are important for a lot of reasons.

Benchmarking might be one of them. You know, the types of data aren't available any place else. There is only a limited amount of funding to do this. Right now we're spread very thin.

Year after year after year on our ongoing surveys we get the same amount of money which is, you know, not even increased for inflation or anything. And so, you know, we have less and less to spread across kind of a broad number of surveys.

If we can be doing the work to say look we can make these surveys shorter because we can find the information in other places and combine it. And we're doing a lot of work with that right now because they want American Community Survey.

How can we make this a shorter survey so it's less costly to do? So people are more willing to fill it out? On the economic side we're looking at, you know, even some of the economic indicators. Something like retail sales, monthly retail sales.

How do we get those numbers out there with more geographic granularity? If we can use other kinds of things like credit card records that would track closely or that we can model. Or maybe we don't have to do monthly surveys. Maybe we can do annual surveys.

That's the kind of research we're looking at. So therefore if we have the same amount of money and we can focus it on surveys that we feel are critical to continue by doing the research to say we'll give other sources of data that we can measure that we think are going to serve just as well.

That's - we have the right balance between surveys. And if there's fewer surveys and they're shorter and we have the resources to - I mean you can get people to respond. It's just very expensive at this point.

You know, and so - but if we can devote the resources we feel like we really have to get up the response rates and get the rest of the data in other ways. I think we can balance that going forward.

Katharine G. Abraham: Great. Thank you. That was - this has been a really helpful session. I would like to thank all four of our speakers. We really appreciate your making time to be here.

And we'll be coming back to your questions. But we'll take a short break now. It's by my watch 10:20. We will reconvene promptly at 10:30.

((BREAK))

Ron Haskins: The idea of the next two sessions is to give concrete examples of what we are really hoping to promote which is research in action. We are going to have two sessions of this, each with two speakers, so a total of four speakers. The commissioners will have ample opportunity to question them. On the first panel, we have Jeri Mulrow from Economic Research Service and Mary Bohman, acting director of the Bureau of Economic Analysis. So now we're about to hear some great example of evidence in action. So we'll begin with Mary.

Mary Bohman: Well thank you very much. I appreciate the invitation to speak with you today about the joint project that the Economic Research Service has with the Census Bureau and USDA Food and Nutrition Service.

Our three agencies have formed a strategic partnership to develop what we call the next generation data platform and has the goal to provide data and research to strengthen the supplemental nutrition assistance program, or SNAP, which is formally known as the Food Stamp Program.

SNAP is USDA's...

((Crosstalk))

Mary Bohman: SNAP is USDA's largest food assistance program surveying 46 million people per month, 14% of our nation's population at a cost of approximately \$74 billion in 2015. So it's a program that's very large and has a long history of investing in evidence for improving its performance.

My presentation starts by describing the program's goals, the data that was linked and the research results we've obtained today. And then I'll conclude with a discussion of what we think of as the keys to success.

Many, many people have been part of this project as you can imagine with three federal institutions and you'll see state participation. And I won't call them out by name. But I just want to recognize their work over many years to make this happen.

Let me start with the project goals. And here we have three agencies, each with distinct missions and therefore the work that we do has to come together and meet our individual responsibilities.

As USDA agencies, ERS with the research mission as FNS with the program mission seek to inform policymakers and others on who participates in the program, how participation effects the lives of those individuals, and who does not participate and why.

On the other hand, the Census Bureau has somewhat different interests and they are to improve their surveys which includes questions about data on food assistance program participation, to improve the 2020 Census. And here they can acquire better information about low income Americans who have participated in SNAP.

And also to strengthen the data linkage processes. So to meet these individual goals we have these activities in the project. And I'll just - next slide - give you a high level overview.

We are acquiring administrative data on SNAP and this program is operated by states so the data are collected from states and where they reside. I think we heard this mentioned by Nancy. But this is an additional challenge that isn't faced by use of administrative data for programs that operate at the national level.

And then we link the data to Census Bureau surveys. And I'll talk about how this leverages the strength of each data source. And finally the payoff is conducting research and providing tools to states that take advantage of this data.

So before turning to - next slide. Before turning to the results from the analysis I want to set the stage by talking about the strength and weaknesses of the staff administrative data records and the survey data.

And if you look at the administrative records the strengths are that they're complete. They contain the universe of participants. They are reliable in that provide accurate information on participation, the duration of that participation, and the amount of benefits each month.

And they're available in the sense that they're already collected as part of the program administration. However, they don't do everything you would want. As you would expect there are weaknesses.

And one is that the program records exclude non-participants. And this is a common challenge in using administrative data. And they also only include the data needed to operate the program.

As you can think about in these state or local offices operating staff, they're not asking extensive questions about socioeconomic variables as part of these records. And they don't include outcomes such as food security or health of the participants.

So that's on the program data side. On the survey side there are large nationally representative federal surveys that cover statistically everyone in the population both participants and non-participants.

And these surveys contain the rich socioeconomic data and often outcome variables that are missing in the program data. Surveys have weaknesses. In the case of food assistance programs it's known that households tend to under report their participation and the benefits they receive in survey data.

And we'll see that in some of the research results I'll present. And depending on the survey there may be no information on the duration of staff participation or the amount of benefits.

So by linking the goal is to take advantage of the strength of each of these records and be able to address some of these measurement issues and expand the amount of information available to do research on the programs.

I'm going to present results from three studies that we've done. And then I'll conclude with the overarching lessons and where we're going forward with this project. Because we're not done. This is at the beginning.

So let me start with the first report on Texas which was also the first state that provided the data and worked with our partnership to create the data linkages. And at ERS we've done research based on food and nutrition service priorities using this data or questions that they've identified.

So the question that we analyze in this report is who does SNAP reach and not reach. And also to be able to look at a sub-state level. The food and nutrition and this serves the food and nutrition service's mission to understand how effectively staff is reaching the people it's designed to serve.

And they wanted to have some data and estimates within a state about participation across different areas that previously wasn't available. So this allows them to focus expenditures on activities such as outreach to targeted populations.

So our project links SNAP data for Texas to the American Community Survey, a very large survey, and had to develop methodologies to be able to create these measures.

For example, who is an eligible participant based on American Communities Survey. So that took advantage of the economic expertise and the staff that we have within our agency.

As - and I'll just give you a brief - this really a case - and all three of these studies where the devil's in the details. I'll give you a very 60,000 foot look at what we did.

But as defined in this research the access rate measures the share of eligible people who receive SNAP benefits. And we use ACS data to model who's eligible for SNAP.

And then taking that ACS data with eligibility defined we then link to SNAP records. And we can see which percentage of people eligible actually participate in the program.

And next slide. For Texas as a whole we can see the average but really wide variation across Congressional districts. And we had some county data where ACS supported those measures.

So the statewide measure is 63% but access ranges from 37 to 77% across Texas congressional districts. And the data and the research in this report also provided access rates for selected demographic subpopulations.

For example, the elderly or different language skills. It was in the families in Texas. Let me move to a second report that we did. And here the question was looking at what percentage of SNAP households have the very lowest incomes.

For this research we used the data from the second state that participated in project, New York. And we were able to - by linking the data address issues of underreporting in surveys and also income variability of the participants.

Here one of the data challenges for research was that the definition of a household for SNAP eligibility purposes is different from the definition of a household in the American Community Survey.

So we had to do work to reconcile those two definitions to be able to create a measure of eligibility in the base survey. So let me show you what we found in this survey.

In using for New York State and here we looked across multiple years we found that when you use the linked data the measurement shows that the targeting towards low income Americans is much higher from the SNAP program than only using a measure based on the American Communities Survey.

And the differences are sizable. For the share of SNAP households with annual incomes less than 50% of the poverty rate the increase in participation is from 18 to 27% of people who receive SNAP benefits.

And while the share of poverty to combine the bottom two categories it increases by 10 points from 51 to 61%. So this provides valuable information to people in the Food and Nutrition Service as well as Congress and others who look at this program about the success that the program's having in targeting.

Let me turn to a third report that we did and here the questions are really also very interesting and technical about how you measure participation. And because SNAP is a program where people can join for a month and they can go in and out of the program depending on eligibility.

And we also have issues of how questions are asked that on surveys people - the questions are often at any time during the past year did you participate in the SNAP program.

So we had an interesting measurement question of making sure we were comparing apples and apples in terms of people who participated monthly to the population of monthly participants. People who participated on an annual basis to those that were eligible in a given year.

And I'll tell you if that's complicated I have gone back and forth with the authors in the last day to make sure I understand it. But I think in terms of the research findings we use the New York SNAP record in how we did this.

We use the New York SNAP records to count the number of people who received SNAP benefits for one or more months in the annual timeframe. And using this data the monthly participation rate for New York was 80%. So based on the monthly count.

We used the annual rate with a different denominator in terms of the universe of those who are eligible on an annual basis is 75%. So you get a different measure. But it's also a value to the Food and Nutrition Service and the community looking at SNAP effectiveness in terms of the kinds of people who participate.

And whether we're - how we're capturing long term eligibles who may have a disabled family member or elderly family member versus the working poor who are the ones that move in and out of eligibility most commonly.

So I think that you see that this research gives you both insights into measurement of participation and targeting but gives you a richer set of information to understand the implications and effects on Americans and how the program is working.

So those are I think in terms of lessons of success is the better measurement and greater availability of data for program managers and researchers.

So let me talk now about some takeaways on how...

Man 6: You have about one minute left. So...

Mary Bohman: Takeaways on how we made all of this happen. I did figure about 17 minutes of remarks. Can I just get a couple extra minutes?

Man 6: You have 15 minutes.

Mary Bohman: Okay. Takeaways on how we made this happen. It's a three-way partnership. And I think the mutual benefits part is important. One can change the legal frameworks. But agencies have missions that authorize our spending.

And we need to have motivations to participate in data sharing that are based on our responsibilities. And here through a lot of work we found that. Let's skip to the next slide.

You can see these different - you have this in your slide, the things that motivated our participation. I think this is important. The Census was a critical partner. We couldn't have done this on our own within USDA.

We needed the infrastructure that they've built in terms of data sharing and security. That's something we couldn't have afforded to do. And they also have the American Community Survey as they developed expertise in linking to that.

And I think the Census will help going down the road and through the data centers to sharing the data with academic researchers as just one of our goals. Within USDA we needed to make program support.

For ERS this is a core part of our mission to provide information, to inform public and private decision making. We've provided both analytical and

financial support including providing funds to the Census Bureau for their work.

And I think looking at the resource side is an important dimension. Because as partners come from mutual benefit this drives their ability to share resources.

State SNAP agencies have been partners. And then it's really taken a lot. OMB, Office of General Councils, sometimes they may seem like your enemy but at the end you really need them to help make sure you're doing all of this appropriately within our statutes and to protect privacy.

And political leadership in each of these departments needs to see the benefit to support the funding and allocation of resources. Let me just...

Man 6: Please sum up for the people.

Mary Bohman: I will sum up. Yes. So going forward we - this map shows you where we have agreements that are signed, ready for signing and draft review and some of the states in white there are already preliminary discussions ongoing.

I think it's been a coalition of the willing as people see the benefits of what we've done in Texas and in New York. There's increasing interest in getting in the queue to participate in the data links.

And I - going forward we want to have more states but also more programs. The WIC program, the Woman, Infants and Children Nutrition Program. And expand access to the academic community via the Federal Statistical Research Centers.

And there's also interest in providing states with data visualizations that they can use for program administration for the state level of combined data.

So let me wrap up. Thank you for your patience.

Ron Haskins: Thank you very much. Jeri Mulrow.

Jeri Mulrow: Good morning. It's an honor and my pleasure to be here in front of the commission at its first meeting. I'm presenting some work several of my current and former colleagues have worked on and as Mary said there are many people that have been part of this.

Gerry Ramker is here today and he has much more detail about it. I also wanted to really acknowledge the terrific cooperation and participation that we have with officials at the FBI's Criminal Justice Information Services Division and the state public safety agencies and NLETS who've all been integral to this work.

Let's hear about them more later in a minute. I hope you find this to be a useful case study. It will show that there are some impactful policy decisions that can effect a large number of people.

And then there are challenges and costs along the way.

Next slide please. First just a bit of - first just a bit of context. The Bureau of Justice Statistics is a statistical agency within the Department of Justice.

We sit within the Office of Justice programs which is the research oriented arm of DOJ. But DOJ is large and diverse, covering a broad complex justice

system. And so you may be familiar with some of the other agencies that fit under the DOJ umbrella.

Such as the FBI, the Bureau of Prisons, the Civil Rights Division and the mission of BJS is to measure the US Justice System which a collection independent system at the local, state and federal levels.

Next slide. This is a picture of what flows through the criminal justice system. And it's broad and it's complex. This is referred to as the crime funnel. It was first developed by the President of Crime Commission on Law Enforcement and Administration of Justice in 1965.

And for the 50 years it has served as a blueprint for developing the National Criminal Justice Statistical System. BJS conducts collections of both victims and offenders and of agencies and facilities that respond to crime.

BJS measures that with victimization, move on to arrest, and law enforcement information and through courts through trial and convictions, then into corrections including prisons, local jails, probation and parole. And then into the criminal's history record and recidivism.

So we measure quite a broad spectrum of what's going on. And the criminal history records I'll be talking about soon can actually be used to link across these various different phases.

Next slide. So there are many criminal justice policy questions that could be studied using criminal history records. For example what are the implications of policy regarding lengths of sentences imposed?

Such as should there be longer or shorter initial sentences for drug crimes? Or should sentences - sentence lengths be extended for those who pose serious threats to the public and reduced for those who pose less serious threats?

It can also be used for evaluation purposes such as looking at recidivism and re-entry in the Second Chance Act authorized in 2009 to provide funding for services and programs for individuals returning to their communities.

Such as substance abuse treatment or career development programs. These data already exist. They are used operationally every day for both criminal and noncriminal justice uses.

For example, they are used for background checks in firearms purchases or for pre-employment screening. There are over roughly 100 million state criminal history records.

And as you can see these can be used for great potential for use in evidence-based policymaking. And as such BJS has actually been working on a new project for about a decade.

Next slide. So criminal history records or rap sheets as they are commonly known are created when a person is arrested and fingerprinted by one of the nation's 18,000 law enforcement agencies.

Years ago these records were kept on paper. Now they are electronic. These records are submitted to a state repository and new records are registered. Subsequent events are linked to the records such as a prosecution, a court disposition, or a sentence to a correctional facility.

Each state maintains its own criminal history record repository. And each state has its own format and specific set of variables. I have a couple of example, mock examples from that sheet.

So next slide. So this is an example. Note that it's a text based kind of example. And that is actually the way the data exists right now. Each state has its own statutes and regulations, criminal codes, and codes of criminal justice procedure.

There is some commonality but a lot of variability. Next slide. And it can be difficult - this is another example from a different state. It can be difficult to decipher the meaning of a transaction.

So for example if the charge in one state is the same as the charge in another state. There's another. Next slide. Since each state maintains its own repository there needs to be a way to connect criminal history records across states.

The FBI maintains an automated pointer system called the Interstate Identification Index or III. That allows authorized agencies to determine whether any state repository has a criminal history record and an individual.

Nlets, the International Justice and Public Safety Network operates a network over which these requests can be asked and receive state records.

So prior to the III system, inquiries had to go to each state individually. So here's how the process works. Next slide. Say a police officer or maybe even DJS submits a criminal history record inquiry to the FBI's III index system via Nlets.

In this example the request is connected to a specific New York State criminal history record. Next slide. The FBI's III system responds with a list of states that have criminal history records on that individual and forwards the request on to the state via Nlets.

In this example the FBI III indicates that there are criminal history records in California, Texas, and Virginia that match the name of this (arrest). Nlets is the go between so FBI sends the information to Nlets who then contacts the state repositories for the information.

Next slide. The states that send back their criminal history records on the individual and the records can be accessed there. Next slide. So as you can imagine there are some challenges to using criminal history records for statistical purposes or evidence-based policymaking.

I'll highlight four of the challenges. BJS has been working on the first two, gaining access and standardizing for statistical purposes. And it's still in the process of understanding the quality. And we're still working on the last piece of for archiving for transparency and reproducibility purposes.

Next slide. So for gaining access. In 2006 BJS began exploring how do we get permissions for these records? Permission is granted by the FBI to authorized agencies.

Based on the FBI regulations and policy BJS had to be deemed a criminal justice agency performing administration of criminal justice functions to be able to meet - and be able to meet stringent FBI security standards before we would be allowed to gain access.

In 2008 we finally felt - we finally signed - excuse me - an MOU with Nlets and FBI which outlines the specifically duties of each agency and terms of use and extent of security requirements needed.

The MOU grants BJS the authority to access the III, specifies how that access has to be operationalized and dictates the security and data protection provisions that BJS has to meet.

It is important that for this work we have to know which rap sheets we want to access. So it's not a blanket provision. But we have to specify which ones we want. We provide Nlets a list of record identifiers and then we (launch) the record request.

Next slide. So with having the ability to launch these record requests we see state responses, BJS worked with Nlets to develop a series of parking programs that take those individual state raps sheets, identify and extract key data fields and create a uniform record.

As noted earlier these rap sheets are mostly text based so the required developing text parsing programs. The original parsed texts are copied into a relational database. But they're still not really ready for use for statistical purposes.

The next step is to take this information and recode it into a common taxonomy. This means coding the state specific statutes into common set of codes.

To do this BJS has partnered with NORC to develop a software system that reads every field in the parsed relational database and converts that entry into a standardized BJS code.

The initial work was extremely labor intensive. Individuals had to review the rap sheet, interpret the specific statutes, code it into a standardized code, develop these programs to automate the coding and then validate the coding and programming.

The software system now utilizes a family of states specific crosswalk fields in order to do this. Every time a parsed database is passed through the standardization process a similar - a smaller percentage of fields are coded unknown and we have to go back and recode and review and look and add on to these libraries.

Next slide. Every time BJS parses a criminal history record we conduct a data quality assessment. States regularly change their individual formats and can add new statutes of crimes which require us to go back and revise the parsing programs and update and validate.

BJS justifies these facts - this information back to Nlets for coding and processing there. And in addition we can mission not all states maintain their criminal history records the same way.

So we check for gaps and inconsistencies and try to provide feedback to the states on those gaps. The third challenge that we are - gaps in quality. Okay. And the fourth challenge which we're still working around is archiving.

The FBI grants permission to BJS to access these records and use them for statistical purposes. And the question is still remaining whether we can archive, standardize, and anonymize statistical records for research access later and that is still being discussed.

And finally I just wanted a word on costs. As you can imagine this word was not free. BJS has estimates that the development cost was about \$3 million and the annual maintenance costs to keep this up is about \$360,000.

At the present day. This presents a case study for the commission to consider as they embark on their work over the next several months. I hopefully illustrate some of the benefits and the challenges to using existing administrative data. Thank you.

Ron Haskins: Thank you very much. So we have about 15 minutes for questions for both of the speakers. So members of the commission have questions? Yes. Go ahead.

Man 7: I'd like to comment on Administrator Bohman's talk just - and then ask a question that actually Katharine asked me.

Ron Haskins: Could you move a little closer to the mic?

Man 7: So I want to comment on Administrator Bohman's talk and then ask her a question that actually Katharine Abraham asked me last night.

So I think either presentation illustrates a good - is a good example of how administrative data can take a long time to acquire. And how there are many different people interested in linking administrative data to survey data and it's use to many different audiences.

So the New York data I started to try and get you about 15 years ago. And, you know, of course it wasn't 15 years of work. It was mostly 15 years of waiting. But I was very glad to see that you were able to put them to use for your agency for the Economic Research Service.

I used them in academic papers. And New York State was willing to supply the data because they wanted to use the data for projects of their own. And they did in projects looking at take-up of staff in their staff by geographic area and demographic group.

So I think it illustrates just how hard and how long of a process it is to get administrative data and link it to survey data and how it even - that there are many different groups that would like to use the end result.

The state, the federal government, outside researchers. So that's a long way of getting to the question which was just how uniform are these state records of periods of receipt and amounts and characteristics of that households.

Because that's impediment if one wants to try and assemble data for many states or even nationally.

Mary Bohman: I'll give a quick answer and we'd be happy to follow-up and provide the commission with more details. But there are differences across the state records both in terms of information but also in terms of how their IT and what's available.

And that is a challenge. And I think this is where having the Food and Nutrition Service as a partner was essential because they allowed us to understand how the program operates but provided access at the state level the people who knew the state records.

And I think will be important going forward. FNS is also investing in more harmonization of the state records and improving their quality. And I think somewhat it's a chicken and egg. You have to show the benefits of having this in order to invest.

But that is something that they're doing. So I - we'll follow up on some of the specifics. But that is a challenge in going state by state. Especially when this is something new that unlike some of the labor records that have been used across states before this is a new initiative to use the SNAP - the state SNAP records.

Man 7: I'm sure one - I'll try to be quick. So just one natural follow up is at some point we'll want to know is there a role for federal legislation in trying to push states to supply the data in a more uniform format.

Just how that could be done. How we could make the data more usable.

Mary Bohman: We'll certainly describe carefully what we find and what the challenges are.

Ken Troske: So I think your presentation I think has raised two issues that maybe you can address. The first is, you know, I've had the pleasure of using Census Data both as an insider and then as an outsider.

With special sworn status. And it often is - somehow you have to demonstrate the benefit of the project to the mission of the agency. And apparently the mission of the statistical agencies aren't necessarily to produce policy relevant research. It's to collect the data.

And so I guess one question is do we think it's worth thinking about making the mission of statistical agencies to be involved in producing the use of their data in producing research so they can comment on policy in general that makes government operate more effectively.

That would be one question. And the second thing I think your slide with you map on it demonstrates is you have to communicate on a state by state basis. I

still must admit I struggle to understand how a federally funded program has to go out and beg the states to return the data that will then collect as part of this federally funded program.

And they have the option of saying no. And I guess there's a - I guess a question there is do we see a way forward in which - is there something in some legislation somewhere that gives the state these rights? Or they just sort of - is this something that a law that has to be created?

Is it when we give you money to collect data you have to give it back to us when you ask for it? Or is there something else that can be done to address this? Because again I think that's a huge question that we're going to struggle with.

Mary Bohman: Well I'll start by saying as someone who's an agency head, you learn quickly that you need to be aware of your authorities. And that is what Congress authorizes you to do as an agency. Both the scope of what you can do and then appropriations law that provides funding for your activities.

And that guides all of us who work in the public sector. And we operate within those constraints but also opportunities. You know, it provides us funding to do the work that I was able to discuss today.

And at the Economic Research Service we have a mission to provide information and analysis to inform program decisions. And I think that will probably vary across the different statistical agencies to the degree to which they're authorizing language gives them the scope to do that work.

But all of us I think have a principle providing relevant information for decisions. And that relevancy is a core principle across agencies.

So the other issues that come up that I didn't talk about. But a lot of these records are collected with their own authorities for what can be done with them. And the research is often not in that list.

It is program administration and enforcement. And this is where you get in discussions about whether research is a part of program administration. And you have to look at the specifics, talk with your general counsel office.

But this takes a lot of time. And I do think work by OMB on the statistical directive 1406 last year provided some general guidance on how to think about this. But it still requires looking at individual statutes and working to make sure you're operating within the authority of not just the agency but the way the program and the survey data was collected.

So I think in that term, you know, this is something where - this is a congressional decision. I won't weigh in on what Congress should do but there's also a lot of very specific pieces that lie within each agency and each data set.

Ron Haskins: Ken, I think we're going to have to get information counsel on this. But I don't see why the federal government can't in fact in any cases which have not been involved in writing statutes in which we specify in great detail what the states must report and how often they must report.

And they normally do it now. There are problems quarrel about the definitions and all that. But they basically have to comply with the federal rules. And the federal government could say you don't get the money then in programs like food stamps.

They pay 100% of the benefit, 50% of the administrative costs. Right? So that gives you a real - a lot of leverage.

Ken Troske: My own experience is that - again, in doing a report for the DOL in which the DOL was unwilling to tell states - well they told states give us data and they got a response from 12 out of 50.

That doesn't seem like a very high response rate.

Ron Haskins: Well I'm saying though that I think - and we should get legal, you know, opinion on this. But the federal government - Congress would have the power to require them to report data, define data, how often, all that kind of thing.
Paul.

Paul Ohm: So for Director Mulrow. I think one issue we're going to tangle with a lot is the problems of mission creates unintended consequences and it's a fine line sometimes between research and operational uses of data.

So I'd love to hear a little bit about in these negotiations with FBI and then the resulting MOU. How do you deal with these issues? I imagine a rationalized rap sheet database is probably a very, very useful thing for the operational mission of the FBI is well.

I know that they have NCIC. They have the Terrorism Database. They have all sorts of other data where they're probably reinventing the same wheel that you are inventing.

How does your agency or sub-agency think about the protections around the database to make sure it doesn't just become an adjunct to law enforcement as opposed to something that creates research?

And in particular what do you do about concerns about future administrations? Attorney General that might think much more aggressively about dipping into the data that you have, for example, to find bad people?

Jeri Mulrow: Well that's a lot there. Let me start in the Gerry might jump in. But for the rap sheet data that we have, the research project. These - every specific research project we actually have to request permission from the FBI to initiate that project.

It has to go through their IRB. And then we at BJS are the only ones that are able to actually right now access that data from the rap sheets. And then right now we have to destroy the data when we're done.

Now for a federal statistical agency that's a little tough. Because you want to be transparent and reproducible right? But if we are the only ones who have access to the data we're producing reports that we aren't able to save those data right at the moment.

So that's one of the negotiations. I want to say that - and so we don't actually have the rap sheet data that does data standardized parking programs are actually now used regularly by the Department of Homeland Security, other parts of DOJ.

So we have authorized access to that. So some of that infrastructure that we help develop is actually being used operationally for other purposes. Other than - it's not being used for research necessarily.

For the - some of the data that we collect not here but say on a correctional data or inmates or something we do use our privacy. We use the statute that a lot of us who collect that data which also protects that data.

So this is why I think when Nancy was talking about access in a controlled environment that's very important to us. Because these data are sensitive in many cases. And we need to be careful about how people can access these data.

And we are living in an age where really we can't de-identify 100% risk anymore. So I think there needs to be a way to allow researchers in a curated fashion on some of these really highly sensitive data. And this is different than what Nancy was talking about.

Sometimes we produce public use data files and we try to do that whenever we are able to. And we feel that we are not exposing too much risk to have a disclosure problem. Does that answer your question?

Paul Ohm: I think it does. I mean I now understand your program differently. I thought you were actually building and maintaining a rap sheet database. It sounds like you're almost a consumer of an external rap sheet database but one that you guys helped conceptualize and design based on kind of some of the - yes.

Gerry Ramker: And maybe just as a way of further clarifying, the database - III isn't a database. III is the way queries for records are routed to the states. The databases exist in each state.

But III is a voluntary system that states, public safety agencies, agree that this information needed to be shared nationwide and internationally now. So that's - it's almost like a contact that the states have agreed to participate in.

So the queries go out and grab the data. But the original data is based in the states.

Paul Ohm: Thank you. Ms. Murlow.

Ron Haskins: Okay. So thank you very much to Mary and Jeri we are going to continue the - continue with our having examples. And so I take great delight in introducing Raj Chetty who will probably challenge our technology to see what...So Raj Chetty, Professor Chetty, you have 20 minutes. And then we go back to ask you some questions. So go ahead.

Raj Chetty: Terrific. Okay. Thanks so much for giving me the opportunity to speak to the members of the commission. And I think the work you're setting up to do is incredibly important for researchers.

And so I'm going to give the researcher's perspective on using administrative data for evidence-based policymaking. Starting by talking about a little bit our recent work in our research on Improving Economic Opportunity in America.

And then I'll try to draw some broader lessons from our experience working with administrative data on what the key challenges are for researchers in the U.S. And where I think we can make improvements.

So to start with the research background. I'm going to start by talking about the American Dream which a complicated concept that means different things to different people.

But to distill it to something that we can measure systematically in the data let's think about the probability that a child born to parents in the bottom fifth of the

income distribution makes the leap all the way to the top fifth of the income distribution.

So the classic version of the American Dream. How common is that in the United States versus some other developed countries around the world where we have comparable data? In the U.S. my colleagues and I estimate that 7-1/2% of children who start out in the bottom fifth of the income distribution make it to the top fifth.

That compares with 9% in the United Kingdom, 11% and .7% in Denmark and 13.5% in Canada. Now initially when people look at this numbers they sometimes react by saying oh even in Canada your odds of success don't look all that high. Right? Only 13-1/2%.

But you have to remember of course that no matter what you do you can't have more than 20% of people in the top 20%. So another way to look at it is if you had a perfectly random distribution where your parents played no role in determining your outcome, you'd have a 20% statistic here.

So that's kind of the upper bound on what you might expect. So these are actually very large differences in rates of upward mobility across countries. One way you might think about it is that your chances of achieving the American Dream are almost two times higher if you're growing up in Canada rather than the United States.

Now these differences across countries have been the focus of much policy discussion and much concern that the U.S. might not be living up to its reputation as the land of opportunity.

But what I'm going to focus on here - and this is where administrative data in the U.S. is so powerful and we can learn a lot from it going forward. What I want to focus on is the fact that upward mobility varies even more within the United States.

So using data from tax records my colleagues and I calculate upward mobility for every metro and rural area in the U.S. And the way we do this is we get information from the IRS from de-identified tax records as part of a broader project on the effects of tax expenditures.

We studied 10 million kids, basically all the kids born in American between 1980 and 1982. And the reason we're focusing on those particular cohorts is because those kids are around 30 years old or so when we measure incomes today which is sufficiently old such that they're out of college, graduate school etcetera and we can measure their incomes reliably.

Using that data we draw this map here which shows you the geography of upward mobility in the United States. So this map shows you the same statistic that I started out with, your chances of reaching the top fifth of the national income distribution.

Starting from the bottom fifth by metro area. So for 740 different metro and rural areas in the US. And the map is colored so that lighter colors represent areas with higher levels of upward mobility.

So for instance if you look at the area where I am right here in the Bay Area, San Jose, 13% of kids make it from the bottom fifth to the top fifth comparable to the numbers we saw in Denmark and in Canada.

But at the other end of the spectrum if you look at Atlanta, Georgia or much of the Southeast, Charlotte, North Carolina the numbers are below 5%, lower than any country for which we currently have data.

So within the United States you see this incredible spectrum in terms of rates of upward mobility with some areas having the rates of upward mobility comparable to the best places in the world and others at the other end of the spectrum.

Now this big map here your eye naturally gravitated to broad regional patterns. The southeast presenting much worse in terms of upward mobility than the mid-west.

But if we zoom in - and we'll zoom into the area where you all are in, D.C. and look at the data at a county level you see that there's lots of variation even at much finer geographies.

So for instance the city of Baltimore is one of the places in the country with the absolutely lowest rates of upward mobility with only 3-1/2% of kids making it from the bottom fifth to the top fifth.

But if you go over to Maryland, to Montgomery County, that number jumps up to 16%. It's a remarkable nearly four, five fold increase in rates of upward mobility across quite short geographies.

Now of course the city of Baltimore is extremely different from Montgomery County. But the point here is that even in relatively narrow geographies where we might possibly expect people to move across areas you see very big differences in rates of upward mobility.

So naturally the question of interest to us as researchers and to policy makers who would like to restore the American Dream and increase economic opportunity in America is to understand why upward mobility is varying so much across areas.

And what ultimately we can do about it. So the next step of our analysis we again exploit the power of having these population level administrative data. And study people who move across areas in order to take a step towards understanding why upward mobility varies across areas.

And so the key point I want to convey here is we find that most of the variation in upward mobility across the area is caused by differences in childhood environment.

And I'm going to spend a bit of time illustrating how we reach that conclusion to you because I think it really illustrates which types of data are extremely valuable for researchers.

So in the next analysis I'm going to show where we look at 5 million families that move between areas. And again we're using information from tax records where we're able to see people's location over time based on where they file their tax returns from in order to identify movers.

So rather than going through the details of the statistical analysis I'm going to give you a simple example that illustrates the idea underlying these results. So let's take a set of families that start out in Washington, D.C. which you saw in the previous map that I showed has relatively low rates of upward mobility.

And let's say just to pick a number that if you grew up in a low income family in Washington, D.C. when you turn 30 years old let's say you have an average income of \$30,000 in your household.

Now in contrast, kids growing up in Montgomery County have - let's say if they're growing up in Montgomery County movers have an average income of \$40,000 if they're growing up in a low income family there.

Now what I want to do is consider a set of families who moved from D.C. to Montgomery County with kids of different ages. And we're going to start by looking at kids who moved from D.C. to Montgomery County when they were exactly nine years old. Which happens to be the earliest age we're able to look at in currently available IRS data.

Now what you end of seeing in the data is that these kids who moved when they were nine years old end up roughly halfway between the kids who grew up in D.C. from birth and the kids who grew in Montgomery County from birth.

That is when we measure their earnings when they turn 30 years old, tracking them forward 21 years we see that their earning roughly \$35,000. So that's for the children who moved when they were exactly nine.

So now let's repeat that analysis looking at kids who moved when they were 10, 11, 12, 13 and so on. And what you see is a very clear and what turns out to be an incredibly robust pattern which is that these gains decline steadily as you move to the better area at later and later ages.

So if you move at age 15 instead of age 10 you pick up less of the gain. If you move in your early 20's you get essentially none of the gain at all. And if you

move after that point things remain at zero and there's no further gain from moving to Montgomery County at older ages.

So what do you see from this chart? First you can see the differences between D.C. and Montgomery County are not merely driven by the fact that there's a different set of people living in Montgomery County and a different set of people living in D.C.

Apparently you can take a given child and move that child from D.C. to Montgomery County you see distinctly different outcomes for that child. That's very important because it shows that place and neighborhoods actually have a causal effect on children's long term outcomes and suggest that place based approaches to make economic opportunity can be very powerful.

Second you learn that childhood exposure really seems critical. If you move as an adult it doesn't do a lot for your outcome. It's really about moving as a child that's critical. And third you see that every extra year of childhood exposure improves your long-term outcomes. Moving to a better area at 10 instead of 11 or 15 instead of 16 has roughly the same effect.

And I want to pause here for a moment and step back to the broader picture of what this - why is this type of analysis only feasible with the population level of administrative data.

This is what I think of as essentially quasi-experimental methodology exploiting variation in the timing of when people move, the exact age of the child to move to learn something about why different places have different effects.

Prior to the availability of large scale data sets like the tax data or social security records it would have been completely infeasible to conduct this sort of analysis because in survey data sets there's no way you'd have an exceptionally large sample to examine people who moved from say D.C. to Montgomery County for example.

So having comprehensive data where we can track people over 20, 25 year periods really allows us as you can see in this picture here to understand precisely what's going on in terms of what's driving these neighborhood level differences.

Now the next step in our analysis. We've shown you that some places seem to generate much better outcomes than others. Naturally the next step is to ask, okay, so what is it about Montgomery County that seems to produce such great outcomes?

What are the characteristics of areas that seem to lead to great economic opportunities for disadvantaged kids? And I'm not going to spend a lot of time on this. But let me just quickly summarize our findings where we look at a variety of different factors.

And we identify five very strong patterns in the data. Which is that first cities that less residentially segregated by race or by income tend to have higher rates of upward mobility.

Second cities with a larger middle class that is less income inequality tend to have higher rates of upward mobility. Third places with more stable family structures - so that is fewer single parents, lower divorce rates and so forth tend to have higher rates of upward mobility.

Fourth places with more social capital. So Salt Lake City with the Mormon Church is a classic example of a place thought to have a lot of social capital, a lot social cohesiveness where other people will help you out even if you're not doing well. Those types of places tend to have high rates of upward mobility.

And finally as you might expect places with better public schools also tend to have high rates of upward mobility. So those factors give you a sense of what types of things we should be looking at if we want to try to improve upward mobility in a given area.

And I'll show you one final set of results which thinks about how we might improve upward mobility in a different direction. So one approach is try to invest in places that aren't doing as well.

Another approach is to exploit the large scale affordable housing policies we have in the U.S. in particular about housing voucher programs to try to give low income families incentives or sufficient money to move to better areas.

And the results I've just shown you suggest that if we give low income families housing vouchers to move to better areas we'll see significant improve to economic outcomes.

Now this result is particularly surprising to those of you familiar with the literatures. It's surprising to us with academics in light of well-known finding from the famous Moving to Opportunity experiment that was conducted by (Hyde) in 1990s which as you might recall gave vouchers to about 5,000 to move to low poverty areas, higher opportunity areas using a randomized lottery.

This experiment was conducted in five big cities in the U.S. There's been a lot of prior research on the effects of the experiment and it as somewhat disappointingly I think for social scientists found very little impact of MTO on economic outcomes.

Now we were puzzled by that in light of the findings I just showed you because we thought we provided quite compelling evidence that where you grow up matters.

And we decided to go back to the MTO data and do a longer term follow up than was feasible based on the survey that the prior researchers did in order to track the impacts of the experiment.

So again illustrating the value of the type of data that this commission will hopefully help facilitate. Linking the MTO data to the tax records in order to track longer term impact on kids who moved at younger ages turns out to be extremely valuable in understanding what's going on.

So just to summarize what we find from that experiment, give you a concrete sense of what's going on here. So in the MTO experiment families received housing vouchers to move to lower poverty areas.

To give you one example in the control group in the MTO experiment in New York lots of families lived in Martin Luther King Towers, a very high poverty public housing project in Harlem.

In the treatment group many families moved to a place called Wakefield in the North Bronx which was about 10 miles away, more of a mixed income area. And so essentially what we're doing is comparing the outcomes of kids who grew up in Wakefield with the kids who grew up in the MLK Towers.

And so to summarize what we find, we link the MTO the test data and track long term impacts we find the kids who moved to low poverty areas when they were young, for instance below the age of 13 do much, much better as adults.

They're earning 30% more. That is - triangulates to a gain of roughly \$100,000 in income over their lifetimes. They're 27% more likely to attend college. They're 30% less likely to become single parents.

However we find that moving has little effect on the outcomes of children who were at older ages, children who moved as teenagers for example. And moving also had no effect on parents' earnings.

So these findings you'll see confirm exactly that declining pattern that I showed you in the previous slides using the 5 million movers. They show that the duration of exposure to the better neighborhood matters.

Explaining why previous studies which we're only able to look at the impacts on kids who moved at older ages and on parents, explaining why those studies did not find any effect of this experiment.

So I think there's an extremely important lesson here for evidence-based policy which is when you track the effect on an experiment you might think, you know, a 10-year follow up or a 15-year follow up, which is what people did, would be adequate to find out the effects.

But apparently not. In this case we would have, you know, falsely concluded that housing vectors have no effect on economic outcomes had we not been able to do this longer term analysis. And that was only possible because we

were able to link the MTO data with the IRS data brought through a collaboration between HUD and IRS.

That evidence has since informed HUD's policymaking decision in terms of where things are headed, providing greater incentives for families to move to higher opportunity areas and so forth.

So showing that this type of analysis can be very valuable for policy decisions almost in real time.

One final point about the MTO experiment which I think is also extremely important in thinking about evidence-based policy which is experiments are very valuable. But they also have significant limitations.

So in this chart here I'm drawing the analogue of the downward sloping pattern that I showed you initially using the 5 million movers. But this is now done with the MTO data. And so we're estimating the effect of moving to a better area by the child's age of move.

And you see the same declining pattern that I showed you before. But you can see the dashed lines here show you confidence intervals. You can see that you have very little precision to really estimate the slope of this line.

Whereas the previous study that I showed you with the 5 million movers you have incredible statistical precision because the sample size, you can see clearly on the pattern of decline linearly.

Here, you know, you have a rough sense that there's a declining pattern but you don't know exactly what that shape looks like. And you don't know exactly what the slope looks like.

So the point here is that I think there's great value in combining different types of methods. The large scale quasi-experimental analysis with smaller experimental analyses.

I don't think we should commit ourselves to any one research methodology. All of the different methods can provide very useful evidence.

Let me conclude by talking about some of the barriers we face in doing this sort of research and where I think we can make significant progress in U.S. data infrastructure and evidence-based policymaking.

So one critical challenge we all face as I'm sure you're well aware is scarce bandwidth and limited access to these types of data. So for instance to IRS data there is often a complex contrasting process that's not really ideally designed for research of this type but has been adapted to make it work for the time being.

But I think ideally you would have a much more streamlined procedure that was really oriented around statistical research. Second you have relatively few physical locations for accessing the data.

You have to go to IRS offices or the U.S. Treasury. This just makes it a lot harder to work with these data. You would much rather work with them in an academic environment where you can collaborate with colleagues.

And now we just need for those who are working with Census data often you hear - I hear my colleagues here at Stanford, you know, basically saying how they have to take turns working with the data because there are inadequate terminals in the RDC.

So very low cost constraint to relax that would have a big impact on research output. Second we're often working with outdated technology due to very scarce resources.

So for example in the type of analysis we do we're constantly figuring out which data sets to delete because we don't have enough space on our hard drives.

There's not enough RAM to open a data set. If two people are working at the same time on a server the processor gets extremely slow. So these are all, you know, in era of modern technology these are extremely minor - we're waiting trying to find data here. And that's a small set of investments on the order \$100,000 to dramatically research with the ability to produce - these sorts of evidence.

Finally it's very difficult to link data sets. So each linkage like the MTO tax data link required a special set of contracts and a whole procedure. And that took a while to work out.

Another example, you know, a lot of our work it would be very useful to do the analysis by race. Because we think that different racial and ethnic groups have different experiences across these neighborhoods.

But it is not possible to do that at the moment because tax doesn't contain information on race and we're not able to link Census data to tax data where of course in the census we would have information on race.

So if we had linked datasets in a centralized, you know, systematic way it would dramatically improve what we can do. Here I would contrast on all three

of these dimensions I think a good model to look is the Scandinavian countries, the Danish Statistical Agency in particular where all three of these issues have been solved very nicely.

Where there's secure access from remote locations. There's modern very high quality technology and everything has been linked already in a very systematic way. So I think that's a good model to keep in mind.

Now going forward I think there are three things we can do to really improve the base for evidence-based policy in the U.S. The first thing I've noted is to create a centralized data warehouse that links data sets.

Here I would note that the IRS data really provides an ideal spine for linking other data sets. It can cover through a combination of information returns and 1040 forms more than 98% of the U.S. population. And they contain a great deal of information.

So they make - I think starting from the IRS and then linking in other data sets possibly serving other administrative data sets would be an ideal data infrastructure that's feasible.

Second I think it's critical to provide secure direct access to the data with simplified access protocols. I think direct access is extremely critical. People often talk about the alternative of synthetic data where you create essentially a fake data set that researchers can use publically which scrambles the information.

Our experience is that synthetic data for the past results I've been showing you really don't work well for iterative analysis where it's not clear what type of synthetic data you need to do the type of analysis I was showing you.

I also think it's extremely important to have relatively rapid retrieval of statistical results. So the current clearance procedure for the disclosing results at the RDC is often a hindrance for researchers because it takes a great deal of time to get one set of results out and then iterate with colleagues and then work on your next set of results and so forth.

So I know the - as since I briefly was with the MTO experiment I think it makes sense to think about randomized experiments as a natural way to obtain evidence on key policies.

But I also think we should be supporting a broader range of methods and that because experiments as I showed you with the Moving to Opportunity Study, are often under powered. You don't have enough statistical precision especially for studying long term effects.

Where you need an enormous sample. The MTO experiment was 5,000 people, it was nowhere close to enough. You need something 10 or 100 times as big to really understand what's going on and with much greater precision.

And so I think the supply of the administrative data that we can harness really lend themselves well to quasi-experimental methods that I illustrated. So it's very important to keep those methods in mind.

And finally I'd note that aside from identifying the causal effects of policies simply constructing descriptive statistics to monitor progress. For instance how is opportunity in upward mobility evolving in each neighborhood in America?

Just having that data is extremely valuable. And so I think creating tools so that we can construct such statistics is extremely valuable. So let me stop there. And I'm happy to take questions.

Ron Haskins: Professor Chetty. Thank you very much. That was very instructive. And thank you especially for talking to barriers. That's one of the things we're most concerned about.

So thank you so much for that. Questions from members of the commission?
Bruce.

Bruce Meyer: Thank you, Raj. Nice to see you. And I think what might also be able to put a little different spin on the access to MTO in that it gives an example of the value of later researchers having access to data that was used in an important study.

And so I think that also points to how we need to figure out protocols that would allow researchers to get access to data from previous studies, to see if they can extend them, reanalyze them, draw additional or maybe even different results from the original data.

Do you find that fair?

Raj Chetty: Absolutely. I'm glad you pointed that out Bruce. You know, I agree completely. I think replication is an important issue in studies like this. And as we see with the MTO experience, you know, you might find different with a longer term follow up.

So again having a centralized clearinghouse where these data sets can be maintained over time would be incredibly valuable.

Ron Haskins: Questions from other members of the commission? Well Professor Chetty, thank you so much for taking your time. It was very, very instructive. We may be back in touch with you, especially about these barriers. So thank you very much.

Raj Chetty: Thank you. My pleasure.

Ron Haskins: I'm going to turn it back over now to our Chairman, Katharine Abraham.

Katharine G. Abraham: Thank you. I want to thank all of our speakers for their thought provoking presentations. You've given us a lot to chew on as we start our work. And we'd also like to thank others who are here in the room.

I know a lot of you have expertise that we're going to want to draw on as we move forward. And we look forward to doing that. I also should - I probably should have done this at the beginning.

But I would like to thank Connie Citro and Brian Harris-Kojetin for making it possible to meet in this lovely meeting space. We really appreciate it.

Connie Citro: Thank Eileen and her geniuses. They actually have been working with the Census Bureau staff and they work quite seamlessly. So thank you.

Katharine G. Abraham: We appreciate it. This brings us to the end of our first public meeting and we will adjourn. The members of the commission who will be reconvening in closed session at 1:00 pm.

We have time for lunch now. There's a cafeteria that people will direct you to in the basement of the building. So we will close the meeting and resume the closed session at 1 o'clock.

END