

**RECOMMENDATIONS OF THE COMMISSION ON
EVIDENCE-BASED POLICYMAKING**

HEARING

BEFORE THE

**COMMITTEE ON OVERSIGHT
AND GOVERNMENT REFORM
HOUSE OF REPRESENTATIVES**

ONE HUNDRED FIFTEENTH CONGRESS

FIRST SESSION

SEPTEMBER 26, 2017

Serial No. 115-48

Printed for the use of the Committee on Oversight and Government Reform



Available via the World Wide Web: <http://www.fdsys.gov>
<http://oversight.house.gov>

U.S. GOVERNMENT PUBLISHING OFFICE

28-506 PDF

WASHINGTON : 2018

For sale by the Superintendent of Documents, U.S. Government Publishing Office
Internet: bookstore.gpo.gov Phone: toll free (866) 512-1800; DC area (202) 512-1800
Fax: (202) 512-2104 Mail: Stop IDCC, Washington, DC 20402-0001

COMMITTEE ON OVERSIGHT AND GOVERNMENT REFORM

Trey Gowdy, South Carolina, *Chairman*

John J. Duncan, Jr., Tennessee	Elijah E. Cummings, Maryland, <i>Ranking</i>
Darrell E. Issa, California	<i>Minority Member</i>
Jim Jordan, Ohio	Carolyn B. Maloney, New York
Mark Sanford, South Carolina	Eleanor Holmes Norton, District of Columbia
Justin Amash, Michigan	Wm. Lacy Clay, Missouri
Paul A. Gosar, Arizona	Stephen F. Lynch, Massachusetts
Scott DesJarlais, Tennessee	Jim Cooper, Tennessee
Trey Gowdy, South Carolina	Gerald E. Connolly, Virginia
Blake Farenthold, Texas	Robin L. Kelly, Illinois
Virginia Foxx, North Carolina	Brenda L. Lawrence, Michigan
Thomas Massie, Kentucky	Bonnie Watson Coleman, New Jersey
Mark Meadows, North Carolina	Stacey E. Plaskett, Virgin Islands
Ron DeSantis, Florida	Val Butler Demings, Florida
Dennis A. Ross, Florida	Raja Krishnamoorthi, Illinois
Mark Walker, North Carolina	Jamie Raskin, Maryland
Rod Blum, Iowa	Peter Welch, Vermont
Jody B. Hice, Georgia	Matt Cartwright, Pennsylvania
Steve Russell, Oklahoma	Mark DeSaulnier, California
Glenn Grothman, Wisconsin	Jimmy Gomez, Maryland
Will Hurd, Texas	
Gary J. Palmer, Alabama	
James Comer, Kentucky	
Paul Mitchell, Michigan	
Greg Gianforte, Montana	

SHERIA CLARKE, *Staff Director*

WILLIAM MCKENNA, *General Counsel*

CHRISTINA AIZCORBE, *Intergovernmental Affairs Subcommittee Staff Director*

KATY ROTHER, *Senior Counsel*

PATRICK HARTOBEY, *Counsel*

SHARON CASEY, *Deputy Chief Clerk*

DAVID RAPALLO, *Minority Staff Director*

CONTENTS

Hearing held on September 26, 2017	Page 1
WITNESSES	
Ron Haskins, Ph.D., Co-Chair, Commission on Evidence-Based Policy Making	
Oral Statement	4
Written Statement	7
Katharine G. Abraham, Ph.D., Chair, Commission on Evidence-Based Policy Making	
Oral Statement	10
Written Statement	12
Latanya Sweeney, Ph.D., Commissioner, Commission on Evidence-Based Policy Making	
Oral Statement	14
Written Statement	18
Robert Shea, Esq. Commissioner, Commission on Evidence-Based Policy Making	
Oral Statement	21
Written Statement	23
APPENDIX	
Letter of September 26, 2017, to the White House submitted by Mr. Cummings	54
Letter of September 26, 2017, to the Department of Health and Human Services submitted by Mr. Cummings	57
September 2017 CEP report titled <i>The Promise of Evidence-Based Policymaking: Report of the Commission on Evidence-Based Policymaking</i> submitted by Mr. Haskins, can be accessed at: https://www.cep.gov/content/dam/cep/report/cep-final-report.pdf	60
Response from Dr. Abraham, Commission on Evidence-Based Policy Making, to Questions for the Record	61
Response from Dr. Sweeney, Commission on Evidence-Based Policy Making, to Questions for the Record	68

RECOMMENDATIONS OF THE COMMISSION ON EVIDENCE-BASED POLICYMAKING

Tuesday, September 26, 2017

HOUSE OF REPRESENTATIVES,
COMMITTEE ON OVERSIGHT AND GOVERNMENT REFORM,
Washington, D.C.

The committee met, pursuant to call, at 10:02 a.m., in Room 2157, Rayburn House Office Building, Hon. Steve Russell presiding.

Present: Representatives Duncan, Jordan, Sanford, Gosar, DesJarlais, Farenthold, Foxx, Massie, Meadows, Ross, Walker, Blum, Hice, Russell, Grothman, Hurd, Palmer, Comer, Mitchell, Gianforte, Cummings, Maloney, Norton, Lynch, Connolly, Kelly, Demings, Krishnamoorthi, Welch, DeSaulnier, and Gomez.

Mr. RUSSELL. The committee will come to order. Without objection, the chair is authorized to declare a recess at any time.

I would like to begin by thanking our Speaker, Paul Ryan, for his commitment to improving the way the Federal Government uses data to solve problems and improve lives. His bipartisan work with Democratic Senator Patty Murray is the reason we are able to hold this hearing today.

According to a 2013 General Accounting report, only 37 percent of managers who oversee 1,500 different Federal programs say their programs had been evaluated in the previous 5 years. Without sufficient data and analytics from programs like these, government agencies are unable to fully assess the benefits or impacts of their decisions and programs.

Although many administrative records and surveys collected by government agencies provide significant sources of data, these records are typically not shared from one agency to another. Oftentimes, agencies either do not have access to or are not aware of data that could contribute to evidence building.

To begin addressing data gaps in policymaking and administration of taxpayer-funded programs, Speaker Ryan and Senator Murray introduced bipartisan, bicameral legislation in 2016 to create a commission to evaluate the current use and availability of data and make recommendations as to how to better improve the process.

On September 7, 2017, the Commission on Evidence-Based Policymaking released its report containing 22 recommendations for improving the use of data to inform government programs and policies. The Commission found the Federal Government lags behind the private sector when it comes to managing and documenting data that could be used for evidence building.

Taxpayers expect Federal policymakers to base their decisions on well-founded evidence and reason. But according to the Commission, there is not enough evidence being produced to adequately inform Federal decisionmakers.

Seeking to balance privacy interests while simultaneously creating a process where data could be used and shared for statistical purposes, the Commission's recommendations fell into four categories: establishing a new Federal entity to facilitate data linkages, increasing access to data, modernizing privacy protections, increasing and strengthening Federal evidence-building capacity.

Government information management, data sharing, and inter-agency coordination are at the core of this committee's legislative jurisdiction. The Commission's recommendations touch on the committee's oversight of the implementation of FITARA and FOIA and other recent data policy reforms, many of which we have worked on in this committee. The recommendations also touch on areas of the committee's jurisdiction that are on our to-do list, such as the Paperwork Reduction Act, the E-Government Act, the Privacy Act, and others.

I want to thank not only the commissioners who are here with us today, but all members of the Commission for the time and effort devoted to creating this report in a relatively short period of time. The information produced in this report will help us strengthen our laws and fulfill our desire for a more efficient, effective, transparent, and accountable government. We look forward to hearing from you today on how we can gain better access to the evidence we need to create informed and effective policies.

It's now my honor to recognize the ranking member of the committee, Mr. Cummings, for his opening statement.

Sir.

Mr. CUMMINGS. Thank you, very much, Mr. Chairman.

Today we will hear from four members of the Commission on Evidence-Based Policymaking. Let me start by saying this.

Frankly, when most people think of evidence-based policymaking, they don't think of the current administration or recent actions by Congress. But that is why today's hearing is so very, very important. Too often, the American people see firsthand how policies that Congress puts in place are completely unrelated to the facts.

Take voter fraud, for example. President Trump claimed that 3 to 5 million people voted fraudulently in the last election. He had no evidence for his claim, none. To their credit, some Republicans have even pointed this out, that is that there is no fraud. Yet now the American people are being forced to spend tax dollars on a new Presidential commission that is trying to hunt for evidence to back up the President's unsubstantiated claim.

Take healthcare. The nonpartisan Congressional Budget Office says that tens of millions of people will lose their health insurance if Republican plans to repeal the Affordable Care Act go through. Again, to their credit, a handful of Republicans have refused to go along with this. Yet, the majority continues to push repeal with no viable alternative for the millions of families who depend upon it.

Take Planned Parenthood. Our committee conducted an investigation of allegations against this group. To his credit, our former chairman, Representative Chaffetz, reported on national television

that we found absolutely no evidence to support those allegations. Yet, despite our findings, the Republican leadership in the House spent millions of taxpayer dollars on a new special committee to continue harassing Planned Parenthood in search of evidence that never, ever existed. Something is awfully wrong with that picture.

Take immigration. President Trump and congressional Republicans have argued that we need to limit the number of refugees we accept into our country because they supposedly utilize too many government benefits. Yet, the White House reportedly squelched an internal report based on economic data showing just the opposite, that refugees provide a net economic benefit to our Nation of more than \$63 billion.

However, instead of changing their policy to reflect these facts, the White House reportedly ordered the report to be stripped of all references to the benefits that refugees provide. Just took them out, throw them away. That left only a biased, inaccurate picture that happens to match their political narrative.

On this topic, we are sending letters this morning to the White House and to the Department of Health and Human Services seeking documents about their actions on this report.

I ask unanimous consent that our letters, Mr. Chairman, be included in the record for today's hearing.

Mr. RUSSELL. Without objection.

Mr. CUMMINGS. Finally, take the most significant long-term challenge of our Nation and the world. We face, for decades to come, climate change. We have just seen massive devastation caused by Hurricanes Maria, Irma, and Harvey which is projected to cost hundreds of billions of dollars.

We've heard for the past 6 years from the nonpartisan Government Accountability Office that climate change is one of the top risks to the financial and national security of our great country. This finding is based on overwhelming scientific evidence.

A few weeks ago, the Miami Herald asked the Republican mayor of Miami about Hurricane Irma. And this is how he responded, and I quote: This is the time to talk about climate change. This is the time that the President and the EPA and whoever makes decisions needs to talk about climate change. This is a truly, truly, truly poster child for what is to come.

Of course, he is saying what we all already know and we know in our hearts: We need to anchor our public policy on sound evidence, not baseless ideology. We cannot pick and choose which evidence we recognize and which evidence we ignore because of politics.

Today's hearing goes right to the core mission of our committee. Collecting evidence and making policy recommendations based on that evidence is exactly what the Committee on Oversight and Government Reform was designed to do. When we do our jobs correctly, it leads to important reforms.

For example, this committee recently approved a bill, by Representative Farenthold and I, introduced to provide Federal employees who blow the whistle the right to appeal their cases in courts other than the Federal circuit. This important reform responds to evidence that the Federal circuit is unfavorable to whis-

leblowers and that expanding the right to appeal to other courts will not result in a flood of appeals.

So I thank our witnesses who are here today. I look forward to your testimony.

And with that, Mr. Chairman, I yield back.

Mr. RUSSELL. And I thank the ranking member.

Before we introduce our panel of witnesses this morning, it's my honor to also note the presence of a multiparty delegation from the Government of Sri Lanka, led by the Speaker of Parliament, His Excellency Karu Jayasuriya. We welcome him and his distinguished guests.

Thank you, sir, for your presence with us here today.

I'm pleased to introduce our witnesses, the following commissioners on the Council on Evidence-Based Policymaking: Dr. Ron Haskins, the Co-Chair of the Commission; Dr. Katherine Abraham, the Chair of the Commission; Dr. Latanya Sweeney; and Mr. Robert Shea.

Welcome to you all.

Pursuant to committee rules, all witnesses will be sworn in before they testify. Please rise and raise your right hand.

Do you solemnly swear or affirm the testimony you are about to give is the truth, the whole truth, and nothing but the truth, so help you God?

The record will reflect all witnesses answered in the affirmative.

Please be seated.

In order to allow time for discussion, please limit your testimony to 5 minutes. Your entire written statement will be made part of the record. And I would also remind you about the microphones, and get them close, and also remember to turn on the button.

Thank you, sir. Please, Dr. Haskins.

WITNESS STATEMENTS

STATEMENT OF RON HASKINS

Mr. HASKINS. Thank you, Mr. Russell and Ranking Member Cummings and other members of the committee. I'm very pleased to be here today in my role as the co-chair of the Commission on Evidence-Based Policymaking. I'm a senior fellow at Brookings, and it made participating in the Commission much better than our members from California who had to keep coming back and forth.

For most of my professional career I have been focused on the importance of generating evidence, especially program evaluation evidence, to support policymaking. So when Speaker Ryan appointed me to serve as co-chair of the Commission, I seized the opportunity.

There are many issues in our country today that generate conflicting views. Some have already been expressed this morning. So we are very pleased that the need for more and better evidence—and the Commission strategy for getting there in a privacy protective way—was unanimously approved by the full Commission. All 15 members voted for the final report and all 22 recommendations that we're discussing before the committee today.

I request that the final report of the Commission be entered into the record.

Mr. RUSSELL. Without objection.

Mr. HASKINS. The Commission was established by the bipartisan Evidence-Based Policymaking Commission Act of 2016, which was jointly sponsored by Speaker Ryan and Senator Patty Murray and signed into law in March 2016. The legislation directed the appointment of 15 commissioners with a broad range of expertise, including academic researchers, data experts, administrators, and experts in computer science, data privacy, and privacy law. The Commission was provided just over 1 year to study and develop a strategy for strengthening government's evidence-building and policy-making efforts.

Decisionmakers depend on having reliable and timely information to guide their examination of how current programs and policies are working and how they could be improved. In establishing the Commission, Congress rightly acknowledged that today too little evidence is produced to meet this need.

Most of the Nation's social programs produce modest or no impacts. Let me say that again. Most of the Nation's social programs produce modest or no impacts on the problems they were meant to address. Wisely, in recent years Congress has been asking for and paying for careful evaluation of some programs. But we still don't know enough about the effectiveness of the many of the Nation's programs.

To help address this gap, the Commission was charged with developing a strategy for increasing the availability and the use of data to build evidence about government programs while also protecting privacy and confidentiality. We took both of these charges seriously.

We wanted to make sure that our recommendations were rooted in, well, evidence. So we conducted an exhaustive fact-finding effort before launching into serious discussion of how to improve evidence building. This fact-finding phase extended for 8 months and included seven public meetings, three public hearings, a request for comments through the Federal Register, and a survey of Federal offices that generate or use evidence. In all, we received input from more than 500 individuals or organizations.

Our report includes 22 recommendations designed to address the barriers to having more evidence available. The recommendations fall into three broad categories. First, improving data access, data access for evidence-building projects. Second, modernizing and strengthening the privacy protections for data used in evidence building. And, third, strengthening the Federal Government's capacity for evidence building.

I'm delighted to be joined on the panel today by my fellow commissioners, who will provide a brief overview of the recommendations that we developed under each of these three categories. First, Katharine Abraham, the chair of the Commission, will highlight recommendations that relate to streamlining and improving data access for evidence building. Second, Latanya Sweeney will describe the kinds of increased privacy protections and transparency that the implementation of our recommendations would yield. And, third, Robert Shea will discuss the set of recommendations related to strengthening the Federal Government's capacity for evidence building.

Their testimony will show that the members of the Commission share a vision in which rigorous evidence is created efficiently, as a routine part of government operations, and used to construct effective policy.

Finally, we hope our recommendations will be implemented as quickly as possible, and we look forward to partnering with the Congress, this committee included, and the administration to advance the recommendations of the Commission and to achieve a future built on evidence-based policymaking.

Thank you, Mr. Chairman.

[Prepared statement of Mr. Haskins follows:]

Testimony of Ron Haskins
Co-Chair, Commission on Evidence-Based Policymaking
House Committee on Oversight and Government Reform
Hearing of September 26, 2017

Committee Chairman Gowdy, Ranking Member Cummings, and Members of the Committee: My name is Ron Haskins; I'm pleased to be here today in my role as the Co-chair of the Commission on Evidence-Based Policymaking. I am also a Senior Fellow and hold the Cabot Family Chair in Economic Studies at the Brookings Institution, where I also co-direct the Center on Children and Families.

For most of my professional career I have focused on the importance of generating evidence to support policymaking. So when Speaker Ryan appointed me to serve as Co-Chair of the Commission on Evidence-Based Policymaking, I seized the opportunity.

The bipartisan Commission has produced a final report that documents a set of recommendations that were endorsed by all 15 members of the Commission. There are many issues in our country today that generate conflicting views – so we are really pleased that the need for more and better evidence, and our strategy for getting there in a privacy protective way, was unanimously approved by the full Commission.

I request that the final report of the Commission on Evidence-based Policymaking be entered into the record.

The Commission was established by the bipartisan Evidence-Based Policymaking Commission Act of 2016, which was jointly sponsored by Speaker Ryan and Senator Patty Murray, and signed into law in March 2016. The legislation directed the appointment of 15 Commissioners with a broad range of expertise, including academic researchers, data experts, seasoned administrators, and experts in computer science, data privacy, and privacy law. The Commission was provided just over one year to study and develop a strategy for strengthening government's evidence-building and policymaking efforts.

Decision makers rely on having reliable and timely information to guide their examination of how current programs and policies are working and how they can be improved. In establishing the Commission, Congress rightly acknowledges that today, too little evidence is produced to meet this need.

Most of the nation's social programs produce modest or no impacts on the problems they were meant to address. Wisely, in recent years Congress has been asking for and paying for careful evaluation of some programs, but we still don't know enough about the effectiveness of many of the nation's programs. To help address this gap, the Commission was charged with developing a

strategy for increasing the availability and use of data to build evidence about government programs, while also protecting privacy and confidentiality. We took both charges seriously.

We wanted to make sure that our recommendations were rooted in—well—evidence—and so we completed an exhaustive fact-finding effort before launching into serious discussion of how to improve evidence building. This fact-finding phase extended for eight months and included seven public meetings, three public hearings with the public presenting information to the Commission, a Request for Comments through the Federal Register, and a survey of Federal offices that generate or use evidence. In all, we received input from more than 500 individuals or organizations.

The feedback we received during our fact-finding process, in combination with the expertise of the Commissioners, enabled us to better understand the barriers to the effective use of government data to generate evidence.

Our report includes 22 recommendations designed to address these barriers. The recommendations fall into three broad categories:

1. Improving data access for evidence-building projects,
2. Modernizing and strengthening the privacy protections for data used in evidence-building, and
3. Strengthening the Federal government's capacity for evidence-building.

I am delighted to be joined on the panel today by my fellow Commissioners who will provide a brief overview of the recommendations that we developed under each of these three categories. First, Katharine Abraham, the Chair of the Commission, will highlight some of our recommendations that relate to streamlining and improving data access for evidence-building. Second, Latanya Sweeney will describe the kinds of increased privacy protections and transparency that the implementation of our recommendations would yield. And third, Robert Shea will discuss the set of recommendations related to strengthening the Federal government's capacity for evidence-building.

Their testimony will show that members of the Commission share a vision in which rigorous evidence is created efficiently, as a routine part of government operations, and used to construct effective public policy.

The Commission was charged by our statute to evaluate “if and how to create a clearinghouse for program and survey data” to support Federal program evaluation and policy as a possible solution for increasing access to data. We considered the establishment of a “data clearinghouse” in the sense of a data warehouse where large amounts of linked data are stored, but we rejected the idea of a clearinghouse, which could raise substantial risks for privacy. Instead, we are recommending the creation of a service to facilitate access to data and linking of data needed for statistical projects relevant to informing Federal programs and policies. This service -- which we have named the

National Secure Data Service -- would build on and reorganize existing resources within government to facilitate secure access by those inside and outside of government to government data, and especially to data sets created by linking information from multiple agencies.

The Commission views the proposed National Secure Data Service as a vital component in support of the Commission's recommendations for improved data access, strong privacy protections, and increased transparency.

Our full report lays out the Commission's vision for the future and the steps we believe are needed to ensure that we maximize every opportunity to produce better evidence in support of government policies and programs.

We hope our recommendations will be implemented as quickly as possible, and we look forward to partnering with the Congress and the Administration to advance the recommendations of the Commission and to achieve a future built on evidence-based policymaking.

Mr. RUSSELL. Thank you, Dr. Haskins.

And the chair is now privileged to recognize Dr. Abraham for 5 minutes.

STATEMENT OF KATHARINE G. ABRAHAM

Ms. ABRAHAM. Thank you, Mr. Russell, Ranking Member Cummings, and honored members of the committee. I very much appreciate the opportunity to speak with you today about the Commission's report.

I will focus on our recommendations related to secure access to confidential data for evidence building. This is an area in which I have some personal experience, having served two 4-year terms as the Commissioner of Labor Statistics in the Department of Labor.

I know that this committee has taken a particular interest in making sure that open data from the Federal Government are available to the public. The Commission, on the other hand, focused mainly on data that are not readily accessible for evidence building because of confidentiality concerns and legal restrictions.

The open data initiative has done a great deal to make government data more available for evidence building, and I know all of us applaud the work this committee has done to make that possible. Our charge was somewhat different, to figure out how to harness the power of data that can't be made publicly available.

One of the central charges in the statute that created the Commission was for us to evaluate if and how to create a clearinghouse for program and survey data to support Federal program evaluation and policy. We had to interpret what the statute meant by data clearinghouse. We understood that term to mean a data warehouse where large amounts of data would be brought together, linked, and retained to be available for evidence-building purposes.

That was an idea that we rejected. We were concerned that such a clearinghouse would create substantial risks for privacy. Instead, what we have recommended is the creation of a service to facilitate access to data and linking of data for specific projects. The data would be brought in, linked for the specific project, identifiers would be removed. Researchers, analysts, would work with the data. When the project was done, the data set would be destroyed.

This service, which we have notionally named the National Secure Data Service, would build on existing resources within government to facilitate secure access to data for analysts inside and outside of government for evidence-building purposes, especially to data sets created by linking information from multiple agencies.

During its fact-finding phase, the Commission heard about several examples of exciting research done using confidential data that has generated valuable information for designing and carrying out programs and policies, such as the path-breaking research of Stanford University Professor Raj Chetty and his colleagues on social mobility.

Too often, however, we found legal and bureaucratic barriers to accessing data have prevented researchers from studying important policy questions. Surmounting these barriers is especially difficult when the researcher seeks to access data from multiple jurisdictions or agencies.

In its review of applicable laws, the Commission found considerable variation in provisions governing data confidentiality and permissible uses of data. The laws that authorize statistical agencies, for example, include varying restrictions on who can access data that has been collected and for what purposes.

Many program agencies' authorizing statutes do not address data confidentiality and the use of data for evidence building at all. Other program agencies' laws establish narrow standards for the acceptable use of administrative data. For example, Title 26 of the U.S. Code generally limits the use of tax data to projects that would improve tax administration, precluding the use of these data under controlled circumstances and conditions for other evidence-building purposes.

To provide clarity about permissible statistical uses, the Commission recommends that Congress build on the legal framework for data protection already established under the Confidential Information Protection and Statistical Efficiency Act. We propose to extend that framework to cover the National Secure Data Service and enable it to acquire and combine survey and administrative data.

The Commission also proposes that the Congress review and, where appropriate, revise relevant statutes that place limits on the use of administrative data for statistical purposes as well as making more of the data collected by the States in the course of operating federally supported programs available for evidence building.

The Federal Government's principal statistical agencies already play an important role in generating and providing access to data. The Commission has a set of recommendations that broaden the role that the principal statistical agencies would play.

Taken together with the Commission's recommendations related to privacy, which Dr. Sweeney will discuss, we believe that our recommendations on data access will allow data that the government has already collected to be safely harnessed to produce the evidence that is needed to make government work better.

[Prepared statement of Ms. Abraham follows:]

Testimony of Katharine Abraham
Chair, Commission on Evidence-Based Policymaking
House Committee on Oversight and Government Reform
Hearing of September 26, 2017

Chairman Gowdy, Ranking Member Cummings, and honored members of the Committee,

Thank you for the opportunity to speak with you today about the Commission's report. I will focus on our recommendations related to secure access to confidential data for evidence building. This is an area in which I have personal experience, having served two four-year terms as Commissioner of the Bureau of Labor Statistics in the Department of Labor.

I know that this Committee has taken a particular interest in making sure that open data from the Federal government are available to the public. The Commission, on the other hand, focused mainly on data that are not readily accessible for evidence building because of confidentiality concerns and legal restrictions. The open data initiative has done a great deal to make government data more available for evidence building and all of us applaud the work this Committee has done to make that possible. Our charge was somewhat different—to figure out how to harness the power of data that can't be made publicly available.

During its fact-finding phase, the Commission heard about several examples of analyses using confidential data that have generated valuable information for designing and carrying out programs and policies, such as the path-breaking research of Stanford University Professor Raj Chetty and his collaborators on social mobility. Too often, however, legal and bureaucratic barriers to accessing data have prevented researchers from studying important policy questions. Surmounting these barriers is especially difficult when the researcher seeks to access data from multiple jurisdictions or agencies. The Commission believes that enabling better use of confidential data the government has already collected will produce substantial benefits for society.

In its review of laws applicable to the collection, protection and sharing of data, the Commission found considerable variation in provisions governing data confidentiality and permissible uses of data for evidence building. For example, the laws that authorize statistical agencies include varying restrictions on who can access data that have been collected and for what purposes. Many program agencies' authorizing statutes do not address data confidentiality and the use of data for evidence building at all. Other program agencies' laws establish narrow standards for the acceptable use of administrative data that limit the availability of these data for evidence building. For example, Title 26 of the U.S. Code generally limits the use of tax data to projects that would improve "tax administration," precluding the use of these data to provide critical insights about the impacts of programs and policies across the government.

To provide clarity about permissible statistical uses, the Commission recommends that Congress build on the legal framework for data protection already established under the Confidential

Information Protection and Statistical Efficiency Act (CIPSEA). Recommendation 2-3 proposes to extend the protective CIPSEA framework to enable the National Secure Data Service (NSDS) to acquire and combine survey and administrative data collected by other agencies. In addition, the Commission addresses legal limitations on the permissible uses of administrative data that preclude their use for evidence building. The Commission proposes that the Congress and the President review and, where appropriate, revise relevant statutes to authorize access and use of such data for statistical purposes, under the stringent privacy protections of the expanded CIPSEA framework. The Commission similarly recommends making more of the data collected by the states in the course of operating federally-supported programs available for evidence-building purposes.

The Federal government's Principal Statistical Agencies (PSAs) are already covered by CIPSEA. They have a long history of providing secure access to confidential data and ensuring that researchers abide by strict rules to protect confidentiality. Despite their central role in providing policy-relevant information and their expertise in data stewardship, the heads of the PSAs generally do not have a seat at the table when other senior agency officials—such as the Chief Information Officer and the Chief Privacy Officer—are considering changes that would have an impact on data resources. The Commission envisions a broadened role for the PSAs in regard to decisions about data across their departments. Specifically, the Commission recommends that Federal departments assign a Senior Agency Official for Data Policy—in most cases the head of the departmental PSA—to coordinate access to data and stewardship of the department's data resources for evidence building.

Implementation of these recommendations will be vital for ensuring that the evidence-building community has access to data well-suited to generating insights about Federal programs and policies. Taken together with the Commission's recommendations related to privacy, to be described by Commissioner Latanya Sweeney, the Commission's recommendations on data access will allow data that the government has already collected to be safely harnessed to produce the evidence that is needed to make government work better.

Mr. RUSSELL. Thank you, Dr. Abraham.
The chair is now privileged to recognize Dr. Sweeney for 5 minutes.

STATEMENT OF LATANYA SWEENEY

Ms. SWEENEY. Thank you.

Congressman Russell, Ranking Member Cummings, and members of the committee, my name is Latanya Sweeney, and my career mission has been to create and use technology to assess and solve social, political, and governance problems. I am a computer scientist, a data scientist, a professor at Harvard, and the director of the Privacy Lab at Harvard. I served as the Chief Technology Officer for the U.S. Federal Trade Commission. And I really thank you for this opportunity to speak before you today about the importance of the Commission's recommendations in protecting the American people's confidential data.

I think it's most directly relevant for you to know that my research and my research team have spent many years showing how data that should be anonymous isn't. That is, how many different ways confidential data sets on health, criminal justice, and income records, and other areas could be re-identified easily using other data sources in today's technological society.

So let me be very clear: We do have a major problem protecting the privacy of confidential data today, and it stems, primarily, because of a mismatch between our historical way of thinking about privacy and privacy protections and today's society being so empowered by technology.

The Commission believes we can securely increase access to confidential data for evidence building based in great part on new advances made in how we think about privacy, in particular data privacy, and new techniques that use the same technology that challenges privacy to help provide protections. If we don't take these actions, we risk exposing confidential data about Americans widely, which leaves us vulnerable to many problems.

The Federal Government collects a lot of information about individuals and businesses during the course of its daily operations, and much of that information is and should be open data. That is, it should be publicly accessible government information. And things like weather forecasts and train timetables do not carry the same privacy burden.

The government says it will keep some of that information confidential, like names and dates of birth of Social Security recipients. And when the government pledges to keep data confidential, the data should have strong protections, and data used should generally be made known to the American public. Versions of the data that can be rendered sufficiently de-identified should also be made publicly available.

I'm here today to tell you about why the Commission's privacy and transparency recommendations are critical to protecting the government's confidential information.

First, and of utmost importance, there's a great variation in how Federal agencies go about protecting confidential data today. Instead, this process really needs to be consistent, it needs to be rig-

orous, and it needs to be able to evolve with new technological advances.

Second, protecting the privacy of the American people means being transparent and open about decisionmaking and processes, and clear about how confidential information is being used, and giving opportunities for feedback and improvement.

So what happens now when a Federal agency wants to release a public use version of some confidential data that it's collected? Well, the 13 principal statistical agencies routinely apply rigorous methods of data masking, and they seek to review and get approval from experts on the way they're going to disclose the data before they actually release public use files.

That's the current best practice we have, and we've had it for a long time and we accept it as pretty sufficient. But the context of public use data releases has changed because the amount of information about Americans and individuals that is publicly available has grown tremendously over the last few years.

In addition, the technology to permit unauthorized re-identification has improved. Financial incentives exist in many data analytic companies to take disparate pieces of data and put them together to build profiles on individuals. Within the Federal Government alone the open data initiative made over 150,000 data sets accessible through a single website, including many administrative data sets never before released to the public.

While releasing these data can generate tremendous value, enabling entrepreneurs to produce better products and departments to understand their work better, it's important to consider how publicly available data could compromise confidentiality if we don't take appropriate actions in adhering privacy.

This is not about whether or not the data is released or not. It's not a binary decision. That's how we've historically looked at it. But today's technology allows us the opportunity to say we can, in fact, provide public versions of data. The question is which version and the techniques used to render that version.

Government agencies follow their own applicable laws and regulations in providing access to their confidential data. These agencies are not necessarily coordinated in the decisionmaking they make. And the fact that they have different policies and different procedures about what it means to be identifiable creates a lot of problems. Sometimes two different agencies releasing the same data make different decisions, and the two pieces can be put together.

Some program agencies use the same best practices that we just described earlier about the principal statistical agencies to assess the risk of re-identification, an ongoing process. The Department of Education even set up a dedicated disclosure review board for its program agencies.

But some program agencies do little more than just remove the explicit identifiers, like name and address, leaving lots of other pieces of information out there available that can be linked to other data to re-identify individuals. Confidential government data collected by program agencies are often subject to FOIA with minimal redaction and also with inconsistent coordination.

The problem is that there are so many sources of data out there today that can be matched to insufficiently de-identify confidential data to re-identify individuals or businesses. And in my resume you will see many collections at both the State and Federal level for where we've demonstrated this.

In fact, we just released a study showing how data on air and dust samples from 50 homes in two communities in California could be combined with data released under the Safe Harbor provision of the Health Information Portability and Accountability Act to uniquely and correctly identify 8 of 32, or 25 percent, by name.

Many people—many of us—many of our policies look at the protections available in HIPAA as being sufficient and Safe Harbor as being sufficiently strong. The fact that we were able to re-identify uniquely and correctly 25 percent of the records is alarming. It, again, speaks to the nature of data that's out there and the old way that we think about privacy.

So I can tell you by name health information about eight people in one community from which that data was released, de-identified. And if those eight people lived in your district and they learned that a Federal agency had just released their data with insufficient privacy protections, you would be likely hearing about it also.

This is what my colleagues and I discover every day with many different types of data from many different venues. And this is a real problem. And it's important that any legislation implementing the Commission's recommendations address the seamless improvements of privacy head-on.

Many programs have released the identified public use data files for decades without being required to ever formally assess its risk. They also often include provisions that make it impossible to actually inform them about the risk that we find.

The Commission's recommendations, 3-1 in particular, will make sure that Federal agencies planning to release de-identified confidential data use state-of-the-art methods to protect individuals and businesses from privacy harm. Instead of being a static way of thinking of privacy in a binary decisionmaking, it's a continuum. And as we get better and better about how to think about privacy and technical tools to release privacy, it enables those new techniques to go right into government use.

Next, I'd like to explain why transparency is so important to privacy. Privacy does not mean secrecy, and there's often a lot of confusion that the idea of privacy is to hide it and not let anyone know about it. But, in fact, the Commission believes that advancing beyond the status quo and achieving unparalleled transparency means first telling the public about how government data are used for evidence building, and, second, regularly auditing whether the government is doing what it said it would do to protect privacy when allowing access to government data for evidence building. Further, transparency means how was the data redacted and learning new ways and encouraging the use of new techniques.

As a first step, the government needs to make clear its decisions about which data are open data and which data are nonpublic confidential data. The Commission calls for OMB to develop a public inventory of data available for evidence building, including a deter-

mination of the sensitivity level of the data. Based on the data's sensitivity, we are recommending—

Mr. RUSSELL. Dr. Sweeney, we—

Ms. SWEENEY. Oh, I'm so sorry.

Mr. RUSSELL. Yeah.

Ms. SWEENEY. And I'll stop right there.

[Prepared statement of Ms. Sweeney follows:]

Testimony of Latanya Sweeney
Commissioner, Commission on Evidence-Based Policymaking
House Committee on Oversight and Government Reform
Hearing of September 26, 2017

Chairman Gowdy, Ranking Member Cummings, and members of the Committee,

My name is Latanya Sweeney and my career mission has been to create and use technology to assess and solve societal, political and governance problems. I am a data scientist and the Director of the Privacy Lab at Harvard University. I also served as the Chief Technology Officer for the Federal Trade Commission. Thank you for the opportunity to speak before you today about the importance of the Commission's recommendations in protecting the American people's confidential data. I think it is most directly relevant for you to know that my research team and I have spent many years showing how individuals in supposedly confidential datasets on health, criminal justice, and other areas can be re-identified by using other data sources and computing power.

Let me be very clear: we have a problem protecting the privacy of confidential data today. The Commission believes we can securely increase access to confidential data for evidence building. But unless we also increase privacy protections as recommended in the Commission's report, we risk a bigger problem than barriers to data access. We risk exposing confidential data about Americans.

The Federal government collects a lot of information about individuals and businesses during the course of its daily operations. Much of that information is, and should be, open data—publicly accessible government information like weather forecasts and train timetables. The government says it will keep some of that information confidential, like the names and birth dates of social security recipients. When government pledges to keep data confidential, the data should have strong protections, and data use should generally be made known to the American public.

I'm here today to tell you about why the Commission's privacy and transparency recommendations are critical to protecting the government's confidential information. First and of utmost importance, there is great variation in how Federal agencies protect confidential data—this process should be consistent and rigorous. Second, protecting the privacy of the American people means being transparent, open, and clear about how their confidential information is being used and giving them opportunities to provide feedback.

So what happens now when a Federal agency wants to release a public use version of some confidential data it has collected? The 13 Principal Statistical Agencies routinely apply rigorous methods of data masking and seek review and approval from experts on a disclosure review board before releasing public use files. That is the current best practice and for a long time we accepted it as pretty sufficient. But the context of public use data releases has changed because the amount of information about individuals that is publicly available has grown. In addition, the technology to permit unauthorized re-identification has improved. Within the Federal government alone, the Open

Data initiative made over 150,000 datasets accessible through a single website, including many administrative datasets never before released to the public. While releasing these data can generate tremendous value, enabling entrepreneurs to produce better products and departments to understand their work better, it is important to consider how publicly available data could compromise confidentiality.

Government agencies follow their own applicable laws and regulations in providing access to their confidential data. The problem is that agencies have different policies and procedures for what it means to release data that is “not individually identifiable” (as it says in the Privacy Act). Some program agencies use the same best practice techniques as Principal Statistical Agencies to assess risk of re-identification. The Department of Education even set up a dedicated disclosure review board for its program agencies. But some program agencies do little more than remove direct identifiers such as name and address, and perhaps remove outliers before assuming the dataset is sufficiently de-identified for public release. And confidential government data collected by program agencies are often subject to FOIA with minimal redaction.

The problem is that there are so many sources of data out there today that can be matched to insufficiently de-identified confidential data to re-identify individuals or businesses. My colleagues and I just released a study showing how data on air and dust samples from 50 homes in two communities in California could be combined with data released under the Safe Harbor provisions of the Health Insurance Portability and Accountability Act (HIPAA) to “uniquely and correctly identify [in one community] 8 of 32 (25 percent) by name and 9 of 32 (28 percent) by address.” Think about it: I can tell you, by name, health information about 8 people in one community from data that was released publicly as “de-identified.” If those 8 people lived in your district and they learned that a Federal agency had just released their data with insufficient privacy protections, you would likely be hearing about it. This is what my colleagues and I discover every day, with many different types of data. This is a problem, and it’s important that any legislation implementing the Commission’s recommendations address privacy head on.

Many programs have released de-identified public-use data files for decades without being required to formally assess risk. The Commission’s recommendation 3-1 will make sure that Federal agencies planning to release de-identified confidential data use state-of-the-art methods to protect individuals and businesses from privacy harm.

Next, I’d like to explain why transparency is so important to privacy. Privacy does not mean secrecy. In fact, the Commission believes that advancing beyond the status quo and achieving unparalleled transparency means first, telling the public about how government data are used for evidence building and second, regularly auditing whether the government is doing what it said it would do to protect privacy when allowing access to government data for evidence building.

As a first step, the government needs to make clear its decisions about which data are open data and which data are nonpublic, confidential data. The Commission calls for OMB to develop a public inventory of data available for evidence building, including a determination of the sensitivity level of the data. Based on data sensitivity, we are recommending that OMB establish standards for appropriate and controlled access. And this is important—agencies should use technology and statistical masking techniques to develop less sensitive versions of datasets when possible and make more information available to the public and to researchers with appropriate safeguards for evidence building.

This idea of multiple versions of datasets, or tiered access, isn't some theoretical concept. Tiered access approaches are practical and we can actually implement them today. Tiered access is an application of data minimization, which is a key privacy safeguard for evidence building. Imagine a system where each dataset is labeled based on whether versions have more or less identifiable or sensitive information. Then, researchers or the public only receive an appropriate level of access to complete their project with appropriate privacy safeguards.

This tiered access approach is a way to increase evidence building and better protect privacy at the same time. And we already have examples of how it is being done today in the Federal government. The Commission found that many PSAs implement tiered access programs that set data access and security requirements based on an assessment of dataset sensitivity. Tiered access is also taking root in Europe in response to the implementation of the General Data Protection Regulations and at home by organizations such as my own, Harvard University, to define sensitivity levels and set corresponding access and data security protections.

Recommendation 4-3 calls for OMB to develop a transparency portal that includes the data inventory with sensitivity levels, risk assessments, and descriptions of projects and approved researchers using confidential data for evidence building. Researchers like myself and my colleagues must have a way to give feedback to the government about potential risks they find. And for the public to provide feedback on data sensitivity. Therefore, the Commission's report calls for a feedback mechanism as part of the transparency portal.

Preventing bad actors from breaking into confidential data they are not authorized to use requires consistent and rigorous processes to assess the risk of release in light of all other sources of data. A disclosure review board has the expertise to determine if agencies are doing enough to protect the privacy of the American people. The Federal government needs to help the public understand how confidential data are being used and conduct regular audits to ensure compliance with privacy laws, regulations, and best-practice procedures.

Thank you for your time. I urge you to move swiftly to implement these changes and improve how the Federal government protects the confidential data it collects from the American public.

Mr. RUSSELL. I've tried to be generous from the chair, but we've given you 10 minutes instead of 5. And, hopefully, we can get to some of these other points during the questioning.

Ms. SWEENEY. No problem.

Mr. RUSSELL. We want to be respectful, also, to Mr. Shea, who it is my privilege to recognize for 5 minutes.

STATEMENT OF ROBERT SHEA

Mr. SHEA. I assure you the committee members benefit far more from hearing from Latanya than they will from me. So I apologize in advance.

Mr. Chairman, Ranking Member Cummings, and members of the committee, I'm Robert Shea, a proud former staff member of this committee, and I'm pleased to be here today with my fine fellow commissioners testifying on behalf of the Commission on Evidence-Based Policymaking.

You've already heard the testimonies of my colleagues that highlight the Commission's recommendations related to data access and privacy protections. I'd like to focus on those recommendations that pertain to the strengthening of the Federal Government's capacity for evidence building.

Today, evidence building by government occurs unevenly. Some departments have robust approaches for routinely generating and using evidence, but these are the exception. We've taken to calling those who generate, manage, and analyze data, those who transform information into evidence, and those who support those functions through the routine processes of government, members of the evidence community.

Principal statistical agencies and other statistical programs, program evaluation and policy research offices, program administrators, performance management offices, policy analysis offices, and privacy offices, all play important roles in evidence building. But many shared with the Commission that administrative barriers hamper the efficient production and the use of this evidence.

To achieve the Commission's vision—a future in which rigorous evidence is created efficiently as a routine part of government operation and used to construct effective public policy—Federal agencies must have the capacity to support the full range of analytic approaches required for evidence building.

To grow the program evaluation function across agencies, we recommend Federal departments anoint a chief evaluation officer who'd be charged with establishing department-wide evaluation research policies, coordinating technical expertise for evaluation, identifying priorities for departmental program evaluation, and adopting human capital strategies that expand the department's program evaluation capacity.

We also recommend agencies develop multiyear learning agendas to support the generation and use of evidence. A learning agenda is essentially a strategic plan for evidence building, identifying important policy questions relevant to the department's mission. The learning agenda can be used by leadership and by Congress to prioritize research investments.

Several of the evidence-building examples referenced in my colleagues' testimony involve linking data sources administered by

different Federal departments. Officially implementing evidence-building activities across government requires strong coordination. This committee has already vested a great deal of responsibility in OMB, but we firmly believe it's the right institution to help coordinate these activities. We recommend OMB facilitate cross-government coordination and consider whether consolidation or reorganization of evidence-based policymaking functions at OMB would accelerate adoption of the Commission's recommendations.

The commissioners also identified actions related to procurement and the review and approval processes for new data collections that would require little cost but offer substantial benefits and savings while making it easier to produce evidence. These are small but important reforms.

The Commission acknowledges and appreciates the role that this committee in particular plays in reforming and overseeing the operations of the Federal Government. Though some of our recommendations will require legislation, others simply require administrative action. You can help us ensure OMB and Federal departments use their existing authorities to begin to increase access, enhance privacy, and expand capacity, and ultimately create a future in which rigorous evidence is created efficiently, as a routine part of government operations, and used to construct effective policy.

No pressure, but we're all counting on you, and we're here to help in any way we can.

[Prepared statement of Mr. Shea follows:]

Testimony of Robert Shea
Commissioner, Commission on Evidence-Based Policymaking
House Committee on Oversight and Government Reform
Hearing of September 26, 2017

Committee Chairman Gowdy, Ranking Member Cummings, and esteemed members of the Committee, my name is Robert Shea and I am pleased to be here today with my fellow Commissioners testifying on behalf of the Commission on Evidence-Based Policymaking. I am currently a Principal and lead the Public Sector Strategy practice at Grant Thornton. I also bring to bear, in my role as a Commissioner, my former experience both as Associate Director for Management at OMB, as well as several years of experience on Capitol Hill, including with this very Committee.

You've already heard the testimonies of my colleagues that highlight the set of recommendations put forward by the Commission related to improving data access in support of evidence-building, as well as modernizing and strengthening the privacy protections for data used in evidence building. Implementing the features of the Commission's vision to improve the use of data for evidence building requires more than just improving access to data and privacy protections, however. The capacity to analyze data inside and outside government must exist in order to apply insights to inform policymaking. I'd like to focus my remarks on the Commission's recommendations that relate to strengthening the Federal government's capacity for evidence-building, which may include partnering with organizations and individuals outside of government.

Today, evidence building about government policies and programs occurs unevenly. Some departments have developed robust approaches and procedures for routinely generating evidence. But these are the exception rather than the rule.

The Federal evidence-building community includes those who generate, manage, and analyze data, those who transform information into evidence, and those who support those functions through the routine processes of government. Within the Federal evidence-building community, Principal Statistical Agencies (PSAs) and other statistical programs, program evaluation and policy research offices, program administrators, performance management offices, policy analysis offices, and privacy offices all play an important role in evidence building. During the Commission's fact-finding phase, both governmental and non-governmental witnesses identified numerous administrative barriers within the Federal government that hamper the efficient production of evidence.

To maximize evidence building, Federal departments must have the capacity to support the full range of analytic approaches required for evidence building, including the development of statistics, program evaluation, and policy research. These functions must be operational, appropriately resourced, and well-coordinated within and across departments. Strong leadership within government that prioritizes evidence building and creates the demand for evidence is vital for institutionalizing these functions. Without a strong institutional foundation, other recommendations

related to improving data access, establishing the National Secure Data Service, and implementing enhanced privacy protections will not have the comprehensive impact desired and envisioned by the Commission.

The Commission's recommendations related to capacity acknowledge the particular need to grow the program evaluation function across Federal agencies. As such, the Commission recommends that Federal departments identify or establish a Chief Evaluation Officer, in addition to needed authorities to build a high performing evidence-building workforce. The Chief Evaluation Officer would be charged with establishing department-wide evaluation and research policies, the coordination and provision of technical expertise for evaluation and research across the department, leadership in identifying priorities for departmental program evaluation and policy research, and the establishment of human capital strategies that expand the department's capacity for program evaluation.

A second capacity recommendation directs Federal departments to develop multi-year learning agendas to support the generation and use of evidence. A learning agenda is essentially a strategic plan for evidence building, identifying short- and long-term research and policy questions relevant to a department's mission and legal responsibilities. The learning agenda can be used by leadership to prioritize the set of research and policy questions to be pursued by the department, and also as a public document, alerting external audiences to the priority research and policy questions of interest to a department. In addition, the learning agenda can be used as a mechanism for organizing various units within the evidence-building community in a department in order to determine the best approach to answering a priority research or policy question and to allocate the work appropriately across the different evidence-building functions. The Chief Evaluation Officer and the Senior Agency Official for Data Policy would provide leadership for the development and implementation of a learning agenda.

Several of the examples of evidence building that have been referenced today involve the linking of data sources that are administered by different Federal departments. Efficiently implementing evidence-building activities across government requires a strong coordination function to address such crosscutting research and policy questions, minimize duplicative efforts, and reduce the burden on the public. The Commission, therefore, recommends improving the coordination of government-wide evidence building. Specifically the Commission recognizes that the Office of Management and Budget is well situated to help coordinate these activities. We recommended that OMB facilitate cross-government coordination and consider how a greater commitment to foundational information policy responsibilities can be achieved, including through any consolidation or reorganization at OMB that might be necessary.

In addition to increasing cross-governmental coordination, to generate a greater volume of evidence in a more efficient manner, foundational administrative processes must be aligned and tailored to better support evidence building. The Commission identified a specific set of actions related to procurement and streamlining the review and approval processes for new data collections that

would require little cost, but offer substantial benefits and savings while making it easier to produce evidence.

Finally, through the course of the Commission's research and deliberations, the topic of resources repeatedly emerged as a major perceived need for the evidence-building community and a challenge for improving the volume and quality of evidence produced. The Commission believes that a responsible investment of resources in more and better evidence holds the potential to yield substantial savings in the longer term as programs that are improved become more cost-effective, and as programs that are not effective are discontinued. The Commission recognizes that resource prioritization is essential to ensuring the goals of the Commission are achieved. Thus, the recommendations of the Commission balance the need to prioritize evidence building while recognizing fiscal constraints. In some departments, sufficient resources already exist to enable evidence building, though such resources may have use restrictions that inhibit the most cost-effective approach for evidence building.

These capacity recommendations are small steps that government can take to ensure the institutional foundations are present to support evidence building and use in policymaking. More broadly, all of the Commission's 22 recommendations present a comprehensive strategy for addressing the greatest problems facing evidence building today: unintentional limits on data access, inadequate privacy practices, and insufficient capacity to generate the amount of quality evidence needed to support policy decisions. The Congress, the President, and the American people are ill-served by this state of affairs. The Commission believes that fully implementing the Commission's recommendations will lead to substantial progress in addressing these challenges, enabling more and better evidence for our society, generated in a more secure fashion.

We appreciate the role that this Committee plays in reforming and overseeing the operations of the Federal government. It occurs to me that the Committee can drive implementation of the Commission's recommendations in both of these roles. Some number of recommendations will require legislation to codify requirements for access and privacy, and this Committee can make sure those requirements have the force of law. But this Committee's oversight function can be also be a valuable tool for ensuring that OMB and Federal departments use their existing capacity and authorities to begin to increase access, enhance privacy, and expand capacity.

We look forward to answering questions from the Committee both today and into the future as you consider the implications of this important report on advancing evidence-based policymaking.

Mr. RUSSELL. Thank you, Mr. Shea.

And thank all of you for your testimony.

I will now recognize myself for 5 minutes as we move to the questioning period.

Dr. HASKINS, you made an interesting comment here in your testimony, and also in your written statements, that most of the Nation's social programs produce modest or no impacts on the problems that they were meant to address. I think most of America would not be shocked by that statement, many of the things that we see. But can you give us some examples of that and, in your discovery, why that is?

Mr. HASKINS. One example that people often mention is Head Start, which most people, if you do a survey of Americans, they think the program is immensely successful. But research shows that it's not necessarily successful.

Let me be clear on this. There are some—and this often happens in programs implemented around the country. There are individual sites that are very successful. But if you average up all the sites and see if they perform better than, say, a control group that did not have the same experience, they usually fail. There are all kinds of evidence that 80 to 90 percent of programs in medicine, in social science, and in business, fail.

That's the main reason that we need to have more evidence. We need to develop these programs. I have a feeling we may have a chance to talk about some of these programs later in this hearing or it would be appropriate for you to look at these programs later. But you would find, if you did this, that most programs, like Head Start, often have a good reputation, but when you look at what they do in the country as a whole they are not successful.

Mr. RUSSELL. And that's one of the key things that this committee does, is to have oversight for the American people so that we can make sure that we're spending dollars wisely instead of just creating new bureaucracies that really don't address the problems that they were designed to do.

Mr. HASKINS. Mr. Chairman, can I add one very quick thing?

Mr. RUSSELL. Sure can.

Mr. HASKINS. Getting from here to there is going to be extremely difficult. We're in the middle of this now with teen pregnancy prevention, home visiting, and several other programs that were initiated 5, 6, 7 years ago. We need a strategy that we can gradually build up these programs. We are not going to go overnight from 10 or 15 percent success to 80 percent. That's just not going to happen.

So we need better strategies for finding the right people at the local level, giving them the right resources, having Federal agencies that can help them implement their programs. It takes all of the above. It will take us years to develop this. We need to be patient and do it right.

Mr. RUSSELL. Thank you.

Dr. Sweeney, recommendation 3-3 in the report is that each Federal department assign a new role of senior agency official for data policy. We already have chief information officers and senior agency officials for privacy, and yet now we want to create new positions when these old positions, when they were first recommended

and implemented, it was to fix problems. And now we're going to fix a problem by having another person to fix a problem. What's the difference between these existing positions and why?

Ms. SWEENEY. You know, that succession of increasing those positions really speaks to the changes in society and our operations based on technology. Many of the Federal information technology officers are primarily focused on just that, the machines themselves, the ITs, the infrastructure of the systems on which work is based. The chief privacy officer in most of the agencies is based on making sure that the agency is in compliance with privacy laws and regulations.

But what we're talking about is a different issue. It's about the data that's on the technology, the data that's being provided within the context of existing privacy laws. That is, the agency has the right to give out the data or not, or is responding to the Privacy Act. That would be in the chief privacy officer. But what version of the data is actually being given? That's a technical analysis that neither of the other two would be able to actually implement.

Mr. RUSSELL. I see.

And, Mr. Shea, in your comments and also your written report, you talked about agencies using learning agendas. And there's been some examples of success on that. I mean, imagine having a strategic vision and laying that out. Could you give us an example of what you're talking about there and who's using it?

Mr. SHEA. Sure. I think there are some agencies, perhaps the Department of Labor, Department of Education are ones, that identify early in the year, or perhaps over a longer term, what are the major questions they want answered. And they contract for rigorous independent evaluations of programs that answer a lot of those important questions. And then they do a better job than most at integrating the evidence that they learned into their decision-making process.

This committee has a long history of trying to inject outcome-based management in Federal departments and agencies. This is a step in that maturity, in our view.

Mr. RUSSELL. Okay. Thank you.

And my time has expired. And it's now my honor to recognize the ranking member, the gentleman from Maryland, Mr. Cummings.

Mr. CUMMINGS. Thank you very much.

Dr. Abraham, one of the primary themes of the Commission's report is the need to provide researchers both inside and outside the government with better access to government data. You noted in your testimony that the kind of data the Commission focuses on is data that the government keeps confidential because it contains sensitive information. Can you give us an example of the kind of data you're talking about?

Ms. ABRAHAM. So we're talking about data on participation in programs and data on the outcomes associated with participation in those programs, data on the earnings of people, such as unemployment insurance wage record data, potentially Federal tax data on earnings that, if made available to researchers under strict confidentiality protections, could let them do a much better job of understanding how well these programs that Ron was talking about were actually working.

Mr. CUMMINGS. Now, the reason this data is not publicly released is because it contains sensitive information about individuals, such as Social Security numbers.

Ms. ABRAHAM. Exactly.

Mr. CUMMINGS. And the Commission recommends that agencies provide better access to data, but, obviously, without compromising the security of the sensitive information in that data. Is that right?

Ms. ABRAHAM. That's right. We're envisioning this National Secure Data Service where data on, say, participation in a program would be brought together with data on earnings outcomes. The staff there would link the data up. The identifiers would be removed. The researchers would be given access to the data within that secure enclave to carry out the analysis. The only kind of results that would be released from such projects would be aggregated information that didn't allow anyone to be identified.

Mr. CUMMINGS. Well, Dr. Abraham, you talked about how confusing and inefficient it can be when laws passed by Congress impose inconsistent restrictions on how government data can be used. For example, the Commission highlighted the Higher Education Opportunity Act. That law limits how the data the Department of Education collects from colleges and universities can be used. That limitation is creating less accountability for the performance of those institutions.

Should Congress amend the Higher Education Opportunity Act to increase access to data from colleges and universities?

Ms. ABRAHAM. The recommendation of the Commission was, essentially, that the Congress take a hard look at that. As I think is clear from reading our report, that kind of limitation on how data can be used does reduce accountability, it does reduce our ability to understand what we're getting for the Federal dollars that we're spending. And we certainly think that another look should be taken at that.

Mr. CUMMINGS. Well, are there other laws Congress should reexamine for reform?

Ms. ABRAHAM. Well, I mean, there are similar restrictions in law on the Federal Government compiling information about people participating in workforce training programs supported by Federal dollars. That would be another example of something we would think the Congress would want to take a look at.

Mr. CUMMINGS. Now, Dr. Sweeney, you said in your testimony that the government should make more information available to the public about its data, how the data is being used. Is that something that the administration could do now?

Ms. SWEENEY. In some situations. Not all data—they could. But there's no incentive in most—there's no one answer because at the last I counted we have 2,167 privacy laws and regulations in the United States, and they're inconsistent with each other. And so there's no one answer, could they just do that unilaterally. But, in certain places, that certainly could be done. But the incentivizes may not be there, necessarily, to give it. And whether an agency chooses to make that decision is not transparent.

Mr. CUMMINGS. Well, I agree that it's important that data collected by the government be as transparent and as accessible as we

possibly can make it. Taxpayer money is spent collecting the data, and taxpayers deserve the highest rate of return possible.

Would you agree, Dr. Sweeney?

Ms. SWEENEY. Yes, I would, especially because versions of the data can be made free of privacy concerns.

Mr. CUMMINGS. With that, I yield back, Mr. Chairman.

Mr. RUSSELL. The chair now recognizes the gentleman from Tennessee, Mr. Duncan, for 5 minutes.

Mr. DUNCAN. Thank you, Mr. Chairman.

I noticed that three of your chapters deal with restricting access to confidential information, and especially with privacy. Most of your recommendations deal with privacy. And then there was an article in The Hill newspaper which said that your goal is to make sure that Federal tax dollars are spent effectively and how we keep up with rapid, advancing technology and still make effective public investments.

I read after the—also, today, just this morning, the National Journal Daily that's on all of our doors each morning said the Equifax hack is going to lead to sweeping cybersecurity legislation.

About 2 or 3 months ago there were some worldwide cyber attacks. And Robert Kuttner, who is the co-founder of The American Prospect magazine and a very liberal columnist who I wouldn't ordinarily quote, he wrote this. He said: Last week's cyber attack could produce the wrong lessons. The immediate take-away seems to be that large institutions need much better cybersecurity systems.

But he goes on. He says: Hackers will always be able to find ways of getting into network systems. The fantasy of ever-better cybersecurity is delusional. We could spend half the GDP on network security and someone will still find a way to breach it.

And so I guess I have two sort of related questions. Number one, do you agree with Kuttner that—I assume you don't agree with Kuttner that cybersecurity is delusional or is a mega-billion-dollar hoax. And I'd like to hear your comments about that.

But, secondly, if the goal of your Commission is more effective spending of Federal money—you know, I drive cars that are several years old. They're still working real well. But yet they say that computers are obsolete the day they're taken out of the box.

So how do we have the state-of-the-art technology that's already been mentioned in your testimony here today, yet we get effective use—we can't just throw away a computer just because the next year they come out with one that's got more bells and whistles on it.

And I sometimes wonder—I know at this committee several years ago, we had a business in here that had downloaded 250,000 Federal tax returns just to show that it could be done. So I sometimes wonder if the technology and the internet that has done so much for us has also almost completely done away with privacy.

So, Dr. Sweeney?

Ms. SWEENEY. Thank you.

So first of all, what the Commission is addressing does not actually have anything to do with cybersecurity. So the relationship between security and privacy is really interesting because usually

when we think about computer security, it's the breaking in of a machine, you know, getting in through illegal means.

The data that we're talking about are data that's given away freely, whether it's through open data, through a public use file, or chosen not to be given away out of claimed fear or something like that.

So these kind of data decisionmaking is not the same as cybersecurity. No one is breaking in. This is data that's usually freely given away or decisions made not to do so at all.

Mr. DUNCAN. Well, if it's given away so freely, why is there so much of your report dealing with privacy?

Ms. SWEENEY. Because the errors are happening on both sides. We have situations where the data is given away, and it leaks private information, sensitive information, health information, income information, and so forth, on Americans. And at the same time there's data that could be incredibly useful for evidence-based policymaking that isn't being given out at all because they say they don't know how.

And so we have problems on both sides, and the reason we're having those problems is because the privacy decision making system doesn't really use any of the technology.

So you talked about the wave of the technology curves and the speed at which technology changes. What the Commission does is it basically says we want privacy decisionmaking to ride the wave of the technology and not be left in its 1970s format.

Ms. ABRAHAM. If I could just jump in with one quick comment.

Mr. DUNCAN. Sure.

Ms. ABRAHAM. In terms of the cybersecurity issue, a point that I would make is that what we're talking about is better use of data that are, in most cases, already collected and being held. So in terms of that risk, we're really—we should do as well as we can on the cybersecurity issues, but we're not increasing that risk in any of our recommendations.

Mr. DUNCAN. What about the second part? How do we have effective use of tax dollars but we still keep up with the rapid advancing technology?

Ms. ABRAHAM. I think that what Latanya's response was, was attempting to answer that, which is we're not so much talking about the hardware and having to buy new hardware. We're talking about applying the right sort of methods to the way that data files that are going to be released get structured.

Mr. DUNCAN. All right. My time is up.

Mr. RUSSELL. The chair now recognizes the gentleman from Massachusetts, Mr. Lynch, for 5 minutes.

Mr. LYNCH. Thank you, Mr. Chairman, and I thank the ranking member for his good work.

And thank you to the witnesses for helping the committee with this task.

Dr. Haskins, we do a lot of work on this committee regarding veterans. And we spend right now about \$14 billion a year on educating veterans after their service. Yet, we don't collect any data on how efficacious or how much good we are doing on behalf of those veterans. The intent is there, but there's no followup. And I'll give you an example.

So we have a GI Bill that we've offered to members of the military. And in the past not only have we given it to veterans, but we've also said to Afghan and Iraqi veterans that if you already have a degree in higher education, you can give this to your child, which is a wonderful, wonderful benefit.

Some of the members of the committee were in Afghanistan not too long ago and we met—we were with a Stryker Brigade from Washington State. And one of their sons was actually—one of the officers there, one of their sons was entering University of Washington.

But when we look at the numbers, we find that only 50 percent—50 percent—of our veterans are actually using that bill. That's unbelievable. With the cost of education, with the benefit that it could provide to them and to their children, there's only 50 percent uptake of that benefit. And then we don't know whether, for those who do choose to go to college, we don't know if those veterans are actually benefiting to the full extent that they may.

So there's sort of a vacuum of usable data. I think that part of the problem is because of the restrictions we put on the use of information regarding—under the Higher Education Act—we have put a clamp on some of that. And I think we've restricted our own ability to collect and to use that information.

It's unbelievable to me that that benefit would be out there, so desperately needed, and left unused at least half of the time.

Do you have any thoughts on that, Doctor?

Mr. HASKINS. Yeah. Several. One thought that I have—first, of all, I want you to know that I went to school on the GI Bill during the Vietnam era. So I am very appreciative of what Congress does for—

Mr. LYNCH. Thank you for your service to your country, Doctor.

Mr. HASKINS. Yeah. I was glad to do it and even more pleased to go to college and have somebody help me.

Second point, we can find out almost anything you want to know about these programs. We have great research designs for almost all education and training programs. Dr. Abraham was the head of the BLS and knows as much about this as anybody in the country.

Mr. LYNCH. I'm sure.

Mr. HASKINS. So if you approve the money and direct the Department of Labor to do careful studies of these veterans programs, they can answer any question that you want.

Now, I'm going to tell you something that's very important. It is not the case that you can just take anybody that comes out of military service and send them to a good university and they'll do well.

I have looked into this issue several times over the years, and I always come away with the same conclusion. There is a recent book by Harry Holzer, who is a very well-known labor economist, comes to the same conclusion. A lot of kids are not ready to go to college. And the GI Bill, they don't wind up with a debt. But on Pell Grants and other means and loans, they can wind up owing money. And they don't get the degree, they owe the money, they're really in a bind.

So those are the kind of considerations we need to take into account. People in the military need to be better prepared for college, and many other people do as well. So the idea that we're going to

send all of our kids to a 4-year college and they're going to make \$80,000 a year, that's not going to work. They need other things. And, fortunately, the Department of Labor specializes in those other things as well.

Mr. LYNCH. Dr. Haskins, let me just reclaim my time. I concede your point. However, I also have examples in my own district where veterans did not know that they could have sent their child to college on their GI Bill. So that was a missed opportunity. And not only that, but a high number of our veterans are coming out of the service with great skills in STEM, you know, on science and math, and those are also missing the opportunity.

But I do appreciate the candor and the insightful answer that you've provided. Thank you.

And, Mr. Chairman, I yield back.

Mr. RUSSELL. The gentleman yields back.

The chair now recognizes the gentleman from Texas, Mr. Farenthold, for 5 minutes.

Mr. FARENTHOLD. Thank you very much.

I want to visit with Dr. Sweeney a little bit. You've kind of hit on a topic that I'm on three sides of. I'm the sponsor of the OPEN Gov Data Act, which makes a lot of government data available that was just included in the Senate version of the NDAA. So it actually has a pathway to becoming law.

I also consider myself to be a privacy advocate. And as someone who runs a campaign, having a lot of information about individual people that's identifiable makes it a whole lot easier for me to get them the information that they need to make an informed choice to vote for me.

But I want to talk a little, you mentioned that—you cite some examples of how anonymized data is de-anonymized. And I wasn't able to find it in any of the material that I have. I wanted to ask that you could get that to me. I assume it's more complex than just finding two data sets that have a common field. But that's probably the easiest way to do it.

Ms. SWEENEY. Are you asking—clarification: Are you asking about re-identifying or de-identifying?

Mr. FARENTHOLD. Yeah, de-anonymizing or re-identifying information. You said some—you had some examples. I'd like to see those, and I couldn't find them in the material that I had.

Ms. SWEENEY. Very easy. I have a long list of them. I'll give you one very simple one. Washington State releases hospitalizations on—all hospitalizations made in the State. It's over a million visits. And we were able to just simply match the de-identified versions of those data against simple blotter stories, the kinds of things, you know, that appeared in the newspapers, just simple matching, no statistics or anything, for 41 percent of the records—

Mr. FARENTHOLD. So is that, is the solution to that just more education or policy with respect to how that information is released, or—I mean, is there a way to solve this, or is this just going to an inherent unsolvable problem?

Ms. SWEENEY. So I'll stick to the Washington State example, because I think it's a great one.

So Washington State responded by getting rid of its 1970s way of thinking about privacy. That is the old way of saying, "Oh, I just

removed these fields, and the rest of the data is fine.” And they went instead to a risk-assessment model, exactly the thing that we’re talking about in the Commission report. And they came out the other where you can still get, for \$50, the 1 million records, but now it has a scientific assessment that’s done that you can’t do the same type of re-identification.

If you still need the more sensitive version, they have a more aggressive application process that you would have to—

Mr. FARENTHOLD. So this is solvable?

Ms. SWEENEY. Yeah, it’s solvable. And risk assessment is a great way to do it, because risk assessment keeps us riding what’s the latest thing that the technology allows.

Mr. FARENTHOLD. And I think this is important, because I think way too many decisions here in Washington are made on anecdotal information and not scientific information.

Now, Dr. Haskins, you talk about—the report talks about creating a whole new government agency. As a conservative who wants to shrink the size of government, that really rubs me the wrong way.

Why couldn’t this be something that’s rolled into—well, let’s pick the Census Bureau, because they’re the first agency that comes to mind for dealing with large amounts of data and has that experience. Can there be something done without setting up a whole new bureaucracy? And could you talk a little bit about why you recommend creating a whole new government agency?

Mr. HASKINS. Yes. As Dr. Abraham already said, we created this—or proposed this new agency called the National Secure Data Service. It would first exist within the Department of Commerce. And one of our explicit intents was not to create some big new agency that would have mountains of data. Rather it was to build on things the Census Bureau is already doing and expand those gradually over a period of years so that primarily we could have a temporary repository for data that is needed for good studies that have been approved through an elaborate process. And then the data would be sent back to wherever it came from, whatever agency it came from.

And over a period of years, I could imagine that we would wind up spending more money on this new National Secure Data Service. But to begin with, I think we’ve been as efficient as you could be to create the ability to do this kind of making the data available and linking the data. And I think I’m confessing it would cost more over a period of years, but I think we’ve done it in an efficient way.

Mr. FARENTHOLD. And finally, one last question.

Mr. HASKINS. Can I just add one thing? If we ask the Census Bureau to do this, they would have to stop doing something else. So we’re hoping that, at least on a temporary basis, certain employees can be borrowed and that they can make a minimum of hiring in order to build this agency. And if you went out and hired all new people and created all these new positions, then it would cost a lot more than under the system that we’re recommending.

Mr. FARENTHOLD. Well, you ran the clock out on me.

Thank you very much. I yield back.

Mr. RUSSELL. The gentleman yields back.

The chair now recognizes the gentleman from Virginia, Mr. Connolly, for 5 minutes.

Mr. CONNOLLY. Thank you, Mr. Chairman.

Dr. Haskins, you've stressed evidence-based policy making, and I certainly agree. I think probably everyone does in the abstract. But it's almost laughable that public policy is, in fact, always consistently based on evidence. In fact, quite the opposite.

Marijuana. Evidence-based policy?

Mr. HASKINS. I don't know.

Mr. CONNOLLY. You don't know.

Mr. HASKINS. I haven't researched marijuana.

Mr. CONNOLLY. Okay. Wouldn't you say metrics tell us evidence?

Mr. HASKINS. Wouldn't I say what?

Mr. CONNOLLY. Metrics has something to do with evidence?

Mr. HASKINS. Yes.

Mr. CONNOLLY. All right. Any metrics on how many marijuana overdoses there are every year?

Mr. HASKINS. I believe there are, yes.

Mr. CONNOLLY. No, I believe—we've had testimony before this committee there aren't. And this is classified as the most dangerous drug in America.

How many marijuana users die on the roads every year? Do we measure that?

Mr. HASKINS. I don't know.

Mr. CONNOLLY. No. The answer is no, sir.

So we have incarcerated millions of people, arrested people, and decided marijuana is the most dangerous drug in the United States since the era of Richard Nixon. And we have very little data, scientific data to justify that. And the damage done—enormous.

Terri Schiavo. Remember that case?

Mr. HASKINS. Yes.

Mr. CONNOLLY. So a Republican Congress and a Republican President, who actually interrupted his vacation, a rare event, to come back to Washington to actually sign into law an unprecedented intrusion by the United States Congress, imposing its judgment, scientific judgment, on the state of the health of that young lady so that her husband couldn't make a private medical decision. Was that evidence-based, do you know, Dr. Haskins?

Mr. HASKINS. I know something about these type of cases. There is evidence, but the doctors often disagree about the evidence. It's pretty murky.

Mr. CONNOLLY. Pretty murky?

Mr. HASKINS. Yeah.

Now, some cases are clearer than others. And my understanding, the Schiavo case was quite clear.

Mr. CONNOLLY. They were quite clear.

Mr. HASKINS. Yes.

Mr. CONNOLLY. Right. Because when she was autopsied after she died, her brain was one-third the size of a normal human brain. It had atrophied. That didn't stop Congress from overriding scientific data.

Climate change. Evidence-based? The decision, for example, to rip up the Paris Climate Accord, was that evidence-based?

Mr. HASKINS. I would say no.

Mr. CONNOLLY. No.

In fact, would you not agree that the overwhelming evidence is climate change right now is certainly being—a key variable is human activity, and the evidence is pretty overwhelming about that in the scientific community. Would you agree with that?

Mr. HASKINS. I don't consider myself an expert in this, but I've read a fair number of things, and I think that is correct.

Mr. CONNOLLY. And certainly, given your position, you would want Congress to base all—as much as possible—its decisions with respect to climate change, not on belief, but on evidence.

Mr. HASKINS. Okay. Let me say this.

Mr. CONNOLLY. Yeah.

Mr. HASKINS. I do not expect that Congress would make decisions exclusively based on evidence. I want evidence to have a place at the table. I want Congress to understand what the evidence is. But they will use other factors to decide how much money they should spend or whether they do anything at all.

Mr. CONNOLLY. Oh, believe me, I know.

Okay. One of the findings in your report, you say that you're worried about data center consolidation and that it might hamper the ability of the evidence-building community to limit access to confidential data.

How many data centers are there in the Federal Government? Do you know, Dr. Haskins?

Mr. HASKINS. We have 13 agencies now, but it's growing.

Mr. CONNOLLY. Data centers?

Mr. HASKINS. Yes, special data centers are 13, and they're growing.

Mr. CONNOLLY. Mr. Shea.

Mr. SHEA. I think there's confusion. We're talking about statistical agencies who have the mission of collecting data for producing statistical information versus the myriad data centers that are created for a wide variety of—

Mr. CONNOLLY. So you're not referring to the latter.

Mr. SHEA. That is correct.

Mr. CONNOLLY. Okay. Because let me just say, this committee has spent years looking at that subject, Mr. Shea.

Mr. SHEA. We think you should divert the savings from data center consolidation to evidence-based policymaking.

Mr. CONNOLLY. We're with you. Because we've got four agencies that have saved \$2 billion. And Mr. Hurd's not here, but he and I and Ms. Kelly and Mr. Meadows have cosponsored legislation to allow agencies to reinvest in themselves pursuant to the savings from data center consolidation.

Okay. That's really important, because we were going to get in a bit of a tizzy about that.

Let me just say in closing, if I may, boy, do I agree with the premise of your Commission. But what is so troubling, frankly, about the era in which we operate is how easily dismissed facts and evidence—measured facts—are because of a priori beliefs or because of denial. I don't want to accept that.

And it's true in creationism versus evolution. It's true in climate change. It's true when conclusions come up from a scientific agency that somebody doesn't like or doesn't want to accept.

And Congress is as guilty of that practice as any entity. And until we try to move more toward the mean, we take Dr. Haskin's caveat seriously, we won't have the best public policy.

Mr. RUSSELL. The gentleman's time has expired.

Mr. CONNOLLY. I thank the chair.

Mr. RUSSELL. The chair now recognizes the gentleman from Montana, Mr. Gianforte, for 5 minutes.

Mr. GIANFORTE. Thank you, Chairman Russell and Ranking Member Cummings.

And I want to thank the Commission for your work on this. It's very important that we have efficient and effective Federal programs.

I want to direct my first question to Mr. Shea, if I could. You mentioned in your comments about outcome-based objectives. And I was just curious, being a business guy, I know that evidence is critically important, especially as you apply it against the goals of a program.

What evidence did you find that the 209 government offices you surveyed had clear outcome-based goals that they were collecting evidence against?

Mr. SHEA. So that's not one of the questions we asked in our survey. But having spent my career trying to help Federal organizations develop those measures, I can tell you they're not on a steady glide path to all have clear outcomes with aggressive measures of their performance year over year.

Mr. GIANFORTE. Okay. So I'd like to follow up on that. So you're saying that these, in your experience, government agencies and programs don't have clear goals for what they're trying to accomplish?

Mr. SHEA. In many cases they don't. I find them much more willing to measure their inputs or perhaps their output. Outcomes, things over which they don't have complete control, things that the programs Dr. Haskins was talking about don't work, they're very reluctant to hold themselves accountable for those kinds of goals.

Mr. GIANFORTE. Kind of hard to hit a target if you don't have one.

Mr. SHEA. That's exactly right.

Mr. GIANFORTE. Yeah.

So do you have any—in your experience, what suggestions you have along these lines?

Mr. SHEA. Well, you can play a major role in assessing agencies' strategic and annual performance planning process. Ask them to what degree they have clear outcome-oriented goals. And if you don't like what they produce, tell them to refine and improve them.

When I was a staff member on this committee, the Government Accountability Office was an enormous help in setting a framework with which to judge those outcome measures.

Mr. GIANFORTE. Yeah. I can't imagine that there's much satisfaction even for our dedicated government workers if they're working hard but not knowing what they're trying to accomplish.

Mr. SHEA. It's a very important point, because to clarify the mission that employees are trying to work to accomplish is a major factor in improving their engagement and improving recruitment and retention, which, as you know, is a big challenge.

Mr. GIANFORTE. Okay. Thank you.

Chairwoman Abraham, just in the work that you've done collecting evidence, I know in the private sector big data analytics, artificial intelligence have been used for a long time, particularly in the financial industry, to uncover fraud and inefficiencies. My question is, in your working with these 209 agencies, what evidence did you find that these government agencies were using these common private sector approaches to uncover inefficiency?

Ms. ABRAHAM. That is, again, not something that our survey really spoke to directly. I guess I should say, when you think about big data, a lot of—there are a lot of data that the Federal Government collects that we could be doing more with. In terms of these agencies that we surveyed, you know, many of them don't necessarily have the capacity to do that.

Part of what we were attempting to accomplish with the recommendations that we were making was making the data that they collect more accessible to people who do have those skills and who could do some of the kinds of things that you're talking about.

Ms. SWEENEY. May I add to that?

You know, big data and AI are basically statistical algorithms. And when you think about where in the government we see that, it's primarily in the statistical offices, the very data that we're talking about. And so we're also saying that kind of innovation use is exactly the kind of thing we want used in privacy in rendering the data confidential. And this would be among the group who's more likely to be able to use it.

Mr. GIANFORTE. And I know we all benefit every day when credit card transactions are identified that might be fraudulent. You know, looking at the volumes of data we have in the government and being able to say, you know, this one is not like the others, we maybe investigate a little further. I think there's opportunities there to uncover inefficiency and fraud.

Ms. ABRAHAM. I should clarify one thing in response to your question. What we are proposing is a facility that would be used for statistical analysis, to identify patterns and the outcomes associated with programs. We are not envisioning that this facility would be used to go in and identify individuals who had committed fraud and then come back and go after them. We're talking about improving the use of data for statistical purposes, not for targeting individuals and taking action against them.

Mr. GIANFORTE. But it's certainly in our interest to uncover fraud where it exists.

So, Mr. Chairman, I yield back.

Mr. RUSSELL. The gentleman yields back.

And the chair now recognizes the gentlelady from New York, Mrs. Maloney, for 5 minutes.

Mrs. MALONEY. I thank the chairman and the ranking member very much. And I thank my good friend Val for yielding.

I do want to share that we have a meeting coming up at 11:30 with the Foreign Minister of South Korea. Everyone has been invited. I need to get to that too.

But I wanted to ask about, really, the largest collector of information in our Government is the Census Bureau. And I'd like to first ask Dr. Abraham, the Commission's report noted that prin-

cial statistical agencies, such as the Census Bureau, have, and I quote, “demonstrated responsible stewardship of data collection through census and surveys,” end quote. It also noted, and I quote, “Not surprisingly, public trust in the accuracy and validity of statistical data reflect the public’s trust in the statistical agencies that produce them,” end quote.

Particularly for the Census Bureau, it seems that maintaining the trust of the public is important to encourage participation in censuses and surveys. Do you agree?

Ms. ABRAHAM. Absolutely.

Mrs. MALONEY. Well, then, Dr. Haskins, one of the most critical surveys of the Bureau is the annual community survey, which is randomly sent to addresses across the country to collect data on an ongoing basis.

How important is the American Community Survey for evidence-based policymaking, Dr. Haskins?

Mr. HASKINS. Extremely important.

Mrs. MALONEY. Some of my colleagues have suggested making the Community Survey a voluntary rather than a mandatory survey. And do you expect that such a change would impact the amount of data the government is able to collect and the integrity of that data?

Mr. HASKINS. Yes. But the more important point is, if you can’t have a random sample of the public, we won’t know the frequency of all sorts of things that we estimate when we get a random sample. So if you do it on a voluntary basis, it basically ruins the American Community Survey, which is the best and most accurate survey of the population we’ve ever had.

Mrs. MALONEY. Well, thank you. And I know not only government, but the business sector relies on it. And I want to note that Canada made its long form census voluntary in 2010. And like you said, Dr. Haskins, the response rate dropped from 93 percent in 2006 to just 68 percent in 2011, and government agencies were forced to make policy decisions based on old data.

So reducing the volume and quality of data collected by the Census Bureau seems to be exactly the opposite of the Commission’s goal of expanding access to and use of federally collected data. Does everyone on the panel agree?

Mr. SHEA. Yes.

Mrs. MALONEY. Okay. Thank you.

Mr. HASKINS. Can I add one thing very important?

Mrs. MALONEY. Sure.

Mr. HASKINS. The decline in the number of responses is not the biggest problem. The biggest problem is that the sample is no longer random. So you can’t draw conclusions from it. It’s ruined. That’s a big point. I mean, the more people you have, the more reliable the numbers are. But if they’re voluntary participants, none of the numbers are any good, they’re not reliable.

Mrs. MALONEY. Thank you. And I want to note on that that the current President’s budget proposal does not include enough for funding for the Census Bureau to perform the work it does for the upcoming 2020 Census. So the point you made earlier, if you add something, they have to stop doing something. So the appropria-

tions bills currently moving through Congress would implement the President's approach. So we don't have enough funding right now.

So do you believe that underfunding the Census could impair the quality of the data collected in the 2020 Census?

Mr. HASKINS. You're asking me?

Mrs. MALONEY. Yeah.

Mr. HASKINS. Yes, I do.

Mrs. MALONEY. And the President and Congress must increase funding. We need to do that. Starving the Census will result in inaccurate data and undercuts the vulnerable populations who rely on programs that are funded on this data. And if you don't have good data, you don't have good policy.

I want to thank my colleague Val so much. I've got 4 seconds left. I'm going to yield it back to you. I know you've got another 5 minutes coming up. But if you've got a point to make in 4 seconds, I'm sure it's an important one.

Thank you so much. That allows me to get to my other meeting. Thank you so much.

Mr. RUSSELL. The gentlelady's time has expired. But we will get to Mrs. Demings.

The chair now recognizes the gentlelady from North Carolina, Ms. Foxx, for 5 minutes.

Ms. FOXX. Thank you very much, Mr. Chairman.

And I want to thank this distinguished panel very, very much for the work that you did.

I feel almost as though I have always been cursed with having real objectivity flowing through my veins and always being concerned about evidence-based decisionmaking. It may be a product of my having grown up extremely poor. You've got to make really, really good decisions in your life. But I have done this, again, all my life.

I also think that Congressman Gianforte opened up a very important item when he talked about whether there are clear and measurable goals and outcomes out there in most Federal programs. My experience has been that there are not.

And I think it was Dr. Haskins who mentioned—someone did—that we have really, really great research designs out there. And maybe I confused that with measurable outcomes. But at some point, I'd like to talk, whichever member of the Commission talked about that, because I would really like to see where those great research designs are in the Federal Government, because I'm not aware of them.

And then the other issue I'd like to follow up on later, but not now, on the what you all call workforce training programs. I try never to use that "T" word when it comes to dealing with human beings, because I think we educate people. And I don't use that. But because that comes in the other committee on which I serve, and, in fact, I chair the Education and the Workforce Committee, I would like to follow through on that.

But I want to ask a couple of questions related to the work that you've been doing. Again, I think it's extraordinarily important work, and I can't thank you enough.

Did the Commission consider how data quality affects the ability to use the data?

Dr. Haskins, I think so you alluded to this.

Mr. HASKINS. Yes, I think we did.

Ms. FOXX. Great.

In the DATA Act pilot report, which was passed out of the Congress about 3 years ago, OMB recommended continuing the effort to standardize data. Would standardized data collections from Federal award recipients improve the evidence-building community's ability to use the data?

Mr. HASKINS. I think the answer is yes, but with a caveat that different programs have different objectives. So you'd have to have measures of those particular objectives, and they would differ substantially across projects.

Ms. FOXX. Again, I'm really well aware of that. And I think one of the concerns I have—I was a reader for programs in the Department of Education. And to the extent possible, readers got information, got evidence about the success or failure of programs. And then the readers would—readers would evaluate, recommend to the staff. The staff would then ignore the recommendations of the readers based on the evidence that we were presented, which we thought was pretty good.

You all have any ideas on how we can make sure that decisions within the agencies are made on evidence other than utilizing our oversight responsibilities here?

Mr. HASKINS. I have two quick recommendations. One, grill the senior officials when they're nominated about their attitudes about evidence and whether they plan to use it. And, second, when you get reports from Federal agencies that are under your jurisdiction, which is huge, as I understand it, call them before the committee and grill them on these issues, because what you described does happen, I know it happens, and the officials that are responsible for it should be called on the carpet.

Ms. FOXX. Right.

And the last thing I would say to you is, in considering the reauthorization of the Higher Education Act we've had probably 20 hearings. What we heard from people who complained about data they submit to the Federal Government is that we have lots and lots of data and little information.

So I would ask you all to keep that in mind as we push the collection of more data, that we try to make the connection between—and I think you are—the connection between data for data's sake and data which provides us information.

Thank you, Mr. Chairman. I yield back.

Mr. RUSSELL. The gentlelady yields back.

And the chair now recognizes the gentlelady from Florida, Ms. Demings, for 5 minutes.

Mrs. DEMINGS. Thank you so much, Mr. Chairman. Thank you, Ranking Member. And thank you to our witnesses for joining us today.

Dr. Sweeney, in your written statement, you stated, and I quote, "Preventing bad actors from breaking into confidential data requires consistent and rigorous processes," unquote. In other words, we basically need to do more to protect confidential data.

The inspector general of the State Department recently issued a report finding that 77 percent of the Department's IT assets are not in compliance with Federal cybersecurity laws.

Dr. Sweeney, does that surprise you? Why or why not?

Ms. SWEENEY. No, it doesn't surprise me. But in terms of the Commission, even though the language I use sounds like security, about breaking into confidential data, I mean exploiting the data as it's given to you, which is not the same as breaking into a computer and breaching a database. But computer security problems are also rampant.

Mrs. DEMINGS. The State Department until recently was making strides in another area of its cybersecurity efforts. The State Department's Office of Coordinator for Cyber Issues serves as liaison between the State Department and the White House, other agencies, and outside stakeholders. The office also engages with international partners.

Dr. Sweeney, do you think it makes sense to have a senior official within an agency like the State Department to coordinate cybersecurity efforts?

Ms. SWEENEY. So the Commission report isn't on cybersecurity. So I'm speaking for myself and not for the Commission. But, you know, cybersecurity problems are huge in the United States, and they do dovetail constantly with these data problems, because they just leave all of our systems vulnerable. And so we don't have a rigorous—we haven't come to a full circle as to how we address that in a comprehensive way.

I could imagine having senior people in the way that you've provided would actually be incredibly helpful. But we still need a comprehensive perspective both on the data side as well as on the computer security side.

We just published a paper showing how data—the data to impersonate a voter can actually—could be used to actually change voter files by making changes. So that's kind of a combination of data that we were talking about in the Commission effecting a kind of security outcome.

Mrs. DEMINGS. Any of the other witnesses would like to?

Mr. Shea.

Mr. SHEA. Yeah. Are you asking whether the State Department should govern the Federal Government-wide cybersecurity efforts?

Mrs. DEMINGS. If the State Department should eliminate or create the Office of Coordination.

Mr. SHEA. You know, you can take the boy out of OMB but you can't take the OMB out of the boy. So having seen the coordinating role OMB can take, whoever has the lead role at an agency level, OMB has an important role to take in this effort, or at least the White House—something that has the imprimatur of the White House.

The Department of Homeland Security and the Department of Defense are also key players in this initiative, National Institute of Standards and Technology. All that needs to be governed somehow centrally. So that I don't really—I don't think the State Department seems to me is the right place for that, but there does need to be central governance of the cyber apparatus across government.

Mrs. DEMINGS. Dr. Sweeney, in the wake of the Equifax breach, one of the worst breaches of personal privacy in history, it is critical for Federal agencies to reevaluate the state of their security protections. And I'd like for you to begin, but any of the witnesses can talk about, in the limited time that we have, what immediate steps can an agency take to improve the security of their data?

Mr. HASKINS. Hire Latanya Sweeney to be in charge of their data.

Mrs. DEMINGS. That's why she was asked the question.

Ms. SWEENEY. So data are just valuable. I mean, they're worth a lot. There are different kinds of actors who are after this data and the mechanisms that they will go through to get the data make it constantly a kind of cat-and-mouse kind of game.

And so how we engage—you know, are we using the latest—are businesses using the latest? And even then, there are all these simple vulnerabilities, like getting someone through email to reveal their password and that particular account having particular access privileges. And that's a human engineering.

So it's not quite as—I didn't answer your question, but I just bring emphasis on security is multifaceted in its complexity. The data problems that we talk about in the Commission are certainly a piece of security, because the more vulnerable data, whether it's health data on Americans, whether it's voter records on Americans, whatever it might be, income records, it allows other kinds of systems to be infiltrated by actors as well, whether it's our tax return system or—and what have you.

So there is a relationship between security and the data things that we talk about. And this problem that we have where the privacy of the data is vulnerable on both sides, and what we choose to give out is not adhering to a kind of risk assessment and improvement model and what we choose not to give to let researchers help us learn more, is also a problem.

Mrs. DEMINGS. Thank you.

Mr. Chairman, I yield back.

Mr. RUSSELL. The gentlelady yields back.

The chair now recognizes the gentleman from Wisconsin, Mr. Grothman, for 5 minutes.

Mr. GROTHMAN. Dr. Haskins, in your testimony you said that most of the Nation's social programs produce modest or no impacts on the problems they were meant to address. I think given the mass of money we've flooded at these programs over the last 50 years, that would be obvious.

But could you—first of all, how do you know they don't produce significant results? And then, secondly, could you give us some examples?

Mr. HASKINS. There's an interesting study of the first 90 studies that were—Congress established the Institute of Education Sciences, and they launched into—they revolutionized research and education. And the first 90 studies, about 85 percent of them failed to produce a major impact. This is entirely consistent with what's in the literature on medical research, on business research, and on research in the social sciences.

So there's—I don't think anybody seriously questions that most of our programs don't work. If we want them to work, we need to

continue to evaluate them and we need to improve upon what the programs are doing. We can find programs that work for almost any problem. And then we need to learn how to implement them widely in the country.

Mr. GROTHMAN. Okay.

Next general comment for Dr. Abraham. Looking at it, you had 16 people on that committee is right?

Ms. ABRAHAM. Fifteen.

Mr. GROTHMAN. Fifteen, okay.

I counted, including the adjunct professor at Notre Dame, 11 of the 15 right now are professors. Do you think there's a diversity problem on the committee, or will that have any impact on—

Mr. SHEA. I was very loud.

Ms. ABRAHAM. No, I don't think there was. Many of the people who are currently academics are people who in previous lives have had extensive experience doing other things; a lot of experience in government. We were asked to tackle a set of fairly technical questions. So technical expertise seemed called for.

And we heard from a—we made a big effort to hear from a lot of people to get input into our deliberations. We ended up hearing from—I want to say over 500 people, including at hearings where we invited anyone who wanted to come talk to us to do so.

So I feel like we had a pretty broad set of perspectives represented, and we did hear a lot of input from people.

Mr. GROTHMAN. Okay. One of my concerns on your recommendations is it seems that if you talk to the right people you can get the conclusion that you want. I mean, we just heard from Dr. Haskins say we need to change programs to improve them. I think there may be a bias out there to how can we improve these programs rather than how can we get rid of these programs, which is a much more difficult thing to do.

But I guess my concern here is that usually there are—people of, let's say, goodwill can come up with different—evidence for different results, you know. And we talk about global warming, and various smart people can say it's not going on, but I think the overwhelming bias in the community is saying it is. We talk about—another one that comes around is early childhood education, very smart people feeling we're not getting a lot of bang for the buck there. But I think the push for more, and people who want more government are always going to be there pushing their people to the fore, you're going to find people to say early childhood education is a positive.

How do you guard against having what I'm afraid would happen any time you set up a permanent committee or permanent commission, they are taken over by people who want more, and they will find the experts who claim we need more, they will twist the data to say we need more?

Ms. ABRAHAM. One of the things that is emphasized in the report is the importance of rigor of the evaluations that are undertaken. So what we are envisioning and what the structure that we've laid out we believe would promote is a system where on an ongoing basis there are rigorous evaluations of what the outcomes associated with different programs are. And there are scientific standards for determining whether a conclusion from a study is valid or

not valid. That still isn't going to tell you what it makes sense to do.

Mr. GROTHMAN. Other than, say—my concern is that, despite supposed rigor, which we should probably have throughout the whole world of academia, the vast sea of academia comes out left. So even though they're all supposed to be bound by rigor, unless our country, our idea of limited government is wrong, it seems—you see where I'm going with this? Is this a concern?

Ms. ABRAHAM. I think that evidence can tell you that if you do A, the outcome appears to be B. That can't tell you what your policy should be. So what we're really arguing for—this is something I've heard Ron say any number of times—is that what we're advocating for is evidence having a seat at the table so that everyone can look at it and then, given that and other value judgments and other inputs, make decisions about policy. But we are just—we're advocating for evidence to have a seat at the table.

Mr. GROTHMAN. You have to trust the people. But thank you.

Ms. SWEENEY. I know you're over time, but I would just add one thing. Clearly, you know, as a scientist, clearly, I believe in the pursuit of truth. And so having more data is better than not having the data.

So we might argue about whether there were confounders in an outcome. But not having the data at all increases the space of uncertainty. And the issue here is making more data available to get rid of some of the uncertainty on which these issues would be based.

Mr. RUSSELL. The gentleman's time has expired.

The chair now recognizes the gentlelady from the District of Columbia, Ms. Holmes Norton, for 5 minutes.

Ms. NORTON. Mr. Chairman, I want to thank you and the ranking member for this hearing.

And isn't it interesting that, by coincidence, of course, it comes at a time when there's an outcry in the country about fake news and alternative facts. And here we're discussing the Commission's work, for which I thank you, on evidence-based data. The uses of data, however, takes into account more than the evidence they purport.

Ms. Foxx on the other side asked this difficult question: How can we make sure decisions are made on the basis of data? Well, decisions are here made on the basis of many factors, on the values of each side, how one construes data, debatable data.

We're in the middle of the Affordable Health Care Act, and even the underlying data is being contested. And one side is actually telling us that if you have a preexisting condition you'll be covered in the same way you're covered by the Affordable Health Care Act. You know, my side can only call that a lie. But, of course, the Affordable Health Care Act has gone down. So somehow or the other, there were enough people on the act or being served who were their own evidence.

But trying not to be cynical, you can see how focusing on data alone may not get us the results we want. We just finished a big debate, and it comes up every single year, on whether Planned Parenthood should be funded. It usually gets funded because there's an

outcry in the country if you don't fund it because of the work it does on so many conditions affecting women.

But there was a whole special committee set up here in the Congress to investigate a Planned Parenthood episode. And now the so-called data from this committee is being used to try to defund Planned Parenthood based on whatever this special committee found out, completely contradicted by the entire scientific community. So you can see there's a little cynicism on my part when I hear talk of data, at least when used in a body like the Congress of the United States.

Mr. HASKINS, I have a question for you, because there's a current controversy as the administration prepares or asks Congress to cut the Teen Pregnancy Prevention Program. In this Congress, on both sides of the aisle, we would like to see at least fewer abortions and some no abortions at all. So we're together on that. But, of course, if teen pregnancy goes up, there will be more abortions. And yet you will find that the side who is against abortions does not move on teen pregnancy. Maybe we'll see what happens in the appropriation.

But this is a very small investment, a small program. And, Mr. Haskins, you recently wrote an op-ed, and it was entitled "Trump Team Doesn't Understand Evidence-Based Policies Regarding Social Problems." And in that, you actually cited that 40 percent of the projects funded through the Teen Pregnancy Prevention Program produced, you said, at least one significant impact.

Do you believe when a program produces at least one significant impact that eliminating 81 percent of the organizations that participate in the program is perhaps not the wisest idea?

Mr. HASKINS. They actually eliminated 100 percent. It was 81 programs, and they eliminated all of them.

Ms. NORTON. All 81 of them.

Mr. HASKINS. Yeah.

No, I don't think it is. And I think in this case it's really important, because teen pregnancy prevention in a way is the most advanced of the big sets of programs that we're trying to improve and develop ways to improve, starting with evidence-based policy, which, as we all know, doesn't always work well. And that's what this study showed.

But we've talked repeatedly that most social programs don't work; 80 to 90 percent is a good estimate. And in this case, only 60 percent failed, 40 percent worked. That's progress. And I think it's because the agency did a good job implementing a program, the people out there in the countryside who are learning this are going along. They were in the program for 5 years.

Ms. NORTON. So you would think that if you were going to cut anything, you would say, okay, those didn't work, maybe you cut those. But those that did work, showed progress, maybe we ought to continue that progress by funding those programs that did work in cutting teen pregnancy.

Mr. HASKINS. That would make sense to me.

Ms. NORTON. Thank you, Mr. Chairman.

Mr. RUSSELL. The gentlelady yields back.

The chair now recognizes the gentleman from Kentucky, Mr. Comer, for 5 minutes.

Mr. COMER. Thank you, Mr. Chairman.

My first question is for Mr. Shea. The Commission found that bureaucracy caused problems related to unevenness and capacity in agencies' evidence-building efforts. Can you explain some of the problems that the Commission found as it relates to this bureaucracy?

Mr. SHEA. I can give you a couple of examples. One was procurement. A lot of the questions about adequate rigor in evaluation are solved by ensuring that the people you contract for for evaluations bring sufficient independence and rigor and experience to conducting the evaluations. The procurement process is very cumbersome, as this committee; well knows.

Likewise, the talent you need to oversee and conduct an evidence agenda is really hard to recruit and retain. There's an enormous competition for this kind of talent. And the personnel system under which we operate today makes it really difficult to hire, recruit, retain that workforce.

Mr. COMER. Why does the Commission suggest that OMB be the agency that coordinates Federal evidence-building efforts, Mr. Shea?

Mr. SHEA. OMB has that tacit threat of budget impact from everything it does. So it's got a lot of juice that other similar entities don't have. And this committee, the Congress, in general, has vested in OMB a great deal of central management responsibility, and we think they can have a real important role to play in coordinating this government-wide.

One of the things we did see is that OMB's efforts are a little disjointed. There's a performance and personnel office. There's an evidence team. There's the office of information regulatory policy, the Chief Statistician at OMB. Those aren't necessarily well coordinated at OMB. So OMB itself could do a better job coordinating its own investment in this enterprise.

Mr. COMER. Let me ask you this, and this will be my last question to you. The Commission recommends considering reorganizing and consolidating aspects of OMB. Can you explain what the Commission means by reorganizing OMB? Is the Commission suggesting a complete overhaul of the agency, or what?

Mr. SHEA. No. Just what I said. These various entities that have responsibility for certain aspects—the office of information—Office of E-Gov, the office of information and regulatory policy, the performance of personnel team, the evidence team—all of those four could be better coordinated. Consolidation is only one option in the various things you could do to improve that coordination.

Mr. COMER. My next question is for Chair Abraham.

How did the Commission weigh privacy concerns against the need to ensure transparency of government information?

Ms. ABRAHAM. So as I indicated in my opening comments, we are all for making publicly available government data where that can be done without violating the confidentiality of individual people. In cases where it's not possible to do that, we adopted an approach that Latanya has laid out for sort of tiered access to data. If you can't just make the data available publicly, maybe you can make available a stripped-down data file that lets people learn things about what the Government's doing but doesn't violate confiden-

tiality. And if you can't do that, maybe you can make the data available through the National Secure Data Service.

So we are really pushing as hard as we can in our recommendations for making data available in the best way that you can without violating confidentiality.

Mr. COMER. In gathering testimony and during deliberations over the report, did the Commission solicit feedback from the transparency community? If so, who and what did they suggest?

Mr. SHEA. We had an open comment period and a number of witnesses, and we received a number of written input from what I would call the transparency community, ways to leverage the amount of data currently being reported and what standardization could do to improve our access to a variety of sources of data.

Ms. ABRAHAM. We do have a list of all of the people that provided input, and we would be happy to call that out and let you know specifically who we heard from.

Mr. COMER. Okay.

Ms. SWEENEY. I would also say too that the report itself is very much about transparency. It makes more transparent arbitrary decisionmaking around what data is being given. It also makes more data possible to be given, which adds to the notion of transparency in oversight. So in that way it's very much, the report itself, very much is a champion for the transparency position.

Mr. COMER. Okay.

Thank you, Mr. Chairman.

Mr. RUSSELL. The gentleman yields back.

And the chair now recognizes the ranking member for any followup.

Mr. CUMMINGS. I just want to—you know, I was just wondering, I was thinking about this whole opioid situation where so many people are dying from overdoses and the President's Commission has, in their preliminary report, has said that this should be declared a state of emergency.

And I'm wondering how we can use the information that our fellow Americans are dropping dead everywhere. And there are no boundaries with regard to race, no boundaries with regard to location, rural, urban. They're dying big time.

And then we have fentanyl, which is really taking a lot of people out, where they say that in some instances if you touch it you die.

It just seems to me, you know, when evidence is in your face, how do we best—I mean, you know, we as legislators, I think it's important that we act on evidence. I really do believe that. I would say you can't make a decent decision without evidence.

But I'm trying to figure out how does information like that, how is that—how do you all see that as best handled? Or you don't see that within your purview?

Hello? Somebody talk to me.

Mr. SHEA. I never miss a chance, Mr. Cummings.

I think what you're talking about is our central motivation. The Nation's challenges demand effective solutions. And our investments to date, as Ron has said, haven't really made much of an impact.

And so the more and more we can learn about what has the greatest impact on solving our biggest problems, we want more and

more of that evidence. And so what we think our recommendations do is create a structure, governance, a cycle in which we are developing more and more evidence of what works so that more and more people can learn and invest in those more proven activities.

Mr. CUMMINGS. Anybody else?

Dr. Sweeney?

Ms. SWEENEY. So in Massachusetts we have an opioid crisis. And I was really taken by the difficulty of getting the data needed to address what was really an acute problem. And, even though it's not the Federal Government, it was very similar in the sense that what did it take to get these different agencies and different groups to share data widely.

And so in very much the same spirit as the Commission report, we had two tools. One is we could try to change a law, or the other one was we could argue and explain how it is that this version of the data is okay to give out even within the structure of your existing rules and regulations. And we were very successful in beginning to put together combined data sets in Massachusetts under those kinds of models. So to the extent that that helps.

Mr. CUMMINGS. Mr. Haskins?

Mr. HASKINS. I just want to add one thing, quickly.

Mr. CUMMINGS. Dr. Haskins.

Mr. HASKINS. Pardon?

Mr. CUMMINGS. I said Doctor.

Mr. HASKINS. Oh, thank you.

Quickly. There are certain problems that are almost impossible to overcome. Addiction is one of them. I come from a family of alcoholics. Some of them drank from the time they were teenagers. Many of them died as a result of drinking. And this is an even bigger problem than opioids. It doesn't have quite the flash that opioids do. But the fact is we do not have good solutions to get people to stop their addictions. And opioids appears to be even more addictive than alcohol.

So even if we were willing to spend money, if you did all the investigation, we had all the evidence and so forth, it would still be a very difficult problem and we would have low success rates.

I think they could be improved. There are some improvements that are better than the other. I'm going to Kentucky in 2 days to see some opioid treatment facilities. And maybe some of them are going to be successful and we will be more successful in the future. But addiction is always going to be a problem, and it'll be very difficult to solve.

Mr. CUMMINGS. I want to thank all of you for being here today. It's been extremely helpful.

And thank you, Mr. Chairman. I yield back.

Mr. RUSSELL. The gentleman yields back.

And I'd like to do a couple of followup questions regarding the placement of this data service.

Dr. Abraham, are you aware of the Census Bureau's 2020 Census efforts which include the creation of administrative records verification and access systems, that these were placed on the GAO's high-risk list? So if they placed it on a high-risk list, why would we want to then move immediately to something like that?

Ms. ABRAHAM. So I think that what we're talking about is a little different than the Census operations issue that you're raising, which I have to say I'm aware of only in quite general terms.

What we are talking about building on is capacity at the Census Bureau that has been quite successful at figuring out methodologies for bringing administrative data in from various sources and linking it for statistical purposes. There's a group at Census that has a lot of expertise at doing that and has been very good at doing that and we think that that's something that would be a good core to build on in establishing this National Secure Data Service.

Mr. RUSSELL. Well, there's no doubt they collect an awful lot of information and have since we've been wearing tricorne hats. But I guess given the longstanding problems with the Census Bureau when it comes to estimating costs of large-scale projects, including the massive technological overhaul for the 2020 Census, I go back to, why would the Commission feel that leveraging Census employees and resources would be a best option? Because, you know, we've got some costs, we've got some security concerns. And this would probably be true of any agency.

But I'm just curious. I mean, you're all highly intelligent people. You've certainly done your homework. So given these issues and problems, why do we think that we would want to do this?

Ms. ABRAHAM. So the Census Bureau is not—it's a big place. There's a lot of different parts of the Census Bureau. Doing the Decennial Census is an enormous undertaking. I mean, every time the Census comes around, there are issues of one sort and another that arise and have to be confronted.

The Census Bureau is hiring hundreds of thousands of people to do the Census. That's not what we're—it's a big operational and management challenge. That's not really what we're talking about. We're talking about starting out on a smaller scale, building the capacity to do data linkage for evidence-building purposes.

And so I guess in my mind it's a very different enterprise. And the issues that you're raising don't really, as far as I'm aware, exist in that part of what they do.

Mr. RUSSELL. Well, and along that line, I guess, you know, the Commission made a recommendation that using Census Bureau systems as models for the data service, and, yet, these processes for verifying this are still—they're still trying to verify that the systems will work in the 2020 Census.

So, I guess, in your opinion, before selecting a system as a model, would it not be better to wait until the model has been fully tested and vetted?

Ms. SWEENEY. Let me say the following. So the operation of the Census, the sort of production system of producing the Census, is not what got them into the report. What got them into the report is, of all of the parts of the government, they are, by far, the best—

Mr. RUSSELL. The best collector.

Ms. SWEENEY. No. They're the best if the world at providing a public version of that Census. You can go on to a website, and you can get information from that Census. And that information that I get on that website doesn't violate the privacy. I can't even do it, you know?

And so this is an amazing feature, that the Census Bureau has had this capacity and has been a leader in Federal statistics offices on this idea of, how do I render data sufficiently de-identified that I can share it publicly? This is the skill that got them into the report.

Mr. RUSSELL. And that's useful. And, you know, we certainly see release of data and release of Census information. You know, you've got to live a long time before you can reach back into a previous Census to start mining the information.

Mr. HASKINS. Mr. Chairman, can I add something?

Mr. RUSSELL. Dr. Haskins, please.

Mr. HASKINS. Yeah. I can tell you why I supported the proposal. I never thought of having a National Secure Data Service before I joined the Commission, so I think I would be like a Member of Congress that would for the first time confront this choice.

And here's why I decided. The Census Bureau is already doing almost everything that we want the National Secure Data Service to do. They're selecting the best proposals. They are helping people analyze their data. We envision that that will be an important thing in the future. They give secure access to data that they bring into the Census Bureau from other agencies.

So those are three big functions that will be required of this agency and—of this new agency—and the Census Bureau already is as good as anybody in government, I would say probably best in some of those cases. And there are other things as well. They are just so experienced and so competent that it makes sense to start with them.

Mr. RUSSELL. And it's useful to hear each of your lines of thinking on why you chose this particular agency.

I guess my last question, and to close out, would be the report suggests that FITARA may hamper agencies' efforts to limit access to confidential data. We've made a lot of progress on FITARA and those that have implemented its measures have actually been pretty successful. It's just trying to get people to comply.

And so I guess it raises some other questions. What evidence does the Commission have to support the assertion that FITARA may hamper agencies' efforts to limit access to confidential data? Whoever would like to take that.

Mr. SHEA. It's not an issue that I can talk very deeply about, but I think generally centralizing a lot of authority in the CIO may conflict with data stewards or evaluation officers implementing the systems getting access to data they need to drive a learning agenda.

Mr. RUSSELL. And in this mining of and, you know, studying the issue, did you share these concerns with Federal CIOs? And, if so, what was some of the commonality of feedback?

Mr. SHEA. We had brought input, as you know. But I wouldn't call what we had a conversation.

Ms. SWEENEY. But I would say that when I was at the Federal Trade Commission the role of the CIO—so take the Equifax breach. The Equifax breach means that if I want to release the data, and I think that Social Security numbers are private, they're not so private, right? Especially if copies of the Equifax data are available for \$500 on the black market, right, on the dark net.

So in that kind of thinking, that's outside of the scope that our CIOs think, this idea of how I do think about the privacy of the data, the decision I'm making. You know, it's not the same as how do I choose a particular technology, or an infrastructure, or even the security of my laptop, or—

Mr. RUSSELL. No, that's useful, and I appreciate it.

And I'd like to echo the ranking member's comments and to thank each of you, not only for the hard work that you've done in trying to improve our own government, but also to protect all Americans, you know, how can we best protect Americans with the data. At the same time, we have a responsibility to the Constitution to have an open government and these two halves trying to balance.

So I really want to thank each of, and all of the members of the Commission, for the work that you've done.

The hearing record will remain open for 2 weeks for any member to submit a written opening statement or questions for the record.

And if there is no further business, without objection, the committee stands adjourned.

[Whereupon, at 12:07 p.m., the committee was adjourned.]

APPENDIX

MATERIAL SUBMITTED FOR THE HEARING RECORD

Congress of the United States
Washington, DC 20515

September 26, 2017

The Honorable Thomas E. Price, M.D.
Secretary of Health and Human Services
Department of Health and Human Services
200 Independence Avenue, S.W.
Washington, D.C. 20201

Dear Mr. Secretary:

We are writing to request information about troubling reports that the Trump Administration—and specifically officials in the White House—may have suppressed an internal study highlighting the tremendously positive net economic contributions of refugees.

Our nation's public policies should be based on a thorough understanding of empirical facts and data, but if these recent reports are true, it appears that the Trump Administration may have rejected these facts in order to present a biased, incomplete, and ultimately false political narrative. It is critical for Congress and the American people to understand who in the Trump Administration engaged in these actions and why.

On September 18, 2017, the *New York Times* reported:

Trump administration officials, under pressure from the White House to provide a rationale for reducing the number of refugees allowed into the United States next year, rejected a study by the Department of Health and Human Services [HHS] that found that refugees brought in \$63 billion more in government revenues over the past decade than they cost.¹

The *Times* later posted a copy of the report itself, which concluded:

Overall, this report estimated that the net fiscal impact of refugees was positive over the 10-year period, at \$63 billion, meaning they contributed more in revenue than they cost in expenditures.²

According to the *Times*:

¹ *Trump Administration Rejects Study Showing Positive Impact of Refugees*, New York Times (Sept. 18, 2017) (online at www.nytimes.com/2017/09/18/us/politics/refugees-revenue-cost-report-trump.html?_r=0).

² Department of Health and Human Services, *The Fiscal Costs of the U.S. Refugee Admissions Program at the Federal, State, and Local Levels, from 2005-2014* (July 29, 2017) (online at <https://assets.documentcloud.org/documents/4056060/Refugee-Report-Draft.pdf>).

The Honorable Thomas E. Price, M.D.
Page 2

White House officials said those conclusions were illegitimate and politically motivated, and were disproved by the final report issued by the agency, which asserts that the per-capita cost of a refugee is higher than that of an American.³

In fact, White House spokesman Raj Shah stated on the record:

The actual report pursuant to the presidential memorandum shows that refugees with few skills coming from war-torn countries take more government benefits from the Department of Health and Human Services than the average population, and are not a net benefit to the U.S. economy.⁴

Although the final report has not been released publicly, the *Times* reported that the Trump Administration achieved its skewed results by instead examining only expenditures by refugees while excluding all of their financial contributions:

The three-page report the agency ultimately submitted, dated Sept. 5, does just that, using government data to compare the costs of refugees to Americans and making no mention of revenues contributed by refugees.⁵

If this account is accurate, the Trump Administration's suppression of this draft report suggests an Orwellian dismissal of fact-based analysis in service of the worst kind of political pessimism. It also represents a potential violation of Section 207(e)(4) of the Immigration and Nationality Act, which requires the Administration to provide Congress with "an analysis of the anticipated social, *economic*, and demographic impact of [refugees]' admission to the United States."⁶

The American people deserve to know who directed and participated in these abuses and whether they were committed to provide illegitimate support to the Trump Administration's efforts to restrict the number of refugees entering the country.⁷

For all of these reasons, we request that you produce, by October 1, 2017, copies of the following documents:

- (a) all drafts of the report, including the final version completed in September 2017; and

³ *Trump Administration Rejects Study Showing Positive Impact of Refugees*, New York Times (Sept. 18, 2017) (online at www.nytimes.com/2017/09/18/us/politics/refugees-revenue-cost-report-trump.html?_r=0).

⁴ *Id.*

⁵ *Id.*

⁶ 8 U.S.C. §1157(e)(4) (emphasis added).

⁷ *Sessions: Obama's Syria Refugee Plan to Cost \$55 Billion, Demands it be Killed*, Washington Examiner (Nov. 16, 2015) (online at www.washingtonexaminer.com/sessions-obamas-syrian-refugee-plan-to-cost-55-billion-demands-it-be-killed/article/2576450); *Mike Pence's Refugee Problem*, Rolling Stone (Jan. 25, 2017) (online at www.rollingstone.com/politics/features/mike-pences-refugee-problem-w462763).

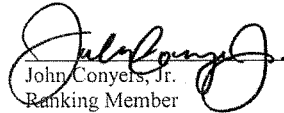
The Honorable Thomas E. Price, M.D.
Page 3


- (b) all communications referring or relating to any draft of this report, its contents, or its findings, including but not limited to communications between or among any federal officials, including officials at HHS, the White House, or any other federal agency, as well as communications between or among non-governmental entities or individuals.


We also request a briefing from appropriate HHS officials by October 1, 2017, regarding: the findings in early versions of the report that were omitted in later versions; the criteria guiding those omissions; the offices and officials involved in producing the report, including any political staff that participated in any way in this process; the methodology underlying the report; and the specific rationale for excluding from the final report the significant economic benefits of refugees.

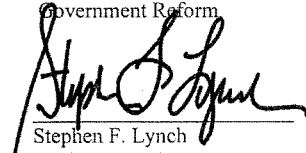
Thank you for your prompt attention to this matter.

Sincerely,


John Conyers, Jr.
Ranking Member
Committee on the Judiciary


Elijah E. Cummings
Ranking Member
Committee on Oversight and
Government Reform


Zoe Lofgren
Ranking Member
Subcommittee on Immigration and
Border Security
Committee on the Judiciary


Stephen F. Lynch
Ranking Member
Subcommittee on National Security
Committee on Oversight and
Government Reform

cc. The Honorable Bob Goodlatte, Chairman
Committee on the Judiciary

The Honorable Trey Gowdy, Chairman
Committee on Oversight and Government Reform

Congress of the United States
Washington, DC 20515

September 26, 2017

General John F. Kelly
Chief of Staff
The White House
1600 Pennsylvania Ave. NW
Washington, D.C. 20201

Dear General Kelly:

We are writing to request information about troubling reports that the Trump Administration—and specifically officials in the White House—may have suppressed an internal study highlighting the tremendously positive net economic contributions of refugees.

Our nation's public policies should be based on a thorough understanding of empirical facts and data, but if these recent reports are true, it appears that the Trump Administration may have rejected these facts in order to present a biased, incomplete, and ultimately false political narrative. It is critical for Congress and the American people to understand who in the Trump Administration engaged in these actions and why.

On September 18, 2017, the *New York Times* reported:

Trump administration officials, under pressure from the White House to provide a rationale for reducing the number of refugees allowed into the United States next year, rejected a study by the Department of Health and Human Services [HHS] that found that refugees brought in \$63 billion more in government revenues over the past decade than they cost.¹

The *Times* later posted a copy of the report itself, which concluded:

Overall, this report estimated that the net fiscal impact of refugees was positive over the 10-year period, at \$63 billion, meaning they contributed more in revenue than they cost in expenditures.²

According to the *Times*:

¹ *Trump Administration Rejects Study Showing Positive Impact of Refugees*, New York Times (Sept. 18, 2017) (online at www.nytimes.com/2017/09/18/us/politics/refugees-revenue-cost-report-trump.html?_r=0).

² Department of Health and Human Services, *The Fiscal Costs of the U.S. Refugee Admissions Program at the Federal, State, and Local Levels, from 2005-2014* (July 29, 2017) (online at <https://assets.documentcloud.org/documents/4056060/Refugee-Report-Draft.pdf>).

General John F. Kelly
Page 2

White House officials said those conclusions were illegitimate and politically motivated, and were disproved by the final report issued by the agency, which asserts that the per-capita cost of a refugee is higher than that of an American.³

In fact, White House spokesman Raj Shah stated on the record:

The actual report pursuant to the presidential memorandum shows that refugees with few skills coming from war-torn countries take more government benefits from the Department of Health and Human Services than the average population, and are not a net benefit to the U.S. economy.⁴

Although the final report has not been released publicly, the *Times* reported that the Trump Administration achieved its skewed results by instead examining only expenditures by refugees while excluding all of their financial contributions:

The three-page report the agency ultimately submitted, dated Sept. 5, does just that, using government data to compare the costs of refugees to Americans and making no mention of revenues contributed by refugees.⁵

If this account is accurate, the Trump Administration's suppression of this draft report suggests an Orwellian dismissal of fact-based analysis in service of the worst kind of political pessimism. It also represents a potential violation of Section 207(e)(4) of the Immigration and Nationality Act, which requires the Administration to provide Congress with "an analysis of the anticipated social, *economic*, and demographic impact of [refugees]' admission to the United States."⁶

The American people deserve to know who directed and participated in these abuses and whether they were committed to provide illegitimate support to the Trump Administration's efforts to restrict the number of refugees entering the country.⁷

For all of these reasons, we request that you produce, by October 1, 2017, copies of the following documents:

- (a) all drafts of the report, including the final version completed in September 2017; and

³ *Trump Administration Rejects Study Showing Positive Impact of Refugees*, New York Times (Sept. 18, 2017) (online at www.nytimes.com/2017/09/18/us/politics/refugees-revenue-cost-report-trump.html?_r=0).

⁴ *Id.*

⁵ *Id.*

⁶ 8 U.S.C. §1157(e)(4) (emphasis added).

⁷ *Sessions: Obama's Syria Refugee Plan to Cost \$55 Billion, Demands it be Killed*, Washington Examiner (Nov. 16, 2015) (online at www.washingtonexaminer.com/sessions-obamas-syrian-refugee-plan-to-cost-55-billion-demands-it-be-killed/article/2576450); *Mike Pence's Refugee Problem*, Rolling Stone (Jan. 25, 2017) (online at www.rollingstone.com/politics/features/mike-pences-refugee-problem-w462763).

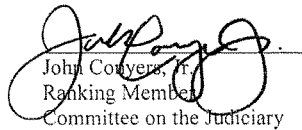
General John F. Kelly
Page 3


- (b) all communications referring or relating to any draft of this report, its contents, or its findings, including but not limited to communications between or among any federal officials, including officials at HHS, the White House, or any other federal agency, as well as communications between or among non-governmental entities or individuals.


We also request a briefing from appropriate White House officials by October 1, 2017, regarding: the findings in early versions of the report that were omitted in later versions; the criteria guiding those omissions; the offices and officials involved in producing the report, including any political staff that participated in any way in this process; the methodology underlying the report; and the specific rationale for excluding from the final report the significant economic benefits of refugees.

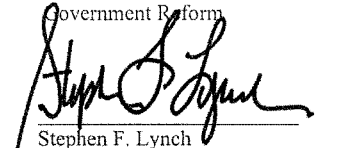
Thank you for your prompt attention to this matter.

Sincerely,


John Conyers, Jr.
Ranking Member
Committee on the Judiciary


Elijah A. Cummings
Ranking Member
Committee on Oversight and
Government Reform


Zoe Lofgren
Ranking Member
Subcommittee on Immigration and
Border Security
Committee on the Judiciary


Stephen F. Lynch
Ranking Member
Subcommittee on National Security
Committee on Oversight and
Government Reform

cc. The Honorable Bob Goodlatte, Chairman
Committee on the Judiciary

The Honorable Trey Gowdy, Chairman
Committee on Oversight and Government Reform

The September 2017 CEP report titled *The Promise of Evidence-Based Policymaking: Report of the Commission on Evidence-Based Policymaking* can be accessed at:

<https://www.cep.gov/content/dam/cep/report/cep-final-report.pdf>

Responses to Questions for the Record

Questions from Chair Gowdy for Dr. Abraham

1. *Definition of Evidence: The Commission defines evidence as information produced by statistical activities with a statistical purpose that is potentially useful when evaluating government programs and policies. What does that mean in plain language? Why did the Commission use this definition? Did the Commission consider other definitions?*

The Commission recognizes that the term “evidence” has broad application, but notes in the introduction of the final report, *The Promise of Evidence-Based Policymaking*, that, for the purposes of the report, the Commission uses the term evidence “to refer to information produced by ‘statistical activities’ with a ‘statistical purpose’ that is potentially useful when evaluating government programs and policies.” The reason why the Commission adopted this definition of the term evidence relates directly to the charge that was outlined in the statute that established the Commission. Specifically, the Commission was directed to review the barriers to using the administrative data that the Government already collects to “facilitate program evaluation, continuous improvement, policy-relevant research, and cost-benefit analyses.” These types of analytic and evaluative activities consider the impact of policies and programs on groups of people, rather than individuals. The essence of a “statistical activity” with a “statistical purpose” is that the results summarize information about a group rather than a single individual or organization.

Administrative data is initially collected to be used by program administrators to operate a program, such as assessing the eligibility of a particular household for a program or reviewing program participation data for evidence of fraud or abuse. This use of administrative data requires knowledge of the individual households that exist in the database. These data are also incredibly useful for research and evaluation, but must only be used for such purposes while protecting the privacy of the individuals and maintaining the confidentiality of the data—using the data in this way is a statistical activity. A researcher might analyze the same administrative data that are collected for the purposes of program administration in order to understand the characteristics or demographics of the program participants as a whole. The Commission was directed to explore how administrative data can be better used for evidence-building purposes, which are inherently statistical in nature.

2. *The Commission’s report says “Placing the directive in law also provides [Principal Statistical Agencies] a stronger basis from which to defend their need for independent information technology resources and tailored procedures to secure the confidentiality of Federal data use for evidence building.”*
 - a. *What does it mean that statistical agencies need to defend their need for independent information technology resources? Who do they need to defend against?*

We consider it essential that principal statistical agencies (PSAs) are able to effectively protect the confidential data under their control, ensuring that it is only used for statistical purposes, honoring the pledge of confidentiality and exclusively statistical use made to American people and businesses when their information is collected. This need only becomes more important in the Commission’s vision when more sensitive data become available to the statistical system to support evidence

building. The very legal framework that requires statistical agencies to protect data collected or acquired under a pledge of confidentiality and limit its use to statistical purposes provides a privacy-protective way to protect newly available data for evidence building. The data protection requirement extends to all of representatives operating on a PSA's behalf, including those working in data centers that store agency data. Under the law, each person with any confidential statistical data access must be made an agent, subject to the law's penalty. Statistical agencies therefore require that non-agency data center employees with access to servers containing confidential data complete training about data stewardship, formally affirm their commitment to keeping statistical data secure and confidential, and pledge not to divulge or attempt to re-identify data. In fact, many statistical statutes provide for notable penalties, including jail time, for breaches of confidentiality. However, at times both historically and currently, statistical agencies have confronted challenges in maintaining these requirements. For example, statistical agencies without authority over their own information technology systems cannot control access to servers that hold protected information and, therefore, may not know who is accessing those data or what they are being used for.

Sometimes the concern about independence for data security is because the statistical agency has lost direct control of information technology systems during required IT and data center consolidation. For example, the need for PSAs to protect confidential data can conflict with other statutory requirements. One example involves FITARA, whose objectives for reform and consolidation are good, but have also created challenges when CIOs have been reluctant to comply with the PSAs' requests to train and qualify IT staff as agents under the Confidential Information Protection and Statistical Efficiency Act or related statistical statutes.

Each of these concerns poses challenges to data confidentiality for evidence building and requires resolution. We do not believe there need be a conflict between the objectives of FITARA and statistical confidentiality, but that there does need to be an explicit understanding in implementing both objectives that one cannot subordinate the other. Improving our acquisition system is essential to good government operations, and equally essential is honoring and protecting the confidentiality of data collected from American people and businesses. By codifying Statistical Policy Directive 1, the Congress is placing in law its clear intent that the government's statistical system has a specific and clear role and responsibilities, and other elements of the government have the obligation to support those objectives.

From page 59 of the Commission report: As Katherine (Kitty) Smith of the Council of Professional Associations on Federal Statistics (COPAFS) noted in comments to OMB about the roles and responsibilities of Federal statistical agencies, "information technology systems that are out of the direct control of the statistical agency can: result in delays in the retrieval of and dissemination of statistical data; impose restrictions on the accompaniment of transparent explanations of methodology with the data; violate the integrity of statistical information; and, very importantly, endanger the statistical agencies' ability to follow through on their pledges of confidentiality and non-disclosure."

- b. *Many of the other recommendations seem to be calling for uniform procedures. What does it mean to say statistical agencies need tailored procedures to secure the confidentiality of federal data for evidence building?*

The recommendations in the report call for uniform procedures beyond the statistical agencies in order to enable the Commission's vision of the Federal evidence building community generating and using data and evidence as a routine part of ongoing government operations. Most of the Commission's recommendations call for consistency and uniformity government wide for evidence building. The primary reason for such consistency is to allow for the efficiency of a well-functioning ecosystem where the authority and roles of actors are clear while also ensuring a safe and privacy-appropriate environment for such data access to occur.

For example, recommendation 2-8 calls for a consistent application process for external researchers that would apply all agencies, including PSAs and recommendation 5-4 calls for the alignment of administrative functions with evidence building purposes across government. Each of these is designed to increase efficiency while maintaining or even improving privacy. While many of the recommendations will help strengthen the PSAs' ability to conduct and support evidence generation, only recommendation 3-4 specifically and uniquely applies to PSAs. The Commission considers it important that the independence and objectivity of Federal statistics, the very foundation of useful evidence building, be codified in law and not simply stated in policy. The Commission included this recommendation in the privacy chapter of its final report because it believes that giving these policies the force of law would help ensure that confidential data are kept safe and secure.

In the context of privacy risk, where a version of data required by law to be protected from unauthorized disclosure is de-identified in order to release it publicly, the criticality of a consistent procedure is particularly important. Any data release can impact a future release of related data. De-identification does not occur in a vacuum, and the actions of one agency can easily but unintentionally provide the means to identify data released by another. The solution to this "mosaic" effect is for all agencies to use common procedures, techniques, definitions, and methods to implement de-identification.

3. *NSDS at Commerce: Why did the Commission recommend the NSDS be housed in Commerce and be modeled on the Census Bureau's systems?*

The Commission considered the institutional placement and governance of the National Secure Data Service (NSDS) in terms of any implications for key objectives: (1) transparency and trust, (2) support for evidence building, (3) strategic coordination and cooperation, (4) confidentiality and security, (5) authority and flexibility, (6) scalable functionality, and (7) sustainability. An institutional placement within an existing Federal department has a practical benefit, particularly the ability to leverage shared services for administrative functions, as well as existing professional staff, established levels of public trust, and operational knowledge. The Commission gave this issue much thought and ultimately determined that the NSDS should be located in the U.S. Department of Commerce because

this institutional placement would allow it to build upon the Department’s extensive portfolio of statistical and data-related bureaus and expertise existing within the Census Bureau, the Bureau of Economic Analysis, the National Institute of Standards and Technology, and the National Oceanic and Atmospheric Administration, among others. The Commission also concluded that the Census Bureau’s current operations, and other capacity throughout the Federal government, form a starting point for implementing the NSDS.

However, to be clear, the Commission emphasized the importance of situating the NSDS in such a way as to provide independence sufficient to set strategic priorities distinct from any existing Commerce agency and to operate apart from policy and related offices. The Commission did not recommend that NSDS be part of the Census Bureau, but rather a sister bureau within the broader Commerce Department. The Commission believes there is key experience and expertise at Census, as well as other agencies, which could either inform, or directly form, the core elements of NSDS. Importantly, the NSDS also must be organized in such a way as to prioritize support for evidence building across government, rather than providing support specific to any one department. On page 83 of the report, the Commission provides a hypothetical sequencing of how NSDS could be setup, although one could imagine a number of other scenarios. At root, the Commission believes that substantial capacity exists within government to do the work of the NSDS, but currently that capacity exists in small groups and offices scattered around various agencies – with the largest element at Census. The Commission believes that by organizing these elements centrally with a mission to support government-wide needs, rather than their current diffused state supporting individual agency needs, will optimize existing capacity.

The Commission does not believe that such an approach would compromise the Census Bureau’s mission related to the 2020 Census, and in fact, the need for Census to be focused on that mission was one reason the Commission did not conclude that expanding the Bureau’s mission to include NSDS was wise. Additionally, as the Commission explicitly concluded that NSDS should not be a “data warehouse,” it then necessarily cannot be housed within any of our statistical agencies, each of which is effectively a data warehouse for their topical areas of focus.

4. *Would any of the Commission’s recommendations need additional appropriations? If so, how much are we talking about?*

The Commission recognizes that resource prioritization is essential to ensuring the Commission’s goals are achieved. The Commission is intentionally not recommending an infusion of large sums of funding to create new agencies or to launch massive new evidence-building endeavors. Instead, throughout the report, the Commission’s recommendations seek to balance the need to prioritize evidence building while recognizing fiscal constraints. In some departments, sufficient resources already exist to enable evidence building, but those resources might be better coordinated or deployed in support of evidence building.

5. *Could any of the Commission’s recommendations help to offset the costs of an increased funding for evidence building or even reduce government spending overall? If so, how?*

The Commission believes that a responsible investment of resources in more and better evidence holds the potential to yield substantial savings in the longer term as programs that

are improved become more cost-effective, and as programs that are not effective are discontinued.

6. *The Commission's report talks about both transparency and restricting access to information. What does the Commission mean by the term "transparency"?*
- a. *Can you help us understand the difference between transparency of records and data, as required by FOIA and other transparency laws, and transparency as envisioned by the Commission?*

The Commission was asked to focus on ways to increase access to confidential data for evidence building purposes. Confidential data refers to data that has been collected or acquired under a pledge of confidentiality (such as the Privacy Act or the Confidential Information Protection and Statistical Efficiency Act (CIPSEA)), and typically contains personally identifying information that would enable the identification of an individual, household, or organization in the resulting dataset. The protections offered under these statutes often make such confidential data inaccessible, including under FOIA, under which personally identifiable information would most often be exempt. The Commission's recommendations are not intended and should not, with one possible caveat, in any way diminish existing efforts to improve transparency under FOIA or related statutes or directive. The one possible caveat is, essentially, data that are not properly de-identified – and thus would not meet the criteria for disclosure under FOIA or other authorities – but is not properly recognized as such. It is true that properly identifying insufficiently de-identified data may result in that data being withheld or further transformed before release, but this should not be seen as anti-transparent. Rather, this is utilizing proper and modern tools to implement transparency laws. FOIA does not intend to release data that breaches the privacy of individuals or businesses.

- b. *How does restricting access to information improve access to information for evidence building?*

The recommendations of the Commission seek to safely increase access to these data for statistical purposes. In particular, the Commission proposes the adoption of a tiered access approach to releasing various data sources. The approach recognizes that there are statistical techniques that enable a highly sensitive dataset to be *transformed* into a dataset with lower sensitivity, and thus making it accessible for statistical purposes. This approach would, in fact, increase the volume of data that is made available for research, as opposed to restricting the amount of data made available for research. In essence, increasing transparency into government programs and operations, while maintaining the privacy of data on individuals and businesses.

The Commission established five guiding principles for evidence-based policymaking. Transparency was identified as one of the guiding principles and described as follows: "Those engaged in generating and using data and evidence should operate transparently, providing meaningful channels for public input and comment and ensuring that evidence produced is made publicly available." The Commission developed recommendations that seek to increase transparency to the public regarding *what* data the Federal government has collected, the *uses* of the data that has been collected, and the *mechanisms* applied by the Federal government that ensure that the public's information will be properly protected.

The Commission believes that the adoption of the recommendations as written would increase *both* the volume of data that is made available for evidence-building, as well as the transparency to the public regarding how this data is collected/acquired, used, and protected.

Questions from Rep. Foxx for Dr. Abraham

1. *The panel referenced that some Federal agencies retained effective research designs in the opinion of the Commission. Please provide some examples of where these effective research designs can be found.*

The Commission believes that evidence building and use must become a regular feature of program oversight and operations to promote continual learning, program refinement, and accountability. Throughout the Commission's fact-finding phase, the Commission identified numerous examples where Federally-funded programs recognize the importance of developing rigorous evidence to inform program administration, though the Commission believes these should increasingly become the norm rather than characterized as exemplars. Several Federal Departments include research and evaluation units which support the development of a broad portfolio of evidence in support of policies and programs, including the Office of Planning, Research, and Evaluation within the Administration for Children and Families of the U.S. Department of Health and Human Services; the Office of Policy Development & Research within the U.S. Department of Housing and Urban Development; and the Institute of Education Sciences within the U.S. Department of Education.

2. *The panel referenced that certain privacy protection provisions in certain workforce training programs were impeding the opening of valuable data. Please provide the specific workforce training programs and any related statutory privacy provisions that should be reexamined.*

During the Commission's fact finding, the Commission learned of a provision embedded in the Workforce Innovation and Opportunities Act (WIOA) that prohibits developing a national database of program participants, inclusive of PII. This ban is modeled on the Student Unit Record ban statutory provision.

"Prohibition on Development of National Database.-- (1) In general.--Nothing in this Act (including the amendments made by this Act) shall be construed to permit the development of a national database of personally identifiable information on individuals receiving services under title I or under the amendments made by title IV." (29 U.S.C. 3341)

As noted in the Commission's discussion of statutory bans (page 29-30), including the WIOA and SUR ban, "within the broad array of factors that influence policymaking, a ban may be an appropriate tool for setting priorities or achieving certain other goals. In the

context of evidence building, however, a ban on data collection or use cannot easily be reconciled with the goal of increased reliance on evidence to inform policymaking.”

Understanding the need to carefully calibrate the value of data for evidence building, and the potential privacy harms, the Commission endeavored to provide Congress a road map to improve the overall privacy and legal framework of data protection and use limitation, allowing for Congress to then consider revising or rescinding such bans in the context of the improved privacy framework.

Responses to Questions for the Record

Questions from Rep. Blum for Dr. Sweeney

Iowa State, University of Iowa, and University of Northern Iowa have advocated for lifting the ban on student-level data in order to provide critical information on higher education outcomes to students and families, policymakers, and colleges and universities.

They are specifically interested in the recommendations which call for Congress and the President to "consider repealing current bans and limiting future bans on the collection and use of data for evidence building."

1. *How would implementing the Commission's recommendations help to provide data for much-needed analysis of higher education outcomes and the labor market?*
2. *How could increasing access to higher education data for the purpose of building evidence potentially improve our school system and other government-funded programs?*

A "student unit record ban" was included as an amendment to the 2008 Higher Education Act reauthorization and prohibits the Federal government from establishing new databases on any individual-level student records, including primary, secondary, and postsecondary education.

During the October 21 CEP public hearing, numerous witnesses asked that the Commission consider overturning the ban. Witnesses in favor of overturning the ban highlighted the inability of current data systems to follow students across multiple institutions and states and to adequately cover nontraditional students, low-income students, minorities, and veterans. Witnesses cited that a national database (or ability to link databases) could help resolve these issues and reduce burden on higher education institutions for reporting to other data collection efforts. These witnesses argued for linking data already held by educational institutions in a privacy-protective manner. Many of the responses to CEP's Request for Comments also addressed the need to measure the outcomes of higher education programs in which the federal government/taxpayers invest billions of dollars. They noted that the questions that could be addressed by such data include: how part-time and older students fare; what happens to the large number of students who transfer from one college to another; how many and which students complete their programs at specific colleges, especially students receiving federal Pell grants totaling in the billions of taxpayer money each year; and whether graduates are able to obtain jobs that pay enough to meet their student loan obligations.

Based on this input, the report concluded on page 30 that: "Prohibitions on data collection and use arguably conflict with the Commission's vision to improve government based on credible evidence. Within the broad array of factors that influence policymaking, a ban may be an appropriate tool for setting priorities or achieving certain other goals. In the context of evidence building, however, a ban on data collection or use cannot easily be reconciled with a goal of increased reliance on evidence to inform policymaking."

Understanding the need to carefully calibrate the value of data for evidence building, and the potential privacy harms, the Commission endeavored to provide Congress a road map to improve the overall privacy and legal framework of data protection and use limitation, allowing for Congress to then consider revising or rescinding such bans in the context of the improved privacy framework.