THE BROOKINGS INSTITUTION

EVIDENCE-BASED POLICY:
HOW IS IT FARING IN THE TRUMP ERA?

Washington, D.C.

Thursday, July 19, 2018

Introduction and Overview:

RON HASKINS
Cabot Family Chair in Economic Studies
The Brookings Institution

Featured Speakers:

THE HONORABLE TODD YOUNG (R-IN)
U.S. Senate

THE HONORABLE THOMAS CARPER (D-DE) (via recorded video)
U.S. Senate

Panel:

MEGAN McARDLE, Moderator
Columnist
The Washington Post

ADAM GAMORAN
President
William T. Grant Foundation

GEORGE OVERHOLSER
President and Founder
Third Sector Capital Partners

GROVER "RUSS" WHITEHURST
Senior Fellow
The Brookings Institution

* * * * *

P R O C E E D I N G S

MR. HASKINS: Welcome to Brookings. My name is Ron Haskins. I'm the co-director on the Center on Children and Families at Brookings along with Richard Reeves. I'm pleased to host this event to mark publication of the volume on evidenced based policy in sponsorship by the American Academy of Political and Social Science.

I'm here with Senator Young, a Republican from Indiana. Senator Young, thank you so much for coming and even more for writing a chapter, a very interesting chapter for the volume. When you find when you read the chapter that there is a rumor in Washington that senators don't know how to write but you proved that wrong. I think you'll find it a fascinating chapter with a lot of stuff about Edison in there. I think it is pretty clear Edison believed in evidence.

Let me just say one word about Senator Young's background and then we'll get rolling. So you attended the Naval Academy, became a U.S. Marine officer in 2005. You were elected to the House in 2010 and the Senate in 2016. That's 11 years from graduating from college in effect to becoming a U.S. Senator. That doesn't happen very often. Married with four children so he had some time to do other things while you were getting and education and running for office. And then your relatively brief time in Washington this is directly pertinent to our event. You've established a reputation as a friend of evidenced based policy. One of the best in Washington, and you played a key role in legislation to pass promotions of evidenced based policy especially, Pay for Success, which maybe you'll talk about a little bit. So welcome. I'm sure our audience will like to hear you and what you have to say about evidenced based policy so the floor is yours.

SENATOR YOUNG: Thank you so much, Ron. I'm grateful to be here and appreciated the opportunity to include an article in the Annals which we're here to celebrate and to promote evidenced based policy making. As opposed to what some

would refer to as policy based evidence making which happens a lot of Capitol Hill.

My interest in this area really was a function of finding myself on the Ways and Means Committee after one term in the House of Representatives. There is a subcommittee of which many of you are familiar on Ways and Means that focuses on many of the social challenges and social programs that are important to the American people. They serve the poor, the vulnerable, the at-risk.

My first sort of threshold question before I tried to develop a familiarity with these programs was which ones work and how do we know that they work. I fumbled around a bit and I was dispirited to learn that the evidence base on so many of our social and our public health programs as well was not robust, was not rigorous. It took me a while to really get a level of confidence that was, in fact, the case. Surely our federal government with roughly 80 different programs had been rigorously evaluating whether or not these programs were effectively serving the least among us. But I discovered that really wasn't the case.

There are a number of reasons for that. One has to do with the fact that many are more psychologically inclined to measure intentions by inputs as opposed to outcomes. I understand that, that's a superficial impulse. But nonetheless, I understand that predilection.   The other is maybe more of a political function of the fact that when we have success, however you might define success when, in fact, it is actually defined for a particular program. Press releases are issued, members of Congress show up, maybe cut a ribbon and celebrate the effectiveness. But when there are suboptimal outcomes when a program has proven to be fruitless or to bear less fruit than might be possible, little is said. That tends to be something that we politicians don't like to associate ourselves with and program managers.

Moreover, they are a vested interest who have a natural inclination to keep particular programs in place without making any sorts of changes. That can result in

a certain stasis or lack of creativity or rigor. Not just in the evaluation of programs but on

implementing any needed changes. But times are a changing in part because the

leadership of people like Ron and the Brookings institution and so many who are in the

audience today and other scholars around town and around the world. You continue to

fight the fight to make sure that we optimally serve those who need to be served.   And

there is a lot of innovation that is occurring. Some of it on account of the way we test

different programs.

In this era of big data, it is now easier to conduct a randomized control

trial across multiple sites then it was a decade or two decades ago. Because we have

large data sets and more people who are sophisticated at manipulating those data sets

and the cost of conducting these trials has gone down so that works to our advantage.

The world has become more globalized and I think increasingly the academic and policy

community is interacting with one another across borders. So we're learning from the

examples, not just of other localities and other states but even other countries. So the

marketplace of ideas, I believe, is forcing change.

I would add that the current fiscal situation, we're well over $20 trillion in

debt. There are a lot of reasons for that but that leads to an imperative during this era of

scarce resources, to make sure that every taxpayer dollar that is spent is spent wisely.

Out of a sobering recognition that there are certain limitations to what we can spend, let's

make sure that we're serving more people, more effectively using the existing dollars

before we start expanding the ambit of dollars that are spent overall. I think that is at

least where we should start with respect to our social programs.

This ought to be natural for us as a country. We are a country of

inventors, of innovators, of doers and dreamers, of pioneers, of titans like Thomas Edison

who tried and failed many times in his effort to commercialize a lightbulb trying hundreds

of filaments. He didn't regard each of those efforts as a failure. He learned from each of

those efforts and iteratively added that bit of knowledge from each of these suboptimal outcomes that informed his future efforts and ultimately he came up with his successful model.

Moreover, we have the most vibrant civil society in all of human history, I would argue. This is something Tote Phil wrote about in the 1830's. We are a country of associations of little platoons, if you prefer Burke. And we gather together, we solve problems collectively at the local level. There are one of the things I've come to really appreciate during my time as first a member of the U.S. House of Representatives and now a U.S. Senator, is how much good just rank and file, neighbors, community groups, not for profit, sometimes for profit groups are doing to address our thorny social and public health pathologies.

Many of these groups do not have, even the modest resources required to effectively evaluate their programs and why they work. Some of their success may not be replicated across geographies in different populations but much of it, I think, can be. So one of the challenges is to scale up things that are working, not just in government but scale up things that are working at the local level, at the state level and also to scale up some of our federal government programs.

I've come up with one model, in fairness, adapted a British model to the United States with a piece of legislation that we have styled. The Social Impact Partnership To Pay For Results Act. For those who are familiar with the social impact bond model, you know it's not really a bond it's a network of contracts between service providers and different levels of government in private individuals, private investors. The idea is to bring private capital into solving our most pressing and social public health challenges. Scale up those things that we know work.

One example would be the nurse family partnership. Many of you are familiar with the nurse family partnership. You take a social worker, you take a public health worker,

you dispatch them into low income areas. They visit with expectant mothers to make sure

those mothers are prepared to lead a healthy lifestyle and make sure that that child is

born without complication. And then a social worker will remain, periodically visiting the

mother and child during that child's infancy. Look, the educational outcomes are

markedly improved through an implementation of the nurse family partnership with fidelity

across populations, across geographies, we know this. They evidence is so strong. The

public health improvements are enormous. We save Medicaid dollars, we save special

education dollars.

So the idea here is to avoid the human cost of suboptimal interventions

to avoid having to put a child into a special education program. To avoid the health

complications that result from not properly caring for ones infant. And, oh by the way, to

avoid those expenditures as well. If we can come up with some level of specificity under

this social impact partnership model of how much money might be saved, serving a given

population through effective implementation of the nurse family partnership. We can

invite in private capital, whether it is through institutional investors or foundations or high

net worth individuals. To scale up something like the nurse family partnership and only if

results are achieved, only if we improve lives are those investors paid back out of future

government savings. Taxpayers win. The recipients win, the recipients of these

services, we all in.

This model can be applied to range of social and public health

challenges. There are many other models out there. But really the emphasis is on

rigorously evaluating what works. On making sure we're rewarding, not inputs but

instead, outcomes and learning from those suboptimal outcomes, talking about those

suboptimal outcomes. And in the end, incorporating that knowledge iteratively into our

next set of interventions whether they're governmental in nature or non-governmental in

nature.

So I'm here because I believe in this whole notion of evidence based policy making. Whether that is governmental policy, non-governmental or what I just described, some hybrid of the two. I'm grateful for you giving me the platform to lay out some of my thoughts on this topic and I really look forward to working, Ron with you. And so many others here who are present or watching this event today to continue to advance evidence based policy in Congress.

MR. HASKINS: Well you certainly have used the platform well both in the written article and in your time you just talked. Let me begin with the family program you just talked about. It is a successful program, we have a lot of evidence but most federal and state programs, we don't have great evidence. When we have evidence in the majority of cases, they do not work. My question to you is, how can you convince your colleagues to pay attention to that evidence and to do something different to figure out a way to use the money more wisely and to produce good results.

SENATOR YOUNG: I can think of a couple of ways. One, I can withhold my co-sponsorship of legislation that doesn't include an evaluative component to that legislation. Set aside a modest amount of money to evaluate the effectiveness of a given program, report on the effectiveness as a trigger for Congress to pull the plug on a given program after a period of years or not. If I do that over a period of time, as a member of the Health Education Labor and Pensions Committee, I think that's one way that I can make a difference. Two, I can continue to communicate about the importance of rigor and evaluation and paying for success on the floor of the United States Senate and other forms like that one and beyond. So that's another thought.

One of the reasons I like the social impact partnership model is because my hope would be that if we can generate enough successes through this model and I would define success here, for the academics among us who like to define everything, I would define success as an improvement upon the status quo. But if we can improve

upon the status quo through the social impact partnership model or other pay for success

models, then we have a new body of evidence that creates a moral imperative for those

who might be ordinarily persuaded not to change a particular program to instead change

their mind. I know you're receiving pressure, colleague A, not to touch this sacred cow of

a program. But we now know that a better way to address prisoner reentry, a better way

to address asthma in low income communities, a better way to get the long term

unemployed integrated back into the workforce is here or here or here. We have strong

evidence, there is a moral imperative that you act. I think that can be really persuasive to

my colleagues and cut through some of the pressures they might otherwise feel.

Because most people, my colleagues on the right and left, run for public office to do good

and they have good hearts, it just takes some really compelling evidence, which often

times doesn't exist, to persuade them.

   MR. HASKINS: Let me ask you one last question on this exact point.

When I was a staffer on the Ways and Means Committee, I sat in many of these

meetings where the members talked and I recall in 14 years, a couple of times when

somebody actually said about a study. So now, if you start talking about a studies and

what pay for success are the things that have shown and you're trying to persuade you

colleagues. If you look at the research, if you look at the studies, here's what they show,

here's how we ought to spend the money. Is that effective? Do they actually listen to

you? Will they even change their vote based on what the research shows?

   SENATOR YOUNG: It depends. In my conversations with colleagues,

we all have varying degrees of confirmation bias. Increasingly, I think scholarship is

improving on behavioral science. We like studies that confirm our existing biases. There

are MSNCB viewers here, there are Fox News viewers here. There is a certain amount

of confirmation bias even in the media that people choose. It is painful to have to change

your mind, especially as you grow older and learn that you've spent many years, many

painstaking hours working on something that really didn't help people, may have even injured them.

So I think one way to get through this is to make sure that there are these joint relationships. I really like how Brookings has partnered with AEI on some different studies and consensus views have been arrived at. I found quite instructive, the work that you did working with Robert Dorr and some others on the lack of economic mobility in this country. And the extent to which our existing social programs have indeed improved the lives of our citizens. That was a consensus view. But the nuance there is there is a huge opportunity cost because they are not well targeted so again, a consensus view.

So I think it is important that people of different minds, academics, policy experts, come up with consensus views, find where there are areas of agreement gives great credibility to findings. Especially in an institution like the United States Senate where it takes 60 votes to get things passed.

The other thing we need to do is make sure that our evaluations are actually utilized. There are a lot of studies out there that are forgotten about. So we have to come up with a way to cut off the funding if the evaluation determines that a particular program is not effective. We need to get in the habit and again, I request that colleagues sunset their legislation often times before it receives my co-sponsorship. I sunset my own legislation, including the social impact partnership legislation, because I want to make sure that that model works and that creates a forcing action for Congress to act.

So those are some of my thoughts and I would welcome others to share with me, if not today than in the future, other ideas about how we might get more of my colleagues who are well intentioned, thoughtful people with few exceptions to focus on the evidence. Just as the private sector so frequently does. They test, they beta test, then they go big.

We here in Washington go big early way too often and I think that's a mistake.

MR. HASKINS: I am getting signals from your staff implicitly threatening me that if you don't leave within 30 seconds you're never coming to Brookings again. So thank you so much.

SENATOR YOUNG: Thanks for having me.

MR. HASKINS: I'm sorry audience, I didn't realize it was going to end this quickly. I was hoping to have at least one question from the audience. As you can see, the senator was late, he was on a very tight schedule so things started off squeezed which often happens.

Now we have a film, you can't squeeze the film because it is already filmed. We're going to show it up here on the screen. It stars Senator Carper and Andrea Cain sitting right there in the fourth row. I ask them questions and that's all I do. When that's over, I'll come back up here and we'll continue the event. Thank you.

So senator, thank you as well as your staff for your enthusiastic participation in this volume. The chapter you, Andrea and Isabel Sawhill wrote on pregnancy prevention was a terrific contribution to the volume. Andrea, from the influential Power To Decide organization, formally the Campaign to Prevent Unplanned Pregnancy is also with us today and will participate in the discussion. I'm going to give you the shortest introduction ever given to a U.S. senator.

SENATOR CARPER: Why?

MR. HASKINS: Because of our limitation of time. So you served in the U.S. military for 23 years including three tours in Vietnam as a naval flight officer. You were elected three times to serve as Delaware state treasurer, five times to serve in the House of Representatives, twice as governor and you're now on your third term for Delaware. I determined that if I added up all the years that you must have been in public service since shortly after the administration of George Washington. All right senator,

let's begin with this. As someone who has spent many years in public service, how have you relied on data and evidence to guide your decision making?

SENATOR CARPER: One of my core values has been, find out what works and do more of that. That's it. Find out what works and do more of that.

MR. HASKINS: How do you find out what works?

SENATOR CARPER: And if you don't have the data, if you don't have the information, you don't find out what works and document that then how do we know. What I do is just that. I work with the General Accountability Office, I work with inspector generals, I work with OMB, a committee on Homeland Security and Governmental Affairs. We relied on data when I was governor. I got to be chairman of the National Governors Association and after that to lead the NGA's Center for Best Practices. We always were comparing, stealing ideas from one another but also looking to find out what works in a wide range of areas.

MR. HASKINS: Did you notice an increase in a number of high quality studies that actually tested ideas for solving social problems?

SENATOR CARPER: Here lately (inaudible) focus so much on actually using good data and good information. There is a lot of fake news and fake numbers as well. But if we're going to actually make progress we've got to find the data, drill down and make sure we know what works.

MR. HASKINS: Okay in your many years of experience in the House and Senate, have you seen an increase, not just yourself but in your colleagues, in the House and the Senate in their use of data and especially in evidence based policy.

SENATOR CARPER: It's a mixed bag. In some cases, you have a great example, Patty Murray wonderful Senator, Democrat from Washington and the Speaker of the House. Them working together in a way that provides a good example for the rest of us. In other cases, we have people just saying whatever they think or whatever they

want people to think so it's a mixed bag.

MR. HASKINS: So what opportunities do you see to stimulate greater use of evidence? One of the purposes of the evidence based moment is to encourage people, especially elected officials, to look for better data, to finance better data. Do you see that that's a realistic possibility, would you support that, do you think your colleagues would?

SENATOR CARPER: We had a hearing just this morning in the Committee on Homeland Security and Governmental Affairs with really focused on Governmental Affairs. We had in one of the top administration officials from OMB. A very bright woman named Margaret Weichert. And she was presenting a government overhaul that the administration was laying at our feet to consider. Not just to do in the next month or even in the next year but something to consider over the next several years. One of the things that we talked about is the need to not just say who thinks this might work but to actually understand what is working in the states. One of the great examples, I like to say states are laboratories of democracy. If we're going to try something big why don't we, rather than just starting at the government wide federal level, why don't we find out what has worked whether it is in Delaware or Minnesota or Wyoming or some other state.

MR. HASKINS: When I first met you we were working on welfare reform.

SENATOR CARPER: Yes we were.

MR. HASKINS: You may remember that 41 states were doing exactly the kind of demonstrations you're talking about here. And doing high quality random assignment studies to figure out what works in welfare reform and how they can encourage mothers on welfare to work. So that has, it's kind of an old story but then it died down and we haven't had nearly so much of it lately. So if the government could encourage it more I think we would be in very good shape.

SENATOR CARPER: We just recently debated in the Senate, agriculture policy. Every five years or so we do a big farm bill. A big piece of the farm bill, as you know, is nutrition. It includes the SNAP program which used to be food stamps. We focused a lot on welfare reform and we wanted welfare (inaudible). We wanted people to be better off and one day they work. One of the nice things when we drill down in the Senate on the farm bill on nutrition and the SNAP program, we actually look at what was working. For people to work, we don't want to take away, we don't want to provide disincentives for people to work. That was actually very encouraging and something that is happening in real time this year.

MR. HASKINS: Unfortunately it often happens that they don't work well so then you've got to figure out new ways to do it so you can actually make them work (inaudible) role of Congress. So Andrea, let me bring you into this. What can we learn from the example of teen pregnancy and more generally, women of all ages, pregnancy reduction, from the chapter that you wrote with the senator?

MS. CAIN: Well I think one of the things we've learned is exactly what Senator Carper just talked about which is that states are following the evidence to reduce unplanned pregnancy and improve outcomes for children and families. There is very encouraging evidence coming out of a number of states. Colorado being the first and then also St. Louis, the St. Louis area and Iowa about when you reduce barriers to contraception for women. And make sure that there is a level playing field so they can pick the most effective contraception based on the evidence. We see really positive outcomes and governors and other policy makers have seized on that evidence across party lines. We've seen this kind of thing happening now in red states, blue states, purple states and it is spreading.

So I think there is hope that when there is good evidence out there, people do try to follow it. It gets shared through different dissemination channels. Just one example,

the now governor of Virginia read about the research from Colorado that showed that their teen birth rate had gone down by half after they implanted a new program to increase access to contraception. So we should do that in Virginia. He is now governor and after a couple of attempts, that is now just passed. So we do see these kinds of things spreading which I think is very encouraging.

MR. HASKINS: You might just mentioned great things are going on in Delaware as well.

SENATOR CARPER: Our immediate past governor, Jack Markell, wonderful guy who was also chairman of the National Governors Association. We've had a high teen pregnancy rate in Delaware for too long. We focused on it when I was governor, we focused on how do we strengthen the basic building block of our societies families. It really started and put together a multi departmental task force for seven or eight years as governor to focus on how do we strengthen families. It started with reducing incidents of unplanned, unwanted pregnancies.

We brought in a whole bunch of kids from high schools all over the state, two for every public high school. Male, female and from all walks of life to say this is what we think would work to reduce the incidents of teen pregnancy. A wonderful woman named Marian Wright Adelman once said and I'll paraphrase. She said when a 16 year old girl becomes pregnant, drops out of school, does not marry the father of her child and does not go back into school, there is an 80 percent likelihood that they'll live in poverty. That same 16 year girl does not get pregnant, does not drop out of school, waits to 21 to have a child, marries the father of her child, there is a 8 percent likelihood that they'll live in poverty. 80 percent on the one hand, 8 percent on the other hand. I've never forgotten that number.

Whenever I speak in high school assemblies I ask the kids, I say how many of you want to live in poverty. There is usually a jokester somewhere who always

raises their hand. Nobody in all the dozens of times I've done this, no one has ever raised their hand one time. Nobody wants to live in poverty. One of the great predictors that families are going to live in poverty is what I just explained.

MS. CAIN: Absolutely and the data are very clear on that.

MR. HASKINS: A part of this story that you go into, very interesting and well put in your article is about LARCs, Long Acting Reversible Contraception. I think we ought to tell just a little bit of that story because it is so fascinating and it has tremendous potential.

MS. CAIN: Absolutely. So LARCs, Long Acting Reversible Contraceptives is not a new idea although it is now spreading because we have more data and evidence.

SENATOR CARPER: There was a good idea a long time ago that got a very bad development in it and we've finally fixed. I think we still have lived with a legacy of that tragedy.

MS. CAIN: Yes, absolutely. So for those who may not know, LARC include intrauterine devices or IUDs or implants that actually go in the arm and they are highly effective. All the scientific research has shown that they are about 20 times more effective than other methods of birth control.

MR. HASKINS: Is that like 99 percent effective?

MS. CAIN: Exactly like 99 percent effective.

MR. HASKINS: Compared to what?

MS. CAIN: Depending on other methods of contraception, 60, 80, 85 percent. Why they're so effective is that once they're in place, a woman doesn't have to do anything. So they are sort of fool proof.

MR. HASKINS: You don't have to remember to take a pill. You don't have to hope that somebody is going to have a condom that works, none of that stuff.

MS. CAIN: That's exactly right. They are not right for everybody but what we want to do and what the states and others are doing now is making sure people know how effective they are, how safe they are, how easy to use they are so that women who want to use them can. What we've seen in places like Delaware is that when you reduce barriers and make them available to everybody, many women are choosing them. So again, it is having the science and evidence drive policy and practice. We still have a long way to go. There is still a lot of barriers and I think one of the things that is really challenging right now is that while we have this forward momentum in places like Delaware and Colorado and other places there is also some real threats to those policies.

MR. HASKINS: There are few areas in social policy where we make as much progress as we are against teen pregnancy prevention. How did you first get interested? I know you've been involved for a long time.

SENATOR CARPER: Really when I ran for governor, I was a congressman. Ran for governor, Mike Castle was our governor. He was running for Congress. We were trading places like the movie, Trading Places. He won, he became a congressman, I won, I became governor. But before we were sworn into our new offices we were at a dinner together in Wilmington. It was a big dinner, a lot of people there. And the fellow who was the emcee introduced both of us to speak and he raised this question. He said in your new job, if you could have a magic wand to start on your new responsibilities what would you choose. I said here's what I'd like. I'd like to walk into my office first day as governor, open the drawer and pull out a magic wand that would allow me to make sure that every child in my state had at least one loving, caring, nurturing parent, at least one, hopefully two. A few months later I walked into my new office, I opened the drawer and there was no magic wand.

About a month later, one of the new cabinet secretaries was a holdover from the Castle administration. They came in and said, we ought to have a statewide

focus for a multiyear effort on strengthening families. I remember being in prison about a month or two later after that where we had 50 inmates in a room with me and the nation's drug czar. The nation's drug czar was there to see what we were doing to reduce the incidents of recidivism. We had 50 mostly young inmates. I remember asking these inmates, tell the drug czar and me, how did you get here? What happened in your life to get here? There was like 18, 19, 20, 25 and one after the other started with this. I was born when my mom was young, I never knew my dad. I started kindergarten behind, I couldn't read or write. Many other kids in my class could.

For me I said bingo. This idea of a Family Services Cabinet Council where we focus on starting on teen pregnancy, we should do that. We should do it for eight years and we did. It was the single best thing we did.

MS. CAIN: And what is dramatic is that since then, the teen birth rates in Delaware and around the country have gone down by well over 50 percent. That really is remarkable.

SENATOR CARPER: And the incidents of abortion. It's one of the things we don't focus on. The best way to reduce the demand for abortion is to do a much better job on contraception. And the top of the line on contraceptives is IUDs and LARCs.

MR. HASKINS: Senator, thank you so much, not just for this interview but for the great article that you wrote with Andrea and Isabel Sawhill here at Brookings. And, of course, as you know, we're going to show it tomorrow's event where we're going to release the journal and we'll have thousands of people crowding the Brookings auditorium.

SENATOR CARPER: Are you thinking of moving it to RFK?

MR. HASKINS: Well, RFK is too old. I think the opening sentence in your article for the chapter shows what I believe most people who are interested in

evidence based policy would want policymakers to say, and here is exactly what you say. A growing number of policy makers and program leaders at the federal, state and community level are committed to using unbiased evidence to guide them in solving tough social problems. I know you've done that throughout your career and we can only hope that more policymakers will do the same thing.

   SENATOR CARPER: From your lips to God's ears.

   MR. HASKINS: Great. Thank you so much.

   SENATOR CARPER: Thank you.

   MR. HASKINS: Because we're not focusing our event on this wonderful journal, 50 of you have copies of this. I'm sorry we didn't have enough for everybody. It is available online. There are instructions on how to get it. It is open to everybody and you can print out whatever you want to. I have noticed in the past that people don't always want to print everything in their journal. You might want one or two or five articles.

   So I want to take just a few minutes to give you a little bit of information about the contents of this volume to promote your understanding of it. After all, that is our major purpose. When Tom Kecskemethy, who is sitting right here, who is the executive director of the American Academy of Political and Social Science, our sponsor, offered me the opportunity to edit a volume on evidence based policy, I immediately accepted. My thought from the very beginning was to get the leading experts in the field to write about the biggest issues. Thereby, creating a volume that would provide a non-technical overview of the field with a substance divided into short, snappy chapters. Perhaps our readers will agree that the authors have achieved this most important goal.

   There are 17 chapters in the volume. I cannot do more than hint at the scope and precision and competence with which they were written. The introduction, written by me at excessive length aims to give readers an overview of the entire volume

and hence, the major parts of the evidence based movement in one single place.

John Barring of the Arnold Foundation provides a fascinating history of the development of evidence based research and policy. Going back to the amazing work of Richard Dole, a British epidemiologist who declared in the 1930's that medical schools were a joke. Because few, if any of the treatments on offer had been systematically tested, so it was evidence based from the beginning. Barring proceeds to bring the story all the way up to the research on welfare and employment policy in the 1980's and 1990's that had a major impact on policy.

As Larry Ore, who is out there somewhere, of Johns Hopkins reminds us, randomized control trials are the foundation of evidence based policy. They are the major tool we use to identify social programs that actually produce impacts. So far, we have found, perhaps to our surprise that few programs are successful. But it does not require great optimism to think that we're moving in the right direction and that we are constantly developing more successful programs. Stay tuned. We're moving in the right direction.

Showing the range of both the volume and the field, we included a chapter by Patrick Lester on the Pew and McArthur First Program which appears to have achieved what was previously thought to have been impossible. Getting states to terminate or modify spending on programs that are unsuccessful and using the money to improve the programs or spend on programs with stronger evidence of success.

In a chapter on the spectacular, poverty action lab at MIT which has grown to more than 150 researchers who have conducted more than 900 RCT's in 90 countries that have impacted over 300 million people. Strong evidence of success as well.

The chapter on behavioral economics written by two of its most notable practitioners Bill Comdum and Mya Shankar, shows many examples of how the theory

and practice of behavioral economics can improve government policy and save money. The chapter by George Overholser which who will be up here talking in a few minutes of Third Sector Capital Partners, will surprise those who are familiar with his role in developing a pay for success program. He presents a substantive revised version of the original design which there is good reason to believe will make it more practical for organizations to adopt and operate.

We have five chapters on the contribution of institutions on the evidence based movement. The Office of Management and Budget by Cathy Stack who is also in the audience, the Institution of Education and Sciences by Russ Whitehurst will be on the panel and the role of non-profit organizations in designing and implementing evidence based policy by Jim Sullivan and a chapter on other federal agencies by Becca Maynard.

Without the work of these institutions evidence based policy would be little more than a concept. With them, we get action, especially implementation and development. I wanted to have two chapters on the significance of the evidence based movement. One more upbeat about its achievements and one more skeptical. As it turned out, the chapters are two of the most thoughtful in the volume and both are a proactive blend of skepticism and optimism. One of the chapters by Ginger Knox, Carolyn Hill and Gordon Berlin, three of the most capable practitioners of both program evaluation and drawing conclusions from programmatic research. They address the Achilles heel of the evidence based movement, namely, program replication. Perhaps none of you will be surprised to discover that if it's difficult and relatively rare to find a program that actually produces impacts, how much more difficult must it be to find a program that produces positive impacts in many settings and with many groups of children or adults. Yet Knox, Hill and Berlin have what looks to me like a possible solution to the replication conundrum.

The chapters follow by a compelling chapter on three orders of

challenges researchers face in producing evidence that could lead to realistic policy solutions by Adam Gamoran who will be on our panel in a minute. He is the president of the Grant Foundation. As much as any chapter in the volume, this one gives readers a concrete view of the challenges facing researchers who want their research to help solve social problems.

Finally, we have a terrific chapter on using administrative data in research and is recently highlighted by the Ryan Murray Evidence-Based Policymaking Commission. And then, of course, we have two chapters written by U.S. Senators, kind of a rare thing to see. These chapters provide convincing evidence that there are influential policymakers who not only know about evidence based policy but who practice it. There you have it with abject apologies to the authors for my poor ability to capture the glories of each of these chapters. I think Tom Kecskemethy is trying to think up ways to punish me for such a pale summary of the wonderful chapters in this volume. I hope he fails.

I now have the privilege of welcoming Megan McArdle, a dazzling new editorial writer for the Washington Post. Megan will moderate our panel discussion. Before turning things over to her, I want to emphasize to people who might notice that no one from the Trump administration is on the panel. We've been talking with officials from the Trump administration for four weeks about sending a representative. We found out Monday that they were not able to work out the details. I invite Megan and our panel to come to the stage. Thank you.

MS. McARDLE: Thank you for coming and staying to hear our great panel. We have Adam Gamoran, who just told me how to pronounce his last name, who is the president of the William T. Grant Foundation; George Overholser, who is president and founder of Third Sector Capital Partners; and Grover J. "Russ" Whitehurst, who is a senior fellow in economic studies at the Center on Children and Families, Brookings, and then there's me.

So everyone has their -- they're going to do brief presentations on their papers and then we're going to have a little chat, and hopefully we'll have at least a few minutes for questions from the audience at the end.

So I'm just going to suggest that we go this way in order, so, Mr. Whitehurst.

MR. WHITEHURST: Well, welcome everybody and thanks, Ron, for inviting me to be part of this and to contribute to the volume.

I was torn between talking about the paper I wrote for the volume, which is about the agency in the federal government of which I was the founding director, the Institute of Educations Sciences, torn between that and talking about evidence in the Trump administration. I'm going to do probably a feeble job of getting those two things together one way or another.

Broadly I think there's an arc of human history that bends towards increasing rationality and use of evidence. That can be reflected in some periods of progress and some periods of retreat, but I think we move forward.

The reason for that I think is not so much an inherent or innate progressive movement of history, but rather the use of evidence increasingly allows more receptive audience for evidence.

I was reading in the Post or the Times, I don't remember which, I'm sorry for that, a piece yesterday about loss aversion.

MS. McARDLE: If it was good, I'm sure it was in the Post.

MR. WHITEHURST: I'm sure it was. I'm sure it was.

Loss aversion is a psychological phenomenon, behavioral economics phenomenon identified about 30 years ago. Basically winning a dollar is not nearly as emotion generating as losing a dollar, so it's more painful to lose than to win.

So as we know that and as people are struggling to create policy and

that policy involves taking something away from someone and giving them something else, it's almost impossible not to use that knowledge as a way of rolling out the policy for political and other reasons to make it more successful.

So my argument is the more we know, the more likely we are to use evidence, because the evidence itself has embedded in it mechanisms for progress.

With respect to the Institution of Education Sciences, I think it is -- I believe the reason I was invited to write about it for the volume is that it is a marked indicator of the ability of federal government to change its way of behaving and to be more active in producing evidence.

So U.S. Department of Education and its predecessor in Health, Education, and Welfare had research offices for a very long time. They were abysmal failures. Up through 2001, the Department of Education had conducted only two randomized trials in its history with respect to the effectiveness of the programs it was delivering. There was little appetite for evidence and little supply of evidence.

My first couple of years in office I'd go up to Congress and they would say, well, what are you doing and what do we know. I would dance around and talk about the promise of the future, but there really was very little to put on the table.

Last time I looked, which was to do this volume, the What Works Clearinghouse, which is a focused effort by the Institute of Education Sciences within U.S. Department of Education to identify What Works and what the size of those effects are and how much they cost, had conducted over 10,000 evidence reviews, had identified from memory 340 different studies that were well conducted randomized trials, or close approximations, and found evidence of what worked and what didn't.

So if I had the job now that I had back in 2001, 2002, I wouldn't have to dance around it, I could talk about a list of things that had been shown to work, a larger list of things that have been shown not to work, and to encourage people to use that

evidence going forward.

The question is what is -- what is the present state of evidence in the Trump administration, what do we need to worry about, what do we need to do going forward.

I think it's very important in that regard to distinguish between the production side, the supply side of the equation, and the utilization or demand side of the equation.

The production side, I'm referring to the people who do the studies, who collect the statistics, who make the reports. On the demand side of the equation, I'm talking about people in the White House who could use that evidence, people in Congress who could use the evidence, practitioners in the field.

I believe at least at present, the production enterprise is not at risk. Congress continues to fund the offices that do research evaluation and publish statistics. Those offices are doing their work. They're publishing the reports. I see no evidence, at least in the part of that world I'm familiar with of any undue or inappropriate political influence in the process. So as long as Congress continues to fund that activity, the offices will continue to do their work.

On the demand side of the equation, as we've heard in some of the comments in the earlier material today, I think there's a tremendous amount of hunger at the state level and sometimes at the local level for evidence that's actually relevant to what those entities are trying to do and it helps them do it more effectively.

So I think the natural supply line, the place where the audience is greatest and I don't see it diminishing; for the federal knowledge production effort is at the state and local level. Again I see that as healthy.

I don't see much demand in the executive branch for evidence. In previous administrations, George W. Bush administration, which I served, and in the

Obama administration, which I followed carefully, OMB was very powerful in advancing

the evidence movement. There were key staff -- career staff there as well as politicals

interested in this activity that moved the ball forward. I don't see that happening now. I

don't see the people there who were doing this. Maybe I'm wrong, I hope so.

I think there's less interest at the executive branch level, the Cabinet

level in using evidence than they're used to be and that's not good. On the other hand,

administrations come and go.

It's very important. And I think just to keep the train moving forward and I

think it will be greater appeals in the future, greater audience appeal, for what the

evidence movement within federal government is doing.

Congress has a critical role to play going forward and that's why I was

very encouraged by the two interviews you did this morning. Not only must they continue

to fund the activity, but they must pay careful attention to the conditions in which that

activity can be carried out well.

That certainly includes protecting the entities in the executive branch that

collect, evaluate, produce statistics, protecting them from undue influence.

The Institute of Education Sciences has a model piece of founding

legislation. In doing that, it operates independent of the U.S. Department of Education

and its publications and priority setting, and I think it's very important for Congress to

extend that kind of independence to other particularly evaluation shops in the federal

government that are under intense political pressure when they show something -- when

they find that something doesn't work to which strong political commitments have been

made by the administration.

So they need to be protected from a natural tendency of Cabinet

secretaries and other people to be very disappointed with an office that's off message.

It's also I think important to (inaudible) Pay for Success as part of this.

They have contingencies in place so that there are ways to reward those who use evidence.

In business there's a very clear bottom line and people's salaries are often connected with how well they conduct an activity whose success can be measured by that bottom line. We often can't agree in the social and behavioral and education areas what the bottom line is. There are many. The effects we get are effects on average. How well does Head Start work on average all across the country, that can't be fed down to rewards for particular program operators.

I think Congress has a role to play in making it to the benefit of those who are implementing the programs and those who are carrying it out when the program actually succeeds.

So again I think we move in the right direction, sometimes slowly, sometimes a step backwards, but we're moving appropriately. The progress sometimes can be amazingly fast and it can slow down for a while.

I'm optimistic. I hope you are. I expect that you are interested in this topic, that's why you're here, and we can share that optimism. Thank you.

MR. GAMORAN: Thank you. Great to follow you, Russ. I think we share some thoughts in common, particularly by the Institute of Education Sciences and perhaps some differentiating.

In my remarks I'll make three points in response to the (inaudible), which is what's the status of the evidence movement in the Trump administration, so I'll make three points.

The first point is bad, the status is bad; the second point is nonetheless there are some perhaps unexpected and surprising points of light; and the third is it's more complicated than you think.

So obviously facts are under assault in this administration. We don't

have to look any farther than the world being transfixed by the sight of a president

ignoring a fact-based analysis of the intelligence community and declaring that a foreign

dictator gave a really convincing denial of that evidence to be concerned about the status

of the evidence movement in the Trump administration.

In his first 400 days in office, President Trump uttered 2,400 falsehood,

that's an average of six lies per day in Twitter, in speeches, and in press interactions. Yet

the counsel to the president says, those aren't lies, they're just alternative facts.

Nothing could be more challenging to the notion that policy should be

based on evidence than an idea that if you don't like the facts, you just make up another

set of facts. Evidence is imperiled across this administration.

A few minutes ago we saw a video about the powerful evidence on

efforts to reduce teen pregnancy, and yet that funding program has been killed by Health

and Human Services under this administration.

I think there -- what you say, Russ, is true to some extent, production of

evidence continues, but it is imperiled in many parts. Just last week Health and Human

Services announced a plan to delete 20 years of medical guidelines known as the

National Guideline Clearinghouse of the agency for Healthcare Research and Quality.

Also last week, the administration promoted adding work requirements to

government-funded healthcare, that's a way of cutting benefits, by declaring that the war

on poverty has been won. That's like holding up -- as one commentator said, I didn't

make this up, it's like holding an umbrella and saying I'm not getting wet, even though it's

raining, again, ignoring -- operating in the face of the evidence.

Earlier the Environmental Protection Agency ruled that its employees, its

scientists, may not site scientific research unless the data are public, that's a phony way

of promoting transparency, a way to ignore scientific evidence.

Similarly staff at the Center for Disease Control were told not to use the

phrases evidence based or science based in preparing their budget request for this year.

Going back to its first -- the first days of this administration, the current administration, the Office of Management and Budget removed its language on statistical policy and standards from its website.

A housing bill that was proposed fortunately not passed, contained the following language, quote, no federal funds may be used to design, build, maintain, utilize, or provide access to a federal database of geospatial information on community racial disparities or disparities in access to affordable housing.

So these are ways in not just the utilization but the production of evidence is at risk, and it will require people working inside and outside government to keep that going. We're not assured of its continuity.

With that said, it's important to recognize sparks of light in dark times. Notably these examples are cases of carryover across administrations from the George W. Bush administration, to the Obama administration, and continuing on today. I'm sure there are more, but I'll give three very quick examples.

The president's Fiscal Year '19 budget proposal, which was not taken seriously, dismissed out of hand, and so you probably missed the important Chapter 6, which is Building and Using Evidence to Improve Government Effectiveness. This is a very strong chapter. It carries over and even elaborates insights developed in budget proposals in the Obama administration.

If Trump had did the recommendations of the Commission on evidence-based policymaking, which Ron Haskins co-chaired, it offered concrete strategies to strengthen the production and use of evidence for smarter policy decisions, and some of these proposals may actually happen and that is President Trump's budget proposal.

One key recommendation was to strengthen interagency coordination

and my second example is an interagency coordinating group on using evidence in government. This is an active group of staff involved in the research side of government who are continually grappling with the challenges of producing credible evidence and then getting it utilized by the policy divisions within their agency, so that is an ongoing effort.

My third example, some federal agencies at least are promoting the production of high quality evidence for decision making. The one I know the best is the U.S. Department of Education. Both the political and the professional appointments are strong in this department, particularly the Institute of Education Sciences, the research arm of the U.S. Department of Education, which Russ founded and is now led by one of your former appointees and I think inspires confidence that in this agency at least the production of credible evidence will continue.

So even in the worst of times, there is some carry through. Will it be enough? Only with continued vigilance and efforts I think.

Third point, it's complicated. In my article, in the Annals volume, I explained that there are three challenges to evidence-based policymaking. The first is producing credible evidence. We've made a lot of progress, identify some challenges, replication is one of them, but we've made a lot of progress in this area.

But the second challenge is getting evidence used and that is the bottleneck today. Russ spoke about the supply side and the demand side. It's the demand side that continues to be problematic. We need a science of evidence use, something that the William T. Grant Foundation, which I lead, has been strongly promoting through our grant making.

We need an infrastructure for evidence use, such as appeared in that OMB budget proposal. We need partnerships inside and outside government between those who produce evidence and those who use evidence.

Now, even if we could produce credible evidence and it gets used by decision makers, we would still have a third challenge, and that is that social problems have deep structural roots.

So when we identify, for example, teen pregnancy programs that work, so we cut the rate of teen pregnancy in half, we still have inequality and economic insecurity in our society, we have multiple problems that's not attacking one problem at a time.

So programs that respond to problems, even if they are effected and if they're implemented widely, do not get at the structural foundations of deep-seeded problems of inequality and economic insecurity in our society.

In my chapter, which I encourage you to read, I lay this out in more detail. For the evidence movement, to get back on track and truly make progress, we need attention to all three of these challenges.

MS. McARDLE: Finally last, but not least.

MR. OVERHOLSER: I'm going to bring more of a practitioner's point of view. I came to this world actually from a business career. I was on the founding team of Capital One, which was right in Virginia actually across the river many years ago.

Our whole approach was to embrace evidence and to conduct randomized controlled trials. We did thousands and thousands and thousands of very, very high quality randomized controlled trials and used that to drive the innovation process for our credit card company, which started as a little credit card company.

We began -- kind of inherited the credit card division. We had about 250 people. Seven years later, we had 20,000 people and had been the number one growth company in the New York Stock Exchange during a three-year period, all because of our embrace of evidence in a world where the competitive backdrop was a kind of one size fits all set of products that had not been based on evidence. What an experience it was

to build the highly decentralized culture of testing and to watch that in action.

One of the first things we learned was the wisdom at the top was completely debunked by the evidence that pervaded the rest of the system. We no longer tried to drive innovation from the top, we instead learned to trust what was bubbling up from the bottom.

The other thing that we learned was (inaudible) to make posters. We put up posters all around the company of pictures of melting ice cream cones and it said, it melts like ice cream. Now, why did we do this, we did this because we found that even when we had the perfect replication of a breakthrough innovation that had been tested and shown to be highly successful, it would not continue to work over time.

That's because in the world of human beings, it's not like physics. In physics you learn that hydrogen and oxygen make water and it stays true. But for social programs, even if they're in social programs having to do with credit cards, they stop working and they melt away over time.

So when I eventually left Capital One, I moved to the nonprofit sector and have been working in the nonprofit sector for the last 15 years. Third Sector Capital Partners is a nonprofit and our mission is to accelerate Americans' transition to a performance driven social sector.

In many ways our strategy is to echo the work that I had seen work so well in Capital One, and that is to build feedback loops within the procurement offices of social services around the country.

So we currently have 50 people on staff. We're growing rapidly, and we're using predominantly a Pay for Success type of model. The idea is to -- in dozens and dozens of different contracting operations to change the way contracting happens such that providers are no longer asked to follow the recipes that were handed to them by government but instead are being rewarded for changing the recipes, but only if those

recipes bring about better results for their communities.

Now, there's a subtle difference between this approach and much of the evidence-based movement overall that I think is worth spending a moment of time.

In general the way that I saw government as I first showed up, using evidence was what you might call a scale What Works idea or a codification and scale type of an idea.

Let's find something that works and if it works great, we'll write it into law and we'll say anyone who follows that recipe will get funding, which is a very exciting thing that you can do if you have a recipe that truly works and that endures over time.

Unfortunately though these recipes melt like ice cream, and so that approach, the generalizability of RCT findings, if you will, the generalizability of these insights often doesn't work. Just because it works in one place or one setting, doesn't mean it will continue to work over time and it will work everywhere.

I often make the joke if we did music the same way we did government funding, we would all be listening to Meatloaf on an eight track recording machine, because 30 years ago someone would have said, wow, that's great music, that's a great breakthrough in recording, there ought to be a law, and it would be written into law. And from that point forward anyone who did music would have to comply to the recipe.

So the second branch, which is Pay for Success, says, let's complement that codification approach, which does work sometimes, for instance, upstream, which is the Lark program. I think that is an incredibly great application of the codification and scale approach.

But the other branch, which is Pay for Success, says, let's decentralize and put -- in a sense the metaphor is, our firm, we're taking this body politic, which is very numb, and we're microsurgeons that are adding little nerve endings all over the place so that the outcomes associated with contracts are sensitive to whether or not those

contracts give rise to better results for communities.

That's more of a bottom-up approach. It's not about codifying What Works, it's more about rewarding innovation. We think in the process actually prompting more innovation to happen as people invest in trying out innovations, which in the past were not being rewarded, but which will be rewarded going forward.

Final thing I'll say is in the chapter in the book, we focus on progress we've made in our own model as practitioners. When we first started, it's exceedingly difficult to find government spending streams that were permitted to make contingent payments. It was written, here's a hundred million dollars, it's for this purpose, but it can only be used for reimbursement.

Yet Pay for Success, the whole concept is to make payments contingent. There's also it must be used within the year. Whereas with Pay for Success, we were saying let's look further out and have multiyear type of arrangements and look at the longer term outcomes.

So we were so thwarted by that that we had to raise private money instead and that -- really the fundraising was the hard part. We would spend years, literally years, putting together one contract for one provider, that's changed a lot.

In recent our most recent project for instance involves a billion dollars in the state of Washington, where we are working on contracts that affect 30 providers at a time and that don't require us to raise external funding.

Without going into the details, it has to do more than anything with changes in the permissible use rules various spending streams that have made it so that instead of having to raise new money, we can take existing money and do contractual -- contingent contracting and that is something that quietly, at the federal level and at the state level, laws are being changed to -- and often just regulations are being changed to make this a permissible use of existing spending streams.

A billion dollar -- we now have two -- we just got going on one billion dollar mandate and we have another one that's getting lined up. So this is really starting to take root and I'm very excited about it, so I better stop there.

MS. McARDLE: The name of this event to me is a little bit funny when I think about it, evidence-based policy, I mean, what's the alternative, evidence free policymaking, but in fact we seem to do a lot of that.

So I want to actually just start by talking a bit about the politics of this. When I look out -- as a political columnist when I look at what are the big things the parties are talking about right now, well, on the Republican side they're talking about trade restriction, which is -- there's huge amount of evidence this doesn't work. They've got one guy who is really a complete outlier among economists who is driving a ton of this.

You've got immigration restriction, I think that's more arguable. The evidence on what the effects of immigration is not what Trump says it is, on the other hand we're only measuring only one very small effect on wages. We're not really measuring cultural effects, proximity effects, et cetera.

On the other side, what is trending on the left? Universal basic income had some interesting evidence for it in Kenya, that evidence just blew up and it now seems like the Universal basic income didn't do anything, increased wealth, but did not produce any improvement in any other outcomes and seems to have made their neighbors so much worse off that in fact the net effect of this policy is negative. I've heard nothing about that from the left.

I'm still hearing about the Kenya study, I'm not hearing about what happened on the -- Early Childhood Education, this is all over every Democrat who wants to talk we're going to do Early Childhood Education. With all due respect to James Heckman who is a very, very smart man, most of the evidence for this is based on

four studies that were done now 30, 40 years ago, had tiny numbers of people in them, and seem to have had some methodological flaws that Heckman says he's correcting for.

What is the total population of students in the studies that we're talking about, it's hundreds, right, below hundreds?

MR. WHITEHURST: Well, 65 for Heckman's favorite study and 120 for the (inaudible).

MS. McARDLE: So we're talking about something that worked with 200 people, and I believe one of them was in the segregated South, which is a different environment. I mean, you're talking about people who didn't have plumbing and running water, just extremely different situations for these things.

This is basically all the evidence, because the evidence from the programs that we're trying to do doesn't really seem -- and every day I have these discussions with democratic policy experts. They say you don't know about the new studies, you're completely wrong, you got to look at the studies. I say, you're actually convincing me that you're not very good at reading studies, not that this is...

I guess the question I have here -- then there is a contrast, mass incarceration, where there is at least a movement to do serious stuff with serious evidence led by NYU now, but it seems to me in general when I look at these questions the evidence-based policymaking doesn't work where you have deep ideological commitments.

At that point, it's like okay, maybe this doesn't work and you get instead into what -- the philosopher Robert Nozick used to joke about normative sociology, which is the study of what the causes of things ought to be.

I think that we're now getting into normative policymaking, which is the study of what the solutions of things ought to be, rather than what the solutions are.

Is there a way around this? How do we get evidence-based

policymaking into the biggest questions and the hardest questions? I think that is

something you see in the Trump administration, but I think it's also something you saw in

the Obama administration.

How do we move beyond that and get these things up to the places

where they matter most? Since I just trashed Early Childhood Education and I'm

probably going to be facing a pitchfork laden mob on my way out the door, I will ask

Mr. Whitehurst to start.

MR. WHITEHURST: I think one of the solutions is to think small. The

more there is a strong political and cultural commitment to what is the right thing to do,

the more difficult it is to bring evidence to bear.

That's why I think across administrations as you -- presidents campaign

on what they're going to do. If they win -- candidates campaign on what they're going to

do. If they win, there's a checklist of making sure that there is delivery.

It's extremely hard to change that in the first couple of years in an

administration when they are doing the policies that they said they were going to do.

But the going small, the going local with the current model of Pay for

Success gives a chance for evidence to work where it makes a difference and where

there's less attached to it politically where the stakes are lower for those who can make

decisions. I think that's the sweet spot in terms of -- in terms of moving forward.

Eventually you hope that will bubble up to the point there's some

embarrassment for people who say on a high level we're going to do this and it's clear

that the evidence on that doesn't work.

To plug myself for a moment, I published a piece last week on whether

state pre-K is actually affecting achievement in elementary school. So if you're interested

in the sideboard, I invite you to go have a look.

MS. McARDLE: I saw some nods.

MR. OVERHOLSER: I do think going small helps and going frequent helps. I sometimes joke and say what we have now is a fund what once worked system. Maybe it's a fund what someone thought once worked system.

You'll see 30 year -- we decided (inaudible), what was that, 45 years ago. I don't think how long ago –- it was pari-preschool. The other trick is I think not only to be highly decentralized -- because what works in one place doesn't necessarily work in the other. Because it melts like ice cream, everything is changing all the time. What you need to do is to increase the cadence of when things are rundown.

MS. McARDLE: Let me ask you about that, though, do you run into a legislative bandwidth problem, right, which is --

MR. OVERHOLSER: So the big reframing that I suggest is that every once in a while something will come along like the Larks and say here is a solution that ought to be rolled out.

By and large, I think it's kind of a fool's errand to attempt to legislate solutions up top in a world where solutions need a tremendous amount of customization locally and they need to change frequently.

So instead I see government's job is rather than to say -- be prescriptive and say here is the recipe that ought to be followed, instead the great opportunity for us is to say here is a way of rewarding you if you come up with something that works in your community.

MS. McARDLE: But how do we change the structure of government financing (inaudible) -- it seems really tough.

MR. OVERHOLSER: With all the progress we've made in producing evidence, the barrier is getting evidence used. The most common use of evidence in government is a tactical use of evidence where decision's already been made and you cherry-pick the evidence that --

MS. McARDLE: Using it the way that drunks use lampposts, for support rather than --

MR. OVERHOLSER: Exactly. There are ways to improve the use of research evidence. It's about developing relationships between producers and consumers of evidence so that the questions investigated are ones where the decision maker has a stake in what the answer is.

We funded a researcher at Penn State named Max Crowley. He's developing a randomized trial of an intervention that brings researchers and congressional staff together.

The researchers are randomized to whether they work with the staff or not, and the congressional staff are randomized to whether they get the intensive relationship or a lighter touch.

This is going to test an approach of relationship building where the researchers pursue questions that will find their way into legislation, depending on what the research is, and to see if that will be impactful.

But it is the going small approach, it's not big, ideologically charged issues like immigration or trade. It's the everyday, but there are hundreds of everyday decisions that are government leaders need to make and these can be informed by evidence.

MS. McARDLE: So it sounds like we've got a little bit of attention I think, because you want to push things down to the lowest possible level, because these effects are often going to be local. This is a point that Jim Manzi who writes about -- wrote a book called Uncontrolled, which I recommend to everyone.

SPEAKER: He used to sit next to me, by the way. We worked together for years.

MS. McARDLE: Did he? He is a very, very clever fellow. I had a

number of really fascinating conversations with him, but he argues that when you do an RCT, he does them now in commercial settings.

He says you do an RCT on whether people like strawberry or raspberry jam, you're not proving like strawberry jam or raspberry jam is better, you're proving that on this Saturday when I did a four-hour test, this was the one that people liked and you have to do them over and over and over again.

So it may always just be informed by local conditions. In the South they like Duke's flour. In the North, they like Gold Medal. So -- I just confused mayonnaise and flour.

MR. WHITEHURST: You did. I noticed that.

MS. McARDLE: White Lily flour in the South.

So you want to push that down as far as possible. But when you want expert evidence, right, there's going to be a benefit to centralization of that. How do you resolve that tension?

MR. OVERHOLSER: Maybe I can give an example. We have to be careful to say expert evidence. Too often as the expertise says if you do this, that will happen as opposed to here's what happened.

The thing that stopped -- we're practitioners, so why aren't we making progress. The reason we weren't making progress is because the rules of spending streams did not permit us to do contingent contracting.

When those rules have been relaxed -- and now billions of dollars have recently been PFS enabled legislatively. We found that actually there's tons of data ready to be incorporated into contracts.

In doing that, you awaken a desire on the demand side, you awaken a desire among practitioners to use that evidence, because they will be rewarded for it. Whereas before their payments were contingent on those outcomes, their incentive was

to comply to the recipe.

So this is a lot easier than it seems. It has to do with relaxing constraints more than it has to do with telling folks what to do better.

MR. WHITEHURST: I think your question is a very insightful one. There is a real tension. It's playing out in policy in the research field between education, where I work, between having every school district have a first class research shop that does the local stuff and answers the really important questions versus having a federal entity that's supposed to be producing generalizable knowledge, and actually you need both.

So for the pre-K issue you talked about, people need to know in general across the nation what's working, what's not, to what effect. And then somebody who's delivering a local pre-K intervention needs to be looking very carefully at their own evidence to perfect and improve their program.

So those things seem to be intention. They are intention in terms of resources, but they should not be intention in terms of the way the world works.

SPEAKER: We need to harness more local firepower, not just within the government agencies, like school districts, but our universities need to support local decision making need to reward their faculty for addressing local problems, for examining issues of social impact, working in partnership with decision makers in nonprofits and agencies.

MS. McARDLE: So I really want to talk about this melts like ice cream metaphor, because I think that's so interesting. I've never heard that before and it's really good. It's something that I've written about. I think it's something everyone here has thought about pretty hard.

So I see three possibilities for why this happens. This is not to say they may all three be in play. One is that it's kind of like antibiotic resistance. As you do something, people figure out how to gain it. So you set up a program, some kind of cash

transfer program, and it works from the beginning and then people figure out how to get around it.

This is something you saw something you saw with Obamacare where (inaudible) we did where people figured out that they could go without health insurance for like -- they could buy health insurance for the first five months of the year, get all of their knee surgery and everything else, and then drop the insurance for the rest. That of course spikes the premiums for the people who are really sick.

The second possibility is just that you got scale problems. There was a wonderful sort of example of this from the Los Angeles Unified School District with their school lunch program where they did this elaborate testing of this beautiful new fresh food program. It was wonderful and then they roll it out.

It turns out that there's a real difference between having a craft team of highly dedicated people who are making this beautiful food to put out for people to sample and then what actually happens in the cafeteria lunchroom.

The cafeteria lunchroom like there's these salads and lettuce is kind of wilted and people complain and they're like, well, it's not unhealthy for you, it just looking brown and disgusting so eat it. The kids didn't eat it and they bought a bunch of Cheetos and you actually had the opposite of the intent you were intending.

The third one is just the statistical anomalies. You take a program that's got 65 kids in it, if you run 20 of those programs and look for statistical significance of the 95 percent confident, you're going to get one on average. One of those programs are going to be a random outlier result.

If you have publication bias, you're going to basically publish that one. The other 19 studies that didn't find anything don't get published and you think, wow, this works great, let's go do it, and then you try to scale it up and it doesn't work.

MR. OVERHOLSER: There's a fourth, which may overlap with your

framework. You mentioned Jim Manzi, this is the language that Jim Manzi uses and it's becoming much more prevalent.

I happen to be married to the dean of computer science at Northeastern, who's a machine learning expert, so I've learned a fair amount from her. It's called causal density.

So if you're in a world where there are many things that can cause a change in behavior, that runs head on with the way we do science. Science is a hold all else equal type of concept, and the problem -- this melts like ice cream thing, we had perfect science. We had no confirmation, no bias, we had N equals a hundred thousand every time and we had absolutely pristine RCT science. We had perfect replication, because it really had to do with maybe the color of the ink on the letter we sent was slightly different, but we had absolutely full fidelity to the model.

We had perfectly allorhythmic approaches of who the letters were sent to. It's all based on computers and data that was highly stable, and yet it stopped working.

So the reason it stopped working is because conditions in the marketplace changed, and there were many conditions that were causal to people's behavior that we could not control that changed.

I would argue that that's why most RCTs -- all the reasons you said are correct. There's a lot of reasons why things you learn that worked are not generalizable, but the fundamental reason is because the world changes so frequently.

So unless you up the cadence and just -- in essence you're just funding what's working now and then changing it and doing the next, doing the next, and kind of following things around. You're kind of doomed.

So writing a law that for 30 years says, everyone do the same thing and do this, are we surprised when we look at the effectiveness of those programs. They're

not very effective.

MS. McARDLE: I understand that you, sir, have to go catch a plane?

MR. OVERHOLSER: I do. Are you trying to get rid of me?

MS. McARDLE: No. I was given a note that said you need to catch a plane.

So whenever you have to leave, we will bid you final -- we should also add there's a fifth one, although I'm not sure it replicated, was this idea when you study people they change.

So they did a study in a factory. They raised the lights to see how productive people got, productivity went up. They lowered the level of lighting, it also went up. It turns out when people know they're being watched, they act differently than they do when they're -- so other thoughts on this --

MR. WHITEHURST: It's also the case that social science research unlike work in physics is about variance and averages. So when you do a randomized trial and Group A does significantly better than Group B, if you look at the distribution of outcomes, they overlap quite substantially. Sometimes they overlap for reasons that are systematic. Certain types of people react well in Condition A and they don't in Condition B, that's just the nature of it. But it does make the replication and extension very difficult, because the local experience is always the local experience. It's not the on-average experience. It's how did the lunch -- how did the wonderful lunch program work in this cafeteria.

That's a great challenge to social science thinking, because the people you want to be consuming the work are experiencing the message and they don't agree with it. It's working for my kid or the salad is delicious here.

We need then to think about what that means, to think about how to hedge bets to allow individual variation and implementation of program characteristics to

better match our local needs and local realities. We're a long way from being there.

I will point out the A/B comparisons done in business are trivially easy. The Capital One stuff is trivially easy. You can do it online, you can change the coloring of your lettering, you can get feedback pretty immediately. If you set up the right infrastructure it doesn't cost much. To try to do most of that stuff with real complex social programs is --

SPEAKER: RCTs are so hard. What we've done -- we used to do thousands of them at Capital One, piece of cake. We've done dozens of RCTs at Third Sector.

The hardest part of all is the ethics of recruiting the subjects, because you are denying care of people if you try to randomize, but that's -- it's as if you were trying to have a scientific laboratory and there's herds of buffaloes running through your laboratory constantly. That's very, very real.

I have not seen a single RCT that did not have implementation problems in the science that made it difficult to interpret the results.

SPEAKER: Yeah, despite all these problems, and I recognize all four or five that have been articulated, we've made tremendous progress.

MS. McARDLE: I'm not in any way denigrating that. I'm merely outlining the challenges.

SPEAKER: One of the Senators this morning said Thomas Edison tried a hundred lightbulbs before he got the one that works. This is complicated business. It didn't surprise us that 80 percent of randomized trials fail, same happens in medicine, same happens in other fields.

We need to do this work more cheaply, administrative (inaudible) helps a lot with that. We need to do more locally. We need to build more local capacity so they can make their own decisions when implementing programs that have worked elsewhere,

but we've made great progress in getting credible evidence produced. We need more takeup of evidence given that it's available.

MS. McARDLE: I'm going to see -- like I am getting nervous about your flight now. We should actually go to the audience and we should let Mr. Overholser go and catch --

MR. OVERHOLSER: As my parting point, that's -- thank you for saying that. That's what's happening. What is happening is that the information revolution is dropping by orders of magnitude the cost of getting very good interpretable data.

I was just complaining about RCTs, but the truth is that there are now -- you can in many cases, you can sense this data. You don't need to worry about drawing samples. You can actually get census data and you can -- and they're natural experiments all over the place. You can randomize in many cases much, much lower cost than before, and that is what's really happening here.

What's really happening here is that the abundance of information at low cost is kind of newly present and just the rest of our systems were not designed to embrace data.

So the reforms have to do with relaxing the constraints that keep us from embracing data and then incorporating that data in as a feedback loop. I would have to agree tremendously. Thank you very much. Sorry to leave.

(Applause)

MS. McARDLE: I think we have time for just a couple questions from the audience, so in the back, you, ma'am. Please state your name so that (inaudible).

MS. DELO: My name is Barbara Delo and I'm a nurse. We did talk about this a little bit, so it's kind of a follow-up question.

How do you ensure the integrity of the experts in politically homogeneous areas? I was in a nursing association in New York state and even obvious questions and

truths would not be heard. Should there be multiple views within the research communities or the community of experts?

MS. McARDLE: Thank you. Gentlemen.

MR. GAMORAN: Well, there's scientific process of peer review, blind peer review, for adjudicating the credibility of research findings, whether it's through an application for a grant or for a publication. I think this is one important reason why it is important for scientists to submit their research to journals and not just issue reports that are not reviewed.

When it comes to their judgment within decision makers, whether it's politics or other context, that's more difficult because the politicians may choose -- may engage in the tactical use of research, instead of the instrumental as I described, so it's more difficult.

MR. WHITEHURST: I think what's very important is I'm not -- a peer review doesn't get us very far. The very strongest journals have inter-rater reliability of like .3 and there's some famous research done on published articles submitted again, which to change the title and the author of a paper submitted ten years ago, you put it back through the process, and they're turned down. So peer review is weak.

I think what's very important in terms of high-stakes decision making is that there be entities that have rules for deciding the quality of evidence.

So it's not the consensus of three reviewers. It was some consensus about what the rules ought to be and then you believe the research or not or you question its quality based on those rules.

So we have that in some federal entities, not in others. We have that in nursing, What Works Clearinghouse in U.S. Department of Education was modeled on the Ontario Nursing Association's evidence guides.

So I think that's a critical thing that federal government can and should

do and is in the business of doing and does it for the FDA, for example, vetting the evidence and judging its quality and its relevance to the issues at hand.

MS. McARDLE: I think we have time for one more. I should take from this side of the room if anyone has a question because I feel like I'm -- no, okay, you, ma'am.

MS. LOUIS: Hi, my name is Gabriel Louis. I'm from the Saint Albans School of Public Service.

So I think it's universally -- we touched on this earlier that corporations in the private sector does a good job of using evidence.

How can the public sector work with the private sector through programs or whatnot to learn how to use evidence better and in a more effective way? Thank you.

MS. McARDLE: Gentlemen.

MR. WHITEHURST: Again the private sector where randomized trial and A/B trials are dominant is in industries where there's a very clearly agreed upon bottom line.

In Jim Manzi's book, there is an anecdote that always sticks with me. It's a guy he's interviewing who's at a casino in Las Vegas. Casino employee is saying there are three ways I can lose my job here.

I can engage in sexual harassment, I can use drugs, or I can run an advertising campaign for this casino in Southern California that's not delivered and it's a randomized trial.

So the consequences go so directly to the bottom line that everybody is motivated to do the work and to figure out what works and to increase at the margin earnings. It's much more difficult in the public -- in the public sector.

So one of the challenges there is if you want to know whether pre-K works, for what. Does it provide useful daycare for parents, does it help kids during the

year they're in preschool, does it make them nicer people when they're adults. There are 15 or 20 outcomes you can think about.

So the idea of doing the -- so it's hard to agree on the outcome, then it's much more difficult to do the intervention itself, because it involves politics, it involves convincing people.

I think there are things to be learned, but there are also some fundamental dissimilarities between business RCTs and social services RCTs that just part of the reality that has to be dealt with.

MR. GAMORAN: Yeah, I agree with everything you just said, Russ. I would add that some of what we can learn are about the value of administrative data, the appropriateness of using data that's already being collected so we can learn, which is something we don't do very well on the public side, and about the importance of doing lots and lots of trials in lots of lots of context, so we learn not just what works on average but what works for whom and under what circumstances.

MS. McARDLE: I think there was a -- David Cutler, the Harvard healthcare economist, once said to me that things were -- the outcomes are most dependent on the people you put into the system of healthcare and education, that's where the outcomes are hardest to measure and it's probably not a coincidence that that's where costs are growing in the most out-of-control fashion.

We can't measure the outputs, so we just keep cramming the inputs into the funnel and hope something comes out.

I want to thank our two great remaining panelists and also our absent panelist for a terrific and provocative discussion. I want to thank the audience for being good questioners and for listening with such alacrity to what our panelists were saying.

* * * * *

CERTIFICATE OF NOTARY PUBLIC

I, Carleton J. Anderson, III do hereby certify that the forgoing electronic file when originally transmitted was reduced to text at my direction; that said transcript is a true record of the proceedings therein referenced; that I am neither counsel for, related to, nor employed by any of the parties to the action in which these proceedings were taken; and, furthermore, that I am neither a relative or employee of any attorney or counsel employed by the parties hereto, nor financially or otherwise interested in the outcome of this action.

Carleton J. Anderson, III

(Signature and Seal on File)

Notary Public in and for the Commonwealth of Virginia

Commission No. 351998

Expires: November 30, 2020