BOND Implementation and Evaluation

Second-Year Snapshot of Earnings and Benefit Impacts for Stage 1

Deliverable 24c.2

June 2, 2014

Prepared by:
David Stapleton
David Wittenburg
David R. Mann
Denise Hoffman
Andrew McGuirk

Abt ASSOCIATES

MATHEMATICA CENTER FOR STUDYING DISABILITY POLICY

Report Context

As part of the Ticket to Work and Work Incentives Improvement Act of 1999, Congress asked the Social Security Administration (SSA) to test alternative Social Security Disability Insurance (SSDI) work rules designed to increase the incentive for SSDI beneficiaries to work and reduce their reliance on benefits. In response, SSA has undertaken the Benefit Offset National Demonstration (BOND), a random assignment test of variants of SSDI program rules governing work and other supports. SSA, in conjunction with several contractors led by Abt Associates, developed the infrastructure and supports required to implement BOND.

The BOND project includes two stages. Stage 1 is designed to examine how a national benefit offset would affect earnings and program outcomes for the entire SSDI population. Stage 2 is designed to learn more about impacts for those most likely to use the offset (recruited and informed volunteers) and to determine the extent to which significant enhancements to counseling services affect impacts.

This report, the second in a series of *Snapshot Reports,* documents Stage 1 impacts on earnings and benefit outcomes during the second calendar year of implementation (2012). As in the previous *Snapshot Report* (Stapleton et al. 2013), we estimate impacts by comparing earnings and benefit outcomes for all Stage 1 treatment subjects (T1) to those for all control subjects (C1). Future annual reports will track Stage 1 impacts through 2017. A parallel series of *Snapshot Reports* will be produced for Stage 2.

Summary of Key Findings

The benefit offset, as administered under BOND, did not have a statistically significant impact on total earnings in 2012—similar to previously reported findings for 2011. Thus, at this time there is no evidence of the increase in earnings that would be needed to yield cost neutrality or cost savings.

There were positive impacts on total SSDI benefits paid *in* 2012. However, the benefit measure used in this report includes only benefits actually paid to beneficiaries through the end of 2012. Since 2012, SSA has continued to make retroactive adjustments to benefits *for* 2012 that are not reflected in the

benefits paid *in* 2012 in this report. Thus, benefits paid for 2012 are likely to differ from the benefits paid in 2012 which are analyzed here. We are unable to predict how impacts on benefits paid for 2012 will differ from impacts on benefits paid in 2012, after all retroactive adjustments are made.

We will continue to document annual earnings and benefit impacts in future reports. Impacts on earnings and benefit outcomes in 2013 and beyond might be quite different from those in this report for two important reasons. The first is the initiation of follow-up outreach to Stage 1 treatment group (T1) subjects in 2013, designed to help them better understand the benefit offset and how to make use of it. The second is that many more T1 subjects had their benefits adjusted under the offset by the end of 2013 than at the end of 2012.

The BOND Evaluation Team

Abt Associates, in partnership with 25 other organizations, is implementing and evaluating BOND under contract to the SSA. To ensure the objectivity of the evaluation, separate teams conduct the implementation and evaluation components of the project. The current report reflects exclusively the views of the evaluation team, led by evaluation co-directors Stephen Bell of Abt Associates and David Stapleton of Mathematica Policy Research. These people have no role in implementing or overseeing the BOND intervention they are studying, nor do any members of their evaluation team. Separation of implementation and evaluation does not extend throughout the project, however. Project Director Michelle Wood and Principal Investigator Howard Rolston of Abt have joint responsibility for coordinating the implementation and evaluation efforts, including, respectively, managing the day-to-day operations of the project and overseeing the effective and efficient implementation of the BOND design. Within this structure, full authority over and responsibility for the content of all evaluation reports rests with the evaluation co-directors. David Stapleton led the writing of this report.

Table of Contents

Abbreviations Used in This Report

AWI	Average Wage Index	**MEF**	Master Earnings File
BODS	BOND Operations Data System	**PHUS**	Payment History Update System
BOND	Benefit Offset National Demonstration	**SGA**	Substantial Gainful Activity
BYA	BOND Yearly Amount	**SSA**	Social Security Administration
CPI	Consumer Price Index	**SSDI**	Social Security Disability Insurance
DAC	Disabled Adult Child	**SSI**	Supplemental Security Income
DWB	Disabled Widow/Widowers	**SSR**	Supplemental Security Record
EWIC	Enhanced Work Incentive Counseling	**TWP**	Trial Work Period
GP	Grace Period	**WIC**	Work Incentive Counseling
IRS	Internal Revenue Service	**WIPA**	Work Incentives, Planning, and Assistance
IRWE	Impairment Related Work Expense		

1.   Introduction

The Benefit Offset National Demonstration (BOND) is a random assignment demonstration that tests one variant of Social Security Disability Insurance (SSDI) program rules governing work and other supports. This Snapshot Report concerns Stage 1 of BOND, which was designed to examine how a national benefit offset and accompanying administrative changes would affect earnings and program outcomes for the entire SSDI population. Stage 2 was designed to learn more about the impacts of the benefit offset for those most likely to use it, and to determine the marginal effects of the delivery of more intensive counseling services than those offered in Stage 1. This is the second in a series of *Stage 1 Snapshot Reports* about impacts on earnings and benefits paid. This introductory chapter provides a synopsis of the demonstration, describes the purpose of this report, and ends with an outline of the rest of the report. Impact estimates for Stage 2 will appear in a separate series of reports.

1.1.   Synopsis of BOND

Under current program rules, SSDI beneficiaries lose all SSDI benefits after a sustained period of substantial earnings and risk potential loss of other (non-SSDI) benefits.[1] Specifically, benefits are lost if an SSDI beneficiary's countable monthly earnings exceed the monthly Substantial Gainful Activity (SGA) amount after completing a nine-month Trial Work Period (TWP) and a three-month Grace Period (GP). In 2012, the SGA amount was $1,010 per month for non-blind beneficiaries and $1,690 per month for blind beneficiaries. The complete loss of benefits for earnings in excess of the SGA amount is sometimes called the "cash cliff." The cash cliff likely discourages some beneficiaries from working at all and encourages those who could work with earnings above the SGA level to keep their earnings below that level.

BOND replaces the cash cliff with a "ramp" (benefit offset) with the policy objective of encouraging beneficiaries who can work above the SGA level to increase their earnings and reduce their reliance on benefits. More specifically, the benefit offset is expected to increase the earnings of those who might otherwise not work at all, or might not attempt to earn more than the SGA amount. If such individuals engage in SGA under the benefit offset, their benefits will ultimately be reduced. Offsetting the possible benefit reduction for this group are benefits paid under BOND to those who would have had earnings above the SGA amount in the absence of BOND. Thus, the direction of the net impact on mean earnings and benefits of all beneficiaries will depend on the size of the impacts for beneficiaries who would not engage in SGA under current law relative to the size of the impacts for those who would. Those in the latter group lose their benefits entirely under current law, whereas under the benefit offset many, perhaps most, will be eligible for a reduced SSDI benefit. While still on the ramp—i.e., while earning above the SGA amount but less than the zero-benefit amount at the end of the ramp—beneficiaries can increase the size of their benefits by working at less than their full earnings potential.

BOND also changes the administrative processes to adjust benefits, including replacing the monthly SGA calculation with an annualized measure of SGA referred to as the "BOND Yearly Amount (BYA)." BYA

---

[1]   Other benefits include Medicare for those on the rolls for at least 24 months, which are extended for a lengthy period following suspension of SSDI benefits, but not indefinitely. Some also receive Supplemental Security Income, Medicaid, or a variety of other public or private benefits that are contingent on earnings in some fashion.

is equal to 12 times the monthly SGA amount (in 2012, $12,120 for non-blind and $20,280 for blind T1subjects). The benefit offset reduces benefits by $1 for every $2 in countable annual earnings in excess of the BYA following the completion of the GP. The change to an annual accounting period was designed to reduce the cost of administering the offset. It can also be very helpful to beneficiaries who have variable monthly earnings. SSA continues to pay benefits monthly under BOND, but the monthly payment amount is based on expected annual earnings. In the following calendar year SSA reconciles payments to actual countable earnings, based on information provided by the Internal Revenue Service (IRS), documentation provided by the beneficiary, or both.

Other administrative differences might be especially important in the early years of BOND, including 2012. The administrative procedures established to initially notify T1subjects of their assignment to the benefit offset, to provide them with information, and to implement benefit adjustments under the offset likely affected the speed with which retroactive payment adjustments were made and improper past payments recovered. By necessity, the impact estimates in this document focus on benefits paid *in* 2012. Impacts on benefits paid *for* 2012, which are not observed in the data available for this report, might be quite different, after all retroactive benefit adjustments and repayments of improper payments have been completed. For future reports, we plan to include estimates of the impact of BOND on benefits paid for the years in the evaluation period.

Finally, T1 subjects have access to counseling services that are tailored to the benefit offset, but otherwise designed to be comparable to counseling services available to all beneficiaries under current law. Although not intended, it is possible that the counseling services implemented are having an impact on earnings and benefits above and beyond the impact of the offset itself.

Although some T1 subjects were eligible to use the offset immediately after random assignment in May 2011, most were not. Prior to using the offset, they had to be engaged in SGA, have completed the nine-month TWP and three-month GP. Some might also have needed to obtain employment services, which could take additional time. Finally, even beneficiaries in position to use the offset immediately had to trust the opportunity presented by their assignment to T1 and understand how to take advantage of it quickly (Wittenburg et al. 2012).

The evaluation team is responsible for all the estimates that appear in this report. In previous reports, we described the BOND design and the framework for estimating the impacts, summarized early assessment activities on the infrastructure to support Stage 1 service delivery, and reported 2011 earnings and benefit impacts for Stage 1 subjects (Stapleton et al. 2010; Bell et al. 2011; Wittenburg et al. 2012; Stapleton et al. 2013, respectively). The team is documenting the outcomes of Stage 2 in a series of parallel reports.

1.2. Purpose

This *Snapshot Report* presents estimates of the combined impact of the benefit offset and accompanying administrative changes for Stage 1 (hereafter referred to as *benefit offset impacts*) from January through December 2012. Random assignment occurred in May 2011, so the 2012 calendar year represents the second year and the first full calendar year of the evaluation period.

This second *Snapshot Report* uses the same evaluation framework and analysis sample that our first *Snapshot Report* (Stapleton et al. 2013) used to estimate impacts through calendar year 2011. Within that framework, the two most important evaluation outcomes—referred to as *confirmatory outcomes*—are

total earnings and total SSDI benefits paid. In keeping with that structure, mean earnings in 2012 and mean benefits paid *in* 2012 are the confirmatory outcomes in this report. However, the discussion in this report places greater emphasis on 2012 earnings because of the expectation that benefits paid *for* 2012, after retroactive adjustments and recovery of improper payments, will ultimately differ substantially from benefits paid *in* 2012, for reasons previously explained. The final impact evaluation will use a measure of benefits paid *for* the years in the evaluation period as the confirmatory outcome—the measure of benefits paid that is most important for policy purposes.[2] Statistically significant findings for the confirmatory outcomes in this report should be interpreted as confirming that the benefit offset had an impact on at least one of two outcomes: 2012 earnings and/or SSDI benefits paid *in* 2012. The findings do not provide information on SSDI benefits ultimately to be paid *for* 2012 because we cannot measure this outcome with the available administrative data for this report.

The report also presents estimates for several exploratory outcomes related to 2012 earnings and benefits paid in 2012, to explore the possibility of other impacts of the offset. These findings do not confirm that the benefit offset had impacts; they suggest only where such effects might have occurred. In providing more information on the potential impacts of the benefit offset these findings receive less weight than the confirmatory findings in assessing the overall success of the tested treatment.

1.3.   Organization of Report

The rest of this report is comprised of three sections. Chapter 2 provides background information on the impact estimation methodology and descriptive findings that provide context for the impact estimates. Chapter 3 presents the impact findings for the confirmatory and exploratory outcomes. Chapter 4 includes a brief discussion of the results and expectations about future impacts.

---

[2]   Bell et al. (2011) identify benefits paid as the confirmatory measure and indicate only the difference between benefits paid *in* a period versus *for* a period in a footnote (footnote 40). It became apparent more recently that the difference between these two measures might be quite large, especially in the early demonstration years.

2.   Methodology and Context

The goal for the Stage 1 evaluation is to learn about offset utilization and key impacts when the benefit offset is offered to all SSDI beneficiaries. The evaluation for Stage 1 compares outcomes for beneficiaries who were randomly assigned to the following groups:[3]

- **T1 subjects** are beneficiaries whose benefits are determined by the benefit offset rules over a period of at least five years and who have the opportunity to use ancillary demonstration services.

- **C1 subjects** are a control group that continues to receive benefits according to current law.

Because only a small fraction of T1 subjects will likely use the benefit offset, the T1 and C1 groups must be very large (tens of thousands each) to detect policy-relevant impacts (Stapleton et al. 2010). The BOND sample for random assignment included all SSDI beneficiaries between the ages of 20 and 59 in 10 randomly selected sites throughout the nation who were receiving benefit payments in April 2011.

For this report, administrative data for calculating earnings and benefit impacts were available through calendar year 2012. Earnings are measured from the SSA Master Earnings File (MEF), which contains longitudinal information on wages and self-employment income reported to the IRS. The MEF records were almost 100 percent complete for calendar year 2012 when SSA extracted them for this report. Benefit outcomes are measured from SSA's Payment History Update System (PHUS) for SSDI and the Supplemental Security Record (SSR, for SSI).[4]

The rest of this chapter describes our methodological approach to estimating benefit offset impacts. We then summarize findings related to benefit offset usage from our previous reports and descriptive data on offset adjustments that are likely related to the size and direction of earnings and benefits impacts.

2.1.   Methodological Approach

We initially specified the methodology and outcomes for the impact analysis in Bell et al. (2011), and later implemented the methodology for 2011 impacts with only minor revisions in Stapleton et al. (2013). For this report, we are estimating impacts for calendar year 2012 using the same outcome definitions and approach as for 2011. We review the outcome definitions, anticipated impacts, estimation methodology, and analysis sample below.

---

[3]   The random assignment process included beneficiaries for Stages 1 and 2. Specifically, beneficiaries were randomly assigned to the T1 group, the C1 group, or a "Stage 2 solicitation pool." The latter initially included all random replicates for purposes of recruiting volunteers for Stage 2. When Stage 2 recruitment was completed, subjects in the unused random replicates were assigned to C1. For details on random assignment, see Bell et al. (2011).

[4]   Because the data are collected by the IRS and are therefore subject to IRS access rules, SSA staff have direct access to MEF data, but contractors do not. Consequently, qualified SSA staff accessed the data, submitted programs developed by the BOND team to estimate impacts, reviewed output to ensure that it complied with privacy requirements, and then transmitted the output to the evaluation team. The MEF earnings data are updated annually. The 2012 earnings data for this report were extracted in November 2013.

2.1.1.   Outcome Definitions and Theoretical Impacts

The nine outcomes for this report include two confirmatory outcomes (total earnings and total SSDI benefits paid in 2012) and seven exploratory outcomes (related to employment and benefits). The exploratory earnings outcomes include indicators for earnings in excess of each of three annual earnings thresholds defined by multiples of BYA (the BYA, and two and three times the BYA) and an indicator for any earnings during 2012. The exploratory benefit outcomes include number of months with SSDI payments, total SSI benefits paid, and number of months with SSI payments—each in 2012.

In the discussion that follows, we consider the expected direction of benefit offset impacts on these outcomes, abstracting from administrative factors that could themselves influence the impacts. We then turn to a discussion of administrative factors and their potential influence on impacts.

Although BOND was designed to test whether eliminating the SGA cash cliff and replacing it with the $1 for $2 offset ramp would increase return to work and earnings, and reduce beneficiary's reliance on SSDI benefits (Bell et al. 2011), the theoretical direction of impacts of the benefit offset on mean earnings and benefits is ambiguous (third column of Exhibit 2-1). As described in detail in Bell et al. (2011), this ambiguity arises because the incentives created by the benefit offset vary with what the beneficiary's earnings would be under current law. T1 subjects who would have had no earnings or earnings below BYA under current law are expected, on average, to have higher earnings and lower SSDI benefits under the benefit offset. Conversely, some T1 subjects who would have had earnings well above BYA under current law are expected to have lower mean earnings and higher mean SSDI benefits under the benefit offset.[5] Positive impacts on the mean earnings for all beneficiaries require that positive impacts for those whose earnings would be less than BYA under current law are sufficiently large to offset possible negative impacts for those who would earn more than BYA under current law.

Similarly, the predicted impact on benefits depends on what the earnings of the beneficiary would have been under current law. For those with no earnings or earnings below BYA, the predicted impact is negative; if they earn more than BYA under the offset than they would under current law, their benefits will fall. Conversely, for those who would have had earnings above BYA under current law, benefits for many under the offset are expected to be higher because they will be eligible for a partial benefit rather than no benefit at all, as under current law. Hence, to generate a reduction in mean benefits paid, the reduction in benefits paid to those whose earnings would be less than BYA under current law must exceed the increase in benefits paid to those who would earn more than BYA under current law.

Theory does, however, predict the signs of the impacts for five of the seven exploratory outcomes. It predicts positive impacts on employment, earnings above BYA, and months with SSDI payments and negative impacts on SSI benefits and months with SSI payments. These predictions can be verified by separately considering the impacts for those whose earnings would be below or above BYA under current law. As indicated earlier, for those who would have earnings below BYA under current law, theory predicts that the offset will increase both the percentage employed and the percentage of beneficiaries with earnings above BYA. Those who would have earnings above BYA under current law will have a

---

[5]   Empirically, there is evidence that some high-earning beneficiaries will reduce their earnings, but not reduce employment. Weathers and Hemmeter (2011) found evidence of a reduction in earnings by beneficiaries earning above SGA before random assignment in the Benefit Offset Pilot Demonstration.

stronger incentive to keep their earnings above BYA under the offset than they do under current law—even though some might work and earn less under the offset. It is not possible to predict the direction of impacts on the percentage with earnings well above BYA (for example, two and three times BYA); however, it is expected that some T1 subjects whose earnings would be well above BYA under current law will reduce their earnings in response to the benefit offset.[6]

Theory also predicts that the impact on SSI benefits paid will be negative. Under current law, any beneficiary who concurrently receives SSDI and SSI (a "concurrent" beneficiary) and is engaged in SGA after completing the TWP and GP is entitled to, at most, only an SSI payment.[7] In contrast, a concurrent T1 subject with the same earnings would likely receive a partial SSDI benefit, and the size of the T1 subject's SSI benefit would be reduced by the amount of the partial SSDI benefit or by the entire current-law SSI payment if the latter is smaller than the partial SSDI benefit. The offset might also have an impact on SSI payments to T1 subjects who are SSDI-only beneficiaries at the outset of the demonstration and whose SSDI benefits are below the maximum federal SSI benefit amount. Under current law, some such subjects are likely to enter SSI after they spend down their assets to the point at which they satisfy the SSI resource test. Higher earnings under the offset might reduce or slow the entry of such SSDI-only subjects into SSI.[8]

2.1.2. Administrative Features of the Offset That Influence Impacts

The previous discussion abstracts from the administrative features of the benefit offset that were designed and implemented to facilitate use of the offset by T1 beneficiaries. As described in Bell et al. (2011), because these processes are necessarily different from current law processes, they are part of the T1 intervention being tested under BOND.

In the first years of BOND, the administrative factors most likely to affect outcomes concern the administrative processes leading to the adjustment of benefits—the special processes implemented for T1subjects and the current processes that apply to C1 subjects. For T1 subjects, that process started

---

[6]   The predicted effect of the offset is positive on the percentage of T1 subjects with earnings above BYA, but is theoretically ambiguous on the percentage with earnings above higher thresholds (for example, two and three times BYA). The variation in the direction of the predicted earnings response by initial earnings level is the reason that the sign of the predicted impact on mean earnings is ambiguous (see Bell et al. 2011 for more details).

[7]   Under the SSI Earned Income Exclusion (EIE), monthly SSI benefits are reduced by $1 for every $2 of earnings above an earnings disregard that is as low as $65. Whether a concurrent beneficiary with earnings above SGA is eligible for a federal SSI payment depends on whether the beneficiary's SSI countable income, including earnings not excluded under the EIE and any other countable income, exceeds the maximum federal SSI payment amount. SSI countable income rules exclude $20 of SSDI benefits unless that exclusion is used against some other form of unearned income. Beyond any exclusion, and holding earnings constant, every $1 of SSDI benefits reduces the SSI payment amount by $1 until the SSI payment amount is zero. At any earnings amount above SGA, any SSDI payment under the offset displaces any SSI payment that is due, dollar for dollar. Under BOND, the benefit offset indirectly affects the SSI payment amounts through the SSDI benefit adjustment. For example, for a concurrent T1 subject with earnings above BYA and positive SSI benefit amounts, a $2 increase in earnings would result in a $1 increase in EIE (reducing SSI) and a $1 decrease in SSDI (increasing SSI), which would leave SSI payments unchanged.

[8]   See Riley and Rupp (2012).

between May and August of 2011 when they were notified of their eligibility for the offset in a letter.[9] The letter explained the offset and urged them to contact the demonstration. Some of those eligible to use the offset contacted the demonstration as recommended and their benefits were eventually adjusted via an administrative process set up for that purpose. Others eligible to use the offset early did not contact the demonstration, however. Instead, SSA discovered their high earnings in its annual review of earnings reported to the IRS, and then initiated the process to adjust their benefits.

The BOND administrative data show that most T1 subjects eligible to use the offset during 2011 and 2012 did not have a benefit adjustment until after 2012. The implication is that a substantial majority of adjustments to benefits paid to T1 subjects for 2012 are not reflected in benefits paid in 2012. This issue is important because our benefit measures for this report are based on benefits paid *in* 2012; it does not reflect the many retroactive adjustments to benefits paid *for* 2012 that were made after 2012.

The direction and size of the impacts of this administrative factor depend on how the processes for the T1group compare to the corresponding processes for C1 subjects. The most striking difference is that T1 subjects had to be notified about a change in the earnings rules before the benefit adjustment process could start, whereas C1 subjects were subject to rules that had been in place for many years. Additionally, C1 subjects could rely on advice from trusted third parties, such as employees of disability service organizations, who had an understanding of existing program rules governing earnings, whereas advice that trusted third parties provided to T1 subjects was not always well-informed (Wittenburg et al. 2012).[10] Finally, T1 administrative processes had not been previously implemented in a large scale, resulting in start-up delays, whereas the C1 processes have been in place for many years.

One other administrative factor seems likely to have a positive impact on benefits paid *for* 2012, and possibly on benefits paid *in* 2012, but an ambiguous impact on 2012 earnings: the change from monthly to annual accounting. Although the purpose of this change was to simplify administration of the offset, it is expected to also help beneficiaries with highly variable earnings (for example, seasonal workers) to a significant degree. Under monthly accounting, earnings above SGA in any month reduce benefits for that month, but under annual accounting the benefit reduction for those same earnings might be smaller or zero because of earnings below the SGA amount in other months of the same year. Holding earnings constant, this administrative change is expected to increase the benefits paid to some beneficiaries; any increase (decrease) in earnings due to this factor will reduce (increase) benefits. The theoretical sign of the impact of this administrative change on earnings is ambiguous.

---

[9]   SSA also sent a letter to T1 subjects that provided a detailed explanation of the effect of the offset on benefits.

[10]   T1 subjects did have access to specially trained counselors, but some also found their way to counselors who were not specially trained or who were not aware of the beneficiary's T1 status.

Exhibit 2-1.   **Definitions of Confirmatory and Exploratory Outcomes and Predicted Signs of Impacts**

	Definition	Predicted Sign
Confirmatory Outcomes		
Total earnings in 2012	2012 earnings	?
Total SSDI benefits paid in 2012	Sum of SSDI benefit payments from January through December 2012; for SSDI workers, this includes benefits for dependent spouses and minor children, but not for DAC[a]; for DAC and DWB, it includes only benefits payable to the DAC or DWB	?
Exploratory Outcomes		
Earnings Outcomes (January–December 2012)[b]		
Employment in 2012	Any 2012 earnings	+
Earnings above BYA	2012 earnings above $12,120 (non-blind subjects) or $20,280 (blind subjects)	+
Earnings above 2 × BYA	2012 earnings above $24,240 (non-blind subjects) or $40,560 (blind subjects)	?
Earnings above 3 × BYA	2012 earnings above $36,360 (non-blind subjects) or $60,840 (blind subjects)	?
Benefit Outcomes (January–December 2012)		
Number of months with SSDI payments	Number of months with SSDI benefits paid above zero	+
Total SSI benefits paid	Sum of SSI benefit payment amounts from January through December 2012	–
Number of months with SSI payments	Number of months with SSI benefits paid above zero	–

Notes: Bell et al. (2011) provide detailed discussion on the hypothesized impacts of benefit offset.

[a] For a description of family benefits, see [http://www.socialsecurity.gov/pubs/10024.html#a0=3]; accessed May 23, 2014.

[b] Earnings relative to BYA is based on earnings reported in the MEF, without adjustment for impairment-related work expenses (IRWE). Less than 1 percent of SSDI and SSI beneficiaries use IRWEs (Livermore et al. 2009), and even when used, they do not appear in administrative records until claimed by the beneficiary and approved by SSA.

2.1.3. Impact Estimation Methodology

The goal of the Stage 1 BOND experiment is to make inferences about what the impact of the benefit offset would be if it applied to all SSDI beneficiaries in the nation who met the BOND eligibility criteria as of May 2011. The statistical design for the demonstration supports the production of unbiased point estimates and standard errors for this population. The standard errors reflect random variation associated both with the selection of the BOND sites and with assignment of subjects in those sites to T1 and C1. As a result, each test of a null hypothesis for "no impact" on the mean of a specific outcome is a test of no impact for all beneficiaries nationwide.

To estimate impacts, we compare mean outcomes for the T1 group to mean outcomes for the C1 group that have been weighted for differences in sampling rates across sampling strata and adjusted for the

effects of small random differences in baseline characteristics.[11] The adjustments for differences in baseline characteristics serve to reduce the standard errors. For each specific outcome, we test the null hypothesis of no impact. Each individual test uses a specified level of significance. For example, a 10 percent significance level means that if the null hypothesis is true, there is only a 10 percent chance that the test will mistakenly reject it.

The impact estimates are "intent to treat" estimates. In other words, they capture the mean impact of the applicability of the benefit offset rules to the earnings of all T1 subjects, whether or not those subjects work and use the offset. Hence, they reflect "no impacts" for the large majority who would not have any earnings under current law or the offset, as well as for those who fail to learn about, understand, or trust the offset.

We make a multiple-comparison adjustment for the two confirmatory outcomes—outcomes selected on the basis of theory and policy interest alone (see Bell et al. 2011). This adjustment is necessary because we are testing multiple outcomes, which makes the probability of a Type I error (rejecting the null hypotheses if it is true) larger than the significance level for the individual tests. To compensate for this effect, we adjust the test statistics for each of the two confirmatory outcomes so that the probability of rejecting the null hypothesis of no impact on either confirmatory outcome is equal to the specified significance level if the null hypothesis is true.[12]

We make no adjustment to the tests for exploratory outcomes. Readers are advised to give less weight to any individual significant result from an exploratory test than they would to an equally significant result from a confirmatory test.

We estimate impacts for the overall Stage 1 sample and two pairs of subgroups identified in our *Analysis Plan*, one defined by duration of SSDI benefit receipt and the other by SSI status in the month prior to random assignment (Bell et al. 2011). The duration subgroups are of interest because prior research and program rules suggest that subjects who have been on the rolls for a *short duration* (defined as three years or less) will respond to the benefit offset differently from those who have been on the rolls for a *long duration* (more than three years). More specifically, we expect more short-duration subjects to work in comparison to long-duration subjects. However, we expect it will take longer for short-duration subjects to actually use the benefit offset, because they will have completed fewer TWP and GP months at the outset of the demonstration in comparison to long-duration subjects. The SSI status subgroups are of interest because concurrent subjects face incentives different from those of SSDI-only subjects. For concurrent subjects, SSI benefits are offset more immediately under existing SSI work incentives, and each additional dollar of SSDI benefits is offset by a dollar reduction in SSI benefits after a small disregard. We treat all subgroup analyses, including the tests of earnings and SSDI benefits paid, as exploratory.

---

[11]   All of the estimates are weighted to ensure that the T1 and C1 analysis samples are representative of the national beneficiary population who met BOND eligibility criteria in the month of random assignment. These weights take account of the differing probabilities of selection into the sample for the different study sites and beneficiary subpopulations (see Stapleton et al. 2013 for more details on the construction of the weights).

[12]   Our approach adjusts the *p*-values for the confirmatory outcomes using the Westfall and Young (1993) method. Details of the *p*-value adjustments for tests of impacts on the confirmatory outcomes appear in Appendix A of Stapleton et al. (2013). See Schochet (2008) for further discussion of the multiple comparisons problem.

2.1.4. Final Analysis Sample Sizes

Exhibit 2-2 presents the sizes for the overall sample and the subgroups. The final Stage 1 analysis sample contains a total of 968,713 subjects, spread across T1 (77,115) and C1 (891,598).[13] By design, the T1 sample is evenly split between short- and long-duration beneficiaries (see Bell et al. 2011).[14] There are substantially more SSDI-only subjects relative to concurrent subjects, because most SSDI beneficiaries do not receive SSI, and sampling rates did not depend on SSI status. As a result, we have less power to detect impacts for concurrent subjects relative to the other subgroups.

The baseline characteristics (not shown) for the weighted T1 sample are statistically equivalent to those for the weighted C1 sample (Stapleton et al. 2013). These descriptive findings give us a high level of confidence that any statistically significant differences in subsequent outcomes between the T1 and C1groups are "internally valid" impact estimates. That is, significant estimates represent real impacts of the benefit offset on outcomes for the T1 group, rather than systematic preexisting differences between the two groups or their environments.

Exhibit 2-2.   Stage 1 Analysis Sample Composition

Random Assignment Group	Full Sample	Duration		SSI Status	
		Short Duration	Long Duration	SSDI-Only	Concurrent
T1	77,115	38,669	38,446	64,709	12,406
C1	891,598	209,790	681,808	694,270	197,328

Source: BOND Operations Data System (BODS).

Notes: The Stage 1 analysis sample excludes (1) subjects who were initially assigned to the sample but were later determined to have died prior to assignment, and (2) any pair of beneficiaries (e.g., a primary and a DAC or two DACs with the same primary) on a common primary record who were assigned to different BOND groups (see Stapleton et al. 2013 for details on this adjustment). Weights are used to ensure that the BOND subjects who meet the analysis criteria in both the T1 and C1 analysis samples are representative of the national beneficiary population in the month of random assignment. The weighted population size is 6,526,888.

---

[13]   The team randomly selected nearly 80,000 beneficiaries for the T1 group. The C1 group included an initial core group (C1-core) that was the same size as the T1 group and a supplemental sample of C1 subjects (C1-supplements) that was added to the C1 sample after completion of Stage 2 recruitment (C1-supplement): those BOND-eligible subjects who were not included in the samples that were released for Stage 2 recruitment. The C1 sample is the combination of the C1-core and C1-supplement samples. As shown in Stapleton et al. (2013), after the application of weights, there were no statistically significant differences between the characteristics of the full C1 sample and those of the C1-core sample. Further, use of the C1-core sample for estimation of impacts on 2011 outcomes yielded estimates that were not statistically different from those based on the full C1 sample. As expected, standard errors were substantially higher when we used only the C1-core sample. Because of the earlier findings for 2011, this report presents findings only from the full C1 group.

[14]   As noted in the previous footnote, the C1 group includes C1-Core and C1-Supplemental subjects. The C1-core group was also evenly split between the short- and long -duration groups. The characteristics of the C1 supplemental sample differ because it includes a larger share of long-duration beneficiaries and excludes SSDI-only beneficiaries in the BOND sites that were in the recruitment pool for Stage 2 of BOND. To account for the differing compositions of the T1 and C1 subjects, we use sample weights that eliminate the effects of compositional differences on comparisons of T1 and C1 means. See Stapleton et al. (2013) for more details.

2.2.   Previous Findings

Previously, we found no evidence that the offset had (1) an impact in 2011 on mean earnings; (2) a very small, positive impact on mean benefits paid; and (3) no impacts for exploratory outcomes (Stapleton et al. 2013). We attributed the minimal evidence of impacts to the very limited use of the offset during the short period for measuring impacts (eight months) available in 2011.[15] SSA had adjusted the benefits paid *in* 2011 to 0.1 percent of T1 subjects (39 subjects), whereas we anticipated that SSA would eventually adjust the benefits of 800 or more T1 subjects *for* 2011. This large discrepancy reflects all of the potential causes of delay outlined in Section 2.1.

---

[15]   As noted in Stapleton et al. (2013), the impact estimates could also reflect issues related to a sluggish economy.

3.   Impact Findings

This chapter presents impact estimates for the confirmatory and exploratory outcomes. As outlined in Chapter 2, each estimate is the difference between the weighted T1 group mean for the outcome and the weighted C1 group mean for the same outcome after statistical adjustments to the C1 group mean for differences in observed characteristics. The T1 sample itself is weighted to reflect the national beneficiary population meeting the BOND eligibility criteria in 2011, so the impact estimates are unbiased estimates of the impact of benefit offset, as implemented under BOND, for the entire eligible population.

Statistically significant impact estimates are identified at the 1, 5, and 10 percent levels. We describe estimates that are statistically significant at the 1 percent level as "strong evidence," at the 5 percent level as "evidence," and at the 10 percent level as "weak evidence." Impact estimates not significant at the 10 percent level or below are considered insignificant. The size of our sample allows our regression-adjusted models to detect very small impacts for several outcomes. As will be seen, some significant impacts are very small as a percentage of the adjusted C1 group mean.

3.1.   Full Stage 1 Treatment Group

Exhibit 3-1 presents the estimates of 2012 impacts on earnings and benefit outcomes for the full Stage 1 BOND sample. Total earnings and total benefits paid are the confirmatory outcomes for January through December 2012. All remaining earnings and benefit outcomes are exploratory.

3.1.1.   Confirmatory Impacts

We find no evidence that the benefit offset had an impact on earnings in 2012. The earnings of both T1 and C1 subjects were relatively low ($1,287 and $1,277 for T1 and C1 subjects, respectively), as few subjects in both groups (about 15 percent) worked during the previous year.

There is strong evidence of a positive impact on SSDI benefits paid in 2012. The estimated impact on mean benefits paid was $69, equivalent to 0.6 percent of adjusted mean SSDI benefits paid to C1 subjects during 2012 ($69/$11,324). The 95 percent confidence interval for this estimate ranges from $42 to $96.

3.1.2.   Exploratory Impacts

Only one exploratory benefit outcome showed strong evidence of an impact. There was an increase of 0.05 in the number of months with SSDI payments, equivalent to 0.4 percent of the adjusted mean for C1 subjects (10.98 of a possible 12 months). This finding presumably reflects retention of partial benefits under the offset in some months by T1 subjects who, under current law, would have lost their benefits entirely because of earnings. As with the impact on benefits paid, the impact on the share of months with benefits is much more substantial for the small percentage of subjects whose benefits were affected by the offset. For instance, if they represent 0.5 percent of T1 subjects, the mean impact for them would be 10 added months with benefits (0.05/0.005). We do not find any evidence of effects on SSI benefits paid or the number of months with SSI payments.[16]

---

[16]   For T1 and C1 subjects, the mean total SSI payment was around $460, and the mean number of months with an SSI payment was 2.0 months. The mean of total SSI benefits paid was small in comparison to the mean of total

There were no statistically significant impacts for the four exploratory earnings outcomes. Our findings indicate that most T1 subjects did not work in 2012, and only a small portion had earnings above BYA. Just over 15 percent of subjects in each group had at least some earnings in 2012, including 2.6 percent with earnings above BYA, about 1.1 percent with earnings above twice BYA, and 0.6 percent with earnings above three times BYA.

Exhibit 3-1.   Stage 1 Impact Estimates on Earnings and Benefit Outcomes

	T1 Mean	C1 Mean	Impact Estimate
Earnings (January–December 2012)			
Total earnings in 2012 **(confirmatory)**	$1,283	$1,277	$6 ($31)
Employment in 2012	15.41%	15.34%	0.06 (0.11)
Earnings above BYA	2.68%	2.60%	0.08 (0.11)
Earnings above 2 × BYA	1.12%	1.12%	-0.01 (0.08)
Earnings above 3 × BYA	0.58%	0.62%	-0.04 (0.03)
Benefits Paid(January–December 2012)			
Total SSDI benefits paid in 2012 **(confirmatory)**	$11,393	$11,324	$69*** ($14)
Number of months with SSDI payments	11.03	10.98	0.05*** (<0.01)
Total SSI benefits paid in 2012	$459	$462	-$3 ($7)
Number of months with SSI payments	1.99	1.99	-0.00 (0.01)

Source: Analysis of SSA administrative records from the MEF, BODS, PHUS, and SSR.

Notes: Weights are used to ensure that the BOND subjects who met analysis criteria in both the T1 and the C1 analysis samples are representative of the national beneficiary population in the month of random assignment. Standard errors are in parentheses. Unweighted sample sizes: T1 = 77,115; C1 = 891,598. See Chapter 2 for variable definitions. Impact estimates are regression adjusted for baseline characteristics. All outcomes are measured during the second calendar year of operations (January 2012–December 2012). Total earnings and SSDI benefits paid are the two confirmatory outcome variables, and statistical tests for the impacts on these two outcomes used multiple-comparison adjustments; see Stapleton et al. (2013) for details on the statistical tests and adjustments to the *p*-values. Tests for impacts on all other outcomes (exploratory outcomes) were conducted independently, without multiple-comparison adjustments.

*/**/*** Impact estimate is significantly different from zero at the 10, 5, and 1 percent levels, respectively, using a two-tailed t-test.

SSDI benefits paid ($11,324 for C1 subjects), reflecting that only a small minority of Stage 1 subjects received SSI benefits in 2012.

3.2.   Subgroups

Below we present the impact estimates for the subgroups defined by duration of SSDI benefit receipt (Exhibit 3-2) and SSI status (Exhibit 3-3). The outcome measures are the same as those in Exhibit 3-1, but stratified by subgroup. For reasons outlined in Chapter 2, we consider all subgroup estimates as exploratory outcomes. Hence, we did not adjust significance tests for multiple comparisons. For each pair of subgroups, we first describe adjusted outcome means for C1 subjects in the two subgroups; these reflect population differences for the subgroups under current law.[17] We then describe impacts within each subgroup of the pair and discuss any evidence of differences in impacts across each pair of subgroups.

3.2.1.   Duration Since Award

Consistent with expectations and our findings from previous reports (Stapleton et al. 2013), there were differences in the levels of 2012 earnings and benefit levels for C1 subgroups by duration (Exhibit 3-2). Short-duration C1 subjects had higher mean earnings ($1,418 versus $1,215) and were more likely to be employed (15.6 versus 15.2 percent) than long-duration subjects. The findings are consistent with past research demonstrating that recent entrants were more likely to have earnings than those who have been on the rolls for a longer period (Liu and Stapleton 2011). In addition, short-duration subjects had higher SSDI benefit payments ($12,269 versus $10,919 for C1 subjects).[18] Total SSI payments were also lower for the short-duration group ($405 versus $485 for C1 subjects).

Within both subgroups there is evidence of impacts on the number of months with SSDI payments, and for the short-duration subgroup there is weak evidence of a decrease in the percentage with earnings above $2 \times BYA$. The impact estimates for the number of months with SSDI payments are 0.03 and 0.06 months for short- and-long duration subjects, respectively. For short-duration beneficiaries, the impact estimate for the percentage earning above $2 \times BYA$ is -0.15 percentage points. The theoretical prediction for the direction of the impact on the percentage with earnings above $2 \times BYA$ is ambiguous. The findings indicate that the number of T1 subjects induced to lower their earnings to less than $2 \times BYA$ exceeds the number induced to increase their earnings above $2 \times BYA$. The value of the estimated impact is equivalent to 7.0 percent of the adjusted percentage of short-duration C1 subjects with earnings above $2 \times BYA$, 2.13 percent.

There is little evidence of heterogeneity of impacts across the two subgroups, based on the difference in impact estimates in column 7 of Exhibit 3-2. The one exception is for the impact on months with SSDI payments, which is twice as large for the long-duration group (0.6 months) as for the short-duration group (0.3 months). This difference presumably reflects the fact that more long-duration subjects had completed

---

[17]   We report only differences in subgroup means that provide at least weak evidence of statistical differences (that is, they are significant at the 10 percent level based on a t-test).

[18]   This difference likely reflects the way that SSA indexes pre-SSDI earnings when calculating benefit amounts. Specifically, SSA uses an average wage index (AWI) to inflate past earnings prior to calculating the initial benefit amount; after that, SSA adjusts benefits for inflation each year using a Consumer Price Index (CPI). As the AWI usually increases faster than the CPI, mean benefits for new awardees typically increase every year after adjustment for price inflation.

their TWP and GP by early 2011 and hence had more months in 2011 during which they were eligible for an adjustment.[19]

Exhibit 3-2.   Stage 1 Impact Estimates for Subgroups Defined by Duration of SSDI Receipt

	Short-Duration			Long-Duration			Difference in Impact (7)
	T1 Mean (1)	C1 Mean (2)	Impact Estimate (3)	T1 Mean (4)	C1 Mean (5)	Impact Estimate (6)	
Earnings (January–December 2012)							
Total earnings in 2012	$1,418	$1,418	$0 ($46)	$1,225	$1,215	$10 ($33)	-$10 ($57)
Employment in 2012	15.60%	15.58%	0.02 (0.22)	15.32%	15.24%	0.08 (0.12)	-0.06 (0.25)
Earnings above BYA	3.02%	2.95%	0.07 (0.14)	2.53%	2.44%	0.09 (0.12)	-0.02 (0.18)
Earnings above 2 × BYA	1.20%	1.35%	-0.15* (0.07)	1.08%	1.03%	0.05 (0.09)	-0.20 (0.11)
Earnings above 3 × BYA	0.75%	0.78%	-0.03 (0.08)	0.51%	0.55%	-0.04 (0.03)	0.01 (0.09)
Benefits Paid (January–December 2012)							
Total SSDI benefits paid in 2012	$12,310	$12,269	$41 ($22)	$10,996	$10,919	$77*** ($22)	-$36 ($31)
Number of months with SSDI payments	11.08	11.05	0.03** (0.01)	11.01	10.95	0.06*** (<0.01)	-0.03** (0.01)
Total SSI benefits paid in 2012	$404	$409	-$5 ($10)	$482	$485	-$3 ($8)	-2.00 (12.81)
Number of months with SSI payments	1.55	1.56	-0.01 (0.01)	2.17	2.18	-0.01 (0.01)	0.00 (0.01)

Source: SSA administrative records, from the MEF, BODS, PHUS, and SSR.

Notes: Weights are used to ensure that the BOND subjects who meet analysis criteria in both the T1 and the C1 analysis samples are representative of the national beneficiary population in the month of random assignment. Standard errors are in parentheses. Unweighted sample sizes: for short-duration, T1 = 38,669 and C1 = 209,790; for long-duration, T1 = 38,446 and C1 = 681,808. See Chapter 3 for variable definitions. Impact estimates are regression adjusted. All outcomes are measured during the second calendar year of operations (January 2012–December 2012). Tests for impacts on all outcomes were conducted independently, without multiple-comparison adjustments.

*/**/*** Impact estimate is significantly different from zero at the 10, 5, and 1 percent levels, respectively, using a two-tailed t-test.

3.2.2.   SSI Benefit Status

Consistent with expectations and our findings from previous reports (Stapleton et al. 2013), there were also large differences in levels of 2012 earnings and benefit levels for C1 subgroups by SSI status (Exhibit 3-3). Relative to concurrent subjects, SSDI-only subjects had higher mean SSDI benefit payments ($12,588 versus $5,651 for C1 and $12,650 versus $5,731 for T1) and higher mean earnings ($1,382 versus $802 for C1 and $1,393 versus $792 for T1), which reflects the fact that SSDI-only beneficiaries generally have more substantial earnings histories than concurrent beneficiaries. The

---

[19]   For example, based on BODS data through December 2013, more than twice as many long-duration T1 subjects than short-duration subjects had started their TWP (22 versus 8 percent).

percentage employed in 2012 for both groups was about the same (about 15 percent), so the large difference in mean earnings indicates that SSDI-only beneficiaries who worked in 2012 earned much more than concurrent beneficiaries who worked. Given that concurrent subjects are classified on the basis of SSI payments at the time of random assignment, it is not surprising that concurrent subjects had substantially higher mean SSI payments in 2012 than did SSDI-only beneficiaries ($2,375 versus $38 for C1 subjects in the two groups, respectively). The fact that some subjects in the SSDI-only group received SSI benefits after random assignment presumably reflects declines in assets or income from other sources of sufficient size to satisfy the SSI means test.

Our estimates for the SSDI-only group provide weak evidence of an impact on employment. The point estimate, 0.18 percentage points, is equal to 1.2 percent of the adjusted mean for the SSDI-only C1 subgroup (15.38 percent).

For both groups, we find evidence of positive impacts for months with SSDI payments. For the SSDI-only group, we also find strong evidence of impacts on benefit payments and weak evidence of a reduction in SSI payments. The benefit impacts for SSDI-only subjects, who make up 83.9 percent of all T1 subjects, mirror the impacts shown in Exhibit 3-1 for all T1 subjects. In addition, there is weak evidence that the offset reduced the mean SSI benefits of SSDI-only T1 subjects in 2012, by an estimated $2 per beneficiary.[20]

There is a general pattern of no differences in outcomes by SSI status, which is consistent with findings from 2011 (see Stapleton et al. 2013). The one exception is weak evidence that the estimated impact on the employment rate is 0.6 percentage points higher for SSDI-only subjects than for concurrent subjects.[21]

---

[20]   This likely reflects their increased SSDI benefits; other things constant, each dollar of SSDI benefits reduces SSI benefits for those eligible for benefits by the same amount. It might also mean that some SSDI-only T1 subjects who would have entered SSI under current law did not do so because of the offset.

[21]   Although theory predicts a positive impact on employment for SSI-only subjects, as well as for other groups, the point estimate for the impact on employment of SSI-only subjects is negative. We regard this as nothing but random noise because the value of the estimate is less than half the size of its standard error.

Exhibit 3-3.　　Stage 1 Impact Estimates for Subgroups Defined by Baseline SSI Status

	SSDI-Only			Concurrent			Difference in Impact (7)
	T1 Mean (1)	C1 Mean (2)	Impact Estimate (3)	T1 Mean (4)	C1 Mean (5)	Impact Estimate (6)	
Earnings (January–December 2012)							
Total earnings in 2012	$1,393	$1,382	$11 ($38)	$792	$802	-$10 ($34)	$21 ($51)
Employment in 2012	15.56%	15.38%	0.18* (0.09)	14.74	15.18%	-0.46 (0.36)	0.64* (0.37)
Earnings above BYA	2.92%	2.82%	0.10 (0.13)	1.57	1.61%	-0.04 (0.09)	0.14 (0.16)
Earnings above 2 × BYA	1.31%	1.30%	0.01 (0.10)	0.26	0.34%	-0.08 (0.08)	0.09 (0.13)
Earnings above 3 × BYA	0.68%	0.73%	-0.05 (0.04)	0.12	0.10%	0.02 (0.05)	-0.07 (0.06)
Benefits Paid (January–December 2012)							
Total SSDI benefits paid in 2012	$12,650	$12,588	$62*** ($10)	$5,731	$5,651	$80 ($52)	-$18 ($53)
Number of months with SSDI payments	11.09	11.04	0.05*** (<0.01)	10.76	10.67	0.09** (0.03)	-0.04 (0.03)
Total SSI benefits paid in 2012	$36	$38	-$2* ($2)	$2,364	$2,375	-$11 ($37)	$9 ($37)
Number of months with SSI payments	0.13	0.13	0.00 (0.00)	10.33	10.34	-0.01 (0.04)	0.01 (0.04)

Source: SSA administrative records, from the MEF, BODS, PHUS, and SSR.

Notes: Weights are used to ensure that the BOND subjects who meet analysis criteria in both the T1 and the C1 analysis samples are representative of the national beneficiary population in the month of random assignment. Standard errors are in parentheses. Unweighted sample sizes: for SSDI-only, T1 = 64,709 and C1 = 694,270; for concurrent, T1 = 12,406 and C1 = 197,328. See Chapter 2 for variable definitions. Impact estimates are regression adjusted. All outcomes are measured during the second calendar year of operations (January 2012–December 2012). Tests for impacts on all outcomes were conducted independently, without multiple-comparison adjustments.

*/**/*** Impact estimate is significantly different from zero at the 10, 5, and 1 percent levels, respectively, using a two-tailed t-test.

4.   Discussion

In this section, we summarize and analyze the impact findings for the confirmatory and exploratory outcomes and compare them to the 2011 findings in the previous *Snapshot Report* (Stapleton et al. 2013).

4.1.   Summary of Findings

The pattern of findings in 2012 for the two confirmatory outcomes is generally similar to the pattern in 2011 (third column of Exhibit 4-1); i.e., we find no impact on earnings, though we do find strong evidence of a positive impact on SSDI benefits paid in 2012 (last column). The dollar magnitude of impact on SSDI benefits paid in 2012 is three times that in 2011, $69 versus $23. The larger value might be attributed at least partially to the greater number of months with the offset in place in 2012: 12 months versus 8 in 2011.

Exhibit 4-1.   Summary of Impact Findings

	Predicted Sign	Year 1 (May–Dec 2011)	Year 2 (Jan–Dec 2012)
Confirmatory Outcomes			
Total earnings in year	?		
Total SSDI benefits paid in year	?	$23	$69
Exploratory Outcomes			
Employment and Earnings in Year			
Employment	+		
Earnings above BYA	+		
Earnings above 2 × BYA	?		
Earnings above 3 × BYA	?		
Benefits Paid in Year			
Number of months with SSDI payments	+		0.05[a]
Total SSI benefits paid	–		
Number of months with SSI payments	–		

Note: The predicted signs are based on economic theory, which abstracts from administrative factors that might have an impact on outcomes (Bell et al. 2011). The estimates from Year 1 are reported in Stapleton et al. (2013). The estimates for Year 2 are reported in Exhibits 3-1 through 3-3.

[a] In additional exploratory analysis, we found that the impact on months with SSDI payments was an estimated 0.03 months larger for long-duration subjects than for short-duration subjects.

For the exploratory earnings and benefit outcomes for the overall sample, we find no evidence of impacts on any of the employment and earnings outcomes, and strong evidence of an increase in the number of months with SSDI payments. The increase in number of months with SSDI payments (0.05 months) is consistent with the finding of a positive impact on benefits paid.

Among some subgroups, we find weak exploratory evidence of impacts on select earnings and employment outcomes, and strong evidence of impacts on benefits paid in 2012 and number of months with SSDI payments. For short-duration subjects we found weak evidence of a negative impact on the

percentage of subjects with 2012 earnings above $2 \times$ BYA (-0.15 percentage points, or 7.0 percent of the adjusted percentage for C1 subjects), and for SSDI-only subjects, we found weak evidence of a positive impact on the percentage employed (0.18 percentage points, or 1.2 percent of the adjusted percentage for C1 subjects). Estimates for the benefit outcomes are consistent with those for all T1 subjects, except that we did not find evidence of an impact on mean SSDI benefits paid to concurrent subjects. We find almost no evidence of differences in impacts across subgroups in 2012. The one exception is evidence that the impact on months in 2012 with SSDI benefit payments was 0.03 months greater for long-duration than for short-duration subjects.

4.2. Analysis

The most important finding is the very limited evidence of impacts on earnings outcomes. Without positive impacts on earnings, especially on the percentage of beneficiaries with earnings above BYA, the main effect of the benefit offset will be to increase the incomes of beneficiaries by allowing T1 subjects to keep a portion of their benefits that they would otherwise lose under current rules.

For two reasons, it would be premature to conclude from the 2012 findings that the benefit offset will never have an impact on earnings. First, many T1 beneficiaries who could potentially gain from using the offset might not have understood, trusted, or effectively taken advantage of the offset during this period. Concern about this important issue was initially raised in Wittenburg et al. (2012). That concern was reinforced by the limited number of T1 subjects whose benefits had been adjusted under the offset by the end of 2012: 297. That figure is up from 39 at the end of 2011 but still far lower than the 800 or more expected to use the offset as early as 2011.

Second, two activities occurred in 2013 that are expected to have a large effect on impacts for 2013 and beyond. First, the demonstration launched a follow-up outreach campaign, designed to increase the awareness of T1 subjects about the offset. This outreach effort was a response to findings from Wittenburg et al. (2012) suggesting that many T1 subjects did not understand or trust the opportunity to use the offset. Second, SSA adjusted the benefits of a large number of additional T1 subjects in 2013, and a large majority of these cases involved retroactive adjustments for 2011 or 2012. By the end of 2013, SSA had adjusted the benefits of a total of 1,031 T1 subjects. These T1 subjects, as well as many others reached through the follow-up outreach effort, presumably have a better understanding of how the offset works and its potential value to them. The knowledge they gained as a result of these 2013 activities might well be having an impact on their work and earnings. That impact might start to emerge in 2013, but is unlikely to be fully realized until 2014 and beyond.

Two factors likely explain why BOND benefits paid in 2011 and 2012 to treatment subjects were larger than those paid to the control group despite the absence of a statistically significant impact for earnings. The first is that partial benefits were paid in 2012 to the 297 T1 subjects whose benefits were adjusted under the offset in 2012 for months in which they would have had no benefits under current law. The second is the administrative factors reflected in the large number of retroactive adjustments to 2011 and 2012 benefits made after 2012. Up to 40 percent of the impact could be due to the first factor. This percentage is based on a comparison of the estimated impact on benefits paid in 2012 to the actual

benefits paid to the 297 subjects whose benefits were adjusted under the offset by the end of 2012.[22] That implies that at least 60 percent is due to administrative factors. The large role played by administrative factors is consistent with the many retroactive adjustments to the 2012 benefits of T1subjects that were made in 2013.

As SSA makes retroactive adjustments to benefits and recovers improper payments, benefits paid to T1 subjects *for* 2012 will diminish. Of course at the same time SSA will be making retroactive adjustments to benefits of some C1 subjects and recovering improper payments, under current law processes. Hence, the difference between mean benefits paid to T1 and C1 subjects *for* 2012 will not necessarily shrink. Nonetheless, the large share of the estimated impact on benefits paid *in* 2012 that must be attributed to administrative factors underlines the importance of the final evaluation's planned focus on impacts paid *for* the evaluation period, rather than *in* the evaluation period.

Two important factors will affect the direction and size of the impact on benefits paid in 2013. First, retroactive adjustments to pre-2013 benefits are likely to have a substantial effect on mean benefits paid in 2013. Second, adjustments made to current-year benefits paid because of current-year earnings will likely play a much larger role in the impact for benefits paid in 2013 than they did in 2012, because of the large increase in offset users by the end of 2013 noted earlier. The direction and size of the latter will depend on the extent to which the offset had an impact on the earnings of these T1 subjects in 2013.

Future evaluation reports will also present estimates of impacts on an array of other outcomes, including those observed via the 36-month follow-up survey of 10,000 T1 and C1 subjects. Findings from the survey, as well as additional process study findings on the demonstration's implementation and further analysis of offset participation, will aid in the interpretation of the impact findings.

---

[22] We estimated the size of these impacts by first calculating the maximum impact on mean benefits paid in 2012 to the 297 beneficiaries whose benefits were adjusted under the offset by the end of 2012. That maximum is the mean amount of benefits paid to these 297 subjects in 2012, $7,591, as the minimum they would have been paid under current law is $0. The maximum contribution of this effect to the mean estimated impact on mean benefits paid in 2012 to all 77,155 T1 beneficiaries is $27 (= $7,591 x 297/77,155), or about 40 percent of the $69 impact estimate for benefits paid in 2012 ($27/$69 = 0.39). Consequently, at least 60 percent of the impact on benefits paid in 2012 must be attributed to administrative factors.

References

Bell, Stephen H., Daniel Gubits, David Stapleton, David Wittenburg, Michelle Derr, Arkadipta Ghosh, and Sara Ansell. *BOND Implementation and Evaluation. Evaluation Analysis Plan. Final Report Submitted to Social Security Administration*. Cambridge, MA: Abt Associates, March 2011.

Livermore, Gina, Allison Roche, and Sarah Prenovitz. "Work Activity and Use of Employment Supports Under the Original Ticket to Work Regulations: SSI and DI Beneficiaries with Work-Related Goals and Expectations." Submitted to the Social Security Administration. Washington, DC: Mathematica Policy Research, 2009.

Riley, Gerald R., and Kalman Rupp. "Expenditure Patterns Under the Four Major Public Cash Benefit and Health Insurance Programs for Working-Age Adults with Disabilities." *Journal of Disability Policy Studies,* 2012. Published online ahead of print as doi:10.1177/1044207312469828.

Schochet, Peter Z. "Technical Methods Report: Guidelines for Multiple Testing in Impact Evaluations." *NCEE 2008-4018*. Princeton, NJ: Mathematica Policy Research, 2008.

Social Security Administration. *Annual Statistical Report on the Social Security Disability Insurance Program, 2012*. Baltimore, MD: SSA, 2013.

Stapleton, David, David Wittenburg, Daniel Gubits, David Judkins, David Mann, and Andrew McGuirk. "BOND Implementation and Evaluation: First-Year Snapshot of Earnings and Benefit Impacts for Stage 1." Cambridge, MA: Abt Associates, and Washington, DC: Mathematica Policy Research, 2013.

Stapleton, David C., Stephen H. Bell, David C. Wittenburg, Brian Sokol, and Debi McInnis. "BOND Implementation and Evaluation: BOND Final Design Report." Submitted to the Social Security Administration, Office of Program Development & Research. Cambridge, MA: Abt Associates, December 2010.

Westfall, Peter H., and S. S. Young. *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment.* New York: Wiley-Interscience, 1993.

Wittenburg, David, David Stapleton, Michelle Derr, Denise W. Hoffman, and David R. Mann. "BOND Stage 1 Early Assessment Report. Final Report Submitted to the Social Security Administration." Cambridge, MA: Abt Associates, May 2012.

Wright, Debra, Gina Livermore, Denise Hoffman, Eric Grau, and Maura Bardos. "2010 National Beneficiary Survey: Methodology and Descriptive Statistics." Washington, DC: Mathematica Policy Research, 2011.