THE BROOKINGS INSTITUTION

REENGINEERING THE SIPP: THE NEW DYNAMICS OF ECONOMIC WELL-BEING PROGRAM: A ROUNDTABLE

Thursday, June 8, 2006

FAULK AUDITORIUM
THE BROOKINGS INSTITUTION

TAPE TRANSCRIPTION

The Brookings Institution
1775 Massachusetts Avenue, N.W.
Washington, D.C. 20036-2103

PARTICIPANTS:

Richard Bavier, OMB
Heather Boushey, Center for Economic and Policy Research
Connie Citro, CNSTAT
David Connelly, OMB
Kathy Creighton, Census Bureau
Jennifer Day, Census Bureau
Michael DePiro, FNS
Elizabeth Down, Representative Moseley
Judy Eargle, Census Bureau
Cindy Fagnoni, GAO
Jason Fields, Census Bureau
Carol Frost, CBO
Nancy Gordon, Census Bureau
Dawn Haines, Census Bureau
Ron Haskins, Brookings Institution
Susan Hauan, HHS
Howard Hogan, Census Bureau
Howard Iams, Social Security Administration
John Iceland University of Maryland
Julia Isaacs, Brookings Institution
Linda Jacobsen, Population Reference Bureau
David Johnson, Census Bureau
Julia Lane, NORC
Michael Laracy, Annie E. Casey Foundation
Kris Moore, Child Trends
Don Oellerich, HHS
Jeff Passel, Population Association of America
Ron Prevost, Census Bureau
Wendell Primus, Office of Representative Nancy Pelosi
Barbara Pryor, Office of Senator Rockefeller
Caroline Ratcliffe, Urban Institute
Andrew Reamer, Brookings Institution
Ralph Rector, Heritage Foundation
Isabel Sawhill, Brookings Institution
Susan Schechter, OMB
Arloc Sherman, Center on Budget and Policy Priorities
Jim Sears, Social Security Administration
Karen Smith, Urban Institute
Ed Spar, Council Pro. Assoc. on Federal Stats
Mark Stephenson, Committee on Government Reform
Chad Stone, Joint Economic Committee
Cynthia Taeuber, CM Taeuber & Associates
Katherine Wallman, OMB
Alan Yaffe, HSS-ACF

CONTENTS: PAGE

Opening Remarks - Ron Haskins	4
Introductions - Andrew Reamer	6
Overview of SIPP Replacement - David Johnson	7
Response by:	
Don Oellerich	24
Howard Iams	30
Heather Boushey	36
Roundtable Discussion	43

P R O C E E D I N G S

MR. HASKINS: Good afternoon. My name is Ron Haskins. I'm a senior fellow here at Brookings. Hopefully you will not get that radio or whatever that thing was.

Is everybody's microphone on? You can tell because it's red. It's not supposed to do that unless we have three or four of these on. So when it comes time to talk make sure you turn your microphone off when you get through, because if we have two or three of these on then we start getting feedback.

I want all of you to know that I learned --

A PARTICIPANT: (Inaudible.)

MR. HASKINS: Do what?

A PARTICIPANT: (Inaudible.)

MR. HASKINS: Oh, they're doing something. Okay, I just want all of you to know I learned a lot about several people in this room because the microphones have been on and I've been able to hear some of your conversations here. I don't know if you all know this, but there's some really amazing Washington stories about members of Congress who made remarks in front of microphones, in one case very nearly ended his political career.

The question I'm going to answer in 60 seconds is why would Brookings sponsor this event. To answer that I'll first observe what Kris said to me just a few minutes ago when she came in. She said, AMy God, it's a full house to talk about data?"

And, you know, that is remarkable. A lot of people wouldn't think that about Washington, but of course we all know that these data sets for some of us are virtually our lives so we really care about having good data. So when the Census Bureau announced that it ws going to terminate SIPP or do something else a lot of people were worked up.

The two things that we try to do at Brookings is select events that are interesting, exciting, controversial and that bear directly on the interests of children and families, especially low income children and families. So this is something that Bill and I often argue about -- does it fit in with our goals, and in this case it clearly does fit in with our goals because as I've said without the data all of us are lost. And not that policymakers are exactly clamoring for better data, but we're going to give it to them anyway and if we don't have better data we can't do it. So it's really crucial.

The Census Bureau's plans for replacing SIPP are a really crucial issue. Just speaking for myself I think they're doing exactly the right thing, which is talking with people broadly about what their plans are and getting feedback, and hopefully they'll have an open mind and they'll hear things that will be interesting and important. They may even modify their plans.

So that's the purpose of this event and I'm pleased to tell you that you won't have to put up with me anymore because Andrew Reamer who is from our urban markets initiative program will moderate today's events, introduce the speakers. So, Andy, thank you very much.

MR. REAMER: Great, Ron, thank you. Thank you and thank you all for being here.

We'll have a presentation and a series of responses and then we'll open it up for discussion.

We'll start with David Johnson who is from the Census Bureau. David is chief of the Housing and Household Economic Statistics Division. He joined the Census Bureau just a few months ago. He'll give us an overview of the re-engineered SIPP in Power point. You all got a copy of his paper previous to coming.

Then we'll hear -- and David will take about 15 minutes, then we'll hear from three respondents -- Don Oellerich with the Office of the Assistant Secretary for Planning and Evaluation at the Department of Health and Human Services, and then Howard Iams who is a senior research advisor at the Social Security Administration, and then Heather Boushey who is a senior economist at the Center for Economic and Policy Research.

So -- and then we'll open it up for discussion until about 2:30. And we'll start, David.

MR. JOHNSON: I moved over here so I can operate the Power point.

Thank you Ron and Andy for hosting this roundtable and for all the Brookings staff who worked so hard on this. I even had an e-mail this morning at

3:30 in the morning asking me about some clarification for the set up. So they've been doing a tremendous job.

I'd like to also thank all of you for coming. I see a lot of friends that I've interacted with in previous lives and I hope that tomorrow I won't be saying that somewhere else when referring today as a previous life.

Finally I'd like to thank, as with all my recent presentations, I'm basically a conduit for a lot of hard work from a lot of great staff. And so the Census Bureau staff who are here today and those of them who are participating via telephone have done a tremendous amount of work in these past three months preparing all this stuff and helping me to present this information to you. They are the experts, so at the end please ask them the questions.

As Ron suggested we basically have two goals for this, is to describe our progress and options for the dynamics of the economic well-being system and to determine your needs for this new system. We're interested in learning from you the components of SIPP that you'd like to retain, but given the budget situation we'd also like to know the things you're willing to give up. I think that's one of the crucial things we need to know from you.

So there's been a lot of discussion about why we're at this point, but I would like to focus this discussion today on our plans for re-engineering. In reviewing earlier SIPP evaluations and studies it looks like the Census Bureau was constantly redesigning the SIPP, evaluating the SIPP and, yes, even re-engineering the SIPP, and even before I got there they were working on the new 2009 panel.

So I think this is part of an evolutionary process and the best analogy I have for this is the story of the mythical Phoenix, and those of you who have kids or those of you who are Harry Potter fans will understand this. I'm both, and so I just want to show you a little clip that sort of illustrates this.

(Video presentation.)

MR. JOHNSON: Even you get it. Okay, fine.

While the burning seems dramatic it is simply part of a life cycle, okay. This doesn't only work for

SIPP, we are simply taking advantage of this opportunity to transform SIPP into the new dynamics of economic well-being system.

So our goals for this, having talked to people inside and outside the Bureau, federal people, other stakeholders, it seems like the value added of this new program and SIPP in general was in providing a nationally represented sample that can be used to evaluate the annual and sub-annual dynamics of income, the movements into and out of government transfer programs, and namely those interactions between these items. I think this is what we're looking at when we design the system that we think is needed.

So we'll compose then things like a new survey data collection, require fewer resources -- this is one of the key components of this -- improve processing efficiency, be releasable to the public in a timely manner, integrate survey data, administrative records data, and make use of the richness of the new data collection of the American Community Survey. So we think all these are things that we think are going to be components of this.

Now we do acknowledge that SIPP has many important if not invaluable uses. SIPP is used by many government agencies, most of you represented here, to measure the effectiveness of government programs. In particular, these are some of the examples: the Department of Agriculture models food stamp eligibility, HHS effects of welfare reform, Social Security modeling SSI benefits, and many researchers around the room in using both the longitudinal and the cross sectional nature of SIPP. This is one of the questions that we have for you, is which of these things are most important.

And finally the Census Bureau produces a bunch of reports. Some of them have -- some of them are in a series called the dynamics of economic well-being. So spells of unemployment, moving up and down the income ladder, those types of things.

The other thing I've learned in reading previous studies of SIPP and other evaluations is everybody seemed to agree on the basic goals of the program -- yeah, ten minutes -- but there's been a lot of disagreement in terms of the specific variables that should be included. The CNSTAT report talked about income and program participation, and a study 20 years ago talked about the relative short periods of time, extended periods of time, and this interaction between household composition income, labor force participation and government participation. So these things I think have always been key.

I think the most interesting quote I found was this one: AThe planners of SIPP are to be congratulated for their intention to combine administrative data with field survey measurements. It is clear that combining such disparate sources of information provides a much richer insight into the status and behavior of individuals."

Now the interesting, the most interesting thing about this quote, it was said over 20 years ago in one of the earlier evaluations of the SIPP. I think it is time now to take this quote seriously to see what we can do to evaluate it and see if we can actually implement this. I think this is what we're trying to do with this new system.

So in the paper that you have we basically had five components or milestones, what have you, and I'm going to go through each of these in turn and talk about where we are in this process, and then obviously end with some questions for you that we'd like to hear from you.

Develop new survey component. This includes content, survey design, survey instrument.

Construct a crosswalk between the admin records data and the needs of the new system.

Develop a prototype, something we can use to see if it works.

Develop data products. What are the products we're going to produce.

Finally, meet with you to see how we need to work this through.

So the first is the new system will eventually include a new longitudinal survey component. We're currently producing what we're calling decision matrixes to evaluate the importance of particular variables such as demographics, labor force, program participation, things that we see as essential components of this new basic system that we're going to operate on.

Now obviously given the budget situation the content that we're going to have in this is not going to be as complete as the entire series of the core and topical modules in the current SIPP, but we hope to retain most of the important information. So we're trying to assess what's in the basic survey.

With this basic survey we're going to try to include questions that are going to act as hooks to other information -- child support, well-being, disability. You can think of this as similar to CPS and supplements or core and modules, but we're trying to say that we know what the -- we want to figure out what the basic topics are and then are there other issues that we can partner with other agencies on to obtain that information.

Obviously after we go through the content we have to figure out what the sample is. So we have another group that's looking at different samples for us. This can range anywhere from the master address file, to directly interviewing ACS cases, to using the matches in CPS from one year ASEC survey to the next year ASEC survey, a variety of options. So we have to evaluate those.

The others, obviously there's different modes of data collection. There needs to be an evaluation of the trade offs between the quality of the responses, the timeliness and the cost. I think we've been very successful at the Census Bureau with the way we do the ACS and having a mail out and then a caddy, and then a cappy (phonetic) approach. So we're going to look at these alternative methods for data collection.

Currently after wave one 50 to 75 percent of the SIPP interview is done via telephone. So this is something we have to evaluate.

Finally, as most people know, we have to worry about the recall periods. This is especially because we understand the importance of these sub-annual measures, and even monthly measures, and we know that as the number of interviews increased the cost increased as well. So currently the SIPP, the current SIPP has about a four month recall, three interview per year, and we ask about the preceding four months. For the past 20 years, and I can see all the documentation, we've always thought, well, what about a six month recall, only two times a year, and we would ask about the preceding six months.

We're also considering a 12 month recall, and in particular we're evaluating what's called an event history calendar. So we're talking with people at PSID, at the University of Michigan. They're having a conference next week. We're going to talk to them.

We've found in one of the methods panel that we conducted one of the results could have been that -- one of these results suggested that using this event history calendar can deal with the seam bias issues that we now know are apparent in the current SIPP.

So with these survey issues, the recall, the mode, the content, we're trying to make sure we address all the issues that are in the current SIPP in this new program.

After the survey, then in conjunction with that, we have to evaluate the administrative records. In particular we have to evaluate what do we know about administrative records, what can be used in this survey instrument. I think during the past few years the

Bureau has been very successful in using administrative data to improve the quality of our statistics. They have acquired -- we have acquired over 50 resource files and the number continues to grow. In particular we use like the Census Newbident (phonetic) file to improve on the census data, to do some imputation. We've used the IRS data in our modeling for small area income and poverty estimates, and the LEHD program is used benefit data -- earnings data to create a synthetic data file.

So we have very, very good success at doing this. One of the main reasons is because of the match rate. So if we take the CPS data and then try to match it using social security numbers through administrative records data we find, for those people who said we could do this, we obtain 50 percent of matches directly from the social security number. Then we can go beyond that and use their address and link another 36 percent, and finally we can even use their name and like another 7 percent for a total match rate of 94 percent. So they were matching 94 percent of the files between survey data and administrative records data. This is true for the SIPP and the ACS as well.

The challenge, though, is we can use the administrative data. We know we can use it to improve the quality, we know we can use it to improve the estimates, can we use it to decrease costs and increase timeliness. Can we get a match of administrative data so that -- let's say we don't have to ask a survey question or we can ask annual information on the data on the survey but use administrative records to fill in sub-annual information. These are the things we don't know and these are things we need to investigate. That's why our next component is developing a prototype.

Now the goals for the prototype are basically to iteratively assess the ability to do this link. So we take the first example, the first cut, and we find two data sets -- a survey data, the CPS, and an ad rec data, this Medicaid file, pick a variable on it and do a match. We find that when we do this the majority of people agree, okay. The people who say they receive it are -- or not receive it match what the Medicaid file says.

There are, however, people who disagree. Now most of the error, if you want to call it error, in the response is that people are saying they receive Medicaid when the ad rec file says they don't receive it. But we think this provides us with an example of, yes, we can do the match, yes, it could improve the quality of the data, and finally we've sent this to the disclosure review board and they said that, well, if you ask the data question on the survey you can release the administrative record file that matches that.

So we've found that we can take this all the way through. Obviously those are just one variable on one data set, but this is our first prototype. The second prototype would go to Medicare, the third prototype could be even more, the fourth prototype, the fifth and so on, until maybe the final prototype is taking something like the CPS file or the CPS ASEC matches over a two year period, okay, matching them and then matching that set to administrative records data to maybe fill in sub-annual estimates or to provide other variables. That would be prototype that we could release and people could evaluate what could be done.

The fourth thing, the component, is the data product, which is a serious issue. Obviously we'd like to have a complete public data file that everybody has. However, with using administrative record data that complicates the files that we can release and make public. We do know that we've been successful at producing internal files and access to the research data centers, the RDCs, and we know we've been successful at producing what we call synthetic data through the OEHD. I think for this program we're looking at somewhere in the middle of this, a mixture.

So we want a release of public data, maybe some components of it you'll have to use internally, maybe some components will be imputed or synthetic. But the key is the combination of these are going to depend on you the users, on what you think is feasible and what you think is possible.

So how does this all work? Well, the products we're going to produce in the new dynamics of the economic well-being system also include the current SIPP products. So as most of you know the current SIPP data collection is scheduled to end in September of this year and those data sets will then be released over a period of time. Wave one has been released, wave two is forthcoming, wave three and so on will be released as well as the topical modules, and this will take us until February of 2008. The data that will be released will cover the periods of about 2004 to 2006.

Simultaneous to this activity we're going to do our planning, as is evidenced by this paper, between now and 2009, and develop this prototype. This prototype we hope will be released sometime in 2008, however it would cover the time period of 2006 to 2007 using the CPS matches.

Finally we start data collection for this new program let's say in 2009 and forward, and if we are successful about using some of this recall that we've looked at this could actually cover the time period 2008. So while the prototype will not provide all the information that users require, it will provide some information so that we can have some information every year between now and 2009.

But we know it's not going to provide everything, but we know that the way we're designing this, because of the budgetary limitations, we're not going to provide every single amount of information that we currently provide. But we have to figure out what those important things are, and that's exactly why we need your help.

So the last component is getting your input. That's why we have this meeting, that's why we have had other meetings, and we also redesigned the SIPP website -- it doesn't like look much, the design -- but we've added this new button, the dynamics of economic well-being. Currently if you go there and push it you'll be sent to an e-mail that can be sent directly to us to provide your comments. Hopefully by next week you'll be able to click there and link to the paper and maybe the slides that we're presenting today, maybe the transcript depending on how those work out, and also maybe one of these decision matrixes that we've talked about that federal users are completing. They're completing a decision matrix that includes a lot of the variables that we have in an instrument, see what's really needed, what we can give up, what we have to add. That could be there as well.

Finally since there's no free lunch, being an economist, you have to do some work. You have to give us some input. So this is sort of the questions that I would like to start you off with. Obviously there's others things that you have, but this is things that we think we need.

We need to know what you really need in this new system, okay. In particular we need to know what you're willing to give up. We need to know the best examples that you have of current researchers in the SIPP and the best examples of research that need to look at these dynamics of economic well-being, and we need you to really take a hard look at the cross section versus longitudinal aspects of the data. We're saying this is a new system of dynamic. That implies longitudinal. When you look at the users and who uses the current SIPP, there are a lot of people who simply use it cross sectionally. So we need to know those trade offs between those two types of data.

The key obviously is our interaction with you -- that's why we want to have these sessions -- and the key is to have a more effective system to measure the dynamics of economic well-being delivered in a more timely manner and at less cost. And this may not be a (inaudible) superior product. Obviously we can't please everybody all the time, but with your help we can determine the critical needs of the data and develop a more effective means to collect, process, evaluate and deliver this data.

Thank you.

(Applause.)

A PARTICIPANT: David, you might want to leave that slide up there.

MR. JOHNSON: Oh, okay.

A PARTICIPANT: You might want to leave that up there.

MR. JOHNSON: Oh, now it's gone. It's not going to go up. We have to flip it back on.

MR. REAMER: We'll start with the respondents. Don, do you want to start us off?

MR. OELLERICH: Well, thank you, David, and thank you all, and thank you to Brookings for pulling this together.

I probably should start off by saying I am from the Department of Health and Human Services, but if I do happen to express an opinion it's my own and it doesn't represent the Department. Hopefully that will cover me.

I was very disappointed when I heard that the Census Bureau had chosen to end the SIPP program. As with many of you I've been a long term user of the SIPP. I started using the SIPP back in the 1980's when it was a relational database at the Institute for Research on Poverty. That was scary. Talk about a difficult database to manage.

But I think what we have is a problem, and as my wife will remind me problems can be turned into opportunities. And so I think we should take this as an opportunity to move forward to design something to meet the needs of the many users of the current SIPP.

I really am glad that David used his analogy of Harry Potter and the Phoenix because I'm a Harry Potter fan myself and I would have been very disappointed if he hadn't used it. But the thing to remember is that the Phoenix that emerges from the ashes is a copy of itself. It's again the Phoenix --

it's not an eagle and it's not a pigeon -- it's the Phoenix. So it's not new and improved, it's not redesigned, it is the Phoenix.

I might also add that a Phoenix is defined as unsurpassed excellence. So if that's what we're aiming for here we have a ways to go and a lot of work to do.

But this is being presented as a work in progress. They're looking for input, which I think is really important and I really applaud their efforts in getting this thing moving in the right direction and seeking all the input that they are.

To give you a sense of where I'm coming from as a data user I think it might be best to quote the administration's Welfare Reform Reauthorization proposal which sought to fund the Census Bureau to continue the SIPP program and expand it, enhance it, with $10 million a year. This legislative language appeared in both Senate and House versions of the Welfare Reauthorization Bill, although it didn't make it into the final passage of the Deficit Reduction Act because hardly anything did.

"The Bureau of the Census shall implement or enhance a longitudinal survey of program participation developed in consultation with the Secretary and made available to interested parties to allow for the assessment of the outcomes of continued welfare reform on economic and child well-being" -- I emphasize that -- Aof low income families with children, including those who receive assistance. The content of the survey should include such information as is necessary to examine the issues of out of wedlock child bearing, marriage, welfare dependency, the beginning and ending spells of assistance, work earnings, employment stability and the well-being of children."

I have two minutes left, oh my gosh. Okay. Too much time on the Phoenix.

In order to do the types of analysis that we do at HHS and at ASPE and work that we support, we need a longitudinal database that continues to maximize the sample of low income families with children. That's very important for the type of work that we do and the focus that we have. We need to maintain monthly income reporting, employment, program participation and transitional events such as marriage, divorce, births, et cetera.

We need quality data, which means we need to maintain a minimum recall period, and the six month recall period sounds like it may be okay given the research and some of the other things we've seen over the past. I would really caution you on this event history approach, look at it very carefully. I sit on the advisory board for the PSID. It took a long time to implement. It took a lot of training of the interviewers, the field staff. It was not an easy thing to do and the quality of the data is still being examined.

I think it's important because most of our work is longitudinal, it's not cross sectional. We need long surveys, long longitudinal surveys, multi-year, three or four years in order to capture the kinds of events that we're looking for, to look for these transitions on and off, in and out, up and down and all around. We need content that includes measures of family, adult and child well-being, demographic characteristics including marital history, fertility, disability, work support -- does it sound like the SIPP -- child support and assets.

I believe we can maximize the use of administrative data, but I don't think we can do it today. I think it's something for the future and I think it's a goal that we should be working for, and I think we have been working for it. I know my department has supported the Census Bureau in a number of efforts to do matching of administrative data.

I think we also need to maximize the samples within states so that we can identify the states. We don't have to have all the states, but we need to be able to identify the states on the public use data so that we can look at state variation and economic and policy conditions.

The prototype proposed by the Bureau is not there yet -- and I need to stop, so I'm going to jump and say I think that there are several other things that they might want to consider. One is to follow one of the rotation groups at the >04 panel through this transition period. They might want to think about using some of that unused sample. This is a Richard

Baverre (phonetic) idea, but if we don't use all of the sample to follow, use some of that unused sample to test out some of these methods because these people are still expecting you to come back.

I have more but I just want to say in conclusion I think we have a problem -- the loss of a valuable source of longitudinal data, but it's also an opportunity to develop something new, innovative and that will meet the users needs. I look forward to continuing to work with the Bureau and it's just time to see whether we'll get a Phoenix or a pigeon. Only time will tell. Thank you.

(Applause.)

MR. REAMER: Thank you, Don. Howard?

MR. IAMS: Thank you very much for inviting me. The Social Security Administration is very encouraged that the Census Bureau is seeking input from the federal user community for a replacement of SIPP.

SSA basically uses cross-sectional data in SIPP with many topics connected to the cross section from topical modules and linkage to SSA administrative records. The very data items in SIPP allow SSA to measure the covariance between important areas of economic well-being, such as pension coverage, assets, lifetime earnings, and income sources. These are used for example in our mint model which projects the future retired population and is used for social security reform estimates.

Another example, SSA's model of financial eligibility for supplemental security income and Medicare, low income subsidies, uses detailed income and asset liabilities for the same month with linkage to SSA administrative records of SSI benefits and Quimby Slimby QI (phonetic) benefits.

SSA's longitudinal analyses usually involve patterns of SSA administrative data on benefits or earnings, but not reported survey data with the SIPP. At this time the Bureau has proposed a prototype of a longitudinal CPS with some linked administrative data. It's very unclear what the new survey of annual data of the third interview includes, but the prototype document does not mention monthly data.

SSA's concerns with the prototype include lack of monthly measures of income, under reported prevalence of asset income in the CPS, omission of a regular receipt of income from defined contribution pensions and retirement accounts in the CPS, high attrition in the longitudinal CPS, disclosure issues of administrative data, and lack of validation of synthetic data.

Currently SSA can take the CPS and match it to SSA's administrative records. The proposed prototype will not meet SSA's needs, replacing what SIPP provides to SSA in terms of monthly income, data nor topical data on assets, pensions, disability, health and marital history.

SSA supports the concept of using administrative data to strengthen or replace survey data. SSA's main concern for the Bureau's plan for administrative data is public release of the data. It is very difficult to release administrative data because of disclosure privacy issues. We know this from experience. Our first priority is protecting confidentiality.

If the Bureau cannot release real data it will be forced to rely on statistically imputed synthetic data. The usefulness of synthetic data is still being established. SSA intends to test a synthetic file of matched administrative records with SIPP in the next year.

SSA has some issues with the administrative data identified in the prototype document. The listing of administrative items omits SSA records of monthly benefits and the IRS records of annual earnings. David Johnson's recent presentation included SSA records of benefits, so maybe it was an oversight in the document.

About one in six Americans get SSA benefits and the entire labor force has records of earnings. These are administrative records of widespread income sources.

Asset income must include 1099s as well as 1040s, all the 1099s. Low income people often do not file 1040s, creating selectivity bias.

The SSA newbident is not a reliable source after 1986 for race and ethnicity due to hospital reporting procedures. SSA looks forward to working with the Census to cover the topics used in SSA's modeling and analysis and in monthly income eligibility determination.

I don't know how much time I have left but I passed out a sheet that had specific items being used in SIPP. Oh, I have two minutes, wow. I got through the prototype, so we're complimentary.

A PARTICIPANT: (Inaudible) would like to borrow two minutes.

MR. IAMS: Oh, please. Let me just highlight a few things.

We use, one, the basic demographics but, two, the thing that's unique to SIPP is the detailed monthly income accounts for individual and spouse. I don't really know of any other nationally representative survey that has monthly income. This is important for SSI, for Medicare Part B subsidies, for food stamps eligibility and for (inaudible) eligibility. We'd get different results if we use an annual base as opposed to a monthly base. So the annual doesn't substitute.

Detailed asset holdings and liabilities in the month, well SSI and Medicare part B subsidies all have limits on resources and we have to have specific disregards that we want to model, and our model for the future retired project the future wealth holdings of the future retired, of the baby boom and beyond, and so we need household net worth, financial holdings and home equity.

The pension coverage and pension characteristics, our retirement model uses coverage for people in the labor force to then create an estimate of what their income will be like when they reach retirement in the future years. So we need to know if a person is on the job with a pension, whether they're participating, some characteristics of the pension, and we need to know a variety of things about that.

Marital history, we need to know whether a woman has a ten year marriage because a divorced person has to have ten years of marriage to be eligible for social security benefits.

Disability and health status, well the Social Security Administration pays benefits to severely impaired persons. I'm supposed to stop.

You can read the rest I guess. Disability, health status, welfare program participation, birth dates, the changing history and labor force, and we need the ability to match our administrative records to this. We've been actively using this for many years. I started using this survey match to administrative data in 1986.

MR. REAMER: Thank you, Howard.

MR. IAMS: Thank you.

(Applause.)

MR. REAMER: Heather.

MS. BOUSHEY: Thank you for all the work you've done putting these together and to David for your memo. The comments are little bit different from the memo, and my comments will touch mostly on what was in the memo, and I'm going to gloss over the things that Don and Howard have already spoken to. There's a handout of my full comments.

At this time we all know that the Census Bureau has not actually fielded the prototype for the new survey and so we do not know how successful it will be. Further, in the interim researchers will have to go without data for many years, possibly through 2009 or 2010, leaving us without data on the dynamics of economic well-being.

The Census Bureau is optimistic that the new survey will be cost effective and provide an improvement, perhaps a Phoenix not a pigeon, but what they're proposing may be more technically difficult, time consuming and expensive than they are right now admitting. With so much at risk and so little certainty about the new survey the Census Bureau should not eliminate the SIPP until the prototype has been fully analyzed and peer reviewed and a plan is in place to ensure comparability across the SIPP and the DEWB. Somebody has got to work on that acronym.

Now there's a number of identifiable problems with the new survey, some of which have already been discussed, but these indicate that it may not be an improvement in terms of data quality and may actually be more expensive than the SIPP in the long run. I'll go through a few of these.

First of all matching administrative records is not that easy, fast or inexpensive, nor will it necessarily solve all the problems identified in the SIPP, including improving benefit coverage, timeliness and survey cost. There's a lot of work out there, many of which has been done by the people around this table documenting that matching administrative records is often difficult. While matching administrative records is fundamental to the DEWB, at this time we know that they cannot, the Census Bureau cannot match administrative data for most if not all of the benefit programs -- many if not most of the benefit programs because the records are either produced too late to be of timely value or are produced at the state level.

My understanding from the memo was that the numbers that David showed for the Medicaid program are three years old, because he says in the memo that those -- that there's a three year lag on the Medicaid program. That's a significant problem for policy analysis if this data is to be important to us.

At this time the Census Bureau does not know whether or not the DEWB will be public use, which jeopardizes its usefulness with government and outside researchers. We policy analysts work on incredibly short time tables. We do not have time necessarily when Social Security is supposed to be reformed to fill out a bunch of forms and go to a special site to do the analysis. We need to have public access data.

One question I had from David Johnson's memo is that they say that to ensure continuity with prior SIPP panels they're going to use the 1993 SIPP. This is very confusing. Since there is a redesign in 1996 why are THEY not going to be using the most recent SIPP panels to look at the comparability issues?

both Don and Howard touched on the issue of sub-annual monthly -- the need for monthly longitudinal data, so I will skip that part, but I do want to underscore that that is of critical importance and we know from prior research that sub-annual interviews are a priority. If you want to look at the dynamics of economic well-being, many of these transitions occur sub-annually.

Now the Census Bureau plan appears inefficient and costly for data users both inside and outside the government. Not funding a true transition from the SIPP may actually cost the federal government more than if continued funding for the SIPP was maintained. Many of the agencies, and what Howard just also noted, use the SIPP in their micro simulation models. Depending on how long this time frame is before we have a new panel they may have to adjust their models in the interim, costing a lot of money, and then they will have to be re-calibrated again. So there will be a two time change for those micro simulation models.

So one of our questions is whether or not that would be more expensive than continuing the SIPP in the short run to ensure that we don't waste valuable government funds.

And just a couple more notes. The plan that has been put forth so far does not make new data available for policy relevant research for many years. The prototype will not be fielded until 2008, that is -- that was my understanding -- and then it will be at least a few more years before researchers understand the differences between the new survey and the SIPP. If all goes well we'll be able to use this data for policy relevant analysis that can be taken seriously by maybe 2010.

But that's only if there are no glitches. As we know, many of the things that they are trying to do are new. We don't know yet how well this is going to turn out. This is an inexcusably long delay and could be easily solved by fielding the SIPP while the prototype is being developed.

It's my understanding that this is a highly unusual event, that typically -- when they re-engineered the SIPP in 1996 they continued to field the survey. When the Census worked to implement the American Community Survey they continued to do the census. This seems highly unusual and we should do something to ensure that we're not jeopardizing our ability to do policy relevant work.

In conclusion, whether or not the Census Bureau continues the SIPP while they study the possibility of a new survey is a critical public policy issue. The SIPP is designed to examine policy changes and without it we will not be able to ascertain the impacts of current and future budgets on U.S. families. This is an important issue, especially because we've made such broad cuts last year and we're planning to this year.

The Census Bureau is putting at risk hundreds of millions of dollars in prior government investments, along with private investments of course, but without having a viable and importantly tested plan for replacing the SIPP. With so much at risk and so little certainty about the new survey the Census Bureau should not eliminate the SIPP until the prototype has been fully analyzed and peer reviewed. We cannot wait years for data on economic well being. Thank you.

(Applause.)

MR. REAMER: Heather, thank you. Thank you to all respondents. Why don't we open it up for conversation.

First I want to do a quick poll of the room. How many people here are now or have been active users of the SIPP data set, SIPP data file?

(Show of hands.)

MR. REAMER: Yeah. Great, thank you. And how many of your are active consumers of other people's analyses of SIPP.

(Show of hands.)

MR. REAMER: Good. So almost everyone in the room is covered.

A couple of rules here. One is this is being transcribed, so we ask everyone to speak into the mike and also say your name and your organization. And two, we're going to -- we have about 90 minutes for conversation and we have over 40 people in the room, so you do the math. So be concise please. I'm happy to open it up for comments to David's presentations, questions, responses to the respondents.

Ed. Press the button.

MR. SPAR: Ed Spar, Copath. Would you elaborate on the phrase or words sub-annual?

MR. JOHNSON: Yeah. I think we used sub-annual because we don't know exactly what the needs are. I think in these decision matrixes that we're going through, and I can sort of address Heather's question as well, we chose the >93 longitudinal panel because that was the lsat time we had a full panel and there's a smaller set of questions.

So we thought, well, let's start with a basic set and add to that. So it wasn't that we thought it was the best or that it was a smaller set.

In terms of annual, if you go through that we can ask some things quarterly, some things monthly, some things every -- so we used sub-annual because -- I'm assuming not everybody needs everything monthly, but there are some things we need monthly. That's why sub-annual.

MR. REAMER: Cindy.

MS. TAEUBER: Cynthia Taeuber, University of Baltimore. I have a couple of questions.

Have you developed any idea yet of what a budget would be like for this, the rough idea of what you've talked about, the survey plus the admin records?

MR. JOHNSON: Well, we have a budget for >07 that's out there and we have proposed budgets for >08 and >09, so we have those budgets. I can't really -- I don't think I can give you what the numbers are on those budgets, but we've tried to evaluate what the costs will be.

MS. TAEUBER: The basics of the question really is the SIPP is not just the SIPP alone. The Census Bureau has an entire data set. And so -- and part of this came about because of the impact -- the Census Bureau having to make choices among critical things for this annual, the American Community Survey, the economic data. How will this fit into -- everybody here is talking about really their interest in one survey not the rest of the data set from -- that the Census Bureau puts out. How is the Bureau looking at that issue now in terms of the priorities?

MR. JOHNSON: Well, I think we're committed to this program, okay. We're committed to the dynamics of economic well-being. So we've put in money that we think is sufficient for >07 to do development for >08, to do development for >09, to do data collection. So that's -- you know, that's where are plans our. This is the plan and we think the funding is sufficient to provide this. Obviously we're not -- it's not as much as -- one of the key things, it's less total dollars than the current SIPP. I guess I can't address the key priorities of --

MS. TAEUBER: Well, then once the SIPP --

MR. JOHNSON: -- the Census Bureau, I kind of just sort of stated for this program.

MS. TAEUBER: More specifically then my question is why -- it seems that including a survey is taken as a given in this plan and I was wondering why

-- has there already been an evaluation that the admin records will not meet alone, will not meet most of the -- many of the needs of the data users here?

And then I also am wondering related to that is what is the lowest geography level for what you're planning now with the survey? Would it only be national?

MR. JOHNSON: Do you want to answer -- do you want me to answer these as we just go along or should I build them up?

We haven't evaluated geography yet. Obviously a sample size has to be much bigger if we need state level estimates. We do think we could use, exploit some of the stuff we've learned on modeling for local area estimates. That might be an idea. We haven't evaluated that.

MR. REAMER: Yes, Ron.

MR. HASKINS: I guess a follow up on this. Let me answer the question because it seems obvious and I'm sure everybody already knows it that SIPP is among the lowest priority surveys at the Census Bureau. You had to save money. I mean, you didn't have a choice so

SIPP got hit. So now the question is, and I think at least my own concern not as a primary user but I, you know -- probably a tenth of the articles I read are in some way related to SIPP. So this is going to be a real loss.

So the question is are you really going to replace it, how serious is this gap, will all the new procedures work and so forth, and I think the point of Cynthia's question is when it comes time to field a full blown whatever you're going to do in 2009, 2010, do we have any assurance at all that the money is going to be there?

MR. HOGAN: Yeah. Clearly I'm not going to say --

MR. REAMER: Howard, say who you are.

MR. HOGAN: Howard Hogan of the U.S. Census Bureau. I'm not going to say, you know, there's a guarantee that money is going to be there. There's Congress and a few other players in this game. From the Census Bureau's viewpoint in the next few years our priorities are going to be planning for the population census -- if this is a surprise to anybody. We also have another priority which is the economic census which is coming up in 2007. That's a very high priority in terms of where we want to go with our budgets.

But then I think right up there is the dynamics of economic well-being. It's maybe, you know, low relative to the population census or the economic censuses, which are, you know -- Census is our middle name. But the Census Bureau's leadership is very serious about putting forward proposals, initiatives, whatever to support this program. We're very serious about that.

MR. REAMER: To give people context, between the two of you say what the budget situation is for SIPP in >06 and what is going -- plan to be for >07.

MR. HOGAN: Yeah. From the money, the direct Census Bureau money for >06, I believe it was 32 million and for >07 about 9.6 million. And then our proposals are for it to start to grow back. So we're putting in a little bit more money in >08 and >09. It's not going to go up to 32 million, but it's not going to stay at 9.6. There's obviously some budget negotiations between now and then, but we are -- we the Census Bureau leadership are very serious about requesting and pushing forward the money to bring this survey back.

MR. REAMER: For >09 what range are you thinking about?

MR. HOGAN: I can't remember.

MR. REAMER: Okay.

MR. HOGAN: Is that a good answer, Cathy?

MR. REAMER: Thank you. Other questions please. Do we have a mike for --

MR. PRIMUS: I guess my question is, just following up on this --

A PARTICIPANT: Give your name, would you please. Who are you?

MR. PRIMUS: Wendell Primus.

I want to follow up on that. You know, I take your word at face value and, you know, I respect what's been said, but I guess from my standpoint this is a very politically risky strategy because you're going to end the survey and then start something that's new and different that's going to build up in funding. It seems to me unless I hear that political leadership is dedicated to this, it seems to me it is a very risky strategy to end something and then start something up. It seems to me a much better strategy would have been to say we need to make some improvements in SIPP and keep it going and keep it funded and then do some of the things you want to do as opposed to ending it, starting it up, and then having to fight a budget pattern that increases with time.

A PARTICIPANT: Did we get your name?

MR. PRIMUS: Yes.

MR. REAMER: Other questions please. Ed, you've already talked. Anybody else back there?

MS. PRYOR: Barbara Pryor with Senator Rockefeller's Office. We're one of the congressional leaders that are questioning the end of SIPP and I think part of it is, following up on Wendell's point, the timing and the time lag. There are major cuts and changes going on in social programs right now and if we have a three year disconnect how is my boss going to make judgments? How is he going to have the analysis he needs to do the right thing and craft the budgets and do the votes?

It's disturbing and scary to us. We don't have good information on child well-being. We need a lot better and maybe this prototype will deliver it in 2010, but what do we do in between? We have re-authorizations, we have policy decisions that have to get made, and I have a Senator who doesn't like working in the dark.

MR. JOHNSON: Can I try to address that one? One of the issues that we've always heard about the SIPP is the lag in the release of the data. For instance the 2004 wave one has been released, but we don't even have wave two and now it's 2006. So I would expect that you wouldn't be able to do a lot of analysis of changes that occurred in policy in 2004 until well into 2007. So even if we didn't cut a gap it's not clear that even if we continued you wouldn't have some of that information until -- at a lag time. So what we hope is that this thing will help us to be more timely in releasing that.

I understand that there's an issue with the gap in the data, but again we have to remember that the analysis that you could with the SIPP came at a three year lag anyway.

MR. REAMER: Okay. We've heard a -- go ahead, Heather.

MS. BOUSHEY: While I respect the fact that the first wave from 2004 was just released, your schedule says that the rest of the 2004 data is supposed to be released in the next two or three months. So it actually is quite likely that we will be able to do an analysis of the SIPP for 2004 before the end of the calendar year. And while the SIPP does have these lags, those of us who get the data the day it comes out on the Census Bureau website know that the lags are typically around 18 months, not around three years.

But just also just to second Barbara's question, that it is -- there's a time lag for the SIPP but there's also -- I mean, there's two questions. Will the time lag for the new data be any better if we're matching to administrative records where there may be a time lag where you can get those, (a), and then (b) there's also the missing years of data and you didn't address that at all.

MR. JOHNSON: I can't address the missing years. I mean, the survey ends in 2006 and the survey -- we will not have enough funding to do both, fielding a survey and planning a new survey. So the idea was to stop and start again.

The idea of the prototype, albeit where we can't provide everything to everybody, we thought that that would be a way for users to understand what we're doing. There's no way it can fill the data gap. I understand that that would be an issue.

MS. BOUSHEY: This is going to probably sound like a very naive or politically incorrect question, but I'm -- I just want to understand the extent to which this is really just a money issue versus a something else issue.

Suppose for example that you could put together a group of users within the government or outside of the government who could come up with the 32 million or whatever the regular SIPP costs, would you take it?

MR. JOHNSON: Howard?

MR. HOGAN: First let me say the Census Bureau leadership is the one who, you know, proposed this given our overall budget situation and given where SIPP was, and it was driven by the Census Bureau's priorities of the housing and population census, the 2010 census, the American Community Survey, and as I said upcoming the economic censuses. So it was driven by our priorities to produce the best data we could given an overall budget. Period.

We did when we found out that there was less money than we had hoped, Nancy Gordon and I did go around door to door, hat in hand, at least within the federal community, to see what we could do to continue the data collection. We did not come up with enough money to do that, but we certainly tried our best to do that given not just the Census Bureau's overall budget situation but other statistical federal agencies.

So I can state that, you know, it was a -- it was driven by our priorities given our budget. It was not driven by any outside political influence that we should stop getting information about poverty or something. That was simply not part of the conversation.

A PARTICIPANT: But maybe this is like discussions with Iran. Now that they are -- we're getting close to your developing your enriched (inaudible) here there might be more willingness to spend some money amongst some of the users.

MR. REAMER: John.

MR. ICELAND: John Iceland, University of Maryland.

I just have a comment in terms of substantive content. I just urge that the focus should indeed be on economic well-being. So when you're thinking of what to keep and what to cut -- for example, you know, there are other surveys that look at labor force, other surveys that even look at health insurance. I suppose people here may disagree on this focus, but I do think that what we're lacking in terms of data infrastructure is like a post-tax, post-transfer well-being measure, and even the current SIPP doesn't even do a great job of it partly because of the types of transfers.

Government transfers are no longer, let's say in the cash type transfers, but have moved to different types of programs. These are challenging to capture and, you know, survey instruments have to be changed to be able to capture these new forms of assistance.

So along these lines, too, in terms of when you're thinking of putting together a survey instrument I think it's important to use a consistent survey instrument, at least, you know, when you're talking about the household survey side of it, because as we know when the SBD was implemented it was very difficult for users -- not only for users to use it but I was at the Census Bureau at the time when the data was being produced and it was a very big challenge even to process it. So it caused big delays in putting out the data.

So at least when you're thinking about survey instruments I don't think you want to mix and match instruments. I think that will cause a lot of problems.

And then just finally in terms of data quality, I don't know what the answer is but with all surveys there are problems with response rates and attrition and somehow this -- somehow it has to receive high priority, maybe at the expense of other things, but in the end if -- if you're not even collecting the high quality -- you know, if you end up having sort of biased data then why bother collecting it. In some ways that's a question that it raises.

MR. REAMER: Ralph.

MR. RECTOR: Yes, Ralph Rector from the Heritage Foundation.

Hopefully short comments and they deal with the data gap, being clearer about the mission objectives, a response to what data are important, or at least a partial response, and also some observations on the planning document, information that was provided in the memo.

First of all with regard to the data gap, I think that to characterize this as part of an evolutionary process or part of a life cycle process, well that might be accurate if you think of the large meteor that striked the earth and ended all life above the surface and ended the age of the dinosaurs. Yes, then I think probably it might be an evolutionary process.

From a data user's point of view I think this is a data crisis and I think that just has to be acknowledged. 2008 we've got the first prototype if things go well. 2010 we've got maybe the second prototype. We're well past 2010 before we really have a useable survey that could I think for a broad range of policy issues be used in the way that SIPP has been used, and that's a huge gap.

The second point, these things may be included in your mission objectives, but I think that you have mission critical objectives and you should be clear about them. A couple of them that I think should be there -- release of a public use version of this file, and that would be true for both the prototypes so that researchers can use it for testing purposes as well as the final product B- your project I think will fail if you can't deliver on that and I think it's important to be clear about that. That is a mission critical objective.

A couple of others that we've talked about or others have talked about, the sub-annual data. One that hasn't been talked about that I think should be included is the ability to use this file to compute poverty estimates using current definitions. It's critical that comparative analysis be available so that the richness of the SIPP --

(End side A, tape 1.)

MR. RECTOR: Looking at the poverty indicators possible with SIPP be, you know, fulfilled on but also the comparisons can be made to the current definitions. And so that element I think is critical to whatever the new survey has.

Now with regard to what data are the most important, again this is just sort of a partial list I guess. But I think that a more complete understanding, appreciation for economic well-being is essential. The economic well-being has both a monetary as well as non-monetary dimensions. In addition, monetary well-being is determined both by current income and by wealth, and so I think what this means is that we're looking at collecting information on income, wealth, non-monetary well-being.

Perhaps we can use a supplemental cross-sectional approach like we've taken with the modules, but I think that it's important to build that in at this stage of the survey design so that monetary well-being includes wealth, asset and liability data and that economic well-being include both monetary measures as well as non-monetary measures. Material well-being, such as the quality of housing, measures of hunger, food and security, other data such as that that was collected in the adult well-being module.

A final comment has to do with the planning. There's two parts to this. I think it's important to include steps in the plan to secure agreements with other government agencies that will permit the release of public use of micro files that can be used for public policy research. New agreements may require the adoption of new or revised regulations and perhaps even legislative changes. You've had some success apparently with the release of the Medicare information, that's encouraging, but prior experience has shown that this can be a long and difficult process, particularly if you have to work with the IRS. Survey planners should include this in the phase -- in their -- as a phase in their planning process. In other words, they shouldn't focus exclusively on survey methodology.

The other planning point that I have is that you should take account of the data integration, and I think this follows up on a previous comment. Potential problems can arise when data from different sources are collected for different reasons and combined in a single file. Data acquisition and processing must view the final product as an integrated system, not just a collection of stand alone items.

So for example you've got data from different time periods. The IRS data is on an annual basis. We're talking about collecting information on a monthly basis. How are you going to combine those? You've got reconciling multiple sources for the same date.

For example, wage information is missing in the IRS database when an employer does not report wages -- the 1099 -- and the employee does not file a tax return, but this is precisely the situation for many low wage earners. And so I really think that you can't treat this as a stovepipe data issue. It's not just your getting information from a variety of sources and putting it all together, but really thinking about how it's going to be integrated.

MR. REAMER: Julia Lane, I want to hear from you. Can you talk a bit about -- yes. I warned you beforehand. I warned you beforehand. Julia Lane is with NORC, an opinion research center, and formerly with -- involved with the longitudinal employment, establishing household dynamics program at Census and familiar with the uses of administrative records and those kind of linkage.

MS. LANE: I think Ralph made a lot of the right points. Users need to be intricately involved in production of the data. It can't just come -- and I've talked to both David and Howard about them and they do understand users need to be involved at every step of the way.

I don't -- I think it would be a mistake to underestimate the difficulty of putting together administrative records from a variety of different sources. Don Oellerich and Susan and Colleen from ASPE were intricately involved, Ron Haskins too, in the LEHD -- in the development of the LEHD data set which was an employment insurance wage record data. We started in 1998 with the State of Maryland and then I think it took us two or three years to get up to five states, and then right now LEHD and (inaudible) may have counted better than -- oh, Nancy --

MS. GORDON: 41 states.

MS. LANE: 41 states and the District of columbia. So it takes a very long time and getting the agreement of the federal agencies was a non-trivial task as well. To get the HHS (inaudible) MOU signed took us two years, didn't it Don? To work with IRS to get the agreement to get the detailed earnings records matched to the CPS in the SIPP took Howard two years.

MR. HOGAN: At least.

MS. LANE: At least. And that -- and many hours a day. So it's a non-trivial task to put them all together and the issue is figuring out what the authorized purpose is for all those different agencies, because you can't just put it all together and put it in a pot. It has to fit the mandate of all the different agencies. I think it -- I actually think it can be done in some form if you work very, very hard for a very long time and have someone who champions it and gets (inaudible) and talks to the users and the producers and can do it.

But I think 2010 is an optimistic date. You might want to start with three companies or states, and we've had this discussion as well. California, Florida and Illinois would be states that have existing data agreements -- you've got Tom McCurdy having data for California, you've got Chapin Hall with data in Illinois, and we have long relationships with Florida because of the ASPE connection, and see what can be put together.

But I am -- what you would want to do is put all the data -- put the data sets together and then get a battery of users to do the analysis on the data that were put together through the administrative records and whatever survey you're matching up with, the SIPP data, and then have the users figure out what differences are there in the outcomes and how might things change -- what's changed as a result of having administrative records rather than the survey. And that goes precisely to the point that you (inaudible).

So you need to have users and producers engaged at every point.

Then the last thing and then I'll shut up --

MR. REAMER: I knew once I got you going you would --

MS. LANE: Yes. You kind of caught me on the hop because I was nodding my head at Ralph and then it was like oh, my God.

But the last thing that I would say is that I did mention that it was so important to get the authorized purpose component in. And so whichever agency that one is working with to get the data you have to think what product can you produce that will make it a value to them. So you have to find the intersection of the interest. So if you're working with state data one of the things that we found was critical in order to get them to ship data to us on a timely fashion, you had to give them something back in return that they wanted to use in a hurry.

I think we've got the key agencies here, obviously HHS and SSA on the federal side, but thinking about from the state side which -- and they're the ones that often have the sub-annual information -- thinking about what you can give them back so that you address that timeliness issue as a major issue.

Having said that, I actually think that it's really -- I think it's really important to have the administrative data component firmly in place. It's really hard to do, but if you take a look at the work that Bruce Myers has done at the University of Chicago and Bob Gurgett (phonetic), a real problem with the estimates that you get off the survey records is that you get one weighted -- when you weight it all up you get the estimates of the usage of different types of programs off both the SIPP and the CPS that is way lower than the estimates that you get from administrative records and it's not stable. It's volatile, and it's unstable, and it varies by program.

So even though I conclude that there is an issue with losing the SIPP, it's not clear that the survey responses were getting at what you wanted to have anyway. And so having -- think of the plus side of having the administrative records is that at least you're measuring what you really want to measure instead of measuring what the survey respondents are telling you which can be very, very different things. So it's kind of half glass empty, half full issue.

So those are things that were coming to mind.

MR. REAMER: Terrific, thank you, Julia. Julia put out a lot of things for people to respond to. Yes, in the back and -- no, we need -- mike and your name please.

MR. SEARS: Jim Sears, Social Security Administration. Merging administrative data sources for an administrative purpose it works. Yes, it's tough, it can be done, but in my experience it can only be done for that administrative purpose, meaning when we have a matched data set I better be using it for that purpose, not sharing it publicly, using it for other purposes.

So the question here perhaps for you Julia, perhaps for David, is realistically how many administrative variables do you expect you will ever, ever be able to release publicly? In our experience we've gotten up to maybe three.

MS. LANE: Since you happened to ask, so Ralph got the absolute critical issue which was data access, because if you don't think about new and creative ways to access data you really are caught, and with all due respect you can't expect people to go to the Census Bureau to work on the data. I think synthetic data has a role, but it's a very niche role.

So where I would like to push the Census Bureau to go is to think of non-statistical ways of protecting the data but releasing it, and this is where Pat Doyle was going very much, and think about remote access procedures. So the thought that this is not feasible in this day and age is -- doesn't make any sense.

So at the national -- I appeal to the Census Bureau to look at the work that the National Science Foundation researchers have been doing in ways in which financial institutions, the CIA, the Defense Department and so on provide remote access to extraordinarily critical data without people physically having to go on site. The Joint Chiefs of Staff, when they want to know about troop movements in Iraq don't all go to an enclave and look at it. What they do is they get on their website that is confidentiality protected and they're in separate areas so they access information about troop movements in Iraq remotely.

I don't see if they can do that remotely in a cyber protected environment why Census Bureau data cannot be accessed remotely. And I know what your answer is going to be, but I would very much push in that direction.

MS. GORDON: So in my new role at the Census Bureau I get to play on many different topics, which is absolutely wonderful, and one of them is remote access because we are actually working to expand our abilities to allow outsiders to have remote access to our data. But I think there's a critical difference between the example that Julia just gave of the Defense Department where the people there are getting access to data that they're allowed to see.

We're talking in the statistical system with these administrative records, with not being able to see the stuff that you're not allowed to see. That's the whole point. So what we're looking at is remote access systems where the analysis comes into the Census

Bureau, is run in the Census Bureau. The results are subject to our disclosure, avoidance, assessment, and if they pass that are returned to the user.

So there's two more points here. One is we're working on some remote access systems where the disclosure review is automated, and believe me that's what we need because the cost of human beings having to do this is extraordinary. So that is one dimension.

That plays into the comments about synthetic data. Synthetic data that you would use and you would trust the results and you'd go off and you would base decisions on them, I'm not sure that we're going to get there in my lifetime never mind my professional lifetime. But synthetic data that's good enough for the researchers to basically do most of their analysis so they're not flooding our system over here, which is the remote access system, that is quite conceivable.

So now the researcher is getting results out of the synthetic data that's good enough for them to do their work and then when they're, you know, pretty far down the road now they come and use the remote access systems where their analyses are run on the confidential data and the results are returned to them. That combination of approaches I think really does have potential. It is in that direction that we've been working, both on the synthetic data front and on the remote access front.

MR. REAMER: Thank you, Nancy. I had David, Howard and Kris all wanted to speak. Go ahead, David.

MR. JOHNSON: One thing. This is a two pronged approach. I mean, it could be -- let's take the premise if we had access to administrative data records that we could release in a timely manner and made them public you would say we should do that and we should evaluate how it compares to the survey data, okay. Obviously if it's something that we've got a three, four year lag and we can't release it, we may have to ask those questions. I think that's why we're investigating.

We think that there are data -- there should be data out there that we can use and -- access and use integrated with the survey data that we can release. That doesn't mean we can't do the survey. We have to -- we're working at both at the same time. It's not all admin records and let's figure out a survey, or all survey figure out records, this is let's do both and figure out what we can release integrated and plan it in such a way that the questions we asked we know we need and we can't get somewhere else and we can't release it.

So it's not -- it doesn't rely completely on the use or non-use of administrative records data.

MR. IAMS: I want to clarify one aspect of what Jim Sears was referring to. This is Howard Iams at Social Security.

Social Security Administration has access to some data that it can use for administrative purposes in administrating particular law, but the research office is not permitted to use that data for any purpose whatsoever, such as statistics or research, because there's no legislative authority for the Internal revenue Service to let us do that.

So I think that there is an issue about the use of the data that you get and what type of use it can be applied to, and this may require some sort of legislative change or permission which probably won't occur quickly.

MS. MOORE: Kris Moore from Child Trends. I think we've struggled for a long time to get a notion out that well-being is more than economic well-being, that it includes the well-being of families and children. I'm concerned that progress is going to be lost. I mean, I'm hearing tight budgets and I'm hearing enormous difficulties with the administrative data, and there's a brief allusion in paper to child well-being data, and I'm wondering if somebody could respond and describe, well, what is being planned, what might happen in terms of collecting data on the well-being of children in particular.

MR. REAMER: Howard or David?

MR. JOHNSON: I think we're looking at the survey to determine the dynamics, okay, of income and program participation and we're modeling this to then allow for these hooks that go along to other aspects, for instance child well-being. It's not clear -- when we look at the uses, a lot of people use the child well-being modules but they don't necessarily use it in a longitudinal framework. So it could that this is one of the things that we could partner with an agency to provide the child well-being.

Being an economist, and this is what I responded the one time when asked if we should extend that, I think -- I'm an imperialistic economist so I believe everything comes down to economics. So I think -- and I agree sort of with John, it's when you focus on these well-being things a lot of these things are economic, and so a lot of the things that I think are in, even in that module are related to child economic well-being.

Now there might be other ones that we'd have to look at, but obviously we'd have to look at those things and consider those things. But I think economic well-being is part of that. So child is in there.

MS. MOORE: Just to follow up. Certainly economic well-being is part of it. Could you just say what you mean by a hook?

MR. JOHNSON: Does somebody else want to take this one? I think what we're looking at is questions -- the idea is you have the survey and you have certain questions that can allow you then to subset another universe, okay. All the people who have childcare payments or all the people who have kids for instance then would hook to another, a supplement or something like that. We could just send the instrument to those particular people. So it might need to be part of this basic instrument that we're developing, but it could be added on at some other time in the future.

MS. MOORE: And that's being actively planned?

MR. JOHNSON: Well, it's being actively investigated. Again, this is all part of -- I mean, this is just what's been done in three months, so we're considering what different options we could look at.

MR. REAMER: We're only 40 minutes into this and there's no hands going up. There we are, here we go.

MS. FROST: I'm Carol Frost from CBO and I just wanted to kind of get back to something you talked about at the end, which is as users we need to start -- you look at your child well-being questions, what can you get -- what can you not use, what can you cut out, how can you make it a better topical module in a third of the questions. That's the kind of thing if you start looking at that and give your feedback to David you might get what you want. But that's the kind of work that only we can do, and I think that's part of the point of this meeting is to start thinking that way and I hope people do.

MR. REAMER: Now we're hearing something about -- from people, what's important to them, and so let's keep going with that. I mean, do other people want to talk about the things that are particular -- really important to their work? Kris, again.

MS. MOORE: Detailed family composition.

MR. REAMER: Anybody else? Okay.

A PARTICIPANT: I have another issue.

MR. REAMER: Go ahead.

A PARTICIPANT: The political part of this, if I were on Colburn's staff I would tell Senator Colburn that there's tons of data out there and there's a fabulous interagency panel that comes out with child well-being data every year and -- so why do they need this SIPP? There's a lot of other stuff out there. This is not really -- it's a bunch of researchers that just want to make sure that they can keep publishing articles. So what is the response to that? What actually are we going to lose from child well-being, thinking, starting with the, you know, this is ought to be an objective answer. What -- how many pieces of information from the 40 or however many measures it is are we actually going to lose if we lose SIPP?

I do think, Barbara, that there will be -- there still will be considerable information about child well-being even if SIPP does not exist. There will be a lot of information about child well-being, and depending on what you expect is going to go wrong because of budget cuts there will be other sources of data. It's not like the whole world depends on SIPP for information about child well-being.

A PARTICIPANT: I think two of the measures in American Children rely on SIPP data and those two measures would probably remain. I mean, one of them is composition. I assume composition would be something important, like Kris said.

Then there's the indicators of welfare dependence that HHS produces, and all those are economic well-being measures. So there's spells of program participation which I assume would be one of the things that wouldn't be included in this new thing. So my assumption of those particular items, we'd still have them.

A PARTICIPANT: I mean, to be straightforward here, you're going to lose them for some period of time, right? You're going to lose them for two years, four years. I mean, obviously we're not going to (inaudible).

A PARTICIPANT: Yeah. But I think if you look at America's Children the current measure in there is 98 -- or do they go to 2001 -- and if you look at the indicators of welfare dependence I don't know how far along they are in the -- which --

A PARTICIPANT: We're currently up to the 2001 panel and we're just waiting for the fuller -- the longer longitudinal data for the 2004, but those measures that we do use for that report to Congress rely on monthly resolution for three or more years.

So if we're talking about an annual CPS, an annual CPS, and then one or event two surveys of maybe at six months with retrospective monthly data, we won't be able to do those. So we'd have to, you know, take out the longitudinal component probably and rely on the PSID for some measures of annual transitions.

MR. REAMER: Kris.

MS. MOORE: I don't think SIPP has been the ideal instrument for indicators, you know, cross sectional indicators of children's well-being. But what you would lose is more the dynamic, the capacity to look at things like work, childcare, program participation, income, and how they relate to child well-being -- also family structure. It has been very important for that.

A PARTICIPANT: If I hadn't been fired yet from Colburn's staff then I would say, well, but then they have the PSID, so why did they need SIPP.

A PARTICIPANT: Because the PSID goes every other year. It has annual measures and very, very little sub-annual interactions, and that's what we're saying is the value -- I mean, that's what our claim is. Now what we can produce obviously is another story, but that's what we're claiming the value added of this new program and SIPP are the sub-annual measures so that, yes, you can get the indicators of welfare dependence. We would expect those to be outcomes.

Now what you can do for >07 and >08 and >09, I agree that's a reasonable question, but the value added are those sub-annual things and I think that's what SIPP and this dynamic (inaudible) well-being can provide.

MR. REAMER: Richard, Barbara, and then back in the corner.

MR. BAVIER: Richard Bavier, OMB. There was also in the American's Children 2005 a special section about family structure and child well-being and that was largely based on SIPP. And I think the answer to your question, Ron, is we think PSIDF -- and there are other longitudinal but more targeted surveys that the feds fund in one way or another, but the answer is you don't -- with the cross sectional survey you see what people are like. You need a longitudinal survey to see what happens to people, and that's essentially what SIPP provides.

MR. REAMER: Barbara.

MS. PRYOR: Well, I think it's the --

MR. REAMER: Barbara, say where you're from.

MS. PRYOR: I'm Barbara Pryor with Senator Rockefeller.

And just one example, and we all acknowledge there's going to be a gap. The question is how long and how well we bounce back from that gap and how well we compare. We just passed a welfare reform bill that we don't even know because we haven't seen the new federal regulations which are coming out on June 30th supposedly. That bill is -- those regulations will have no calming period and will take effect October 1.

The boss is going to -- is supposed to re-authorize it in 2010. What am I going to have for him without this type of data? And last time we did welfare reform I think, Ron, you probably know off the top of your head, 60 -- West Virginia was a 60 percent declining case load. This one has potential to have an even greater declining case load and one of the reasons why my boss is concerned about some of the discussion is it seems like we're going to administrative data.

In (inaudible) we're not going to have much because we're cutting the caseloads. So how do we figure out what happens to the families or as HHS calls them the levers? It's a concern if you care about kids and poverty.

MR. REAMER: Back there. Yeah.

MR. SHERMAN: Arloc Sherman at the Center on Budget and Policy Priorities.

I just wanted to flesh out that kind of example, just as one of many areas where the SIPP data I think are incomparable and we stand to lose quite a lot. You know, as Ralph was saying I think giving up comparability really means we give up the past. It really means that we give up the ability to learn from the past.

SIPP had and has data on a cluster of things that it's hard to imagine getting anywhere else. For example people are only starting to hit time limits at a rate of -- you know we're getting -- there are probably a few hundred people, a few hundred thousand who have hit time limits so far, but it's steadily accruing. How would we know how many families spend a couple of months with no work and no welfare from anywhere else or if it has significantly changed over time and if it has any correlation whatsoever with their material well-being? I think we would project different things to result from that, and the only way we'll be able to resolve that amicably is if we have a survey like SIPP.

MR. REAMER: Richard. Ralph, did you have your hand up? No, okay. And then Wendell. Richard?

A PARTICIPANT: (Inaudible.)

MR. REAMER: Oh, you did, okay. Ralph, go ahead.

MR. RECTOR: This is just a very quick comment, but I do want to emphasize the fact that we do -- our organization uses it as a longitudinal file and I think that the response to Colburn is that it is a longitudinal file and then at the other time dimension is that it's sub-annual. I think that that's critical. So that in combination with the cross sectional, the special modules -- PSID has a wealth module and it's good, but the PSID covers a different universe, it's for a different purpose, the size different, and a combination of the special modules that SIPP has in combination with the longitudinal and the sub-annual aspects I think really are the strengths that need to be emphasized. I know other people have said that but I just wanted to say that from -- we all want to reinforce that.

Also the fact that the family structure information is critical to us. We use that on a regular basis.

MR. REAMER: Wendell.

MR. PRIMUS: Wendell Primus, Representative Pelosi.

I think what I've heard here today is that we're here because you've got a budget mark back to the Census Bureau that was a lot smaller than they anticipated, and I guarantee you that if you looked at the trajectory we're on in terms of discretionary approps that's going to happen again.

But having said that, the good sense of staff then have scrambled and have given us a very enticing formulation of how SIPP could be re-engineered, but there's a lot of bumps. I mean, maybe this data matching doesn't go well. I mean, that's going to increase the political risk of having it fall completely apart.

But my final question, and you can comment on that, is if this data matching from administrative records is such good a thing to do, you know, why don't we do it with the CPS? I don't want to break up the longitudinal sense of CPS because it would be a big methodological thing, but again if we're being enticed here that, you know, this re-engineered SIPP is going to be so much better why aren't we doing it to some other surveys with the administrative records?

MR. JOHNSON: Well, I think what we're saying is we need to investigate this, okay, and the SIPP is the best place to do this. We've done some matches with other data sets, with ACS, with CPS, to see how this works, and it does provide some useful information.

So I think that's, you know, where we're going. In terms of how that's going to save us money in the future is something that we have to investigate. But I think that's -- that's what we're trying to do. We think this is something that has to be done. Why we haven't done it, you know, we've done some, but I think this gives us an opportunity to do some more.

A PARTICIPANT: Well, there's another obvious answer which is that several people here who have done it say it's really hard to do and it's very limited in how you can -- you know, the restrictions on use of the data. So that's another reason we haven't done more of it.

MR. REAMER: Heather.

MS. BOUSHEY: Heather Boushey, the Center for Economic and Policy Research.

I just wanted to go back to Ron's question for what we would tell Tom Colburn. We've used the SIPP data for a number of longitudinal uses that are important for policy questions, but I'll point to John Iceland's work first on the dynamics of poverty. We know when we look at sub-annual poverty that spells of poverty are shorter than if you look at annual poverty spells. That's a very important policy question and someone who is concerned about not over spending on government programs might be interested in that.

Second, we've used it to look at the dynamics of unemployment spells. The SIPP is the only data that you can use to understand eligibility for unemployment insurance. You can't use any other longitudinal survey to do that. So we've used it extensively to look at eligibility for unemployment insurance and the probability of re-employment and how UI fosters better job matches after unemployment insurance spells.

Finally the SIPP because of it's monthly capacity allows us to uniquely look at maternity leave patterns, which is actually quite fascinating, to look at re-employment spells after maternity during the panel. It's only three years, but you do have enough data to look at that and, by the way, the SIPP is the only data set that I'm aware of that actually asks women questions about their leave taking at maternity in one of the topical modules. And so it's the only data set beyond the special data that the DOL did that actually allows us to look at the Family and Medical Leave Act, which is one of the, you know, a widely used and very important policy measure.

So there's three areas of, you know, where we spend a lot of money on programs that only the SIPP will allow us to understand. One more that came up recently over the past couple of years is that -- was in the area of the school lunch nutrition program. There were questions about whether or not we were overspending on school lunch programs because there was some analysis with the March current -- the March CPS that showed that not enough children were in poverty. And of course when you look at sub-annual spells you find that, gee, many -- you know, if you've got your old fashioned, traditional household, let's say you've got a man, a wife and two kids, she's not employed, he loses his job, his $40,000 a year job halfway through the school year, those kids are eligible for school lunch halfway through the school year when they weren't, but it looks like that family earned $20,000 a year over the course of the year. So that's another important policy area. Thank you.

MR. REAMER: Connie and then Julia.

MS. CITRO: I'm Connie Citro with Committee on National Statistics at the National Academies.

I was involved in the early days with Pat Doyle, the early days of the development of SIPP, when it was the income survey development program and a number of years were spent at Social Security and ASPE and the Census Bureau on the development work.

SIPP is very valuable. I would add to what's been said about family structure, longitudinal, sub-annual, something that's been mentioned but not perhaps pulled out enough, which is eligibility. You cannot really understand program dynamics if you just know about participants, and particularly because actually I think it is true in surveys that the participant data are not as good as you would like. But you have to put that together with the data that will let you figure out who was eligible and not participating and so on.

I think SIPP has been very valuable and we definitely need something like SIPP. I do think SIPP has -- it's a tough -- it's a tough survey, not just for the respondent but for the Census Bureau. It has been a tough survey for them from the beginning with a very demanding processing cycle and schedule. I think truth be told that the Census Bureau didn't, for a variety of reasons never quite got on top of it in terms of the processing and we're locked into a system that didn't work all that well. So I do think that it needs to be re-engineered in a number of ways, but I do worry what's been said about if it's sort of cold turkey off and then a new thing and I do think -- I know the Census Bureau staff, Howard and David and everybody working them are, you know, wonderful people, they're working hard, but there are a lot of risks, a lot of unknowns here, a lot of testing that needs to be done, should be done.

In an ideal world, yes, I guess we probably can't get SIPP to continue at its full bore and have money to do re-engineer, but is it totally off the table to think about, you know, cutting the current SIPP sample size or something or doing something so that you would have -- and then have the money for re-engineering? Because I also think the re-engineering is going to be challenging. I think it's going to be fascinating for people working on it, but it's going to be challenging to get a product out. I think there's a real risk that under the pressure, the understandable pressure to get a product out, that again we will get locked into some less than optimal procedures and designs and so on and there will not be then the ability to fully move forward that you would have if you could fund at least something, some part of the current SIPP and have enough money then for groups off to the side who are working on the new thing.

I do think a new thing is needed, but I do really worry that this gap is going to be a problem. Of course you run the risk in this climate of losing the political support for anything at all. You know, that's the -- there are lots of people out there who really don't want to look at data, they just want to do whatever they want to do.

The other thing is whatever we come up with for replacement -- again I'm thinking of Pat Doyle. You know, the one thing she really pushed was the methods panel, and again before you go totally back to >93, yes, maybe it had less content, I haven't followed the details, but she put a lot of work into a methods panel about trying to get better data in several areas. I think if there had been the methods panel research component going along with SIPP over its history we might not have quite wound up quite where we -- quite where we are here.

So I would hope that there will be a constituency among the users to support the Census Bureau to -- even when they get to where they think they have something that's good for production that, you know, that they've always got a methods panel and a bunch of research, a bunch of research going on, because again the world changes.

There was serious work in the 70s that said, yes, asking more questions got more income reported. That's why we've got, I don't know, 60 questions on income or something and 50 on the CPS and so on. The world has changed in terms of what people are willing to tell you and so on, and it may be that that's no longer -- that's no longer the case. It may be that our beliefs that personal interviews are, you know, the gold standard to start out with this sort of thing may no longer be true. It might be we could be creative about not everybody has to answer every questions that address -- you know, I've heard about child welfare, I've heard about unemployment spells for maternity leave. You know, maybe there would be more creative ways of breaking up the questionnaire to get the burden down on any one household. There are lots of things to think of.

The problem is of course we're faced with constrained funding when in fact we need more funding to keep something going and to really do the kind of very intensive design effort that led up to this survey in the first place, and it was intensive and it was a big, big effort with a lot of field tests and so on.

So I'm not quite sure what my bottom line is except that I am nervous about it, I am hopeful that, you know, everybody in this room will be working together to try to move it forward because we clearly need -- we clearly need the data, and that even though we may be upset that, you know, here's this path where we're cutting it off and redesigning something we're not sure about, that at least we try to keep pushing -- you know, that we do try to keep pushing forward.

I do think in the future if we can solve some of the data access problems, that administrative records for some of these variables are a good thing and ought to be considered for all of our income surveys. But my biggest worry is I think the Bureau is probably right now over promising, and I know why they have to do it. You know, they can't come in here and tell you it's going to take ten years to get something suitable back on the table, but I do think there is a big challenge and I'm hoping that resources can be found so that they can, you know, put on a full court press about their development and also so that again we don't get locked into, in the rush to get something out we don't get locked into something that's -- that then people say, oh, we can't give you any more money to keep developing it.

MR. REAMER: Thank you, Connie, very much. Julia?

MS. LANE: I just wanted to correct the record. The SIPP was not the only source of information on employment and unemployment labor market dynamics. I thought David was going to jump in there, but obviously the BLA (inaudible) the national longitudinal surveys of youth, right? So they have SIPP type questions and financial asset questions like that from the survey touching on finances.

So Kris might be able to add something on the (inaudible).

MR. REAMER: Cindy Taeuber.

MS. Taeuber: This is Cindy Taeuber. This is related to what Julia is saying. I personally don't see that the Bureau had many choices given the very large budget that they were -- the budget cut that happened. It seems to me that rather than just hanging onto the old ways of doing things that this is a real opportunity to push using admin records which are longitudinal, which provide data at lower levels of geography.

I have and others have looked at the quality of data on surveys in terms of program participation. It's not good. We did a study on food stamp comparing Maryland administrative records to a survey. The survey got half, half of the food stamp participants saying yes. That's -- you're getting numbers, but what are you using, what are you getting? Public assistance, the same problem.

So this is a real opportunity to use a much less expensive source of data. It is difficult.

Julia, I'm surprised she's still living with all she had to go through --

MS. LANE: (Inaudible.)

MS. TAEUBER: -- to get that labor force data. But when you're talking about periods of not working, periods of not welfare, those are both sources of administrative records that we could really be following month by month over the years, not with the long data gaps.

So I -- what I think is that researchers should start pulling things apart. What things can you do right now with administrative records or with pulling it togther and what can you absolutely not.

MR. REAMER: Howard.

MR. IAMS: Well, I'll tell you one thing you can't do with administrative data, at least for SSI, is now who is eligible for the program and isn't participating, and that presents a difficulty. However, I will put in a plug for our retirement model at Social Security where we're projecting what happens to the baby boom when they reach retirement, and it's being used in planning reform of the Social Security program with the information going to the executive branch and to the legislative branch. Since we pay one in six Americans checks this is not a trivial amount of money nor a trivial effect when deciding what the reform might.

A PARTICIPANT: Just to follow up on that. I don't know specifically that case, but for example food stamp eligibility, Mathematica (phonetic) has done a model that uses both administrative data and survey data. We just did one that used administrative data, and American Community Survey you get a much simpler method and get roughly the same results. Again I think it's an area of research that we could really be pushing ourselves on.

MS. RATCLIFFE: Caroline Ratcliffe, the Urban Institute.

I wanted to follow up on something that Connie said and also what Don said in his comments. I guess the question is, what had been asked was is continuing some part of the SIPP off the table or is it still on the table? Don had mentioned maybe following up with one of the rotation groups. So, I mean, is it decided that there will be nothing or is that still a possibility?

MR. HOGAN: Is it still a possibility? Yes, a remote one. When looking at the possibility of continuing collecting SIPP under the old way there's the fixed versus variable costs which play themselves out in two ways, one of which is simply the field work. Whether you're doing a quarter of the SIPP or one-tenth you still have to program the instrument, hire the field staff, supervise the field staff, retain a processing staff, and pay that. So the cost savings of cutting it as opposed to doing nothing is a big, big threshold effect.

In addition to just the dollar savings we have a corps of headquarter staff experts on the dynamic of economic well-being, and if we told them spend next year trying to run a mini SIPP, process a mini SIPP, spend your time on that, that drains their energy away from trying to make the new thing fly. So our goals are to really make this new thing fly. If we can do that and we can retain some data collection from the old SIPP that would be marvelous, but I've pretty much been, you know, been fairly clear that focusing our attention on getting this new thing off the ground and not trying to jeopardize it.

So that's where we're coming from. We're still looking at the possibilities, but I haven't been able to figure out how to square that circle.

MR. REAMER: Linda.

MS. JACOBSEN: Linda Jacobsen with the Population Reference Bureau.

I wanted to follow up a little bit on this discussion about potential partnerships at the state level, and you noted David in the memo that we got that there is not currently, at least in what you put there, any national level source of Medicaid data that aligns with SIPP except for one that has a three year lag in obtaining the data. Then you noted under demographic characteristics that many programs are administered at the state level and there are limited state data to address receipt and we'll look closely at partnerships.

So I guess I'm interested in two things. One, what specifically are you planning to do with respect to Medicaid, which strikes me as a rather critical program, you know, both for -- well, in many age ranges, and also it's hard for me after having followed LEHD and listening to Julia's comments to think about establishing state partnerships as being a cost effective way to get information. So I wondered if you could comment a little more on both of those.

MR. JOHNSON: I'll say a couple of words and then Julia and Ron might be able to add. Again if we can't get timely Medicaid data it's not going to be, oh, we're not going to collect the data. It would then be we would have to include that on the survey. So Medicaid was the easiest one to do a first prototype in the time frame that we had, but I don't -- but in terms of state partnerships you're right, it would be a lot of effort. But I think it's something that needs to be investigated. I think Ron can speak to some of those partnerships.

MR. REAMER: Ron Prevost.

MR. PREVOST: I'm Ron Prevost, Census Bureau. I just wanted to let you know that relating back to a couple of things that have been said here, we've worked on a number of studies, I know with Cindy, also with

Bob Gurga (phonetic), that we are actively looking at studies that both measure eligibility as well as participation. We've been very successful in working with states on this. Yes, data access is an issue. It takes a lot of effort in which to secure the files.

And, you know, we keep talking about administrative records and matching issues, matching issues aren't the problem. The technical issue has been solved. It's the data access that is the real issue and the ability to use the data files properly.

From the Medicaid perspective we have been working very closely. We have formed partnerships with folks at ASPE, with folks from Centers for Medicaid and Medicare Services, with folks from the National Center for Health Statistics, on a very specific project that's looking at the quality of the survey data versus the quality of the administrative records that are coming in.

Yes, we are looking at this time lag. We are looking at both information on enrollment from the MSIS (phonetic) file as well as information on pay outs from the max file. We are looking at ways that we can perhaps shorten the time that it takes to process all this information because currently right now the information is collected centrally through the Centers for Medicaid and Medicare Services. They have to do a number of things and one of the -- and in order to get the files prepared and sent on for their use and for others use.

One of the things we're looking at is can we use data earlier on in the process. And these are the types of things we have to work together, partnerships with federal agencies that can -- I think one of the models might be you can't go with every state agency for every file. The model we have to look at is we have federal stakeholders. How can these federal stakeholders garner the partnerships of their state entities that are administering the programs in order to come up with a better process for bringing the information into the federal sector, otherwise you'll spend your entire life doing nothing but contracts and Julia and I have -- she gained weight and I lost hair doing it so --

Thank you.

MR. REAMER: Thank you, Ron. Julia, did you want to add anything?

MS. LANE: That's not polite, Ron. I know, I know.

No, I think you said -- I would go with a model -- or again I'm an economist so you do a benefit cost analysis. You go with the big states and do a proof of concept and see what you can produce for the big status. But, you know, I said this before, you figure out where -- we can sit here and talk about it all day, but you need to be doing the work in parallel, you need to be looking at the administrative data, figure out where the gaps are, figure out what uses can be put to the combined file with the new proposed new file and compare it with the analysis on the old file.

I agree, I'm very concerned that the old SIPP is going to disappear, so you're not going to be able to make those comparisons and that's seems to be a real problem. But you're not going to know what can be done until you start pushing the envelope. I don't think -- I've never thought that administrative records alone can replace survey data and it would be the combination of the administrative and survey data, so you figure out what we can the administrative record support and then plug the hole using the survey data. Or, alternatively, start from the survey and say, you know, what can we afford to do on the survey and then plug the holes with the administrative records. One way or another, but you need to know what the holes are and the only way you know where the holes are is by doing the analysis that the different agencies need and researchers need done.

MS. PRYOR: While I understand the economic point and the --

MR. REAMER: Say your name again for the transcription.

MS. PRYOR: Barbara Pryor with Senator Rockefeller who is the Senator from West Virginia, a small state.

And if you --

MS. LANE: An important state.

MS. PRYOR: And if you look at the Senate Finance Committee, which the chairman is from Iowa, the ranking member is from Montana, and then we have my boss and Kent Conran and Orrin Hatch, there's also a political dynamic here. It's a political dynamic and if you want your survey and you want to build it back up with money you've got to ask rural what's happening.

MS. LANE: Okay, let me rephrase it. Big states and important states.

MS. PRYOR: Because there is a huge dynamic.

A PARTICIPANT: The total is 50 I believe.

MS. LANE: You should have stayed with the British system, dammit.

MS. PRYOR: But this is part of this dialogue and, you know, I think it's a good dialogue and I am delighted to be down here, but I hope you're going to schedule another one on the Hill because there are many congressional staffs that should be part of the dialogue and should be part of the debate. I just happened to have a generous boss who lets me take, you know, three hours out of the office to come down. But it is a very important date and if you care about -- it's not only the politics of the small states, there's different policy. Welfare reform in New York City is very different than welfare reform in West Virginia.

The same with CHIP, the same with child welfare, and we really do need to figure out a way to reach all of the areas of our country. It's very important, especially if you want your money.

MS. FROST: I second that for health insurance. You have to --

MR. REAMER: Carol --

MS. FROST: I'm sorry, Carol Frost from the Congressional Budget Office. We do micro sim models. You've got to have state information, not just big states and not just important states.

A PARTICIPANT: (Inaudible.)

MS. FROST: No. Then the model is no good until you get them all in there. It just has to --

MR. REAMER: Julia, use your mike.

MS. LANE: I'm sorry. The problem is is that you've got to start -- so your point is well taken. So the problem is is that you've got to start with a subset of states --

MS. FROST: Well, let's back up a minute.

MS. LANE: -- and show that it's viable.

MS. FROST: Okay, but we have data sources that are national already gathered from the states. The MSIS data which does come out earlier if you get it before they clean it the fifth time, because we get it earlier. You've got your SAF data. You've got a lot of data sources that are already covering all the states. I say start with that.

MR. PREVOST: Ron Prevost, Census Bureau, and by the way resident of the great state of West Virginia.

I wanted to let you know that, yes, the way that we have taken a look at these programs is that we have looked at national files, but often the national files don't have the breadth of information that you need to look at for specific states. So I agree that you have to -- you have to do a combination. You have to look at the national files and you have to look at a scattering of files across different regions and different sizes, but you do have to have some big states in there.

MS. LANE: Let me explain. The big states are the ones that have had the resource to build large data infrastructures. So the reason I said California,

California has had the funding to develop with Tom McCurdy and Henry Brady, to develop an infrastructure almost to statistics for California. The same thing is true with Illinois. They -- Chapin Hall has invested heavily in developing a very large scale data set. The same -- you know, with all due respect the smaller states haven't had the capacity to build that infrastructure.

So you don't know what can be achieved until you start with an already existing infrastructure in those three states -- California, Florida and Illinois -- where that infrastructure has already been developed, all the MOUs have been signed, and then that shows to the other states what can happen. But that's why I was saying you have to start -- you have to start somewhere and you're better off starting where the infrastructure already exists, and that was our experience with the LEHD program.

MS. FROST: Maybe it's a combination of things depending what data you want. If you just want very low level things, which in some SIPP instances it's not really in depth information that we're trying to collect. I mean, you kind of should tailor it to what you're trying to get. I mean, I don't see SIPP trying to get as much as it can on all issues because then it turns into this big problem that we already have.

MS. LANE: So that's why I was recommending that you start taking where you've got users looking at what can be done using all the different sources of data.

MS. FROST: And comparing them.

MS. PRYOR: Okay. Another way is to start saying what's the minimal data set that we can live with and trying to get that.

MR. REAMER: I'll move to Don and then in the back of the room.

MR. OELLERICH: I just wanted to add to this conversation that one of the comments that I didn't get to in terms of things that the Census Bureau might do in terms of an interim product is instead of going forward with the CPS, CPS match and administrative data is to look backward, because if they look backward they go to earlier CPSs, they go to the LEHD program. Here you have UI records, you have TANIF (phonetic) records, you have the access to the Medicaid records, you have probably IRS record, you have some records from SSA. You have a massive amount of administrative data at your fingertips. You can deal with many of the technical issues that you have to deal with then. I mean, in terms of data disclosure issues, et cetera, and I think that might be a better way of moving this whole thing forward is to look back and go forward.

I'm sorry. I also wanted to add that for most of our public assistance type programs TANIF, food and nutrition programs of various kinds, there are no national databases and many of these programs do not collect Social Security numbers. So there's minimal identification information in these databases and so their use is going to be very limited even if they come from the individual states.

And if you're talking about all these different programs, you're talking about MOUs with different agencies within the states. So you're taking what Julia and you folks did, you know, ten years from now.

MR. REAMER: In the back.

MR. SHERMAN: Arloc Sherman. I was wondering just as a point of information, and it may be that no one person in the room knows this, but if anyone is able to go around and for several of the key data sources how far back we think we have longitudinal administrative data from either the nation or, you know, the 40 states or something where we could reconstruct that longitudinal admin record. You know, what year are we talking about? Are we going back to 2000 or are we going back, you know?

MR. REAMER: At this moment in time rather than -- because that would take a lot of time, I actually was --

(End side B, tape 1.)

MR. REAMER: A lot of I think very helpful comments, discussion points, observations here that people (inaudible) and encourage also to put their observations in writing and send them to you including the point that you make about if you know information or facts about particular data sets and how far back they go in time.

So -- but we still have time for --

A PARTICIPANT: (Inaudible.)

A PARTICIPANT: I could do Social Security.

MR. REAMER: Go ahead.

A PARTICIPANT: SSI starts monthly in >74. Title Two starts monthly in >64. The earnings records start annually in >51. The detailed W-2 earnings, Social Security earnings, start in >51. The detailed W-2 earnings start in >90 -- no -- yeah, >82, >82, sorry.

A PARTICIPANT: This is limited but the data set at the University of Chicago, at Chapin Hall on state administrative records on child protection cases is quite extensive. I think they're up to something like 18 or 19 states now and some of them go back more than a decade.

MR. REAMER: John.

MR. ICELAND: Just to comment for David. It sounds good anyway. This is John Iceland, University of Maryland. It's good that you're in a way thinking about the administrative data part of the project is sort of occurring almost at the same time as the SIPP. I think it's -- I wonder if they should be thought of as really two different programs. I don't know if you can -- if there's really -- certainly in the time frame you're talking about sort of thinking of the SIPP as having an integrated administrative component yet. Maybe if you're thinking of replacing SIPP with something it has to be some sort of household survey and over the long run the second type of program is developing the administrative data infrastructure that over time SIPP -- you know, any survey could have this data pulled into it, supplementing the data. It just seems to me it's going to be a very big challenge to really come out with an integrated product in the near future.

MR. JOHNSON: I think definitely a challenge, but I think that the idea was to design the survey in conjunction with using administrative records. We have to understand what's really collected. I mean, I could take the FAR for example of the IRS data, even though that's a harder one to match, but that's collected annually with filing status. So that would suggest that we need on the survey to get filing status as a variable to match.

So that's the idea of designing this together, is that we can ask the right questions to then link up to administrative records. So I think that's why we're doing this in parallel but together.

In terms of moving forward, a couple of things I forgot to mention. We are scheduling a meeting sometime in July out at Census Bureau so that we can invite more people, so other people can come, and we'll do this again and give you more updates. We'll post that date on our website or send it to all of you. We are going to schedule small group meetings on topics with federal agencies and some particular users. So that's sort of how we want to get.

The other thing moving forward is if -- from what I hear, if we didn't have this gap issue, you know, what could we move forward. Let's say we did it a different way. We can, but let's say we did it a different way where we're continuing to collect. Am I hearing you that if we just started -- just did the SIPP all over again everybody would be fine and satisfied? I don't think that's what we were hearing, okay.

So let's remove the gap. Let's say it's -- we're going to do something about that or not, what can move us forward? And this is what I want to encourage you to go to those variables that we have on these decision matrixes, forget about the fact that they're >93. This was just the starting point.

You know, if there's some child well-being stuff that you think needs to be added, put them in a list and let us know. If you think that some of these variables aren't needed we need to know that. If out of the 12 topical modules -- how many topical modules? 12 -- the total number of topical modules, right. A lot of them, if you think, why, we don't need this, the well-being topical module, we need to hear that.

This is the idea that we need to hear what you really, really need for this system because there's no -- I don't think it's possible to produce all the information that's currently there and I don't believe that all the information is used in every single year and every single wave and every single topical module.

So that's sort of what we're -- what we need to do to move forward, because we would need to do that even if nothing happened. We'd still need to do that for >09, okay, and we'd still be doing the same exact stuff for >09, figuring out what really needs to be in the survey. And so I think that's really what we're asking, what really needs to be there.

But in any case, thank you very much. I mean, it has been a very interesting and enlightening, and I do hope that we don't have a pigeon -- although pigeon are very useful, especially homing pigeons. They always come back.

(Applause.)

(The meeting was concluded.)

* * * * *