

THE BROOKINGS INSTITUTION

SHOW ME THE EVIDENCE:
OBAMA'S FIGHT FOR RIGOR AND RESULTS IN SOCIAL POLICY

Washington, D.C.
Monday, December 1, 2014

Introduction:

JASON De PARLE
Senior Writer, *The New York Times*

Overview of Book:

RON HASKINS
Senior Fellow and Co-Director, Center on Children and Families
The Brookings Institution

Keynote Address:

PETER ORSZAG
Vice Chairman, Corporate and Investment Banking
Chairman, Financial Strategy and Solutions Group, Citigroup
Former Director, U.S. Office of Management and Budget

Panel:

JASON De PARLE
Senior Writer, *The New York Times*

JON BARON
President, Coalition for Evidence-Based Policy

MELODY BARNES
Chief Executive Officer, Melody Barnes Solutions, LLC
Former Director, White House Domestic Policy Council

ROBERT GORDON
Former Executive Associate Director, U.S. Office of Management and Budget

ROBERT SHEA
Principal, Global Public Sector, Grant Thornton
Former Associated Director for Administration and Government Performance,
U.S. Office of Management and Budget

GREG MARGOLIS
Co-Author, *Show Me the Evidence*
Law Student, University of Michigan

* * * * *

ANDERSON COURT REPORTING
706 Duke Street, Suite 100
Alexandria, VA 22314
Phone (703) 519-7180 Fax (703) 519-7190

P R O C E E D I N G S

MR. DEPARLE: Good morning and welcome to the Brookings Institution and to an exciting discussion of what Ron Haskins and Greg Margolis call the historic fight for rigor and results in social policy. I'm Jason De Parle. I'm an Emerson Fellow at the New America Foundation and a reporter at the *New York Times* and I'll be moderating today's discussion. The size of the audience here testifies to both the importance of the topic and to the definitive nature of the authors' accounts. Show me the Evidence examines six Obama administration initiatives that have attracted little public notice, but have the potential the authors argue to revolutionize domestic policy. They cover areas as diverse as teen pregnancy, workforce training, and the improvement of low income schools. What's new is that all six efforts direct most of their funding towards models shown by rigorous evaluation to produce positive effects. In addition they require grantees to build rigorous evaluations into their own programs. About 700 such studies are currently underway promising a trove of information to guide the policy makers of the future. The long-term goal sounds simple, to fund programs that work and stop funding programs that don't. Much of the discussion employs the jargon of social scientists who weigh the merits of randomized control trials against other types of evidence. What's at stake however isn't technical at all; it's a question that has loomed over American life since at least the New Deal, whether the power of the Federal Government can be harnessed to solve the nation's most pressing problems. FDR called for bold persistent experimentation. Haskins and Margolis want something in addition, evidence that the experimentation works.

On the surface it's hard to object. Few policy makers call themselves anti-evidence, but neither party in Congress has ever felt constrained by the findings of social science. It's unclear whether the new House and Senate will even sustain the Obama initiatives never mind build on them. And at the policy level there are complex

questions about what constitutes legitimate evidence. The authors plead for the superiority of random assignment experiments. Others may find their embrace of that methodology too limiting. Random assignment can reveal whether a program had an impact, but not why it had an impact, or whether it will have the same impact in a different place and time. An overly strict adherence to so-called proven models could end up stifling creativity, handcuffing the very efforts it hopes to enhance. It's worth noting that the great breakthroughs in American policy, Social Security, Medicare, the GI Bill, did not spring from the world of random assignment. It's also worth noting that the literature of random assignment is largely a literature of disappointment. It finds that most programs fail and those that succeed do so only modestly. Ron Haskins argues we should cherish modest successes and build on them. In an anti-government age however others may seize on disappointing data to abandon efforts rather than improve them.

The authors hope that the careful sifting of what works and what doesn't can win over a skeptical public and build support for the bold, persistent experiments of the future. The vision of the evidence-based movement they write is that the nation will have thousands of evidence-based social programs and that under their onslaught the nation's social problems will at last recede. To succeed even partly in that aspiration would be a great achievement indeed.

We have a distinguished group of people here to discuss these issues. Their biographies are in your packet, so I'll be brief. We'll start with an overview of the book by Ron Haskins who holds the Cabot Family Chair in Economic Studies here at Brookings, who has been working at the intersection of social science and public policy for nearly 30 years. After Ron's presentation we'll hear remarks from Peter Orszag, Vice Chairman at Citigroup who played an important role in launching the Obama initiatives as the President's first Director of OMB. A panel of other key policy makers will follow with plenty of time for provocative questions from the audience.

Please welcome Ron Haskins. (Applause)

ANDERSON COURT REPORTING
706 Duke Street, Suite 100
Alexandria, VA 22314
Phone (703) 519-7180 Fax (703) 519-7190

MR. HASKINS: Well, I'll thank you for that nice introduction; I appreciate that, Jason. Consistent with the season I have a lot to be thankful for this morning. First I'm thankful for the Grant Foundation who did in fact give me a grant to do this work. Without that we clearly could not have done it. I want to thank Jason, a wonderful reporter for the *New York Times* for agreeing to host and moderate the event. Thanks to Peter Orszag for agreeing to keynote despite his busy schedule and life in a different city and problems with airlines that emerged. You notice he practiced just in time delivery and arrived about one minute after the event began. I wanted to also thank a terrific panel of Washington insiders, all of whom played some role in the issues we're going to talk about in just a minute. And I especially want to thank Greg Margolis, my young friend. Unfortunately Greg -- fortunately graduated from University of North Carolina, but he is probably the only person in the history of the world to graduate from UNC and grow up to be a Duke fan which Jason is okay with, but (laughter) his parents are thoroughly ashamed of him, but only for that one thing. Everything else his parents, who are here today in the audience, are grateful that he turned out so well. And so am I because he actually ran the project, made all the big decisions, told me what to do, where to go, when to be there, and so forth and I just did what he said. And I think it came out pretty well, but I'll leave it to you. And then finally I want to thank the audience. Very big audience, bigger than our normal audience. Not necessarily the best possible way to end a nice four day Thanksgiving vacation, but I'm especially grateful to you for showing up.

I'm here today because our nation's social programs, especially the ones that affect children and families do not work. John Baron once wrote an article about 10 huge social programs that had good evaluations, scientific criteria and so forth, and 9 of them were shown to produce modest results or not work, and so it goes. Jim Manzi, a business owner who decided he was interested in social science wrote a magnificent book that I highly recommend to all of you called Uncontrolled in which he estimated that in both medical research and in research on social programs 80-90 percent of the

programs are shown not to work. We spend about a trillion dollars a year on these programs and we are not getting the returns. So we need to do something serious. And I think that the Obama administration from the moment that I heard about this initiative, I think that this holds very serious promise and I heard about it in 2009 and started looking into it, got the money from the Grant Foundation, fortunately met Greg and hired Greg, and we conducted this study because I think it's so important to trace exactly what happens in this initiative. This I think is the best chance we have to base our policy on evidence in ways that I'm about to describe to actually improve our programs and to have an impact on the problems, especially problems having to do with economic opportunity and poverty that we should be able to solve with the amount of money we spend and with the ingenuity that we have in this nation.

I begin by explain what the Obama based initiative is and then I want to discuss its prospects for success. Let's see if the Power Point comes up here. Oh, there it is; okay good. (Laughter) They changed the auditorium on us. So it's all right there. Okay. Good. All right. So let's see here. Now if I can figure out to change it. This is always a mistake to let me do Power Point. The arrow does not seem to be working here so. Now it will work. Okay. Great. Thank you very much. I always need that. That counts against my time now unfortunately.

So the Obama strategy had two essential parts, one has worked quite well to this point, the other one is just beginning to build and it will require further elaboration. The first one is with federal agencies. So that's trillion dollars I mentioned, we have grant programs galore. Some areas we have 10, 20, 30 grant programs, GAO it seems like every month tried to report about all the programs we have that are addressing the same problems. So the idea is to get federal agencies to be aggressive in making sure the programs are evaluated and putting pressure on grantees to change. And in fact the administration would like to severely cut back on grants that are determined by formula because you can't build competition of those grants, so they

would much rather have competitive grants. Just about all the programs I'm going to talk about here are competitive grants. So the administration has conducted and continues to conduct several initiatives to try to get the federal agencies to be more focused on data and I would say that with some success. I would point to HHS, and the Department of Education has great examples. And to a substantial degree both the Department of Labor and the Committee for National and Community Service all have done a lot more with evidence-based policy in the past. And then the second leg of the approach by the Obama administration is federal grant programs, new money where they could control how the money was distributed. And almost all the money, which amounts to around \$5 billion in six separate initiatives, almost all that money is awarded by competitive grant. Some of it expands each year so that new money can be awarded each year. So those are two of the essential parts of the initiative and I want to talk a little bit more about the competitive grant process, first of all to note that it is competitive, this is an essential element for any future administration that wants to put pressure on states and localities both private and nonprofit and for profit groups to produce evidence that they're having effects. And the best way to do that is with competitive grants because you can make sure that there using a good evidence-based program in the very beginning. That would be a huge step forward.

And the reason, the second point here is that the administration wants the maximum amount of money to go out to the states to spend on evidence-based programs. That also would be a huge advance for our social policies. So you don't want to do all of it, I'll talk about this more in just a minute, but most of the money should go out the door to programs that are using policy that we have very good reason to believe will be successful. And then, third and critical element of this is we should evaluate everything we do. The money should be in the budget to do it, it can no longer be an afterthought. We need to know if these programs work. If they don't we should change them. Many of you may recall the initial inauguration address, the President said that in

his administration they were going to expand things that work and they were going to reform or end things that don't. We've had a lot more of the former than the latter, but I think we've made good progress there as well. It's very difficult to end programs as we'll talk about in a few minutes.

So here is an essential part of the Obama approach, namely to define what they mean by evidence-based. I do not intend to bore you with the details of these, I just wanted you to see it. We'll make this Power Point available on the Brookings website so you can get it afterwards if you want to look at this carefully, but the point here is that either in the statute or in the regulations and especially the funding announcements where the availability of federal funds are made available, there is a precise, careful definition of what it means to be evidence-based. And so if you want the money show me the evidence. That's the administration's approach.

As I said there are six evidence-based initiatives. Here they are, one on teen pregnancy, one on home visiting which is a program that tries to help poor mother, usually single mothers, sometimes single mothers with a first birth, to learn more about child development, about breast feeding, about not drinking and smoking; a lot of very specific pieces of advice about how to proceed, and the how to interact with the child once the child was born and especially the importance of language. There are lots of different model programs. I think the last I looked 11 different programs were considered by HHS to be evidence-based. And in the teen pregnancy case there are now I believe 36 programs that are considered to be evidence-based. So about 75 percent of the money is distributed only to people who are using these evidence-based programs that have been determined by the administration, following a procedure I'll lay out in great detail in book, to be evidence-based. The other four people nominate their own evidence. So there are two different ways that the administration does it. And then investing in innovation which is an education end initiative, the Social Innovation Fund which pays for all sorts of different initiatives. And the fascinating thing about this is that

the money goes through intermediary organizations who in turn make decisions about who actually gets the money. So the federal government in that case does not actually decide. You sacrifice some control if you approach it this way, and in the first round of grants there was a little publicity caused by a former Brookings Senior Fellow who criticized the process and said it showed favoritism. This is the typical thing with federal grants. And then the two Labor Department grants that both are quite fascinating. This area is a really important area to illustrate that we can have evidence-based policy even in social policy areas where there isn't much evidence-based policy, where there aren't very many intervention programs. So I'll show you in a minute that there are a lot of areas I've already mentioned. Teen pregnancy had 36 evidence-based programs that could be adopted by grantees, and home visiting has at least 11. So there are a number of areas where we have lots of high quality programs and if we use those programs we'll be better off. In other areas not so much. So we need innovation, evaluation, we need to develop these new programs; it's critical to be able to do that.

As a result of these six initiatives there are now at least 700 programs around the country. Our estimate of home visiting is really an undercount because at the time that we did this we couldn't make a very accurate estimate. The money goes to the states, that's what the 48 stands for, but then they support local programs. So there are many more than 48. So imagine this now, we have 700 programs all around the country address these various problems that I just mentioned that are based -- at least most of the projects are using evidence-based models that have been shown in random assignment or quasi experimental designs to produce important impacts. So that in itself is a great advance. And then most of them are being evaluated, many of them by random assignment designs, so we're going to learn a lot about these programs contrary to many programs of this type that have been passed previously. Most grant programs we hardly have any evaluation if any. And yet now we have these 700 programs, even more that are taking place. Many of them are based on models that have been shown to

produce impacts and they're almost all of them being evaluated many by high quality evaluations so we can learn a lot.

I say here that the Obama evidence-based initiatives worked. Let me be cautious about this, it's worked so far. We don't know yet what the impacts are and that obviously is critical. So we're going to have to wait and I'll come back to that in a few minutes. But get this far and to initiate these six evidence-based initiatives here's what was required. First of all stellar leadership starting with Peter Orszag who was the head of OMB at the time, and Robert Gordon who is on the panel whom I dubbed the king of evidence-based policy in the Obama administration, Melody Barnes who was the head of Domestic Policy Council, and many others who both were familiar with social science, many of them intimately involved in conducting high quality studies themselves, and the administration in every corner had people who jumped on and said this is really important, we can do this, and they could and did. So leadership was important. Second, a relentless focus on evidence-based policy. This leads to kind of a culture in an organization I'm going to come back to in a few minutes. Third, clever and persistent legislative strategies for those of you who like to be regaled by tales of Washington; there's a lot about that in the book. We very carefully trace how the administration got these initiatives through the Congress and it was really a fascinating thing. Some of them are amazing. As always happens in legislative experience, very clever. And the administration was fully prepared to send whoever was necessary up the Hill, including people from the White House, including in one Rahm Emanuel to support and get Congress to pass the initiatives.

Competitive grants, already talked about; that's a crucial part. And all the initiatives involve at least some competitive grants, and most of them are entirely competitive grants. And then tier grants. I haven't mentioned this. This is a very good solution to a problem that we face and that the White House faced, but in the competition between wanting to have evidence-based programs and wanting to have innovation. If

it's innovating then presumably it has not yet been evaluated. Something new, something dramatic, and politicians love to talk about innovation. So there had to be some compromise and tiered structure is the answer so you can give most of the money - I mentioned the home visiting initiative and the teen pregnancy prevention initiative, 75 percent of the money has to go to proven programs, but that leaves 25 percent for new programs and new ideas and that is a very wise thing to do because I think of this as a string of new ideas and new development and then leading to random assignment or quasi random assignment experiments to lead to more programs and then decent review panels.

There are several signs of success. The initiatives first of all and most importantly have to survive the 2015 and 2017 transitions. Congress has already tried to end some of these initiatives so it's really important that the Congress get on the bandwagon here.

All right, let me conclude. The name of the last chapter in the book is "So Far, So Good", and that's my judgment at this point. The new model for federal funding with competitive grants is being implemented aggressively and well, but we still need to find out if we're actually going to have impacts and we're waiting to see. The results will start coming in, especially 2015 and 2016. Second that Congress must continue these initiatives. If you're interested in evidence-based policy and improving conditions for children and families I think it's really essential to do everything possible to make sure that these continue for at least another four or five years. Third, we're going to get the results from several hundred high quality evaluations of a whole range of social programs. Nothing like this has ever happened before, so we'll really learn a lot about whether these programs are actually working. And then finally we'll contribute to a growing culture of evidence in which everybody knows from the federal government and the agencies down to the local level that evidence is really important and evaluation is important; without those two things that you're not going to get federal dollars. You must

do these things to get federal dollars. And then finally the Obama evidence-based initiatives I think may be shown to be the most effective path for delivering on America's long standing promise to offer equality of opportunity to the nation's most disadvantaged children.

Thank you. (Applause)

MR. ORSZAG: Well, good morning, everyone. I'm delighted to be back at Brookings. As Ron had mentioned I had a little bit of airport drama this morning. And it's also the case that evaluation and evidence-based policy making involves to some degree a clash of culture between the analysts and the researchers and the program officers and people who are implementing. The confluence of those two things reminded me of a story I just recently read involving Brian Schmidt who is a Nobel Prize winning physicist from Australia whose grandmother lives in Fargo, North Dakota and she actually wanted to see the gold medal, the Nobel Prize, so he brought it to her. And it all went well until he tried to leave, and I'm just going to read to you his rendition of what happened. "It was in my laptop bag. It's made of gold, so it absorbs all the X-rays -- it's completely black. And they had never seen anything completely black before. They're like, sir, there's something in your bag. I said, yes, I think it's this box. They said what's in the box? So as one does when walking around with a Nobel Prize I said a large gold medal. So they opened it up and they said what's it made out of? I said gold. And they're like, Uh, who gave this to you? (Laughter) The King of Sweden. (Laughter) Why did he give this to you? Because I helped discover that the expansion rate of the universe was accelerating. At which point they became completely confused and lost their sense of humor. After I explained the whole thing to them their main question was, what are you doing in Fargo?" (Laughter)

Ron and his coauthor have written a terrific book that goes in detail into the efforts made in evidence-based policy making in the Obama administration, but I just want to give a little bit of context -- if I can now find my Power Point; there we go; okay;

perfect -- about why we need this. So here's a GAO study that went out and surveyed federal managers. Thirty seven percent of managers reported that an evaluation had been completed within the past five years of any -- and I put the word "any" in italics -- any program operation or project that they were involved in. So there was another 40 percent that reported that they didn't even know if an evaluation had been undertaken. That suggests if it were they weren't using it. So we have a fairly significant evidence gap in terms of doing lots of stuff that we don't really know whether it works or not. This has been noticed before. I put up a quotation here that I thought was just extremely well written and very trenchant. It happens to be from a blog that I wrote in 2009 (laughter) that too often evaluations -- or actually Robert Gordon wrote and I put my name on it -- that too often evaluations don't happen and that we need to do better. And so the point of this book is really to document how the Obama administration continues to try to do better even though there is still obviously a lot left to be done.

So as the book documents there has been -- and this is what I think might be surprising to many people, there's actually been a fair amount of progress here. A lot more is happening than is commonly appreciated and that reflects a whole variety of factors, one of which -- and by the way this is all documented in the book -- but one of which is just the stellar work of lots of people throughout the administration, many of whom are in the audience. I didn't name all of the people so don't be offended if I didn't put your name down, but a lot of other people are noticed and highlighted in the book and it really does require those unsung heroes to make progress on anything that's as hard as introducing a new culture of evaluating what you do because it's an awkward thing to undertake.

So why is this particularly important today? I'd say it's particularly important today because of two phenomenon. One is that we are in an era of hyper polarization in which it is very difficult to get anything at all done, but -- and I just -- because we're evidence-based put up the documentation for that, basically the middle

has disappeared in Congress, what is unique about the effort to drive more evidence-based policy making is there is at least some bipartisan support for it. So Ron Haskins who worked for a republican administration supports this type of effort, the Obama administration supports this type of effort. Another effort that I'll mention in a little bit, Moneyball for Government involves a bipartisan group of former policy makers and researchers. So this is one of the few areas where there is at least bipartisan support at least for the concept. Implementation becomes tricky sometimes. But at least for the concept progressives should be very invested in making sure that we are getting our money's worth out of government programs to make sure that we are achieving social or other objectives as efficiently as possible, and republicans want to make sure that our tax money is well spent, and those two forces come together more important today than ever before given that there's very little that the two parties agree on otherwise.

The second reason that this is really important today is both parties unfortunately in my opinion, but in any case both parties have agreed to discretionary spending caps over the next decade that I think are going to be extraordinarily challenging to meet and probably would be unfortunate if we did abide by them, but we have put ourselves on a path where spending constraints will become tighter and tighter. This is a chart from the Center on Budget and Policy Priorities which shows that if we do stick to the caps on non defense discretionary spending we are going to be at least relative to the economy well below more than a half a percent of GDP below the lowest level of such spending relative to GDP since 1962 when these data began. So in that context where there's tighter and tighter constraints on spending, and this is in the discretionary part but there will also be at least greater attention paid to the mandatory part of the budget also, it makes sense to make sure that you're cutting the stuff that isn't working very well and keeping the stuff, and maybe even expanding the stuff that is.

And we have an example of -- we have many examples, but I've pulled out one example, of how we can do this wrong. So youth opportunity grants were in

existence between 2000 and 2005. They were intended to increase education and employment skills including life skills, including through supporting services like mentoring and what have you. And there was an evaluation done of what the effects of youth opportunity grants were. And you can see it here, they basically reduced the number of out of school youth, they reduced the number of high school dropouts, they increased the percentage of kids with at least an eighth grade education, they increased labor force participation rates, and so on and so on and so on. If you go to the clearing house the Department of Labor has set up for evidence on how programs work this was rated as a fairly well undertaken evaluation, so the quality given the structure was decent. Here's the problem, this program ended in 2005, the evaluation was done a few years later, so we had already gotten rid of the program by the time we even figured out that it worked, okay. So if we're going to go through the next decade of constraining not defense discretionary and other parts of the budget flying blind and just cutting haplessly programs that may or may not be working we are not going to be serving the public very well.

One of the complaints about even the youth opportunity grant evaluation however is even though it was very well done it was not a randomized control trial. And the book deals with the tension over randomized control trials I think very well. There are many examples, some of which I was involved with in which randomized control trials helped to clarify what effects actually do occur, and that's the whole point is that only through a randomized control trial can you provide assurance that you're not just picking up a spurious correlation. And I agree with Jason that it doesn't necessarily tell you why the effect is happening, but at least you know the effect is real and not a spurious correlation. So multiple examples of where it does help. One that I was involved with involved low income saving for retirement. The thesis that was often put forward was that low and moderate income workers would not respond to a match on any contributions they made because they didn't have sufficient disposable income or for other factors. So

we went out and tested that at tax preparation time and it turned out that both low and moderate income workers were very responsive to matches, something that would have been very difficult to disentangle from the non randomized evidence. And that had been attempted to be studied in the non randomized evidence analysis, but just -- there wasn't very much clarity about it. Another example with real time policy applications involves the free application for federal student aid which historically has been a disaster zone in terms of the complexity. It historically was more complicated to fill that form out than a tax return and yet a lot of the items on the FAFSA form weren't actually used to figure out how much federal aid you were to receive. And so one randomized experiment that provided just assistance in pre populating the form and kind of filling it out was shown to be very effective and we're now moving toward dramatic simplification of that process in part because of this evidence. So it's not just that randomized control trials can help, it's also that occasionally at least policy makers will respond to the evidence. And Ron highlights in the book one of the other ways in which randomized control trials can matter a lot which is that the weight that the Congressional budget office will put on evidence from an RCT, from a randomized trial is higher than it would put on other type of evidence, and by the way that makes sense, but it also means that you're not only influencing the broader policy debate you may well be influencing the budget score of what a proposal does by undertaking an RCT if it shows an effective X or Y, it's going to feed into what the Congressional budget office analysis suggests.

That all having been said I think the conclusion which is also in the book is absolutely right, which is that if we only said RCTs are the gold standard and nothing else counts we're going to be making very big mistakes. And so I think this was very well said, claiming that RCTs are the best way to definitively establish causality, which they are, does not imply that other evidence has no value. And I'd also point out to Jason's comment which I think is right that we, you know, Social Security and Medicare were not predicated on randomized control trials -- a true statement, however I would say very

strongly it would be great if in particular Medicare had a bit more dedicated funding within it to be evaluating what works or what doesn't because if you take that institute of medicine or the Rand or other estimates of the share of healthcare spending that does not improve health outcomes and apply that only to Medicare and Medicaid you get the result that something like one and a half to two percent of GDP is spent each year in Medicare and Medicaid on healthcare services that don't improve health outcomes. Think about how much good you could do with \$250 billion a year or \$300 billion a year dedicated to other purposes. And that I think is the point, not to preclude things from happening in the first place, but to try to especially -- and I want to highlight this also, especially large scale things. There was no way someone was going to properly evaluate the delivery system reforms that were embodied in the Affordable Care Act based on randomized control trials because you were attempting to do bigger things than any randomized control trial would have studied. There was no way that you would be able to study Social Security reform or even Medicare using randomized control trials before they existed because you were going to be extrapolating to a system wide change in a way that was out of bounds of the RCT evidence. So when we're doing big things we have to recognize that you're not going to have randomized evidence in most cases and that's okay if you also build into the design of the program some ongoing feedback or evaluation loop so that as we progress through time we have better hope of developing the evidence base to make those programs work better.

I had mentioned that a lot is happening. I just want to highlight some of the things that are happening and then turn very briefly to the suggestions for how to move forward from here. First I think for the first time ever evaluation got its very own chapter in the Economic Report of the President in 2014. So for those of you who are CEA fans that's a pretty big deal. There's also the Coalition for Evidence-based Policy Making which has been conducting a series of low cost randomized control trials for policy interventions that can address various problems. They picked three winners from

the previous round, and I list them up there, and they're conducting a second request for proposals this month in which the number of winners will double to six. And so that's an example of private foundations and organizations getting into the game of trying to more directly measure what works and what doesn't. J-PAL North America which is led -- or this part of the effort is led by Amy Finkelstein who has done path breaking work in healthcare is focused on trying to do more evaluations of healthcare delivery system reforms. So in particular, this I think is striking, she notes that they did a study, 80 percent of the published results on drugs and devices in healthcare are from randomized control trials. With regard to delivery system reforms, so copayments, organization of hospitals, those sorts of things which often are more important in terms of driving the cost-quality trade off, 20 percent of published studies are from randomized control trials. And J-PAL North America is trying to fix that. I mentioned earlier Moneyball for Government which is a bipartisan group of policy analysts, former policy officials, academics, who are trying to increase awareness of evidence-based policy making. I see Michelle Jolin here is in the audience who is leading that effort. We have a new book out called Moneyball for Government that is available through a variety of different eBook type of channels. And again I think highlights the point that there is at least some pulse here. Despite a polarized Washington there is a pulse here in terms of promoting this agenda.

And then the question becomes what more do we need to do. Ron said so far, so good. So what should be the future? And here I crib liberally from a chapter in that Moneyball for Government book that Ron wrote with Robert Gordon and pick out some of the things that I think are particularly important. It would be a good thing for us to set a goal of dedicating one percent of discretionary funding to evaluation. So particularly for new programs, but I would apply it across the board. But any new program should have an evaluation piece built into it so that we have a better sense of again evaluating what works and what doesn't. We can promote more cross government

prizes for innovation, in evaluation, how you should undertake them, innovative ways of using the rise of big data and the decline in the costs of data processing in analytics in order to evaluate what works and what doesn't. We are starting to build what works clearing houses. This is a very good thing because it allows you to quickly go -- you don't have to spend hours Googling results; you can go to one place where the results have already been gathered, some sort of structure has been assigned to the studies that have been done. Best example of this is probably the Institute of Education Sciences, but there are now a growing number of these clearing houses and that is a positive development that should be encouraged.

Ron mentioned the Social Innovation Fund but there's a related effort of paying for success in social impact bonds which is taking hold at the local level including in New York and in Massachusetts. We could be promoting that even as Washington is polarized. More federal agencies could have evaluation officers, chief evaluation officers, like the Department of Labor did. And we could be promoting that function, that role within each agency so that how you conduct an evaluation and then how it can be applied is part of the day-to-day existence of the department. And then finally but perhaps most importantly none of this will work well unless -- especially as those budget constraints that I mentioned earlier become tighter -- we truly take things to the next level and start cutting the things that have been shown not to work and promoting the things that do. That is where the rubber hits the road, it's where it often gets -- everything -- everyone is in favor of evaluations until there is a consequence. That's where the consequence is and unless there is that consequence a lot of this is happy talk. With that step, and that's where it becomes very challenging, but with that step that's where you really start to change the dynamic of federal spending and what we're getting for our money.

So again I'm delighted to be at Brookings, thrilled to be here, and I think Ron is right that progress so far has been good, but we need to be doing yet more.

Thank you very much. (Applause)

MR. DEPARLE: Thank you, Ron, thank you, Peter. Ron, a question I'm dying to ask you, okay. You're famous for many reasons, you're famous in Washington above all for the role you played in the 1996 overhaul of the welfare system, replacing the old program of AFPC with a new program of time limits, work requirements, block grants to states. The Welfare-to-Work literature was extensive and employed random assignment and found very modest effects that did nothing to predict the caseload declines or work increases that resulted after the new law was passed. So if Ron Haskins the congressional aide had been listening to Ron Haskins the author of Show Me the Evidence he would not have advocated those reforms. So why do you feel so passionately then in randomized control trials?

MR. HASKINS: I would still have advocated for those reforms because in Washington, especially when you're working for a committee on the Hill, etiology trumps data every time. (Laughter) However, I do think there's actually a better answer. First of all welfare reform did just about everything that any study suggested might work and did it all at one time. So many states went from a system in which people really didn't really have serious requirements to a system in which they had work requirements, if they didn't meet the work requirements they could lose their benefit, and they had a five year time limit. And the states were given the block grant which gave them motivation to get as many people off the rolls as possible. And no experiment tested all those things at one time. Furthermore I think the import of your question is if random assignment shows that something will work sort of and then you implement it aggressively on a grand scale and it really works even better, I count that as a triumph for random assignment studies.

MR. DEPARLE: What is it in the Welfare-to-Work literature that you think even modestly predicted what happened afterwards?

MR. HASKINS: All the studied showed -- here's how I would summarize the studies and people out there can have their own opinion, but what the studies show

was that simple change the message that people have to work, if they don't they could lose their benefits, and then help them get a job. That's it, help them get a job. Not training, not education, help them get a job and devise a system to do that. And the states devised a good system to help them do a resume, practice interviews, give them the actual notices and send them out to do interviews. And a lot of the mothers turned out to be employable, and not only that they turned out to be good employees. And so within a four year period forty percent -- the work rate among never married mothers, the most disadvantaged group and the most likely to be on welfare, increased dramatically. So that was suggested by several experiments that the real key was job search and that's what the states did. They followed the results of the random assignment studies. They did other things too that were not random assignment at the time, like the five year plan.

MR. DEPARLE: Do you think other people could reasonable have had a different interpretation of that literature?

MR. HASKINS: I think they could say -- not really. I think they could say well the results are not that great, they're kind of small. You could say that kind of thing, but republicans want to say what works is work and the literature supported that claim.

MR. DEPARLE: Peter, you seem to be --

MR. ORSZAG: No, what I was going to say is again I come back to any time you think you're going to or you're intending to or the actual impact is to change social norms, you're going to have a very hard time lining up any randomized control trials beforehand with the actual effects. So, you know, the example I like to use is the adherence to the financial penalties for wearing your seat belt, not wearing your seat belt, and for speeding in most states are about the same, and the adherence to seat belt laws based on my personal experience is dramatically higher than it is to speeding laws, and that's because the social norm has changed. If I got in the car and the driver wasn't wearing a seat belt I might say something. If the driver is driving seven or eight miles or

ten miles an hour faster than the speed limit I wouldn't, and that's -- generally (laughter) -- and --

MR. DEPARLE: And say speed up. (Laughter)

MR. ORSZAG: Especially if I were coming from the airport and were late. And that is very hard to have picked up in, you know, a randomized control trial. And we see this, it comes up with the individual mandate, it comes up in a whole variety of important public policies. If you're trying to change the social norm then the randomized control trials or any other evidence is going to give you only a very limited base for extrapolating to what that new world would (inaudible). Randomized control trials do not do well with big structural change because we don't generally do big structural changes in a randomized way.

MR. DEPARLE: But isn't that another way of saying we should put all our faith in randomized control trials except when our hunches tell us to do something different?

MR. HASKINS: For me this is not a huge issue. It's been made to be a big issue and there are all kinds of fight in Washington, but I think generally they're pretty silly. There is a hierarchy of evidence. Randomized control trials at the top, quasi experimental designs next, pure correlations. So there are other ways to get at the truth in randomized control trials, but if you want to know if a program works and you want the most reliable answer, randomized control trials is what you should use. If you can't for whatever reason then let's use other evidence. It's not deprecating other evidence to say the random assignment is the best.

MR. ORSZAG: But again I want to -- you're not going to do anything very big in terms of changing policy. So you're not going to make large scale big changes based on a clean extrapolation from existing data in most cases because the whole point of a big change is we don't have what we want to get to. I mean the most recent example of this is it comes back to the Affordable Care Act. The Congressional

budget office did not put very much weight on the various different delivery system reforms and the general kind of atmosphere of moving towards value based payment as part of the health legislation and it was, you know, effectively given no credit. The results since then even outside of Part D, so in Part A and Part B have been dramatic. And I think a big part of it is you go out and talk to practitioners is it was changing the kind of environment in a way that none of the pre existing evidence spoke to. And so that's a great example. We did it anyway and that's another point. We did it anyway because to Ron's point the available evidence that we had suggested that it would help. There would be some effect, not as large as what I thought at the time it would turn out to be and what history I think is starting to show has actually been happening and that's okay. In other words I would not want to hold up policy interventions, especially ones that are involving systematic change, to this, you know, absolute adherence that you have pure evidence that speaks directly to all the effects because you're generally not going to have it. But there's -- I guess what I want to come back to is there's no excuse not to build evaluation into the program itself. So we can go do Welfare-to-Work, that's fine, but build evaluation in. We can do the Affordable Care Act, build evaluation in. So often you have to take a risk in doing big policy changes beyond what the evidence shows, but part of what you're doing should be to build the evidence so that you have better sense going forward of what you're doing makes sense.

MR. HASKINS: Which, Jason, by the way we did. There was a lot of money in the welfare reform legislation for evidence, including (inaudible) evidence. We gave some money to the Census Bureau to a whole new big study on whether the states were making progress and helping mothers work. And we gave HHS I think \$15 million that they could assign any way they wanted to address the issues that were raised by welfare reform. So we did exactly what Peter just.

MR. DEPARLE: One more question and we'll go to the audience. As the book documents extensively, the history of random assignment is to either show the

things -- mostly to show the things that don't work and then to show that even when work they don't work all that well. So what are the political risks -- let me ask Peter, start with Peter, that these 700 studies may just be laying a foundation for further reductions in government efforts to address social programs rather than an attempt to improve them? It seems to me that putting out all these studies assumes a certain amount of good will on the part of the political process to work towards improvement rather than eradication and good will would seem to be somewhat in short supply at the moment.

MR. ORSZAG: I think it's a good question, although I would come back to the youth opportunity grant as an example of evaluations don't always show negative effect. And in fact if we went in thinking that evaluations were just going to show negative effects pretty much across the board we should all just go home and give up because it means we're completely wasting our time and money. So cognizant of the risk of, you know, it's awkward to expose things that aren't working so well, I think the best way of trying to combat that is to really put emphasis on expanding where we see positive results. So if this were all just a game of cutting and it were all just, you know, any time a result was published you just have downside risk because the only thing that might happen is your program is either continues if it's shown to be neutral or positive, or it's killed if it's shown to be negative. That's not a very good structure to be encouraging more evaluations. But if you can credibly show that you're willing to expand the stuff that does work that changes the calculus a little bit. The risks are still there and, look, results can always be abused and misapplied and misinterpreted and used for political purposes.

MR. HASKINS: Irresponsible headlines in the *New York Times*.

MR. DEPARLE: Irresponsible headlines in the *New York Times*.

MR. ORSZAG: Especially that, especially in the *New York Times*. But I don't really see the alternative, especially because both parties, again unwisely in my opinion, but both parties have locked themselves into a very tight set of budget constraints, so things are going to be cut back.

MR. DEPARLE: Unless you have something urgent to say I'm going to go to the audience. Do you have something urgent to say?

MR. HASKINS: Just a very quick point. We have lots of evidence on many, many good interventions. If you look on Jon Baron's Coalition for Evidence-Based Policy website there are 19 intervention programs in a whole range of areas that have been shown to have impacts. So we don't -- and those are under very strict criteria. You could get to 30 or 35 or 40 pretty easily. So we do -- if they start with that, if the grant is given to people who are starting with an evidence-based program it will greatly increase the probability that it works and it will reduce that 80-90 percent failure rate that is -- from Jim Manzi's book. So there's a lot of hope out there.

MR. DEPARLE: Ma'am? I think there's a microphone coming right behind you. There you go.

MS. RADIN: Hi. Beryl Radin from Georgetown University Public Policy Program. Two questions for this panel. One is all of the discussion has really left out the reality of federalism. And as we have a system in which increasingly there's discretion given to states to really make the determinations that get the funds to a local or nonprofit level, how are you dealing with that? Because it does sound like we're really thinking about authority being in at the federal level to make those kinds of determinations.

The second question is whether you actually have thought about other definitions of evidence. You're really using a fairly traditional kind of positivist view of evidence as social scientists define it, but there are at least two other definitions that really might be useful to think about this stuff. One is what the kind of the general scientist looks at when they think about the scientific method. They're usually very modest, they're always opening themselves to continual evaluation and to determine whether something continues to be accurate or positive. And the second is I'm always intrigued by the fact that we're surrounded by lawyers and yet we're not using the

definition of evidence that lawyers use, which is information that can help you make your case. And when we start thinking about --

MR. DEPARLE: Okay, Ron.

MS. RADIN: Okay.

MR. HASKINS: Yeah. Okay, on federalism, that is a very good point. I'm a republican, republicans generally believe in state's rights. We often have big arguments with democrats about state's rights and this is definitely setting the tone and very specific requirements from the federal level. In this case I'm willing to accept it because I think it will improve the programs. But secondly I think in the future as we adopt what I called -- or create what I called an evidence-based culture the states will want to do it. I have been especially interested in seeing the reaction of many states to these initiatives. They want to do it. They wanted to use evidence-based policy because they want their programs to work too. So they are willingly following -- maybe not in every case, I can think of a couple, North Dakota is one, where they didn't, but generally they have. In the future it should get better because I think the states will agree with -- the federal government is doing exactly what the states want to do and they can show they can do better and they will.

MR. ORSZAG: And by the way I would just also identify, I mean the Social Innovation Fund is all about -- I mean it's a different level of federalism if you will, but devolving a lot of the decision making to other organizations that are not the federal government for evaluating the specifics of what works and what doesn't. So we're already walking down that path.

MR. DEPARLE: Sir?

MR. BRASS: I'm Clint Brass from the Congressional Research Service. So there is a lot of enthusiasm with more impact evaluations coming to bear to improve the evidence-based about efficacy effectiveness, but there's also been somewhat of a backlash because there is another issue out there that we call external validity where an

extrapolation gets made from impact evaluation findings in particular time and place to other times and places. And many of the definitional efforts for evidence sort of occurred with the tiered evidence initiatives have focused primarily on evidence that comes from impact evaluations. And perhaps related to Beryl's question, there are scholars out there that believe that evidence for external validity comes from more than just impact evaluations. Former member of Congress Brian Baird critiqued kind of this movement as making a plug and play audio device assumption that just because an intervention works in time and place X that it will work elsewhere. And that, Dr. Haskins, you've talked about creating strong incentives with the tiered evidence initiatives to get people to replicate and scale up, but less effort seems to have occurred in building kind of evidence architecture for external validity. Could you address those critics?

MR. HASKINS: I can give you a quick answer. First of all even if it's the case that there are other evidence to bear and that you have problems especially when you're trying to generalize other circumstances, the Obama administration is well covered in these initiatives. Why? Because they call it scale up in several initiatives and they're studying exactly what happens. Look at the MDRC evaluation of (inaudible). They're studying exactly what happens at sites all over the country that may be related to whether it will generalize to Des Moines on an experiment that was done in New York City or some such thing. So the administration, we're going to learn a lot about implementation of these programs and it may turn out that it's correct, that it doesn't work in Des Moines because they have different kinds of teachers, because the problem is deeper, because they don't have good -- I mean there could be a million reasons. There are many ways to err. And I think that the scale up evaluations in the Obama initiative will show when that happens and give very good clues about why it happens, and then if these last long enough to try to correct it.

MR. DEPARLE: If we could have quick questions maybe we can get a couple of more in. Ma'am, on the aisle right there.

SPEAKER: I'm surrounded by microphones.

MR. HASKINS: But you could talk out of both sides of your mouth.

(Laughter)

SPEAKER: Welcome to Washington.

SPEAKER: Ron, your book, which is fantastic, focuses a lot on changing the culture of the administration. I think your point a few minutes ago about etiology trumping data in Congress makes me wonder two or three things do both of you think would help to change the culture in Congress when it comes to evidence-based policy making?

MR. HASKINS: Oh, good question. Okay, first of all I have in the book a little chart that I often use and it shows all the factors that I watched on the Hill that influenced policy, the members of the committee, constituents, all kinds of things. And the teeny sliver for evidence that has an impact. And I think that is generally the case. My personal goal has always been to get some impact for evidence. Evidence should be at the table. It should not dominate. We have a democracy and there are always going to be politicians that say I don't care. I had a meeting with seven members of Congress who didn't like something I wrote about college prep programs, and I started laying out the evidence for them. They were all democrats and I said President Obama wants this and so forth and one of the members held up a letter from a constituent and said I know this thing works because I have letters from my constituents; they love it. And what do you say to that? So this is always going to be the case. Evidence -- if we can get it to the table and Congress knows what the evidence is and they still decide no, we're not going with the evidence, that's the way the system should work.

MR. ORSZAG: And I would just add the two things -- it's never going to be the panacea -- but the two things are one, you need to build out evidence so that it's there, and the second is all these various efforts that I highlighted, the book helps -- Moneyball for Government, what have you -- in highlighting the existence and what the

studies show because -- it's not always the case but it's sometimes the case that there's an information problem too, so. But no silver bullet.

MR. HASKINS: So, Andrea, I would say it's clear that we're making progress. As you probably know and many people in the is audience, Paul Ryan on a bipartisan basis with Patty Murray for god's sake, introduced legislation that they're trying to pass in the continuing resolution which may pass before Christmas, probably the 24th at 11:59, and they will establish a commission -- get this -- on evidence-based policy. A republican, a prominent republican, vice presidential candidate. So there's a lot to be said for progress.

MR. DEPARLE: I want to squeeze in one last question and leave plenty of time to buy the book, right? (Laughter) Do we have a quick question?

MR. HASKINS: If we have another question or not we need plenty of time to buy the book.

MR. DEPARLE: An urgent question here. Yes, please. Quick question.

MR. OAKLEY: Sure. Tom Oakley, National Economist Club. One tool that has come out a lot and I'd love to get your opinion on is social impact bonds and how those work, and especially Peter being from Wall Street. I'd love to get your take on that.

MR. ORSZAG: Well, the thing about a social impact bond is that it's not actually a bond. So let's just be clear that the nomenclature is wrong. The short answer is we don't know yet because the programs are still very young. Even the New York program on recidivism, you know, it's still too early in its development let alone the stuff in Massachusetts. So I would say the concept is promising in terms of basically -- I mean it's very simple, you pay when something works period. And so all the fancy financial whatever is kind of misleading. It's just basically paying for outcomes, but it's too early to know.

MR. HASKINS: I would add just one thing and that is this is another example of when you start to change the culture you get these kind of innovative ideas. Evidence is really prominent now. I hope it continues and grows.

MR. DEPARLE: Thank you. (Applause) And with a quick switch of the lavaliers we'll have our panel come up.

(Recess)

MR. DEPARLE: While everyone's getting settled, let me go ahead and do some brief introductions. Again, you have the full bios in your packets, so I'll be very brief. Jon Baron is President of the Coalition for Evidence Based Policy, a non-profit, non-partisan organization. They were founded in 2001. He plays an important role, both as an advisor on the Obama initiatives and as an advocate for them. Melody Barnes is CEO of Melody Barnes Solutions, a domestic strategy firm and Vice Provost at New York University, as the first director of President Obama's Domestic Policy Council, she played an essential role in bringing the evidence based initiatives to light. Robert Gordon spent four years as a senior official at the Office of Management and Budget, where he served as the champion of the Obama initiatives and a skillful negotiator on their behalf. If you read the book, you'll find he's the unsung hero of the narrative. Robert Shea is a principal at Grant Thornton. He's served as a senior member of the Office of Management and Budget during the administration of George W. Bush and is also a former Senate staffer with an expertise in issues of government performance. Greg Margolis is the co-author of "Show Me the Evidence" and a former analyst at The Brookings Center on Children and Families, where he served for three years. I'm not sure exactly what this says, in front of an audience of Ph.D.'s, but I will note that four of our panelists are lawyers and the fifth is a law student. (laughter) We'll hear from each of them for seven minutes. I'll ask a few questions and then we'll have more questions from the audience. I think we're going to start with John.

MR. BARON: Thank you. The evidence based initiatives that Ron and

Greg discussed in their book, I'd suggest are not just the desirable policy change. They may be essential if our country's going to make progress against problems like poverty, educational failure and other social problems. Here's why. If the history of randomized control trials or rigorous evaluations generally has taught us anything, it's that many widely held beliefs about what works in medicine, in social spending and other fields are not accurate. Ron touched on this briefly, but let me give you the numbers. In business, Jim Mancy reports that Google and Microsoft have conducted 13,000 randomized control trials in the last two years, of different business strategies, 80 to 90 percent of which found no significant effect, in Mancy's words, innovative ideas rarely work. Similarly, in reviews in four different areas of medicine have found that 50 to 80 percent of promising findings from phase 2 studies, which are usually small clinical trials, sometimes randomized, are overturned in larger more definitive randomized evaluations in phase 3 which are required for FDA approval. So even in this case, where you have promising initial evidence, in most cases, the intervention in medicine is still found not to work. In education policy, since 2002, the Department of Education's Institute of Education Sciences has sponsored large randomized trials of approximately 90 educational interventions, everything from school choice programs to different curricula, teacher training programs and so on. Eighty to 90 percent of those initiatives were found to produce weak or no positive effects compared to what schools were doing anyway. A similar pattern occurs in other areas of social spending, like employment and training and in crime prevention. In short, an absolutely clear lesson from past rigorous evaluations is the need for humility in our beliefs about what works and the fundamental importance of rigorous testing if our efforts in medicine, social spending and other fields are going to improve the human condition.

Now in medicine, we've seen amazing progress over the last 50 years, even though as I just mentioned, only a small fraction of rigorously tested interventions are found to produce meaningful effects. How can that be? The answer I'd suggest is

that the medical system -- in medicine the system is set up to reward evidence of effectiveness. Since 1962, the Food and Drug Administration has required two randomized control trials showing clinically meaningful effects before it will allow a new drug or medical device onto the market, which creates an enormous incentive for the development of strong experimental evidence and the mechanism to put proven replicated interventions into widespread use. Findings from those studies and others sponsored by the NIH and others, have profoundly improved life and health in America. Such trials have definitely established the effectiveness of vaccines for measles and hepatitis B, treatments for hypertension and high cholesterol, which have helped cut the incidents of coronary heart disease and stroke in this country by more than 50 percent since the 1960's, and of cancer treatments that have dramatically improved survival rates from leukemia, Hodgkin's disease and many other cancers.

Now social spending is similar to medicine in that while most rigorous evaluations produce disappointing results, there are some clear examples of demonstrated effectiveness. New York City Small Schools of Choice Initiative is a recent example. A large randomized trial, more than 20,000 students in high poverty communities, randomly assigned in the study -- there was a six year follow up by MBRC, which found that the program increased students likelihood of graduating from high school and enrolling in college in the next year, by 15 to 20 percent compared to control schools, control -- students in the control group, and this was achieved at a lower cost than large traditional public high schools. As my mother would say, what's not to like about this? However in contrast to medicine, such rare and important evidence of effectiveness is generally not rewarded in the current system of social spending.

Most large federal social programs like Title I at the Department of Education, Headstart at HHS, the Workforce Investment Act at DOL, are essentially set up as faucets in which each program allocates large streams of funding to state and local organizations, sometimes through a funding formula or a competition, but where

evidence about which activities are effective or not effective plays very little role in what gets funded. So interventions with strong evidence of large impact, like small schools of choice, are not prioritized for greater funding under these faucets and may never expand. And in New York, under a new mayor, the opposite is actually happening. The small schools approach is not being continued with federal funds or other funds. And so it goes in social spending. Interventions come and go. There's no systematic mechanism in place to identify what works and put it into widespread use. Little progress is made.

The six evidence based programs discussed in Ron and Greg's book are a fundamental departure from the faucet approach. They are the first social programs in history to recognize what is demonstrably true, that many interventions don't work. So each program in its own area, home visitation, teen pregnancy prevention, K-12 education and so on, funds both evidence building, innovation coupled with rigorous evaluations, plus the scale often the top tier of interventions that do have strong evidence of effectiveness. In all of the program tiers, scientific evidence is the central funding principal. Since 2001, our organization has worked closely with both the Bush administration and the Obama administrations to advance evidence based reforms, and with Congress. Our experience is that evidence based policy offers a place of common ground between the two parties that transcends endless partisan debate about more government versus less. The evidence based initiatives described in the book are an important start and we've encouraged and pushed them every step of the way but what would fundamentally shift the social spending landscape, is if similar evidence based funding criteria were incorporated into billion dollar federal programs including entitlement programs, doing so would create a powerful new incentive for the development, rigorous evaluation and if effective, dissemination of new programs, strategies and models. It would bring to social spending the same kind of similar dynamic for evidence driven improvement that has produced amazing progress in medicine over the past half century.

MR. DEPARLE: Melody?

MS. BARNES: First of all I want to congratulate Ron and Greg. It seems like it was not too long ago that we were having lunch here talking about this book so congratulations on having it published. And what I want to do is talk a little bit about my experience both in the Obama administration, but also my experience now since leaving the administration with a piece of work that was germinated during our time there, Michelle Joel's, and a lot of roads lead back to Michelle Joel and was actually instrumental in this as well. And some, and connect a few dots and talk about a couple of things that I think are critical, if we're going to make sure that this work is sticky, because we're talking about how do we make sure that this lasts.

Back in 2010, the President told us as an administration, his staff, that he wanted to look at, as he said in his inaugural address, why certain communities are successful in taking on big social changes. What makes them successful? And in doing that, he asked us to look at several different communities and we formed a council for community solutions that was chaired by Patty Stoneseifer, Vice Chaired by Michelle Joel and had business people, foundation leaders, social scientists, grass roots leaders and others as a part of it. There are several different things that came out of that.

One, the council determined that they wanted to look at the kind of change that was taking place in community that represented about a 10 percent change off of whatever the norm was in the community, that the community wanted to change. So you take the well-known STRIVE work, the cradle to career work that STRIVE is doing in Kentucky and in Ohio. You take the now, I think, more well-known work that was taking place in Milwaukee on teen pregnancy. Also the work in Philadelphia on high school drop outs that project U-turn was doing. Those are some of the examples that came out of the work that the council was doing. And the President also asked them, in addition to identifying that work, tell us why those communities are successful. That led the council to lift up this strategy or model that was known as collective impact, that includes five different elements, one of them being the use of data in evaluation, as being

extremely important. It also led the council to think about and specifically target a big challenge that we were looking at more broadly, that of what commonly called disconnected youth, but now becoming more commonly known as opportunity youth -- 16 to 24 year olds, disconnected from education and employment, and how do you build pathways to bring those young people on -- back into the social norm and to connect them to education and to employment. That collected impact model has gained more prominence now and the use of data, and the use of the evaluation in communities is being spurred by this work, I'll touch on that a little bit more in a second. We also have recommendations that were made to the administration that are starting to be carried out, including the fact that the most recent round of grants made by the social innovation fund have a collective impact base to them, and we also -- that also led to the Aspen Forum for Community Solutions and the Opportunity Youth Incentive Fund that Michelle and I, and I think some others in the room, have been working on, that I serve as the chair of.

The Opportunity Youth Incentive Fund has as its goals one, to use collective impact, to study collective impact as a model for change, but also to study collective impact as a model for building out those pathways that we were just talking about. To date, we have funded about 21 communities, starting with development grants and going on to implementation grants, and we've also brought on a third party evaluator, OMB, to look at several different things as we are starting to do this work. One, to look at the Aspen Forum and jobs for the future as intermediaries, two, to look at this learning community of collective impact sites, these 21 communities, mostly urban, some rural, some tribal, and to determine how successful they're being, and then finally, to determine collective impact as a model for the kind of change that we're talking about. Does it really work or, as we talked about before, is it some hunch that people have that when you're dealing with cross sector collaboratives, something better might be taking place than we thought before?

We're right now in a design phase or completing the design phase with

ANDERSON COURT REPORTING
706 Duke Street, Suite 100
Alexandria, VA 22314
Phone (703) 519-7180 Fax (703) 519-7190

the evaluator and about to move on to the implementation phase. I'm blowing through this very quickly because we have seven minutes, and there's a lovely young woman that's flashing numbers at me. (laughter) But I'm more than happy to go into more detail. But I just give all that to you as background to tell you what some of the things are that we found as a result of doing that work, again, germinated because of the President's desire to identify what works, but also now, having had the experience of working with communities on the ground, as they start to work through efforts to build policy off of this, to build out these pathways, to deal with data, to deal with evidence, and how that can hopefully become more sticky.

One thing I believe is critical is that this cannot be a D.C. centric conversation. And too often, while this is very very important and we have to think about ways that we're going to, as Andrea said, encourage Congress to take on evidence, even when constituents and others are saying no, I've got this one example of cute children doing something important, and we have to move beyond that and say, is this really working, because we want those cute children to be well educated, to have good health outcomes, et cetera. So we have to move beyond just educating Congress, but also to work in communities, because right now we know there's a great -- there is friction that occurs, and in some of the communities that we're working with, they've embraced a data driven culture. Some of them have much more sophisticated data collection systems, and that culture that has already existed, for many of them, this is something that's brand new, and they're trying to figure it out, but we've heard the concerns about lack of creativity. We've heard the concerns about, how is this affecting those that this is supposed to help. We've also heard concerns about cost. And we know, when the rubber hits the road, when those two things meet, that friction starts to occur, that often the good policy takes the hit. And you can look at some of the work on education and common core. You can look at the work that we've done and Peter was very instrumental in with regards to Let's Move and some of the work on childhood obesity to

see where some of those things are getting a backlash. The quantitative and the qualitative experiences need to be brought together.

Secondly, we talked about innovation and continuous improvement. And I would say the work that we've done with learning communities has been very very important. What we often found was that two communities within 20 miles or 200 miles were doing the same kind of work and had no idea that the other existed. So how can we bring and leapfrog over some of the mistakes and bring into play what the successes are and what's being evaluated and what works, in a faster and more beneficial way for the beneficiaries? I would also say, I think that Ron and Greg have it right, with regard to the RCT and not letting the perfect be the enemy of the good, and making sure that as we tout the favorables of RCT, that we're also looking at other forms of evaluation as we're trying to move the ball forward. So I've got the hook. I'm more than happy to talk about this a little bit more later, but I just wanted to give you some sense of some of the experiences that we've had.

MR. DEPARLE: Thank you. Robert Gordon.

MR. GORDON: Hi everyone. I thought I'd talk a little bit about how this work came about and then the biggest challenges that I think it faces moving forward. I think when you see the chart in Ron's block of the six programs, it all looks very tidy and orderly and well-conceived and I think you know it is a very rationalistic technocratic approach to policy that interestingly came about in a very -- in a bottom, within the government, came about in a bottom up happenstance, cobbled together way that Ron accounts for -- Ron and Greg account for beautifully, but is worth underscoring and I think it came about because of the quality of people in the federal government, and in particular, the quality of the civil servants in our government that were a good resource that is not well understood, I think in politics, and not adequately respected. So I'll just do a few very quick stories. And the home visiting program ultimately takes shape -- the shape it ultimately takes in law bears a lot of resemblance to ideas that two people, Jack

Smalligan and (inaudible) put down in a briefing paper for Peter and the OMB political team coming in, and got a lot of things right, from the beginning, that I3, Investment Innovation Program, you know, we set up, we want to have these evidence tiers and that's great, but then it had to be a grant program, and the political people coming in, most of them had never run a grant program. And so Margaret Anderson, who's here, actually figured out how to do that with these tiers, and made it work.

The teen pregnancy prevention program -- we saw the evidence tiers as a useful -- it was a good idea for evidence. It was also a good idea for dealing with political challenges on the left, on the right. And on the hill, interestingly, you know, David Obee's lead approach staffer -- you don't think of David Obee as Mr. Evidence. But his lead approach staffer was Cheryl Smith, and she said oh, this is a good idea. We're going to -- let's get this done. And we got it done, and then it had to be implemented and Naomi Goldstein, who is also here, led the way in figuring out how to do that. So I could go on with stories like this.

I think that there's a kind of -- to use a trendy term, there's a distributed intelligence in the federal government that has built up over years and survives from administration to administration. And there's a level of trust between staff and a willingness to work hard. So last story that just comes to mind is like the -- and Ron tells the story -- well, the community college program which Melody and I worked really hard on, was a kind of -- went on this crazy legislative path that was changing committees, changing agencies, sort of struggling to figure out how to get it done. And at a key moment, had to figure out where to put it and the Labor Department branch chief at OMB, Melissa Bomberger figured it out on a Saturday with us calling her and harassing her. And so, I tell these stories probably because they're fun, you can read more about them in Ron's book, but partly because I worry deeply that the demonization of the federal government is going to make this kind of thing impossible and that these are not -- this is not partisan work. This is just trying to make things work better and trying to get more

bang for the taxpayer buck. And whether you want a big government or a small government, you should want a sensible smart government. And ultimately, you know, the people -- the political people who come in are going to have ideas. This was one of them. And they're going to -- some will be good and some will be bad, but they will be implemented better or worse and that will make all the difference by people who are not political, and so I do think this story is ultimately about the intelligence of a lot of people who you never heard of an you're never going to hear about and it's just -- my hope would be that we can do a better job honoring all that work.

Second thing I'll say is, what do I worry about moving forward. I don't actually, you know, this is just picking up on Brad and Peter's -- I don't actually worry about evidence based policy sort of totalizing and driving out everything else. These approaches are not appropriate for everything. And they're also not really in any risk of taking over, like I think we should be so lucky as to have that problem. (laughter)

But you know, the I3 is, I think, a tenth of a percent left, something like that, of the Department of Education's budget. The home visiting and teen pregnancy, they truly are a rounding error within the amount that we spend on that healthcare. So the takeover is not the worry. I do think the worry, and this goes to the point about scaling, is that we continue to get the details right. And that is really hard, so that -- I'll throw out two quick issues -- one, is one that John always stresses, is that you know, are we being careful in defining what works, and are we actually demanding real evidence that something works. So you can have different home visiting programs. One can affect peoples' levels of work. One can affect a child's birth weight. One can affect the number of referrals that are made to programs. Those are all effects and this can all be called successes I think under federal law, but we probably ought to weigh them differently, and we don't always. And that's partly a technocratic challenge, partly a political challenge.

Second example is this scaling question. I do think that the issue of, does a program that works 25 years ago in city A -- is it right for city B, and how do we make

sure that we are adjusting as we need to, to change circumstance, allowing for those adjustments and learning as we go -- that's really important questions people are right to raise and people are right to be skeptical that if we are overly cookie cutter, that this is going to work. But having said all that, to me it is undoubtedly the right framework as to say, what do we know, realizing that knowledge comes in lots of different forms, and then, given how the world is changing, how do we need to adjust? And so I think that's all that we're trying to do here. And it's a big step forward and its success ultimately will depend not on any of us but on this kind of thousand decision makers all over the place, doing the best they can and doing well.

MR. DEPARLE: Robert, thank you. Yes, when you read the book, you'll, just to underscore what Robert said, there's lots of intriguing cameos by otherwise anonymous members of the federal bureaucracy doing very interesting and creative things to keep us moving forward. Robert Gordon -- or Robert Shea, excuse me.

MR. SHEA: Robert Shea. Not at all. They get us confused sometimes. I want to add my congratulations to Ron and Greg on a great book. It's as close to a page turner as we'll get in our business. (laughter) You know the quest to get evidence on which to base decisions in government's a long one. In my experience of about 20 years working on this, it began with the Government Performance and Results Act, which sought to get agencies to focus on outcomes instead of activities. In the past, they had been really counting the things they had been doing, not what they had been accomplishing. And though agencies made significant progress as a result of this important law, they have for the first time, for instance, long term plans with outcome goals. They didn't produce a lot of information that facilitated decision making at the program or any level for that matter. And Congress was an afterthought in their planning as well. During the Bush administration, when the management agenda was being designed, it was designed in part to address the shortcomings of the Government Performance and Results Act. Agencies were asked to assess and report the

performance of their programs, because that's where OMB and Congress made funding and other decisions. Whatever you thought of the tool, the program assessment rating tool, the need to develop standards to judge the adequacy of evidence of program effectiveness is a major reason we're here today. At the time, we were asking agencies whether their programs had been subjected to independent rigorous evaluations. For many agencies and programs, so long as the evaluation was conducted by an independent entity, GAO, Inspector General, others, the evaluation results were deemed to be adequate. Then came my friend Jon Baron, of the Coalition for Evidence-Based Policy, probably the most effective public policy advocate I've ever known, probably helps that he's pedaling motherhood and apple pie, or so it would seem. John advocated a more robust assessment of evaluation methodologies. Were they rigorous enough for policy makers and other stakeholders to rely on for results? When we modified our standards for what constitutes rigorous evidence, the uproar was surprising, I think you'll agree. Whether it was because people didn't trust what we do with the evidence, or feared we were putting a huge swath of evaluators out of business, it's hard to say, but there was palpable resistance to what we were trying to do. One forum, in which we made a targeted effort to insert evidence, was the Academic Competitiveness Council. Led by the Department of Education, OMB was a major partner in the effort, not only to inventory science, technology, engineering and math programs, but to assess what evidence we had of those programs' effectiveness. Precious few of the programs had been evaluated at all, much less with standards any of you would consider rigorous, and the resistance to any effort to evaluate the programs was strong, getting so many agencies and programs to agree on goals and evaluation methodology standards is an enormous challenge for government and one that this initiative will continue to face. That's why where we are today is so significant in my opinion. The efforts ran two administrations and appear to be thriving, enjoying bipartisan support in the Congress. Ron mentioned earlier the recently introduced legislation by Congressman Paul Ryan

and Senator Patty Murray. This bill would inventory data in the custody of federal agencies, and establish policies that would make it easier for evaluators to access this data for the purposes of conducting evaluations. I think it has enormous promise. It's likely to be enacted in the next several weeks. I don't know how long this drive for evidence will last. We still have a long way to go before we have what I would call a sufficient basis on which to make a lot of decisions about many of the programs aimed at our nation's biggest challenges. We owe a great deal to the major champions of the evidence agenda, many of whom we've talked about today, without whom we couldn't have gotten even this far, but I do want to give special recognition to OMB's Cathy Stack. She's a tireless, effective advocate for rigorous evidence and she is without peer in her ability to navigate the bureaucracy to accomplish important public policy objectives, whatever they may be. Lucky for those of us who crave more rigorous evidence, she's dedicated much of her career to advancing the cause.

But no one can accomplish this alone. You know, I'm reading Walter Isaacson's "The Innovators" right now, about the invention of many of the modern technologies that drive our lives today. There are a lot of management insights that Isaacson teases out of these innovations, but one I think is particularly applicable to the drive for more rigorous evidence of program effectiveness -- "the main lesson to draw from the birth of computers", Isaacson writes, "is that innovation is usually a group effort, involving collaboration between visionaries and engineers, and that creativity comes from drawing on many sources. Only in story books do inventions come like a thunderbolt or a light bulb popping out of the head of a lone individual in a basement or garret or garage." So this has been a group effort, but with some truly exceptional committed leaders. It will take their long perseverance to realize the many benefits of the evidence agenda.

MR. DEPARLE: Thank you Robert. Greg.

MR. MARGOLIS: All right. It's an honor to be back here at Brookings today, and obviously to be on the panel and at the event with such distinguished people.

What I want to do is explore a pretty simple question, that nonetheless is fundamental to our whole endeavor here, and that is, what do we mean when we say, evidence-based initiative. And then, as a corollary, I want to explore a little bit how the Obama administration was successful in translating the idea of an evidence-based initiative into concrete initiatives that use rigorous social science research systematically but recognize the inherent limitations. And you have the evidence itself, the political reality and other practical considerations. To set the stage, one thing that I think is important to keep in mind, when we talk about evidence-based initiatives is that we should really be thinking about them as occupying an evidence spectrum, rather than being a binary. In politics especially, we don't like a lot of nuance. We like things to be a clear choice. You know, are you going to cut my taxes, or are you going to raise them? And I think the temptation is to treat evidence-based initiatives as fitting into this binary initiative -- an initiative is either evidence-based or it's not. But this of course, is an oversimplification that masks the underlying challenge that we should really be concerned with, and that's how to translate evidence based initiative from concept into practice, crafting a social policy initiative that has, to borrow a phrase from Peter Orszag, "rigorous social science and research and evaluation baked into its DNA." All of the Obama initiatives fall somewhere onto an evidence spectrum. They have different levels of commitment to incorporating evidence standards into their DNA that reflect the nature of the evidence base behind the various interventions in the issue are, and the political realities inherent in moving the initiative from conception through implementation. But as we move forward, it's helpful to keep the spectrum in mind, that the goal should be to continually re-evaluate the best way to systematically apply rigorous research to social policy initiatives rather than coming up with the one size fits all label of evidence based or not evidence based, to slap on an initiative to make things easier. The approach to social policy making should be a commitment to rigorous evidence and all of our social policy making should be evidence-based and then where it falls on the spectrum will of course depend upon other things, on

the quality and quantity of the evidence. So with that context, I want to highlight four things that I think were fundamental to the success of the Obama administration's approach.

The first one is defining the parameters of evidence and what works. Evidence and what works are words that get thrown around a lot and mean different things depending on who's saying them and what the forum is. For example, one Congressman at a hearing on the appropriations bill that funded the teen pregnancy prevention initiative, in talking about abstinence only education explained, it works. Parents prefer it two to one over other sex education type programs. (laughter) Well, for his audience, what works is what parents prefer. But of course, for the Obama administration's initiatives, we're talking about social science evidence again, on a spectrum of rigor. These initiatives take, as a given, that there's reliable social science research that can help give a meaningful and empirical lens to address social problems. What's important to note about what the Obama administration did in crafting these initiatives is that there were people at the highest levels of the administration that understood what rigorous evidence meant and understood the importance of setting meaningful evidence standards for choosing which program models to fund. And that there was a commitment to upholding those high standards even in the face of political or advocacy pressure, which leads to the second point which is that the top officials in the administration set the tone for these initiatives, that the goal was to think systematically about what really does work and how you would build that into the initiative, even though as Robert Gordon mentioned, the administration didn't necessarily come in with a cohesive or coherent strategy for doing evidence based policy per se, top administration officials did set the tone that rigor and results, as we name our book, were to be a fundamental part of the administration's approach. Rigorous evidence as the North Star, if you will, helped define the debate and provide the vocabulary and context for discussing each of these initiatives from conception through implementation.

Then, third, the administration wisely reframed the initiatives in terms of outcomes and issue areas, rather than ideology or specific intervention program models, and then used evidence as the lens to examine the particular social program or issue area. The key insight that the administration had was that rather than merely promoting particular program models per se, for example, nurse family partnership, or a particular approach that the evidence might suggest works, for example, comprehensive sex education, you could take an issue area and focus on understanding the evidence base and how you might make meaningful funding decisions based on what that evidence base was. The administration realized you could in a sense depoliticize the issues and reframe them around the evidence and outcomes, hence, teen pregnancy prevention initiative, rather than the old abstinence only versus comprehensive sex education dichotomy. So the administration developed a thematic approach to evidence based policy. Once they found an area that they wanted to do an initiative in, they applied this evidence based thinking into the issue in forming the policy. And you can see this approach clearly in both home visiting and I3.

For home visiting, initially the idea was simply to scale up a particular program model, and FP, nurse only partnership, but those plans quickly ran into a political backlash from other competing home visiting models. This forced the administration to rethink the strategy. By turning the focus generally to home visiting, it allowed the administration to use the evidence lens to view the problem and attack it systematically. The thinking became, let's build in standards of evidence and then let the chips fall where they may in terms of which of the specific program models get funded. I3 is another example that a body of this general evidence lens, this time from the beginning. There the focus was never on a particular program model. The idea was to do an education initiative and then use it as an opportunity to think about meaningful standards or tiers of evidence that you could apply in awarding money. That is, to apply evidence, the evidence lens to the issue of education. So the administration embraced

this important shift in thinking about evidence-based policy merely as a way to fund particular social intervention programs that have some evidence of success, to an evidence-based initiative as one that tackles any given social problem with a systematic and thoughtful look at the evidence base.

And then the final key in the administration's approach that's rightly been highlighted here, is that top administration officials empower those who knew the evidence best to craft meaningful standards that recognize the limitations. The administration wisely made sure that the debate about evidence and the setting of evidence standards for each initiative took place largely outside of the political sphere. Instead, top political figures empowered the dedicated and expert career staff, many of whom are in the audience today, and other researchers at the federal agencies, to think systematically about the evidence base in each of these issue areas and come up with a way to craft meaningful and defensible standards. Evaluation experts were largely insulated from Congressional pressure, advocacy lobbying and partisan politics. They were free to think systematically about the evidence in a particular area, what the most important outcomes were and what realistic and meaningful distinctions could be created in designing the initiative. Thank you.

MR. DEPARLE: When you read most Washington narratives, there's a central figure driving the story. Somebody has inspiration, an idea, an unlikely quest overcomes odds, but it's driven by a central figure. What I was puzzled by in reading "Show Me the Evidence", is how -- what a decentralized narrative it is. There are six programs. Each program arguably has five or six authors, people in control of it. It's -- somebody used the phrase, bottom up initiative. It's a very unusual Washington story and it's not built around a single inspiration or a single figure. I wanted to ask Melody and Robert and anybody else who has an insight into it -- can you tell us more about the genesis -- where it came from and who exactly was driving it. In particular, how much did the President know and care about it? Maybe Melody, do you want to start?

MS. BARNES: (laughter)

MR. DEPARLE: Well some initiatives come from the President of the United States and some come from inspired staffers. And it wasn't entirely clear where this came from.

MS. BARNES: Sure. And the six programs that Ron and Greg identified all have their genesis, and to your point, different places. I mean, we knew coming in, with the teen pregnancy program, that that was an issue that was just sitting there, that needed to be addressed. It had been an issue that had garnered a lot of concern and generated a lot of concern in previous years, so that was something that we had to take on, because of the economic downturn, given impetus or momentum because of the recovery act, because of the President's interest, Rahm's interest with regard to the economy, and the underutilized asset that is the community college and the number of people that are touched by community colleges, that this is an important moment for greater innovation in our community college system, and we have the opportunity to be able to do that. I mean, I think the book really documents well where some of these things come from. I would say the President said in overall tone and tenor of, let's do what works, and we have to get rid of some of the things that are not working. And that automatically drove us toward evidence as a way, and I think Greg just said it well, to, in a highly political, highly partisan environment, to use evidence as the wedge to try and drive smart policy forward. But it came from a line of thinking --

MR. SHEA: But do you think of it as one initiative or do you think of it as six initiatives that have some relationship to each other?

MS. BARNES: You know, I'm curious what some of my other administration or former administration colleagues would say. I think of it as, we have that overall theme and we all started working on different pieces and then it was all of a sudden, it was like, oh, you know, we, there's a narrative, there's a light motif that cuts through all of this that can be articulated. But I wouldn't say, or maybe it was a meeting I

wasn't invited to, but -- (laughter), that we, you know, there was a moment when we sat down at the table and said, let's stake out an evidence-based policy. You know, I wouldn't say that that happened, but evidence was clearly a theme and sat at the root of a lot of the work, and then a narrative was built up around it.

MR. SHEA: So there is one interesting glimpse in the book of where the President gets involved. There's a meeting in the oval office where there's a debate about, I forget which of the programs it is but whether to make the granting competitive, i.e. evidence-based --

MS. BARNES: (inaudible) race to the top and (inaudible).

MR. SHEA: Or formula, and the political, some unnamed political advisors argue that it's more politically wise to do a formula so that everybody gets some part of the pie and President Obama said no, to do it on a competitive basis, so that's our -- I think the most interesting glimpse we get of his role in it. Robert, do you want to talk about where it came from and what his --

MR. GORDON: I mean, just to echo what Melody said, I mean, Melody was working in the administration in a higher level than I was. From where I sat, the President had made doing what, more of what works and less of what doesn't, a clear priority, and I talked about that with my boss, Peter Orszag, talked about that a lot. I've worked with Peter at, here, actually, and knew that this kind of thing was important to him. And that was kind of enough. And after knowing those high level priorities and knowing that we would -- the people above us, including Melody, would have our backs, was enough to drive a set of conversations and unleash a lot of good thinking, and I do think as Greg described really well, the fact that evidence tiers are both a good substantive solution and a good political solution to the problem of having lots of people wanting you to fund their thing, and you don't structurally, it shouldn't be that Congress is writing into laws or administration is writing into laws, this is the program, shall have these 12 elements. That's not what you want to do. And then actually having a system

that's open ended and responds to evidence is actually a good response to the political and the substantive problem. I think that was the insight that we all had and that made this something that we wanted to reproduce from initiative to initiative.

MR. DEPARLE: So the other thing that struck me as a non-expert reading the book was how many near death encounters (laughter) these various initiatives had in the political process. It's the community college one that one moment was at 12 billion and the next minute was at nothing, and then was back up to two billion.

MS. BARNES: Yeah.

MR. DEPARLE: These were arguably ideal conditions, right? You had a democratic Senate, a democratic House, a newly elected Democratic President and also a time of crisis, where the attitude as the book notes, puts Rahm Emanuel saying, "let no crisis go to waste", so you had large -- you had the Recovery Act, the Healthcare bill, large legislative vehicles moving through where you could stick these things on and nobody would notice. And yet even under these very ideal circumstances, it almost seems miraculous at times that various ones of them got entered into law. Robert, John, Greg, do you have any thoughts about what that suggests for future initiatives?

MR. BARON: Well, I think, to pick up on comments that Robert made and Peter made, if you're looking for some sort of common theme, ground zero, I think the closest that comes to it is OMB career staff, people like Cathy Stack and others. In the Bush administration, there was a very important change that, which initiated with career staff, but also Robert Shea was who was leading the, President Bush's performance and management agenda, which was that they began to elevate rigorous standards of evidence for the first time, as official OMB policy. And they were trying to figure out in the Academic Competitiveness Council that Robert talked about, how to imbed these concepts. They were essentially laying the ground work, the exploratory work to figure out how to imbed these evidence based concepts into legislation and policy. When the Obama administration came in, there were people who were in the

leadership, like Robert, Jeff Liebman, Peter Orszag, who were very interested in evidence-based policy. It was a perfect storm at that point. Their interest and leadership combined with a career staff that OMB was thinking about these issues and had built up capacity, I think was the perfect storm that enabled many of these initiatives to go forward.

MR. SHEA: You know, as I mentioned, I think there's a risk that our progress stalls. It's a -- there's no one central person and I mentioned Walter Isaacson's book. His whole book is about the fact that even though the narratives suggest there's a single leader, innovator, in no situation is that the case. All of these are drawing on multiple contributors and that's clearly the story here. This has been a loose coalition that has morphed over time as people enter and leave government, you know, and there are a lot of these potential mines that you mentioned. I remember vividly celebrating the enactment of the law that funded one of what we consider a proven practice, funded a small pilot. Career staff came to me and said, you're going to get a call from an official at the agency responsible for managing this initiative and they're going to want to completely water down the criteria with which we judge applications for funding. So my office was literally roughly this size when I was at OMB, if you can imagine it and these two career staff came over and I had one line, two phones. They were on the other side of the room listening in on this conversation because they were so nervous that the whole thing, after all this work was going to be completely abandoned. Pick up the phone and this guy, this fellow over at this agency said, "I understand you consider this to be a very important program. I understand also the legislation gives me great authority and my staff is urging me to open applications up to a wide number of people. I'm going to make sure that the funded application -- the applications that get funding are the ones that adhered to this proven practice and not to this -- not to the many other unproven practices. But I'm going to need your help." And so it was this coincidental quiet conspiracy. Luckily this fellow over at this agency saw what we were trying to accomplish

and agreed to conspire to help achieve it, even though there are all these other forces were at play, to undermine what we're trying to accomplish, and so, each of these stories has these minefields littered through it, whether or not they are captured in the book.

MS. BARNES: Also, I want to reiterate something I mentioned when I was kind of going quickly through my short statement. That's why I really believe that we have to make sure that this isn't just a D.C. conversation. Because often what we recognize as being an important set of advances to do good, to accomplish these broader social goals that we have, are being perceived very differently by program operators and parents and, you know, the whole range of beneficiaries on the ground. And unless we start to engage and have a conversation about why this is important and why we're trying to achieve those same kinds of goals with people who are the needed beneficiaries of them, it increases the political pressure and it increases -- it raises the hurdle, it raises the bar on being able to continue the work that we're already doing. It leads to the letter I got from my constituent. It leads to the formula versus competitive grant debate that we have with Race to the Top and Title I funding, and while the President made the stance -- took the stance that he did, when you look at the legislation, you recognize there are still -- much more money that went through that formula program that went to Race to the Top, and the backlash against Race to the Top that exists today. So unless we have this conversation, yes, important, and use the page turning book that we have today, to broaden our conversation beyond Washington, these kind of challenges will only increase for us I believe.

MR. MARGOLIS: Yeah, I think it's really important that you get the buy in from the program operators on the ground that are running these programs. I mean, it's the culture of evidence we talk about. You have to get people at the local level who are actually implementing these programs to believe that there is a reliable way to measure what they're doing and to have outcomes. And we can't sit here and just talk from a 30,000 foot level. You have to engage the people that are running these

programs to understand, help them to understand the importance of using evidence to make these decisions.

MR. DEPARLE: I'm going to reserve the remainder of my time and recognize members of the audience. There's a man in the back who has his hand up.

MR. REEG: Good morning, I'm Bob Reeg with Healthy Teen Network. I wanted to launch off the last comment about how this works at the ground level, a couple of observations. One is, what's your all's belief about, like the need to constantly re-evaluate, so like if I'm a program operator and I picked one of these interventions off the clearing house or whatever, and it's been told that it works, what level of evaluation would I need to do in terms of scaling and replication? Do we have to keep repeating like, that, wow, this is a new discovery? I want to hear some thoughts on that. The other is, just an observation that, I mean, these programs are silos basically that don't account for the fact that the populations we're working with are whole people. I consider like silos in factories. I mean, if I'm a person, I'm a factory. I need all sorts of things in my life to be successful, and typically these programs address one component of one's whole life. And how does this work that -- do you have to assemble all of your life in evidence-based programs, or -- (laughter). How do we recognize that a lot of these are oriented around solving some common sense stuff like, I look in the area of housing for example. If it was good to be not housed, most of us wouldn't be housed. And the fact that most of us are --

MR. DEPARLE: Let's go to an answer -- Who's got a comment?

MR. REEG: Yeah, so there's a lot of complexity that's not about science, it's about the environment.

MR. DEPARLE: Jon?

MR. BARON: Well, I think two parts. Number one, I think the question about how much to evaluate -- there is a lot of evidence that a single site, single evaluation, even a single site randomized trial, no matter how well it's conducted,

probably is not enough to have strong competence that the intervention works. Ideally what you want is not only that, but also a replication in another setting showing that it generalizes, to get back to the earlier point. And there are some important examples of that, in Welfare to Work. So MDRC, and this gets a little bit to comments from Jason earlier. MDRC did a study in the early 1990s in California where they found randomized trial in Riverside, California of a Welfare to Work program produced very large effects, much larger than had been seen anywhere. It was of a work focused Welfare to Work program and produced an improvement of 40 percent in employment and earnings and large savings to the government, about 6000 dollars per welfare recipient over five years. And people thought that it's got to be a fluke. Maybe it's a mistake. Maybe it's one of these spurious findings. Los Angeles had a program that had been evaluated at that time that focused more on education, basic education, which had been found in another side of the trial to have no effects. Los Angeles ended its program, adopted, based on those disappointing results, adopted key elements of the Riverside program, implemented it, it was evaluated in a randomized trial of every welfare recipient in Los Angeles County -- 80,000 people over a period of time, and was found to replicate the Riverside, not perfectly. The effects were not quite as large, but it was still like a 25, 30 percent increase in employment and earnings. I think that kind of evidence of reproducibility and generalized ability across different -- across different sites, is what's needed, to have strong evidence. Replication is a central concept in all fields and social policy should be the same.

MR. SHEA: One thing I would add on your question about the person being a whole self. There are a lot of -- there are initiatives underway to try and bring multiple programs together to help individuals. The Performance Partnership Pilots in particular, which are aimed at disconnected youth, seeks funding from grant programs that the Department of Education, the Department of Labor, the Department of Education, Corporation for Community and National Service, state and local jurisdictions

are asked to collaborate among different elements of those communities, access funds from these multiple sources to develop a holistic evidence-based or evaluated program to get at that whole person, in this case, disconnected youth. So I'm very excited about the initiative and eager to see what results will be accomplished. It's going to be messy, I know, because collaborating among, you know, just here among the panelists would be difficult enough. But among complicated programs, multiple jurisdictions, multiple agencies, multiple Congressional committees, going to be very difficult. But if we let it just stay for several years, we're likely to see some very interesting results and perhaps the foundation of a new program approach in the federal government.

MS. BARNES: And I would just add that those performance pilots -- part of the work that came out of the council that I mentioned, and that collective impact strategy, which is a cross sector and taking multiple programs and looking at the way that they come together, kind of like if you flipped your watch over and you looked at all the different wheels and pulleys that were coming together, that evaluating the efforts underway down to evaluate that as a model for big social change is part of what's underway to see how successful it can be. But it does require an urgent patience to it, to let it gestate, which you were just saying.

MR. SHEA: Well said.

MR. DEPARLE: I think if we keep our questions terse, we might have room, time for a couple more. There's a question in the back.

MS. SHORE: I'm Elizabeth Shore, Friends of Evidence at the Center for the Study of Social Policy. Everybody here has talked, with the exception of Melody Barnes; everybody has talked about how we evaluate the unit that is a program, or even many programs. So much of what is promising today is not going program by program. The promised neighborhoods, the choice neighborhoods, the promised zones -- those are all efforts to change whole communities, not program by program but with -- and it isn't even a joining of programs, it's doing, going beyond programs. And the issue of how

we evaluate something that isn't a program, that is circumscribed and stands by itself and can therefore be evaluated with random assignment studies and other experimental studies -- how do we go about evaluating those initiatives?

MR. GORDON: I think that there are lots of ways, obviously you can make decisions about which communities you fund. You can make decisions within communities about who receives funding, and the example of promised neighborhoods, I know that there's a very good interesting evaluation that was done a long time ago now by Roland Fryer of Promising New York that used of mixture of evaluation strategies and had an interesting finding. I'm not sure that we have done all that we could have done to be learning from promise in the year since that study. I think clearly it's the case that we need to be open to multiple strategies for evaluation and that's something that I think everyone has emphasized. And I guess my last thought is that it clearly also is the case that there will be -- there are plenty of times when we want to use holistic strategies of -- that are multi-faceted and it will be harder to use RCTs and then there are plenty of times when we don't actually, where we're looking for an intervention that targets a particular outcome and if we can achieve that outcome it's a massive accomplishment. I think a lot of us feel that way about a lot of the programs that have been funded here. So I do think it's a challenge. I'm not sure that it's an insurmountable one.

MR. MARGOLIS: Yeah, I would just add that one answer to the question is that you can randomly assign whole communities. It has been done. And it's been done sometimes on the cheap. Let me give you a quick example. There was a CDC funded study of the Triple P Parenting system which is implemented on a whole county level. It includes multiple components, parent training, intervention with families that are particularly at risk to provide them coaching, media strategies at the county level to try to encourage positive parenting practices. So the researchers who got the funding didn't have enough money to implement it in all counties. They were located in South Carolina, so they took 18 counties in South Carolina, randomly assigned them to match them and

then randomly assigned them to either get, implement Triple P on a county-wide level or not. The other counties, control counties, implemented usual services, and then they measured outcomes over time using administrative records in each of the counties including child hospitalizations, foster care placements and so on and they found a fairly sizable effect. That whole study, incidentally, was done for about 300,000 dollars, because even though it was a huge study, a huge number of individuals, because all outcomes were measured with administrative data, that were already being collected anyway. So it's an example of how you can -- it is feasible to randomized counties, to evaluate the kind of system level changes that you're talking about.

MR. DEPARLE: Last question, sir.

SPEAKER: On this last point, and on the general block grant or formula versus individual projects, I have been studying child protective services. Now that's a very complicated area. It relates to housing, to mental health, to Medicaid, to God knows what, and we've developed a system now where waivers were granted by Congress and to get to the point quickly, about half the states now have waivers, half don't. The waiver authority is lapsed. The question is where do we go from here, but if you look at the waivers that were granted, the ones that have really worked were statewide waivers, and that consisted of one state -- Florida, and Los Angeles County. Where you had the teeny weeny little county by county or even lower projects, this was great for researchers. It was not so good for the children. It was for the states. But in Florida you had a system wide thing which precluded by its very nature, random control, because it was for the whole state. But there were some ways you could get evaluation into it. The HSS said okay, give us an evaluation. So they gave an evaluation but it could not be random, but I would urge that politically and socially and program effectiveness -- the more you can get formula funding into these social areas which have so many intricate parts, the more discretion at the state level, the more chance you are going to be effective.

MR. DEPARLE: Melody?

MS. BARNES: I'm not familiar with that particular study. It certainly -- it's a really interesting proposition as a way to try and look at these issues and move forward. Now when we were all working together and we were trying to address, you know, the formula versus competitive grants challenge, in part, we came to conclusion that for a policy and political reasons, there were formula grants that needed to continue and that would allow for some of what you were just talking about. At the same time, we wanted it to try and move the ball forward with interventions that we had recognized had evidence supporting them as a way to drive, in this case, context education policy forward and that the competitive grant would allow us to do that, and we could learn a lot from the winners of those grants, and we could build a learning community of those winners that would also help inform the work of those who had not won, but who had competed but had come up short in their proposals. So there was also a lot to learn in that instance as well.

MR. DEPARLE: Ron -- 24 shopping days until Christmas? Books for sale? You'll sign as long as the audience demands?

MR. HASKINS: It was written with (inaudible) in mind. (laughter)

MR. DEPARLE: Round of applause please for our panelists. (applause)

Thank you. Thank you for coming.

* * * * *

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, 2016