

Promoting Opportunity Demonstration Technical Expert Panel Report

Promoting Opportunity Demonstration Overview

The Social Security Disability Insurance (DI) program provides benefits for certain qualifying disabled workers and their families. DI beneficiaries generally stop receiving their benefits if they complete a Trial Work Period (followed by a three-month grace period) and continue to work and earn over the Substantial Gainful Activity (SGA) amount. The Promoting Opportunity Demonstration (POD) changes these rules by instead applying a gradual reduction in benefits through an offset. Benefits will be reduced by \$1 for every \$2 earned above the greater of an inflation-adjusted Trial Work Period (TWP) level or the amount of a participant's itemized IRWEs.

Section 823 of the Bipartisan Budget Act of 2015 (BBA) instructs the Social Security Administration (SSA) to carry out this demonstration. The BBA states that once benefits reach \$0 under the offset, a beneficiary's entitlement to disability benefits may be terminated, but Medicare coverage will continue for 93 months. The demonstration must begin no later than January 1, 2017 and will last for five years. Participation must be voluntary, based on informed written consent, and participants may withdraw at any time.

SSA plans to award separate contracts for the implementation and evaluation of POD. The contractors will implement POD in multiple sites (various geographic areas such as States, regions, metropolitan areas that have a sufficient population to achieve the number of participants required for the national evaluation). The sites will include a diverse group of participants (e.g., race, gender, etc.) from a variety of settings (e.g., urban/rural, region, etc.) and meet the standards for the sample population. The target population will include a random sample of SSA's DI beneficiaries. Contractors will invite targeted individuals to participate in POD and randomly assign those willing to participate and who provide written consent to participate into one of three groups of 5,000 each (two treatment groups and one control group). Recruitment and enrollment will occur on a rolling basis beginning in October 2017 and will continue throughout 2018.

There will be two treatment groups in the demonstration. The following rules apply to beneficiaries in both treatment groups (subject to change as the demonstration design is finalized):

- Participants will have varying lengths of participation depending on when they are randomly assigned, a minimum of 2.5 years and a maximum of 3.5 years;
- Benefits are offset \$1 for every \$2 earned each month above the greater of 1) an inflation-adjusted level (\$810 in 2016) or 2) the amount of the participant's itemized IRWEs;
- The Trial Work Period and Extended Period of Eligibility do not apply;

- Participants can withdraw at any time, but not retroactively;
- A modified expedited reinstatement is available to those whose entitlement to benefits is terminated; upon re-entry, the benefit offset still applies; and
- Auxiliaries are eligible for the offset, but if the Primary (Worker) has his or her entitlement suspended or terminated, so does the Auxiliary.

The two treatment groups differ in their rules regarding termination. All participants are placed into suspense in months they earn enough to bring benefits to \$0. Participants in one treatment group will have their entitlement to benefits terminate if they are in suspense for 12 consecutive months. Participants in the other treatment group will not have their entitlement to benefits terminated, regardless of the how long they are in suspense.

The contractors for the implementation will be responsible for the collection and coordination of participant earnings and IRWE information. Participants earning above the threshold are required to report this information on a monthly basis. The contractors will assist the participant in the reporting of this information and then transfer it to SSA for the benefit adjustment. The implementation contractors will also provide work incentives counseling to the treatment groups that mirrors what is available under current law through the Work Incentives Planning and Assistance (WIPA) program, with the addition of counseling on the benefit offset and POD rules. Services will begin in October 2017 and conclude in May 2021. Enrolled participants will begin reporting earnings in December 2017 for their wages earned in November 2017. As these reported earnings are transferred to SSA and processed, the benefit payments will reflect the changes in POD rules in their January 2018 checks.

The contractors completing the evaluation will conduct an early assessment, a participation analysis, process analysis, impact analysis, and cost-benefit analysis. The main research questions in the POD evaluation include the following:

- Does POD increase employment?
- Does POD increase beneficiary income?
- How does POD affect the SSDI Trust Fund?
- What interest is there in POD?
- What are the differential effects of POD?

POD Technical Expert Panel

SSA convened a Technical Expert Panel (TEP) during acquisition planning, as part of market research, to assist in refining the requirements for the demonstration. The purpose of the TEP is to provide independent guidance on the design of the implementation and evaluation of POD, to gather recommendations that will strengthen the project, and to gather lessons learned from relevant experts outside of SSA. The group was charged with reviewing the design of the demonstration, its anticipated outcomes, and identifying potential barriers to a successful

demonstration. SSA sought specific feedback on the proposed POD policies, services, evaluation, and lessons learned.

SSA invited individuals with a variety of expertise and knowledge in economics, vocational rehabilitation, work incentives counseling, disability and return to work policy, Federal demonstrations and evaluations, labor and employment policy, and research and evaluation design and methodology. The panel included representatives from academia, advocacy agencies, Federal agencies, State Vocational Rehabilitation agencies, and WIPA agencies. The TEP included the following individuals:

- Lisa Ekman, Director, Government Affairs, National Organization of Social Security Claimants' Representatives
- Ellie Hartman, Senior Scientist, University of Wisconsin-Stout Vocational Rehabilitation Institute
- Jennifer Kimble, Project Director, Division of Vocational Services, Maine Medical Center
- Nicole Maestas, Associate Professor, Department of Health Care Policy, Harvard Medical School
- Jennifer Sheehy, Deputy Assistant Secretary, Office of Disability Employment Policy, US Department of Labor
- Kimberly Vitelli, Deputy Administrator, Office of Workforce Investment, Employment and Training Administration, US Department of Labor

The TEP members convened in-person for an all-day meeting in Washington, DC on April 25, 2016. In addition to the above individuals, the meeting included Nicholas Hart and Jamie Wilson from the Office of Management and Budget, Stephen Goss, Robert Weathers, and Mary Kemp from SSA's Office of the Chief Actuary, and representatives from SSA's Office of Research, Demonstration, and Employment Support facilitated the meeting. Prior to the meeting, TEP members were provided an overview of the POD project and brief scopes of work on the implementation and evaluation. The meeting included four sessions on POD policies, services, evaluation, and lessons learned. The TEP was charged with providing their insights on each topic. SSA asked TEP members to provide written comments after the meeting if they had additional thoughts on any issues; these were incorporated into this document as well.

SSA posted the same materials and discussion questions provided to the TEP in a Request for Information (RFI) on FedBizOpps.gov. The [POD RFI posting](#) opened on April 13, 2016 and closed on May 4, 2016. The remainder of this report shares the input and recommendations made by the TEP and the responses received from the RFI.

POD POLICIES

SSA requested the panel focus on the proposed POD policies and discuss potential alternatives or modifications to these policies.

1. What are the pros and cons of using the TWP monthly amount (\$810 in 2016) as the threshold?

- There are pros and cons associated with using the TWP monthly amount as the threshold. A pro is that beneficiaries are familiar with the TWP level and they understand that number, especially work-oriented beneficiaries. The familiarity could also be a con since beneficiaries know this number for a different purpose, with different implications. However, this threshold amount seems to be a rational option, as there are no other notable amounts lower than the SGA amount.
- Another pro is that even at current minimum wage, beneficiaries would have to work 28 hours per week in order to reach the TWP level.
- There was concern about whether the beneficiaries would actually understand POD rules.
- The panel requested clarification on the difference between POD and the Benefit Offset National Demonstration (BOND). SSA noted that with POD, there is no TWP or Extended Period of Eligibility (EPE), so the beneficiary can begin offset as soon as earnings are above the threshold. Additionally, beneficiaries will not need a Continuing Disability Review (CDR) to begin offset. A panel member wondered whether we would see more employment effects just because it is an immediate offer.
- Study the distributional effects, not just the effects on the average group. Some beneficiaries working below SGA could decide to work more, some would terminate more quickly than under current law, or some could work less to keep partial benefits.
- From a participant's standpoint, a higher threshold could be more advantageous, but in order to see impacts, a lower level would be more useful. There are individuals that have advocated for using a lower amount, but the panel doubted anyone would volunteer for the demonstration with a lower threshold.
- Beneficiaries think in terms of limits and not opportunities. They are also used to these numbers changing, often on a yearly basis.
- Beneficiaries may park just below the threshold, regardless of the amount.
- It will be critical for the success of any offset model to have prompt and predictable reporting mechanisms and DI check adjustments. This is not effective in the current BOND model. In addition, under normal program rules when individuals earn SGA, DI checks are not stopped promptly, causing overpayments. To effectively test the impact of the offset, the offset must be accurately and effectively applied. The demonstration should use an electronic or telephonic wage reporting system.
- An advantage to using the TWP amount is that it is a known number. The threshold also has implications for the two treatment arms. Participants may not earn enough to lose benefits with a higher threshold, but with a lower threshold it will be easier to lose benefits.
- Equity is also a consideration as the full offset amount increases with increasing benefit amounts. This will especially be an issue with the termination arm. One idea is to have a threshold linked to the Primary Insurance Amount.

- The TWP level threshold coupled with the offset should encourage work effort for beneficiaries capable of earning over the current SGA earnings threshold.
- The offset softens the cash cliff effect and extends the time that beneficiaries can receive at least a partial DI benefit while their income increases.
- Using the average DI monthly benefit amount of \$1,166, the POD threshold coupled with the offset would allow the beneficiary to nearly triple their earnings before benefits are suspended or terminated.
- The TWP level threshold coupled with the benefit offset should result in greater benefit savings to the Social Security Trust Fund and increased Federal Insurance Contributions Act contributions from beneficiaries who work above the current SGA threshold.
- Beneficiaries who are working at maximum effort, but with monthly earnings below the current SGA level will not be attracted to POD.
- If a beneficiary can work part-time and cap earnings at the current SGA threshold in monthly earned income while retaining their full benefit (e.g., \$1,166 a month), he or she may not increase work effort.
- Beneficiaries who have HUD housing, Medicaid Buy-in and other public benefits with values that are eroded or costs that are increased by increases in earned income may not participate in POD, especially if their DI benefit is terminated when they reach the breakeven point.
- From an evaluation perspective, setting the threshold at the TWP amount is preferable to any threshold below that amount. We would expect fewer volunteers and higher attrition for lower amounts, increasing concerns about external and internal validity. Starting the offset at the TWP amount may not be high enough to compensate for the loss of the TWP, Grace Period, and EPE for some beneficiaries. Only beneficiaries who expect to earn above the SGA monthly amount—and are sufficiently confident that they could sustain that level of earnings for a long enough time—will be willing to forego the TWP, Grace Period, and EPE and volunteer to participate in the demonstration. This is likely a small minority of DI beneficiaries.

2. What would the evaluation issues be for testing more than one threshold?

- SSA could have a group that has already used the TWP and compare it to a group that has not used the TWP, to see if there are effects on work behavior or the use of the offset.
- Qualitative data will be important for POD. From experience with BOND, the topic of overpayments is a key issue. An overpayment is very stressful for any household, but especially so for households with a single earner, or those without steady income. SSA responded that POD's key differences from BOND would help with the overpayment issue. BOND participants estimate earnings on an annual basis and POD participants will report earnings on a monthly basis. If a participant reports January earnings in February, the changes will be in the March check. If participants forget to report, they can report that information at any time, potentially leading to smaller overpayments in those cases.
- Twelve months in suspense before termination will not be a strong trigger for termination. SSA replied that there have been conversations with Congressional staff and

SSA's Office of General Counsel on this feature, and both agree with the suspense provision.

- The highest predictor of employment gains is working with a benefits counselor. It is vital to have that one-on-one counseling service. Face-to-face communication is also necessary.
- It is a good idea to test more than one threshold. Without testing other settings, it may be hard to know if certain outcomes were due to other effects. However, this causes concerns for statistical power. SSA noted that with this project, there are concerns about recruiting the number of people needed for impact estimates.
- There will be more confusion with more than one threshold. If there is more than one threshold, each should be conducted in different locations.
- Testing two thresholds is a good idea but the itemized IRWE threshold could have a negative impact on the demonstration. Streamlining the reporting and check adjustment process is critical to the offset's success and the IRWE-specific threshold could have the opposite impact. When deductions come into play, the check adjustment process necessarily slows down, as Claims Representatives need to review and approve these IRWE expenses. SSA should use two thresholds, one at the TWP level and one at 50% of SGA.
- Testing multiple thresholds is a good idea. A limitation though could be the sample sizes needed for adding more treatment groups. SSA could consider adding just one additional treatment arm for another threshold or consider an orthogonal design. Additional treatment arms add to the complexity for counselors and beneficiaries. However, the impact on the volunteer rate could be low with the option to withdraw at any time. SSA could also consider assigning different thresholds to different states, but this would also reduce the comparability of the impact estimates.
- Testing more than one threshold would make the provision of benefits planning services more complicated and would require more in-depth training for the staff providing these services. However, it would provide more meaningful information because of the ability to compare outcomes based on different test models without having to implement another demonstration project.
- Examining the impact of different offset thresholds and different offset ratios are both important aspects of the concept of an offset. Assessing these two aspects in combination as a single variable, and at various thresholds and offset ratio levels, would result in the most meaningful research findings. Assessing the impact of two or more combinations of thresholds and offset ratios would be simpler if not also testing the impact of benefit suspension versus benefit termination. The concept of suspension versus termination is difficult for many people to understand and the mere mention of benefit termination is likely to result in potential study candidates declining participation as study volunteers or dropping out of the study if assigned to the benefit termination group. Testing more than one threshold amount and one offset ratio would provide significant information for SSA to make an informed decision about modifying the DI program to include an offset.
- The impact of various DI thresholds should be evaluated in the context of the value of the DI monthly benefit in combination with the value other public benefits (e.g., HUD housing, Temporary Assistance for Needy Families (TANF), Supplemental Nutrition

Assistance Program (SNAP), Medicaid Buy-In eligibility income limits and monthly insurance co-pays, etc.).

- Adding an additional treatment arm to test more than one threshold level would have significant implications for the scale of the demonstration, holding constant the current sample size needed to detect impacts. With multiple thresholds, the evaluation would not be able to disentangle the effect of the threshold difference from the effect of the termination difference. Regardless of the number of treatment arms, a disadvantage of testing more than one threshold is that attrition is likely to be more prevalent in the treatment arm with a lower threshold—potentially including high rates of immediate withdrawal once the result of the random assignment lottery is known—making direct comparisons between the active treatment arms more problematic in terms of internal validity.

SSA asked if there are other groups to target.

- Newly awarded beneficiaries. New beneficiaries would not already be influenced by DI limits on working. Getting to them at the beginning of their entitlement would help them to view the policies from the start and understand it is more like the Supplemental Security Income (SSI) model.

3. Are there other policy modifications that could be tested, for example, related to the offset rate, suspense, or termination?

- POD policies are more in line with SSI rules and that may help with simplification for concurrent beneficiaries. Concurrent beneficiaries would not have two sets of reporting and rules.
- There is considerable research on TANF, unemployment, and return to work, and the services involved. There are several modifications to consider but it is important to ensure the treatment and control groups have the same supports.
- How many people after the EPE earn above SGA -- even 2 or 3 times over SGA? There may be negative impacts and they may cut back on work.
- WIPA services are demand responsive. The control group would only have access to the WIPA services and may not be engaged at all. In BOND, they have not seen different effects from the two types of benefits counseling, regular and enhanced. Is there a need for another control group, or a group with some intervention but not all interventions, or a group with extra services?
- Work affects Medicaid in different ways in different states. There needs to be a way to mitigate Medicaid concerns for participants. Certain states have an unearned income tax that has been a barrier for individuals that could benefit from Medicaid but are unable to get it. Other states have lower earnings requirements for Medicaid and some people restrict their earnings in an effort to avoid the risk of losing Medicaid. It is important to measure the effect of Medicaid loss.

- Given the short timeframe for the demonstration, there will likely be so few terminations that the usefulness of the termination treatment arm is a question. It is likely that beneficiaries in that treatment arm will withdraw prior to the termination event out of fear. The fear of having to reapply is huge and very real. Continued DI eligibility should be designed like SSI 1619b where work doesn't trigger termination, only medical CDRs can cause termination.
- SSA should consider multiple offset rates as well as multiple thresholds allowing for an understanding of how different combinations will affect earnings and benefits.
- SSA could also allow the states to propose creative options for policies and for administrative duties regarding wage reporting and benefit adjustment.
- Use the HUD family of one (i.e., the beneficiary) annual income limits as the threshold for each participating state and the average DI benefit at the end of 2016. The HUD income limits have the advantage of controlling for cost of living differences across states.
- Higher thresholds could be combined with the elimination of IRWE disregards which would simplify the administration of work incentives.
- A goal should be to move beneficiaries from low income to median income status by changing the offset rate if the threshold does not move beneficiaries beyond low income status in the state in which they reside before benefits are terminated or suspended.
- Provide treatment and control groups enhanced benefits planning services so that the variances in the offset can be better evaluated against current DI work incentives.
- Study different levels of earnings for the offset and if permitted, do not include the variable of benefit termination versus benefit suspension. The idea of suspension of benefits versus termination of benefits is confusing to beneficiaries and the idea of terminated benefits is likely to both suppress and skew enrollment.
- Certain segments of the beneficiary population (e.g., individuals with chronic mental illness and individuals with disabling conditions that are deteriorating) will most likely decline the opportunity to participate in POD at the mention of benefits termination.
- There must be a requirement that all potential participants be thoroughly briefed on the purpose of the study and the implications of participating, including what happens to their benefits at the end of the study.
- The demonstration would teach us more about DI policy if it tested offset, suspense, and/or termination parameters that reduce DI benefits less relative to current law rules. In all these dimensions, the proposed parameters for testing will result in "faster" reductions in benefits than under current law if beneficiaries volunteer for the demonstration and then earn above the threshold level of earnings at which the offset begins. While some beneficiaries may be confident of sustaining employment at a level that assures they will come out ahead in the long run by entering the demonstration and earning more, most will be more concerned about the reverse—receiving lower benefits for a given earnings level in future months. These beneficiaries likely will choose not to volunteer or, once experiencing benefits lower than what current law would provide, will withdraw from the demonstration and drop out of the evaluation research sample. To limit this problem, if SSA has a way to test offset rates and suspense and termination policies with less rapid or extensive potential to reduce benefits—and thus encourage beneficiaries to volunteer—it

should. However, simply adding a treatment arm with a more appealing set of benefit rules to a lottery that already contains a less generous treatment arm may do little to induce a beneficiary to volunteer—and certainly will not stop him or her from withdrawing from the demonstration (and thus be lost to the research) should the sharper benefit reduction package be the one assigned to him or her by the lottery. To avoid beneficiary concerns about the “worst case scenario” when deciding whether to enter a random assignment lottery as part of the demonstration, policy parameters with less potential to reduce benefits (relative to current law) would ideally be the only treatment tested.

SSA asked about the length of the participation period - whether it should be two years or whether it should vary, and how important that is to the policies.

- In the Benefit Offset Pilot Demonstration (BOPD), after two years, there were differences in employment outcomes.

SSA asked if surveys could help answer some of these questions. What value might surveys offer for these types of questions?

- Surveys using hypothetical scenarios were used for the topic of induced entry. It can be an inexpensive way to answer questions about intent. They used hypothetical questions to ask what people would do if they had certain incentives. It should, however, be viewed as supplementary evidence, not primary evidence.

POD SERVICES

SSA requested feedback on potential demonstration services for participants, such as the support necessary for monthly earnings reporting, what services are most appropriate, and how best to provide these services.

1. Are services that mirror the current WIPA program sufficient to provide support for the benefit offset and earnings reporting? Are there additional types of assistance that a WIPA-like entity could provide that would be valuable, e.g., outreach, reminders, follow up?

- Some states have fee for service contracts that may be able to ensure contact with benefits counselors if necessary. In another demonstration project, all participants are Vocational Rehabilitation (VR) customers, which can offer an advantage.
- The VR advantage may depend on the particular state and that VR system. There is poor performance data on VRs nationally serving individuals with mental health issues.
- VR contact is usually face-to-face contact. VRs have strong connections to employment services.
- SSA should determine if VRs could prioritize demonstration participants for VR services.
- WIPAs operate with limited resources even for their current duties. Additional resources will be needed if WIPA duties will be expanded to provide assistance for POD

participants. If new duties will be added to the current WIPA tasks, additional resources will also be needed, as well as training for the new tasks. It is important to note that counselors could assist beneficiaries in wage reporting, but this should remain only the responsibility of the beneficiaries.

- A critical piece is a computer application that assist participants or counselors calculate benefit amounts under various earnings scenarios.
- SSA should consider counseling with a proactive approach to outreach with beneficiaries rather than the current demand responsive model.
- WIPA services should include outreach, reminders, and follow-up for POD participants. The same type and level of services should be available to those in the treatment and control groups.
- Rather than testing the value of benefits planning assistance, SSA should test simplification of work incentives within POD (e.g. eliminated IRWEs, TWP, and EPE).
- It will be important for the benefits counseling services provided under POD to provide the same range and intensity of services that WIPA provides under current law but the content of the benefits counseling services offered will need to vary from what is provided currently.
- POD beneficiaries will require assistance with three things that would involve expansions of current WIPA operations. This would include 1) assistance to prepare and submit earnings and IRWE documentation to SSA, 2) assistance to withdraw from POD, if requested, and 3) provide ongoing information to participants to remind them they are in POD, how the POD offset works, and the importance of reporting their earnings and IRWEs monthly to SSA through POD. SSA should consider centralizing these tasks in a centralized work unit rather than having them handled locally. Centralization will improve implementation fidelity and monitoring of implementation, and thus facilitate the evaluator's efforts to ascribe impacts to a particular implementation. With current sample sizes, it would be difficult to detect impact variation in cross-site analyses (for example, due to varying levels of implementation fidelity) at a satisfactory level of statistical precision.

SSA asked if VRs would have the capacity to prioritize demonstration participants for VR services, even with more funding to provide these services.

- The panel did not know the answer.

2. What are the implications of increasing the capacity of current WIPAs, compared with completely separate entities, to serve POD participants?

- There are also other entities that provide employment assistance other than VRs and WIPAs.
- An advantage to using WIPAs is that their content is already SSA specific.
- In BOND, we learned that the processes involved in wage reporting, CDRs, and benefit adjustment need to be easier for the beneficiaries. BOND participants have often been

frustrated with the lack of timely adjustments resulting in overpayments and underpayments. They were also unable to reach anyone at SSA for assistance, leaving only counselors. The important task of benefit adjustment should remain with SSA and not an outside agency that is potentially not trusted by beneficiaries. SSA operational staff, or contractors embedded in SSA operations should calculate the benefit adjustments. This should happen in the SSA field offices. This demonstration will require a sufficient number of well-trained SSA operational staff.

- Counselors should be integrated with the local benefits counseling resources with a sufficient number of staff to provide these services. Local services are critical in understanding differing state Medicaid rules, which is important to many DI beneficiaries. Separate or additional resources will lead to confusion for beneficiaries as the rules and resources are already difficult to navigate.
- WIPAs vary in consistency and quality and increasing the capacity will not necessarily result in increased quality of services. Experience from BOND shows that a centralized entity to provide supports should be considered for POD. This could enhance the consistency, quality, and speed of services, and may require fewer staff resources. A drawback is the lack of local knowledge that is important for beneficiaries and a lack of comfort for beneficiaries.
- Benefits planning assistance should be provided through local WIPA providers rather than through newly established separate entities. It will require significant time and resources to set up and adequately prepare separate entities to serve POD participants. If benefits planning assistance is made available to POD participants through a centralized source, the counselors providing those services would need access to a database on and significant training on all of the State-specific programs and policies that may come into play when providing benefits counseling.
- A requirement could be included that POD WIPA services also be located in the organization receiving the funding to implement a pilot site or within a local Independent Living Center, a VR agency, or an Employment Network (EN). Housing WIPA POD services in these organizations could result in better response to recruitment and ease of access to benefits planning service for participants.
- Staff on SSA's national toll-free hot line should be trained to respond to questions about POD.
- There are several advantages to increasing the capacity of the current WIPAs to provide POD benefits counseling and earnings and IRWE reporting services including:
 - Staff who provide WIPA benefits counseling services are Certified Work Incentive Counselors (CWICs).
 - Staff who provide WIPA benefits counseling are required to have extensive knowledge of state-specific programs and rules.
 - WIPA organizations that currently serve an entire state have established procedures and networks that cover the full state service area.
 - Utilizing current WIPA providers could ease the transition of POD beneficiaries back to regular DI program rules at the end of the demonstration.
- Disadvantages to consider are:

- WIPA organizations would need to differentiate counseling provided for POD beneficiaries from the counseling provided to current law beneficiaries.
- WIPA providers would need to take on new responsibilities for POD that are not part of current WIPA operations. These activities require specialized skills and procedures that may, or may not, be compatible with skill sets of the current WIPA staff. In addition, asking WIPA agencies to expand their scope of work to include supporting POD could disrupt the delivery of core WIPA responsibilities.

3. The WIPAs negotiate a multitude of state and local entities and environments and reflect these variations in their services. Could a national provider by telephone/call center deliver these differing, individualized services?

- A combination of a WIPA services delivered via a national call center and at the State and local level is probably the best approach. A national provider of WIPA services could provide on-demand assistance for generic questions about disability benefits and work but it would not provide the level of assistance needed for state or local information. Centralized services would also mean few if any face-to-face consultations and possibly limit the ability to review necessary documentation without inconveniencing beneficiaries in terms of having to submit the information to a centralized location.
- It would be possible to deliver POD benefits counseling services using a telephone/call center staffed by CWICs and clerical staff. The CWICs in this type of remote, telephonic model collectively would need to have detailed knowledge of the state-specific programs for all of the POD States in addition to nationally uniform POD program rules. This service delivery model could be based on SSA's Employment Success Advisor (ESA) program that operated nationally during a period when WIPA was not available. SSA provided ESA services remotely to beneficiaries that were similar to those available through WIPA. Staffing efficiencies could be realized from this type of service delivery model, such as potentially fewer full-time CWICs employed centrally compared to the number of staff required if based locally in each POD State. However, the model utilized by ESA can only work for POD if there is a sufficient number of centralized staff trained for each POD State so beneficiaries do not experience a burdensome wait-time.
- SSA should consider centralizing the collection and reporting of earnings and IRWEs to SSA.

SSA also asked whether the use of one implementation contractor would help to provide more uniform services.

- Multiple contractors would not allow for consistency across the sites if that is important.
 - If participation is limited to states with certain Medicaid characteristics, it could be possible to control for that type of variability.
- It is possible for a centralized staff unit to provide specialized services. In the BOND project this has been a positive development, although there were concerns about limited local knowledge and the effects on the counselor-participant relationship. The idea of local versus centralized support is something that could be tested in POD.

4. Are there other services widely available that would be beneficial for POD?

- Accurate forecasting is a key service. Beneficiaries need to know what their income will be in the future and what their wages could be using accurate wage information. WIPAs know and understand the labor market but do not steer people into it. WIPAs do not advise on employment, they only refer to places where they can get labor market information. VRs are able to do job matches.
- There is a cost-benefit analysis to working. Beneficiaries need to understand what happens to their income if they go into offset or if their benefits end. They need to know how much income they will have and how they will access healthcare. The healthcare piece is just as important for beneficiaries as the income.
- SSA could focus on people who have the highest prospect of getting and retaining a job, such as people who just started a TWP.
- BOPD targeted Medicaid buy-in volunteers. Those in the buy-in had lower earnings and were constrained by the buy-in.
- It is important to know who has work capacity and how big that population is. SSA beneficiaries do not have work capacity, so you cannot expect positive results. SSA should target predicted work capacity, which are new beneficiaries. One in ten may have work capacity among beneficiaries.
- A certain number of beneficiaries are working. Among those who are not, given the right services and supports, many could work. Linking to effective services, such as Individual Placement and Support (IPS), could make a difference. IPS is available through Medicaid.
- POD participants would benefit from knowing how their earnings would affect their Medicaid eligibility, which will vary by state. Given the nature of the intervention, some POD participants may experience an overpayment if the adjustment in the \$1 for \$2 offset is not applied, or if changes in earnings are not reported to SSA, in a timely manner. For such cases, these individuals would benefit from having assistance with understanding and appealing the overpayment determinations. The WIPA project, the state or state's implementation contractor, or a centralized entity could provide this support. States could also be invited to propose innovations in this area.
- Participants will need to understand how their earnings will affect their benefits and Medicaid eligibility. If wage reports are not timely, participants may experience overpayments or underpayments. These participants would benefit from assistance with the appeal process.
- SSA currently provides several services to DI beneficiaries that may be beneficial to POD such as the Ticket to Work Help Line, SSA's toll-free number, and SSA field offices.
- Other state and local agencies and organizations may also provide assistance to the demonstration such as VR agencies, advocacy organizations, and local social service agencies.

- Many beneficiaries already access other resources such as HUD low-income housing, SNAP, community mental health services, and the Medicaid Buy-in program. The availability of resources such as these must be taken into account because changes in a person's financial situation can affect access to these programs.

SSA asked what services should be mandated.

- SSA's Mental Health Treatment Study used IPS and did not see earnings high enough to go into benefit offset.
- SSA could target those most likely to exhaust unemployment insurance.
- The Department of Health and Human Services will release some TANF research soon that may be helpful.

EVALUATION ISSUES

This session focused on evaluation issues related to POD such as outcomes, participant criteria, and generalizability.

1. What primary research questions best reflect the legislative intent for the demonstration?

- A critical question is how many beneficiaries are going to volunteer.
- SSA should compare the termination and suspension groups.
 - Will the termination policy keep beneficiaries from volunteering for the demonstration?
 - There may not be many individuals that terminate from entitlement, so is it worth testing this treatment arm?
 - There should be no suspension, just an immediate termination of benefits.
 - If the suspension period is shorter, will anyone volunteer?
 - Perhaps those in the termination arm should not be told that there are other treatment arms.
- The expedited reinstatement policy should remain but be modified to be a quick decision. This would only be an issue for the termination arm.
- If no one volunteers or the volunteers dislike a particular treatment arm, that is important for policymakers to know. Even if the numbers needed cannot be recruited, it is still worth testing.
- An additional research question to consider is whether POD reduces SSA and beneficiary administrative burden, such as overpayments.
- Did the time-limited nature of the demonstration adversely affect beneficiary participation?
- Did the information on what beneficiaries should expect at the conclusion of POD adversely affect beneficiary participation?

- Does POD result in more DI beneficiaries achieving earnings above SGA?
- Is there a difference in the percentage of POD volunteers achieving \$0 cash benefit between those who are assigned to the suspended versus the terminated benefit research groups?
- Does the availability of a \$1 for \$2 offset at various earning thresholds result in significantly different gains in earnings among the threshold groups?
- Would a variance in offset ratio (e.g. \$1/\$2, \$1/\$3, \$1/4) at the \$810 threshold result in significantly different gains in earnings among the offset ratio groups?
- What combination of POD DI work incentives produces the highest increases in earnings?
- Does POD simplify the DI work incentives for both beneficiaries and for POD program/site administrators?
- Is there a difference in willingness to volunteer for POD between DI beneficiaries who only have a DI benefit and Medicare as compared to those with additional public benefits such as a Medicaid Buy-In program, HUD housing benefit, and SNAP?
- To what extent were the services and support provided by WIPAs effective?
- To what extent were the services and support provided by State VR agencies, ENs, and other types of service providers effective?
- What is the effect of the offset on average beneficiary earnings?
- How does the offset affect average SSDI benefits?
- How does the offset affect other means-tested benefits, such as SSI, SNAP and Medicaid?
- SSA should study economic and other effects on DI beneficiaries such as on individual and household income, and other aspects of well-being.

2. What measured outcomes best answer the primary research questions?

- Economic environment
- Service environment
- Geography including both urban and rural environments
- If individuals are working and not on DI, possibly through unemployment insurance
- Minimum wage in each state
- Beneficiary earnings and benefit amounts paid in the year and benefit amounts paid for the year. These categories take into account the presence of overpayments, underpayments, improper payments, and retroactive payments.
- Employment and Earnings:
 - Annual earnings (mean \$)
 - Earnings between TWP and SGA (proportion of beneficiaries)
 - Earnings above SGA (proportion of beneficiaries)
 - Employment during the year (proportion of beneficiaries and beneficiary-years)
- Benefits
 - Total SSDI benefits paid (mean \$)
 - Number of months with SSDI payments (mean)
 - SSDI benefit receipt (proportion of beneficiaries with SSDI benefits paid)

- Redetermination and Continuing Disability Review (CDR) outcomes
- Income
 - Total income (mean \$ of SSDI benefits paid + earnings)
 - Total income (mean \$ similar to above, but including broader measures of other sources of income obtainable through survey)
- Well-being
 - Material hardship (various measures: scales, proportions)
 - Food insecurity (scale measure)
 - Mental and physical health, including mortality (various measures: scales, proportions)
 - Health spending and utilization (various measures: mean \$, proportions)
- Is there a difference in the percentage of people in the three groups who have:
 - Volunteered for POD and were in the workforce with earnings at the time of enrollment?
 - Entered the workforce after enrolling in POD?
 - Were hired with earnings at or above TWP/SGA?
 - Accepted a job offering employer-sponsored health insurance, paid sick leave, annual leave?
 - Achieved SGA level earnings after enrolling in POD?
 - Moved from part-time to full-time employment while enrolled in the program?
 - Chose to voluntarily drop out of POD prior to the program ending?
 - Voluntarily quit a job while enrolled in POD?
 - Received WIPA services at the local or national level or a combination of local and national level assistance?
 - Received services from an EN, a VR agency, a American Job Center, or a Veterans Program versus those who did not?
 - Actively sought benefits planning assistance while participating in POD? How often did they seek these service?
- Did the time-limited nature of the demonstration adversely affect beneficiary participation? Measured outcomes include:
 - Percentage of invited POD beneficiaries who responded with interest to POD but declined to volunteer due to the time-limited nature of the demonstration.
 - Did the percentage of those who declined volunteering for POD due to the time-limited nature of the demonstration increase for those who were recruited later in the demonstration time period as compared to those who were recruited earlier?
- Did the information on what beneficiaries should expect at the conclusion of the POD adversely affect beneficiary participation? Measured outcome is percentage of invited POD beneficiaries who responded with interest to POD but declined to volunteer after learning about the what would happen to their benefit status at the conclusion of POD.
- Does POD increase beneficiary income? Is there a difference in beneficiary earned income among the three groups in terms of:
 - Moving from one job to another to increase hours of work and/or earnings?
 - Earnings at time of POD enrollment and annually until POD ends or the beneficiary voluntarily withdraws from POD? Consider looking at whether

- earnings are greater or less than TWP, SGA, minimum wage (federal and state if applicable), poverty line, etc.
 - Percentage increase/decrease in earnings from time of enrollment to program end or voluntary withdrawal?
 - Increase in earnings negatively impacting any other public benefits the beneficiary was receiving at time of POD enrollment?
- How does POD affect the DI Trust Fund? Measured outcomes include:
 - Number/percentage of POD participants experiencing a reduction in benefits while enrolled in POD? Consider breaking this information down by size of reduction.
 - Number/percentage of POD participants reaching zero benefit status resulting in benefit suspension versus termination.
 - Number of months of benefit cessation while enrolled in POD across and among the research groups.
 - Amount of savings generated to the Trust Fund as a result of reductions in benefits to benefit offset and cessation.
 - Amount of savings generated to the Trust Fund as a result of benefit suspension and termination across and among research groups.
- What interest is there in POD? Measured outcomes include:
 - Percentage of invited POD beneficiaries who volunteer for the demonstration.
 - Percentage of volunteers in the experimental groups that enter employment above POD threshold.
 - Percentage of volunteers in the experimental groups who maintain their POD engagement until the demonstration ends.
- Does POD result in more DI beneficiaries achieving earnings above SGA? Measured outcome is the increase in the number/percentage of POD beneficiaries who increase their monthly earnings above SGA after engaging in POD as compared to the control group.
- Is there a difference in the percentage of POD volunteers achieving \$0 cash benefit between those who are assigned to the suspended and terminated benefit groups? Measured outcome is the difference in achieving \$0 cash benefits between the two groups of POD volunteers.
- Does the availability of a \$1 for \$2 offset at various earning thresholds result in significantly different gains in earnings for those earning above SGA among the various earnings threshold groups? Measured outcome is the difference in earning gains of POD beneficiaries with monthly earnings exceeding the SGA level, for higher threshold groups as compared to lower threshold groups.
- Does a variance in offset ratio (e.g. \$1/\$2, \$1/\$3, \$1/4) at the \$810 threshold result in significantly different gains in earnings among benefit offset ratio groups? Measured outcome is the difference in the earning gains of POD beneficiaries with monthly earnings exceeding the standard SGA level, between benefit offset ratio groups.
- What combination of POD DI work incentives produces the highest earning gains? Measured outcome is an appropriate statistical analysis to determine the correlation, if any, between various existing and POD work incentives and earning gains.

- Does POD simplify the DI work incentives for both beneficiaries and for POD program administrators? Measured outcome is the difference in the level of perceived complexity of the work incentive rules between POD and the current DI work incentives as reported by beneficiaries and program administrators.
- Is there a difference in willingness to volunteer for POD between DI beneficiaries who have only a DI benefit and Medicare as compared to those with additional public benefits such as a Medicaid buy-in program, HUD housing benefit, SNAP, etc? Measured outcome is the difference in willingness to volunteer for POD between DI beneficiaries who have only a DI benefit and Medicare as compared to those with additional public benefits.
- Is there a difference in the percentage of POD volunteers who maintain their POD engagement until the demonstration terminations between DI beneficiaries who have only a DI benefit and Medicare as compared to those with additional public benefits such as a Medicaid buy-in program, HUD housing benefit, SNAP etc? Measured outcome is the difference in willingness to maintain POD engagement until the demonstration ends between DI beneficiaries who have only a DI benefit and Medicare as compared to those with additional public benefits.
- To what extent were the services and support provided by WIPAs to all POD participants effective? Measured outcome is beneficiaries' perceived effectiveness of WIPA services in decisions to volunteer for and participate in POD.
- To what extent were the services and support provided by State VR agencies, ENs and other types of non-WIPA service providers effective? Measured outcomes include:
 - Beneficiaries' perceived effectiveness of the services provided by State VR agencies, ENs and other types of service providers in decisions to volunteer for and participate in POD.
 - The difference in earning gains for POD participants who were not employed at the time of POD engagement and who received service and support from State VR agencies, ENs, and other types of non-WIPA service providers compared to beneficiaries who were not employed at the time of POD engagement and who did not receive additional supports or services.
 - The difference in earning gains for POD participants who were employed at the time of POD engagement and who also received service and support from State VR agencies, ENs, and other types of non-WIPA service providers compared to beneficiaries who were employed at the time of POD engagement, but who did not receive additional supports or services.
- Costs and benefits of health insurance

SSA asked about the best ways to obtain the health insurance information.

- Surveys, Medicaid data, or self-reporting for private insurance.

3. What modifications to criteria for the participant population should be considered?

- Age
- Current WIPA participants
- Beneficiaries who are ready to work, identified through VRs and ENs. The voluntary nature of the demonstration would naturally engage those who want to work.
- SSA should be careful about self-selection and consider stratification. It is possible to get an entire sample of individuals that cannot work. If SSA stratifies, at least there will be some beneficiaries that would be able to work.
- SSA should consider beneficiaries in their first year of DI eligibility, those that have not already used their TWP or grace period months. This provides better impact estimates as the impact estimates will be for only those under POD rules and not a mix of POD and current rules. This could create challenges for reaching necessary sample sizes.
- Consider potential participants who have been receiving benefits only within the last two calendar years.
- Potential participants should have a current or past wage history of earnings at or above \$810 a month.
- SSA could exclude concurrent beneficiaries because including them would make recruitment even more difficult, as it would require more complex analysis and explanation about how work would affect benefits to inform individuals prior to their enrollment.

SSA asked about the stratification categories and what are the predictors of work?

- It may be better to apply that type of weighting until you ask beneficiaries about their interest in work.
- Screening is very expensive.
- Another demonstration project sent mailers that included an application and almost half of participants enrolled via this method.

SSA asked about allowing a consortium of states, considering medical services are different in each state. We need to ensure consistency in the evaluation.

- States could use fidelity measures.

4. What are the pros and cons of having the States recruit participants vs. the evaluation contractor conducting recruitment?

- Local providers serve as entities that beneficiaries know and respect. Beneficiaries would view them as more trustworthy, although an evaluation contractor could become a more familiar face throughout the recruitment process.

- A single entity doing the recruitment may not provide the face-to-face contact needed. Yet multiple contractors working separately could be challenging too.
- How to address issue of states volunteering to participate versus those who won't apply - this is a difficult task and the use of econometrics may not be able to help with this issue.
- An advantage to having an evaluation contractor conduct recruitment is uniformity across the sites. This also allows for consistent messaging to potential participants and could allow for targeting beneficiaries that have an interest in working.
- There are three advantages to having the evaluation contractor conduct recruitment:
 - A centrally organized process for recruitment with uniform messages, outreach contacts, scripts, and procedures will help to further a rigorous evaluation.
 - Obtaining truly informed consent will require careful presentation of POD's rules and should be conducted in a highly uniform and careful manner.
 - A centralized recruitment approach will also help assure that the aggressive enrollment targets (15,000 beneficiaries) are met within the one-year period specified in the draft SOW.
- State VR agencies, ENs, and local employment and long-term support agencies within each state have established and trusted relationships with viable POD volunteers. Prospective volunteers may be more responsive to recruitment efforts coming from a known and trusted service organization.
- Individual states or a state consortium would be better able to enroll eligible participants from current state programs.
- National recruitment of POD participants may ensure that evaluation findings are nationally representative. If recruitment is done through individual states or state consortium, additional controls and oversight may be needed.
- Recruitment by individual states or a state consortium may be more resource intensive than recruitment done at the national level.

SSA mentioned pooling all the states for one estimate as opposed to six separate estimates.

- Would a national policy be able to implement this in a consistent manner? On a national scale, states may implement this differently, so many would want to look at the states separately.

5. Are there ways to make the evaluation more generalizable, given the conditions set out in the legislation?

- Knowledge of POD should be similar to how it would be under a national policy. In BOND there were issues with lack of knowledge by local or SSA staff that resulted in beneficiary concerns of legitimacy for the project.
- Random assignment at the site level rather than individual level could also increase generalizability as it would be more similar to a national rollout of the policy.
- Another way is by conducting a strong process study using tools from implementation science.

- Given the parameters set out in the legislation, a non-generalizable group of DI beneficiaries is likely to volunteer for the demonstration—probably a very small proportion of all the SSDI beneficiaries who would be affected if an offset of the type the legislation envisions were adopted as national policy. This places a severe constraint on options for learning reliably about the potential effects of such a national program, which should be the goal of any federally-initiated policy evaluation.
- The results may be more generalizable if higher thresholds are tested and IWREs are not part of the threshold

6. What are the benefits and drawbacks of the alternative ways to achieve generalizability?

- With site-level random assignment, the rollout would be similar to a national rollout. All staff at field office will be trained in POD and all beneficiaries will be subject to POD. This provides a nationally representative sample. A concern is that it reduces the evaluator’s ability to detect effects large enough to be of interest.
- Individual-level random assignment could increase costs. There is a potential for spillover effects and lessons may be more limited for the operation of administrative processes.
- A wide variety of approaches to improved generalizability allow inference from a small non-representative sample to a larger target population, and there are tradeoffs in terms of assumptions required and bias-variance tradeoffs, but essentially none of these approaches can generalize from a population with a narrow range of potential earnings to a population with a larger range.

LESSONS LEARNED

SSA also wanted to know important lessons learned from prior demonstrations and how to increase beneficiary understanding of POD policies and rules.

1. What have we learned from prior SSA demonstrations that is important for POD?

- From the BOPD we learned that manual processing is very labor intensive and that as much as possible should be automated.
- When conceiving the demonstration design, SSA should not be concerned about scaling up but rather just about providing evidence.
- Eliminate the need for work CDRs, as it is challenging and difficult.
- Avoid any project element that could result in large overpayments or underpayments.
- Medicaid is important. Look for a mix of states with strong Medicaid buy-in and those without.
- A longer intervention period is better.

- A lesson from BOND is that SSA needs to sufficiently staff resources that support POD. It is important for the reporting and adjustment process to be timely and avoid overpayments and underpayments.
- The volunteer rate for POD could be low unless outreach is intensive or there are features added that make it more attractive to beneficiaries.
- Prior SSA demonstrations have shown advantages of a centralized project data system that supports both the evaluation and implementation contractors. Having one system to manage outreach, enrollment and random assignment processes, along with participant information, is more cost-effective than supporting multiple systems, allows for the uniform collection of participant data across demonstration sites, and supports the sharing of data across contractors and with SSA.
- Prior experience has also highlighted that SSA field offices are uniquely positioned to have an integral role in demonstrations. Because of the field offices' locations throughout the country and beneficiaries familiarity with them, field offices can be an asset to both the enrollment of participants into a demonstration and in supporting them once enrolled. SSA may wish to consider having the field offices actively support the evaluation contractor's outreach efforts by educating beneficiaries about POD and, if the field offices are able to confirm their inclusion in the solicitation pool, encouraging them to contact the demonstration to enroll. Once enrolled, the field offices could accept earnings and IRWE documentation directly from beneficiaries who may be more comfortable interacting with their field office rather than the Implementation contractor. Further, the field offices could be responsible for making the benefit adjustments and adjudicating earnings and benefit decisions for POD participants.
- Previous SSA demonstrations involving a SGA threshold and a \$1 for \$2 offset have shown that the benefit reduction rate has no significant impact on earnings outcomes.
- Confusion for demonstration participants occurred when benefits for those receiving the offset are not adjusted in a timely manner following wages exceeding the earnings threshold.
- SSA should include more targeted comparisons including, but not necessarily limited to:
 - Comparisons between those in the treatment and control groups who have completed the TWP.
 - Comparisons between those in the treatment group who completed a TWP and used the offset and comparable treatment group members who did not.

2. How can we enhance beneficiary understanding of the benefit offset and POD rules?

- Use low-tech technology for wage reporting, such as a simple app.
- The BBA requires a hotline for reporting earnings. SSA explained that this will not be ready in time for use in POD.
- There is not enough time to build something onto any of the BOND systems.
- A step-by-step wage reporting tool for families, in the form of an app or a website created by the contractor.

- SSA could use the Work Number for wage verification. Beneficiaries opting into POD would also opt into the Work Number. This could help to prevent overpayments.
- Give all POD beneficiaries a smart phone with access to an app or to use to take a photo of their monthly checks to send in.
- What will happen if a participant does not report? SSA responded that a wage report will carry over until a participant reports again. If a participant does not report for three months, we will send a letter.
- When does a failure to cooperate status occur?
- Beneficiaries who are not working should not be asked to report anything, only those with earnings.
- SSA should try to learn about those participants who do not report.
- Annual reconciliation does not seem necessary considering the frequency of the reporting requirement.
- Beneficiary understanding could be enhanced by using communications and messages that don't use official or legal language or complex jargon. Clear and concrete examples should be used to convey scenarios around earnings and the offset. Information should also be available in different formats – email, letter, website. Reminders could be sent for reporting via text or email to increase reporting. Also, a computer application or calculator should be used to view various scenarios.
- Recruitment efforts could include letters to beneficiaries with detailed explanations of the POD rules and examples of how benefits would change under different scenarios. Other written communications in conjunction with outreach letters are brochures, frequently-asked questions, and diagrams to illustrate POD rules. Enclosures should be written in plain English with easy to follow explanations of POD, augmented by graphics to illustrate the scenarios.
- SSA should consider developing a public website to which frequently-asked questions, scenarios, and other materials could be posted to explain the demonstration. SSA may want to consider developing a short video presentation to explain the demonstration. In addition to “static” printed and video communications tools that explain POD, it is also critical to provide beneficiaries with options for direct contact with demonstration staff to ask questions they have about the demonstration. This service could be provided by a POD call center, a POD email address, and social media.
- SSA will need to ensure that all staff providing benefits planning assistance are certified benefits planners, that earnings are recorded as they are reported, that work CDRs are conducted in a timely manner, and that benefit cessation (whether suspension or termination) is done in a timely manner.
- The support for this study could be designed to include one or two POD liaisons in each local SSA field office within the geographic areas served by each demonstration site. These staff would need to prioritize the processing of work and earnings information received from POD participants, the timely implementation of the benefit offset, and the timely conducting of work CDRs.
- Outreach activities must provide information to organizations and professionals that are likely to have regular contact, not only with Social Security disability beneficiaries, but also with those in this population.

SUMMARY

The TEP and RFI provided valuable information and guidance on SSA's plans for the implementation and evaluation of POD. SSA will review the recommendations on the proposed services for demonstration participants, the evaluation, POD policies, and lessons learned and incorporate them whenever feasible.