



BOND Implementation and Evaluation Sixth-Year Snapshot of Earnings and Benefit Impacts for Stage 1

Deliverable 24f.2(1)

September 11, 2019

Submitted to:

Social Security Administration

Attn: Ms. Joyanne Cobb

Office of Research, Demonstration, and

Employment Support

500 E Street, SW

9th Floor, Rm. 905

Washington, DC 20254

Contract No. SS00-10-60011

Prepared by:

Denise Hoffman

David R. Mann

David Stapleton

Daniel Gubits

Judy Geyer

Mark Sussman

Serge Lukashanets



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 a variant 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 study 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 final in a series of Stage 1 *Snapshot Reports* and *Interim Reports*, documents Stage 1 impacts on earnings and benefit outcomes during the sixth calendar year of implementation (2016). As in the previous *Snapshot Reports* (Stapleton et al. 2013, 2014, Wittenburg et al. 2015) and *Interim Reports* (Hoffman et al. 2017; Croake et al. 2017), we estimate impacts by comparing earnings and benefit outcomes for all Stage 1 treatment (T1) subjects with those for Stage 1 control (C1) subjects. A parallel series of reports exist for Stage 2, including a final *Snapshot Report*.

Summary of Key Findings

Similar to findings from all previous years, the BOND benefit offset did not have a statistically significant impact on total earnings in 2016. In contrast to the 2015 findings, where we found statistically significant effects on the proportions earning more than one, two, or three times the BOND Yearly Amount (BYA—the annualized substantial gainful activity amount), there were no significant impacts on these outcomes in 2016.

To better understand the lack of significant impacts in 2016 on these proportions, we conducted a new subgroup analysis and found that the insignificance was not readily explained by the fact that some T1 subjects reverted to current law in 2016. We also found that the proportions earning more than one, two, or three times BYA grew for both treatment and control subjects, presumably due to improvements in job opportunities.

Without a larger positive impact on earnings—especially on the percentage of beneficiaries with earnings above BYA—the benefit offset will lead to higher SSDI benefits. Consistent with previous years, we find a positive impact on total SSDI benefits paid in 2016. The impact on benefits paid in 2016 does not account for retroactive adjustments made after 2016. We expect impacts on SSDI benefits due for 2016, after all retroactive adjustments are made, to be modestly different.

The BOND Evaluation Team

Abt Associates, in partnership with 25 other organizations, is implementing and evaluating the BOND under contract to the U.S. Social Security Administration. 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 Daniel Gubits of Abt Associates and David Stapleton of Mathematica Policy Research. These individuals 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. The Abt Project Director (Michelle Wood) has 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.

Table of Contents

Acronyms Used in This Report.....	iii
1. Introduction.....	1
1.1. Synopsis of BOND.....	1
1.2. Purpose.....	2
1.3. Organization of the Report.....	3
2. Methodology and Context	5
2.1. Methodological Approach.....	5
2.1.1. Outcome Definitions and Theoretically Expected Impacts.....	6
2.1.2. Administrative Features of the Offset That Influence Impacts	8
2.1.3. Accounting for Reversion to Current Law in Impact Estimates	9
2.1.4. Impact Estimation Methodology	10
2.1.5. Final Analysis Sample Sizes	11
2.2. Previous Findings.....	13
3. Impact Findings.....	14
3.1. Full Stage 1 Treatment Group	14
3.1.1. Confirmatory Impacts.....	14
3.1.2. Exploratory Impacts	17
3.2. Reversion to Current Law	19
3.2.1. 2016 Earnings, Employment, and Benefit Outcomes for TWP Subgroups	20
3.2.2. Changes in Outcomes from 2015 to 2016 for TWP Subgroups.....	21
3.3. Subgroups Based on Benefit Status at Baseline.....	24
3.3.1. Duration Since Award.....	24
3.3.2. SSI Benefit Status.....	26
4. Discussion.....	27
References.....	28
Appendix A. Additional Information on TWP Completion Subgroups.....	30
Appendix B. Impact Estimates for Changes in Outcomes from 2015 to 2016.....	33
Appendix C. Benefits Due Impact Estimates	36

List of Exhibits

Exhibit 2-1.	Definitions of Confirmatory and Exploratory Outcomes and Predicted Signs of Impacts.....	7
Exhibit 2-2.	Stage 1 Analysis Sample Composition.....	12
Exhibit 3-1.	Trends in Annual Earnings in the First Six Years of BOND.....	15
Exhibit 3-2.	Impact Estimates for 2016 Total Earnings and Total SSDI Benefits Paid	16
Exhibit 3-3.	Trends in Monthly SSDI Benefits Paid in the First Six Years of BOND	16
Exhibit 3-4.	Impact Estimates for 2016 Employment and Benefit Receipt.....	18
Exhibit 3-5.	Impact Estimates in 2016 for Subgroups Defined by TWP Completion in or Before April 2011	21
Exhibit 3-6.	Impact Estimates for Change in Earnings and Employment Outcomes from 2015 to 2016 for Stage 1 Subgroups Defined by TWP Completion Pre-BOND.....	23
Exhibit 3-7.	Impact Estimates for Subgroups Defined by Duration of SSDI Receipt.....	25
Exhibit 3-8.	Impact Estimates for Subgroups Defined by Baseline SSI Status	26
Exhibit A-1.	Reversion to Current Law among TWP Completion Subgroup Pairs	31
Exhibit B-1.	Impact Estimates for Change in Earnings and Employment Outcomes from 2015 to 2016 for All Stage 1 Subjects	33
Exhibit B-2.	Impact Estimates for Change in Benefit Outcomes from 2015 to 2016 for All Stage 1 Subjects	34
Exhibit B-3.	Impact Estimates for Change in Benefit Outcomes from 2015 to 2016 for Stage 1 Subgroups Defined by TWP Completion Pre-BOND	35
Exhibit C-1.	Impact Estimates for 2016 Benefits Due	36
Exhibit C-2.	Impact Estimates for SSDI Benefits Due Compared to SSDI Benefits Paid.....	37

Acronyms Used in This Report

BODS	BOND Operations Data System
BOND	Benefit Offset National Demonstration
BPP	BOND Participation Period
BYA	BOND Yearly Amount
CPI-W	Consumer Price Index for Urban Wage Earnings and Clerical Workers
DAC	Disabled Adult Child
DAF	Disability Analysis File
DCF	Disability Control File
DWB	Disabled Widow/Widowers
EIE	Earned Income Exclusion
EPE	Extended Period of Eligibility
eWork	Electronic Work System
GP	Grace Period
IRS	Internal Revenue Service
IRWE	Impairment Related Work Expense
MEF	Master Earnings File
MBR	Master Beneficiary Record
PHUS	Payment History Update System
SGA	Substantial Gainful Activity
SSA	Social Security Administration
SSDI	Social Security Disability Insurance
SSI	Supplemental Security Income
SSR	Supplemental Security Record
TWP	Trial Work Period

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 report is the final in a series of Stage 1 *Snapshot Reports* and *Interim Reports* about impacts on earnings and benefits paid. This introductory chapter provides a synopsis of the demonstration, then describes the purpose of the report and outlines its remainder. Impact estimates for Stage 2 appear in a parallel series, including a final *Snapshot Report*.

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, after completing a nine-month Trial Work Period (TWP) and a three-month Grace Period (GP), an SSDI beneficiary's countable monthly earnings exceed the monthly Substantial Gainful Activity (SGA) amount. In 2016, the SGA amount was \$1,130 per month for non-blind beneficiaries and \$1,820 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 do work to keep their earnings below the SGA level.

In contrast, for beneficiaries subject to the alternative rules test in BOND, SSA replaces the cash cliff with a “ramp” (benefit offset) and a policy objective of encouraging beneficiaries to increase their earnings and reduce their reliance on benefits.² The benefit offset is expected to increase the earnings of those who might otherwise not work at all and those who already work but might not attempt to earn more than the SGA amount. Those who engage in SGA under current law eventually lose their benefits entirely, whereas under the benefit offset, many (perhaps most) will be eligible for a reduced SSDI benefit. While still on the ramp—that is, while earning above the SGA amount but less than the zero-benefit amount at the end of the ramp—beneficiaries can increase their benefits by working less than their full earnings potential.

There are also two key differences between the way SSA administers the benefit offset and current law. First, the benefit offset replaces the monthly SGA calculation with an annualized measure of SGA, referred to as the BOND Yearly Amount (BYA). BYA is equal to 12 times the monthly SGA amount (in 2016, \$13,560 for non-blind and \$21,840 for blind treatment subjects). The benefit offset reduces benefits by \$1 for every \$2 in countable annual earnings in excess of the BYA following completion of the TWP and GP. Use of the annualized measure requires changes in SSA's administrative procedures for adjusting benefits. For treatment subjects who have completed the TWP and GP, SSA pays benefits monthly under

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.

² See Exhibit 1-1 of the *Stage 2 Early Assessment Report* (Gubits et al. 2013) for a detailed comparison of current SSDI program rules with BOND rules related to work.

offset rules and 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) and documentation provided by the beneficiary.

Second, BOND treatment subjects can use the benefit offset during a 60-month participation period. For beneficiaries who completed the TWP before random assignment, the BOND participation period starts the month after random assignment. For other beneficiaries, the BOND participation period begins the month after the TWP is completed. In both cases, the offset applies once all GP months are used up. SSA does not terminate SSDI entitlement because of work during this period, even if benefits fall to zero because of earnings that are well above the BYA.³

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 2016. Random assignment occurred in April 2011, so the 2016 calendar year represents the sixth calendar of BOND.

This report is the final in a series of BOND Stage 1 evaluation reports. In the initial reports, we described the BOND design, the framework for estimating the impacts, and early assessment activities on the infrastructure to support Stage 1 service delivery (Stapleton et al. 2010; Bell et al. 2011; Wittenburg et al. 2012, respectively). Annual earnings and benefit impacts for Stage 1 subjects from 2011 through 2015 are reported in a series of *Snapshot Reports* and *Interim Reports* (Stapleton et al. 2013; Stapleton et al. 2014; Wittenburg et al. 2015; Hoffman et al. 2017; Croake et al. 2017). A series of parallel *Snapshot Reports* and *Interim Reports* describe annual Stage 2 outcomes. The *BOND Final Evaluation Report* describes a comprehensive set of outcomes for both stages of BOND from 2011 through 2015 (Gubits et al. 2018a, 2018b).

The purpose of this report is to examine the magnitude of impacts for the Stage 1 sample in 2016, the sixth year of implementation. We did not include 2016 in the main follow-up period covered by the *Final Evaluation Report* (Gubits et al. 2018a, 2018b) because some treatment subjects ended their BOND participation period and reverted to current law during 2016. Nevertheless, most treatment subjects remained under the treatment condition in 2016, which makes impact analysis of data from this year potentially informative for policy.

The analysis sample and methods in this *Snapshot Report* match those of the *Final Evaluation Report*. In the previous *Snapshot Reports* and *Interim Reports*, we excluded Stage 1 subjects known to have died before random assignment based on mortality data extracted in April 2012. For the *Final Evaluation Report*, we updated the Stage 1 analysis sample to exclude all those known to have died before random assignment based on mortality data extracted in May 2017. We use this updated analysis sample for this report. The analysis methods in this *Snapshot Report* are identical to those used in the *Final Evaluation Report* and first used in the *First Stage 1 Interim Report* (Hoffman et al. 2017).

The two most important evaluation outcomes in this report—referred to as *confirmatory outcomes*—are total earnings and total SSDI benefits paid in 2016. Statistically significant findings for the confirmatory

³ SSA will apply current rules at the end of the 60-month participation period and will terminate the benefits of those engaged in SGA after all remaining GP months have been used.

outcomes in this report should be interpreted as confirming that the benefit offset would have had an impact on at least one of two outcomes if implemented for the national beneficiary population: 2016 earnings and/or SSDI benefits paid in 2016. The report also presents estimates for several exploratory outcomes related to 2016 earnings and benefits paid, to explore the possibility of other impacts of the offset if implemented for the national beneficiary population. These findings do not confirm that the benefit offset had impacts; they only suggest 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.

The impact estimates in this *Snapshot Report* focus on SSDI benefits paid in 2016 rather than SSDI benefits due in 2016. All of the *Snapshot Reports* and *Interim Reports* for Stage 1 and Stage 2 have presented impact estimates for benefits paid, because it is the most readily available benefit measure. Benefits paid is the sum of the monthly benefit checks a beneficiary received in a year. The measure is limited because it includes reconciliation of improper payments in earlier years and does not include retroactive adjustments to benefits. The *Final Evaluation Report* (Gubits et al. 2018a, 2018b) at least partially addressed these limitations by presenting impact estimates through 2015 for benefits due, which is the sum of monthly SSDI benefits that SSA should have paid to the beneficiary based on eligibility and earnings in the corresponding month.⁴ The benefits due measure excludes reconciliation of earlier years' improper payments and, for the *Final Evaluation Report*, incorporated retroactive adjustments through May 2017 (17 months after the end of the analysis period). In this report, we again focus on benefits paid, because insufficient time had elapsed to create a reliable measure of benefits due. For the sake of completeness, however, Appendix C presents results based on benefits due in 2016, including retroactive adjustments for a follow-up period that is shorter than for the final report (through January 2018, 13 months after the end of the analysis period).

In addition to full sample results, this *Snapshot Report* presents Stage 1 2016 impacts for three beneficiary subgroup pairs. Previous Stage 1 evaluation reports have reported impacts for two of the pairs—defined by SSDI benefit duration status and SSI receipt status. The new subgroups are defined by whether or not the beneficiary had completed the TWP before random assignment. For treatment subjects, pre-BOND TWP completion determines whether the BOND offset rules reverted to current law rules in May 2016, or continued thereafter. The benefits of the overwhelming majority of treatment subjects who did not complete the TWP before BOND random assignment continued to be determined by the offset rules for the remainder of 2016. We introduce this additional subgroup to explore an issue that first emerged in 2016: reversion to current law earnings rules for some Stage 1 treatment subjects (see Section 2.1.3).

1.3. Organization of the Report

The remainder of this report comprises three chapters. 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 full-sample confirmatory and exploratory outcomes, describes results for subgroups defined by pre-BOND TWP completion status to help interpret the full-

⁴ For more information about SSDI benefit measures and concepts, see Appendix A in the *Final Evaluation Report, Volume 2* (Gubits et al. 2018b).

sample impacts, and presents impacts for subgroups defined by duration since SSDI award and SSI receipt. Chapter 4 includes a brief discussion of the results.

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:⁵

- **T1 subjects:** 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:** a control group that continues to receive benefits according to current law.

Because only a small fraction of T1 subjects were expected to use the benefit offset, the T1 and C1 groups had to 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 2016. 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 2016 when SSA extracted them for this report.⁶ Benefit paid outcomes are measured from SSA's Payment History Update System (PHUS) for SSDI and the Supplemental Security Record (SSR) for SSI. Benefit due outcomes are measured from SSA's Master Beneficiary Record (MBR) for SSDI and the SSI-Longitudinal File for SSI.⁷

The next section of this chapter describes our 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 might influence the size and direction of earnings and benefit impacts presented in Chapter 3.

2.1. Methodological Approach

Below, we review the outcome definitions, anticipated impacts, estimation methodology, and analysis sample used in this *Snapshot Report*. The analysis methods used in this *Snapshot Report* are identical to those used in the *Stage 1 Interim Reports* and the *Final Evaluation Report*. Relative to the analysis

⁵ The random assignment process included beneficiaries for Stages 1 and 2. Beneficiaries were randomly assigned to the T1 group, the C1 group, or a "Stage 2 solicitation pool." The latter initially included all randomly selected outreach cohorts of SSDI-only beneficiaries for the purposes of recruiting volunteers for Stage 2. When Stage 2 recruitment was complete, subjects in the unused outreach cohorts were assigned to C1. For details on random assignment, see Bell et al. (2011).

⁶ Because the earnings 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 as new information is provided by employers and self-employed individuals. The 2016 earnings data for this report were extracted in February 2018.

⁷ We extracted the PHUS, MBR, SSR, and SSI-Longitudinal File data in January 2018.

sample used in all previous *Stage 1 Snapshot Reports* and *Interim Reports*, the analysis sample used in this *Snapshot Report* is slightly smaller because it excludes some additional subjects who we now know died before random assignment.

2.1.1. Outcome Definitions and Theoretically Expected Impacts

The eleven outcomes for this report include two confirmatory outcomes (total earnings and total SSDI benefits paid in 2016) and nine 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 (BYA, and two and three times BYA) and an indicator for employment (measured as any earnings) during 2016.⁸ The exploratory benefit outcomes include an indicator for at least one month with an SSDI benefit paid, number of months with SSDI payments, total SSI benefits paid, an indicator for at least one month with an SSI benefit paid, and number of months with SSI payments—each in 2016.

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 consider administrative factors and their potential influence on impacts.

The theoretical direction of impacts of the benefit offset on mean earnings and mean benefits is ambiguous (third column of Exhibit 2-1). As detailed in Bell et al. (2011), this ambiguity arises because the incentives created by the benefit offset vary with what a beneficiary's earnings would be under current law. Opposite effects are expected for treatment subjects who, under current law, would have had earnings below BYA and those who would have had earnings above BYA:

- The offset is expected to induce treatment subjects who would have had either no earnings or earnings below BYA under current law to have on average higher earnings and lower SSDI benefits than they would under current law.
- Conversely, the offset is expected to cause treatment subjects who would have had earnings above BYA under current law to have on average lower mean earnings and higher mean SSDI benefits than they would under current law.

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 more than counteract possible negative impacts for those who would earn more than BYA under current law. Similarly, for the impact on SSDI benefits to be negative, the reduction in benefits for those who would have had earnings above BYA under current law must be larger than the increase in benefits for the balance of treatment subjects.

Although the effect on the amount of SSDI benefits is theoretically ambiguous, theory does predict positive impacts on months with SSDI payments. Those who would earn under BYA under current law are expected to have no change or a small decrease in months with SSDI payments. Those who would earn above BYA under current law, however, are expected to have an increase in the months of SSDI payments because some will receive partial benefits rather than benefit suspension.

⁸ The administrative earnings measures include only “Social Security earnings.” See Gubits et al. (2018b) for a detailed discussion of the measure.

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

	Definition	Predicted Sign
Confirmatory Outcomes		
Total earnings in 2016	2016 Social Security earnings	?
Total SSDI benefits paid in 2016	Sum of SSDI benefit payments from January through December 2016; 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 2016)^b		
Employment in 2016	Indicator for any 2016 Social Security earnings	+
Earnings above BYA	Indicator for 2016 Social Security earnings greater than or equal to \$13,560 (non-blind subjects) or \$21,840 (blind subjects)	+
Earnings above 2 × BYA	Indicator for 2016 Social Security earnings greater than or equal to \$27,120 (non-blind subjects) or \$43,680 (blind subjects)	?
Earnings above 3 × BYA	Indicator for 2016 Social Security earnings greater than or equal to \$40,680 (non-blind subjects) or \$65,520 (blind subjects)	?
Benefit Outcomes (January–December 2016)		
Number of months with SSDI payments	Number of months in 2016 with SSDI benefits paid above zero	+
Total SSI benefits paid	Sum of SSI benefit payment amounts from January through December 2016	–
Number of months with SSI payments	Number of months in 2016 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/EN-05-10024.pdf>; accessed March 29, 2018.

^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).

Theory also predicts the signs of the impacts for two of the four exploratory earnings outcomes. It predicts positive impacts on the proportion employed and the proportion with earnings above BYA. As indicated earlier, for those who would earn less than 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 at least as strong an incentive to keep their earnings above BYA under the offset than they do under current law—even though some might choose to earn less under the offset. Hence, overall theory predicts positive impacts on employment and the percentage with earnings above current law. The same line of reasoning does not lead to an unambiguous prediction for the direction of impacts on the percentage with earnings well above BYA (for example, two and three times BYA), because predicted impacts on earnings are in opposite directions for those whose earnings would be above or below BYA under current law.

Finally, theory predicts that the impact on SSI benefits and months with SSI payments 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.⁹ 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.¹⁰

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. Because these processes differ from current law processes, they are part of the T1 intervention being tested in BOND (Bell et al. 2011).

The administrative adjustment of benefits for treatment subjects—the special process implemented for T1 subjects in contrast to the current law process that applies to C1 subjects—has the potential to affect the measurement of impacts on SSDI benefits, but not on annual earnings. The benefit measure used in the body of this report, benefits paid in 2016, includes reconciliation of improper pre-2016 payments undertaken in 2016 and excludes retroactive adjustments for 2016. There will be some difference between the mean benefit outcomes reported in the body of the report (benefits paid in 2016) and mean benefits due for 2016 after all adjustments are made. Appendix C reports impacts on benefits due, but these should be considered preliminary because they are based on adjustments made through January 2018 only. We expect impacts based on both benefits measures to be generally similar. The magnitude of the differences in estimated impacts will depend on how adjustments for the T1 subjects compare with the adjustments for C1 subjects.¹¹

The change from monthly to annual accounting is also likely to influence benefit outcomes in another way. 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

⁹ 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. In 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.

¹⁰ See Riley and Rupp (2012).

¹¹ See Stapleton et al. (2014) for descriptions of the differences in the adjustment processes for T1 and C1 subjects.

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.

2.1.3. Accounting for Reversion to Current Law in Impact Estimates

The use of the benefit offset is a time-limited opportunity. For reasons discussed below, the opportunity ended for some subjects in 2016. As a result, some beneficiaries may have made employment decisions in 2016 on the basis of current law rather than offset rules, potentially affecting the 2016 impacts presented in this report.

Treatment subjects may use the benefit offset during a 60-month period known as the BOND Participation Period (BPP). For beneficiaries who completed the TWP before BOND random assignment, this period began in the month following random assignment (in May 2011) and continued through April 2016. Beneficiaries who completed the TWP after BOND random assignment and before September 30, 2017 began the 60-month eligibility period in the month after TWP completion. After completion of the GP and BPP, SSA terminates benefits of those engaged in SGA in any month thereafter in accordance with current law.

The first treatment subjects to revert to current law at the end of their five-year BPP did so between May and December 2016. Reversion to current law would influence 2016 impacts if these beneficiaries change their earnings in response to the change in work incentives under current-law relative to offset rules. Some beneficiaries who were earning above BYA under the offset rules may reduce their earnings to below SGA when they revert to current law to maintain benefit eligibility.¹² A subset of this group may stop working entirely. Other beneficiaries who were earning above BYA under the offset rules may increase their earnings when they revert to current law to compensate for the loss of partial benefits (an income effect) and in response to the fact that increases in earnings are no longer partially offset by a benefit reduction (a substitution effect).

To explore the effects of reversion to current law, we examine a new subgroup pair (see Section 3.2). Specifically, we compare changes in impacts from 2015 to 2016 for: (1) those who completed the TWP before BOND random assignment, and therefore reverted to current law in May 2016, and (2) all others. The motivation for this grouping is twofold. First, we define subgroups based on baseline characteristics, to ensure that differences in outcomes across T1 and C1 members of each subgroup are internally valid impact estimates of the subpopulations represented.¹³ Second, we expect reversion to current law to only affect outcomes for the treatment subjects who reverted to current law in 2016. That presumably includes every treatment beneficiary in the pre-BOND TWP completion group—apart from those whose benefits were terminated for some other reason before 2016. The much larger complementary group does include a very small number of treatment beneficiaries whose rules reverted to current law later in 2016—those who completed the TWP between May and November of 2011. Our expectation is that any substantial effect of reversion to current law would be manifest by changes in impacts for the pre-BOND TWP

¹² All BOND treatment subjects who completed their TWP before BOND random assignment are considered to have completed their EPE at the end of the 60-month BPP. Hence, if they engage in SGA after reversion to current law rules and have used up all GP months, SSA will terminate their eligibility for SSDI.

¹³ In principle, random assignment could have had an effect on TWP completion in any month after May 2011.

completers relative to the corresponding changes for the complementary group. We provide more details on subgroup construction in Appendix A.

2.1.4. 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 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 on outcomes observed in administrative data, we compare the mean of a given outcome for the T1 group to the mean of the same outcome for the C1 group. We weight the sample means to account for differences in (1) site-selection probabilities, and (2) sampling rates into T1 and C1 status across sampling strata. We also adjust the means using (weighted) linear regression methods to account for the effects of small random differences in baseline characteristics. The adjustments for differences in baseline characteristics also reduce the standard errors of the impact estimates.¹⁴

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.¹⁵

We make no multiple comparisons adjustment to the tests for exploratory outcomes. In this analysis (as in any analysis with many significance tests), some statistically significant results may be spurious. Readers are therefore advised to give less weight to a 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, two pairs of subgroups identified in our *Analysis Plan*, and one new subgroup pair. The two subgroup pairs examined in previous *Snapshot Reports* and *Interim Reports* divide the sample by duration of SSDI benefit receipt and SSI status in the month prior to

¹⁴ See Appendix B of the *Final Evaluation Report* for a full description of the Stage 1 impact methodology.

¹⁵ 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 B of the *Final Evaluation Report*. See Schochet (2008) for further discussion of the multiple comparisons problem.

random assignment. We treat as exploratory all subgroup analyses, including the tests of earnings and SSDI benefits paid.

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.

In addition to the two subgroup pairs examined in the previous *Snapshot Reports*, this report contains subgroup impact estimates that explore the potential effects of reversion to current law at the end of the 60-month period in which treatment beneficiaries could use the offset (see Section 2.1.3).

We present impact estimates that estimate the change in impacts from 2015 to 2016 for the full sample and for the subgroups defined by TWP completion dates. To do so, we define new earnings and benefits outcomes that are the changes in these outcomes from 2015 to 2016. For total earnings, a continuous variable, this represents the increase (or decrease) in earnings from 2015 to 2016. For employment, a binary variable, the change is 1 if the beneficiary had no earnings in 2015 and had earnings in 2016, 0 if there was no change (either no earnings or positive earnings in both years), and -1 if the beneficiary had earnings 2015 but not in 2016. The impact estimation methodology for these outcomes is the same as for all outcomes.

2.1.5. Final Analysis Sample Sizes

Exhibit 2-2 presents the sizes for the overall sample and the subgroups. The sample sizes are different than in previous *Snapshot Reports* and *Interim Reports*, but identical to the sample used for the *Final Evaluation Report*. The sample sizes in the previous *Snapshot Reports* and *Interim Reports* excluded BOND subjects known (as of April 2012) to have died before BOND random assignment. The updated sample excludes an additional 183 subjects who were known to have died before BOND random assignment based on data extracted from SSA's Numident file in May 2017. The final Stage 1 analysis sample contains a total of 968,530 subjects, spread across T1 (77,101) and C1 (891,429).¹⁶

¹⁶ 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): 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 Stapleton et al. (2013) show, 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

The baseline characteristics for the weighted T1 sample are statistically equivalent to those for the weighted C1 sample (Gubits et al., 2018b). These descriptive findings increase our confidence that any statistically significant differences in subsequent outcomes between the T1 and C1 groups 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 pre-existing differences between the two groups or their environments.

Exhibit 2-2 presents the sample sizes for each of the three subgroup pairs. By design, the T1 sample is evenly split between short- and long-duration beneficiaries (see Bell et al. 2011). For other subgroup pairs, the distribution of the sample reflects both the distribution of the characteristics in the beneficiary population from which they were sampled and the fact that relatively large numbers of concurrent beneficiaries were included in C1 because concurrent SSDI and SSI beneficiaries were not eligible for Stage 2. As in all of the impact analyses, we used analysis weights to account for site-selection probabilities and sampling rates into T1 and C1 across sampling strata. (See Appendix B of Gubits et al. [2018b] for more details.) Concurrent DI and SSI beneficiaries and pre-BOND TWP completers each represent small minorities of the Stage 1 sample. As a result, the power to detect impacts for the concurrent and pre-BOND TWP completion subgroups is lower than for their complementary subgroups.¹⁷

Exhibit 2-2. Stage 1 Analysis Sample Composition

Random Assignment Group	Full Sample	Duration		SSI Status		TWP Completion	
		Short Duration	Long Duration	SSDI-Only	Concurrent	Pre-BOND	Not Pre-BOND
T1	77,101	38,663	38,438	64,696	12,405	7,263	69,838
C1	891,429	209,756	681,673	694,110	197,319	109,577	781,852

Source: BOND Operations Data System (BODS), SSA administrative records from the DCF, MBR, eWork, and DAF.

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 (for example, 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 6,453,285.

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.

¹⁷ Unweighted, the pre-BOND TWP completers include 9.4 percent of the T1 sample and 12.3 percent of the C1 sample. Weighted, the groups each represent 11.6 percent of the T1 and C1 samples. The differences in the unweighted percentages for the two groups of TWP completers reflect differences in the ways these samples were selected—particularly the relatively large share of short-duration beneficiaries in the T1 sample, who had less time to complete the TWP before BOND. By design, the analysis weights correct these imbalances.

2.2. Previous Findings

Chapter 7 of the *Final Evaluation Report* summarizes the key findings from BOND years 2011 through 2015 and their policy implications. The Stage 1 analysis finds no confirmatory evidence of an impact on total earnings and a positive impact on SSDI benefits due over the five-year period. The year by year earnings and benefits impacts in the previous *Snapshot Reports* and *Interim Reports* were qualitatively similar to the aggregated results described in the *Final Evaluation Report*. The *Final Evaluation Report* presented exploratory evidence of an increase in the proportion with employment in at least one year during the five-year period and an increase in the proportion with earnings above BYA in at least one year during the five-year period. However, these increases were relatively small (1.6 percent and 7.4 percent of the C1 means, respectively) and not large enough to lead to a decrease in total SSDI benefits.

In the *Final Evaluation Report*, the BOND evaluation team names four factors that plausibly led to the no earnings impact result. First, because of their limited work capacity, too few SSDI beneficiaries may have been able to respond to the benefit offset and increase their earnings for their behavior to have measureable impacts on earnings for all T1 subjects. Second, the benefit offset may not have been a large enough change in work incentives to substantively increase beneficiary earnings; the \$1 for \$2 offset on earnings above BYA is, in essence, an implicit 50 percent tax on gross earnings above BYA. Third, the complexity of the benefit offset (and the SSDI work incentive rules that were unaffected by the benefit offset) may have limited treatment beneficiaries' understanding of (and therefore response to) the change in earnings incentives. Finally, treatment subjects who otherwise would have increased their earnings in response to the offset might have been poorly informed or confused due to limited initial outreach and administrative delays in benefit adjustments.

3. Impact Findings

This chapter presents impact estimates for the confirmatory and exploratory outcomes. As described 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 entire sample 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 in BOND, for the entire national 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.

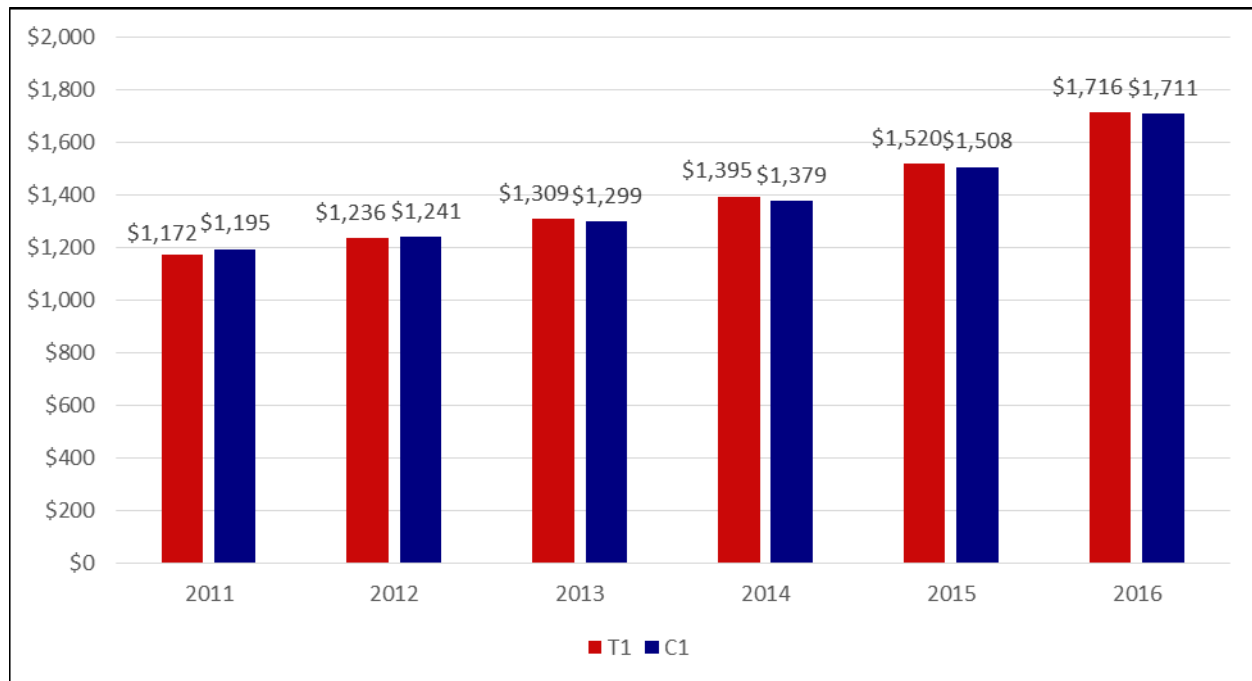
3.1. Full Stage 1 Treatment Group

Exhibit 3-2 presents the 2016 impact estimates on confirmatory outcomes—total earnings and total benefits paid. Exhibit 3-4 presents all other earnings and benefit outcomes, which are considered exploratory.

3.1.1. Confirmatory Impacts

To provide context for the impacts on average annual earnings, we first describe the earnings levels. The level of average annual earnings increased from 2015 to 2016 for each of the Stage 1 random assignment groups considered separately, continuing the trend of rising earnings over the six-year analysis period (Exhibit 3-1). Across all years, the largest annual increase in the level of earnings was from 2015 to 2016: a 13 percent increase for both groups. Economic growth seems the most likely explanation of earnings growth for the control group. According to statistics from the American Community Survey, between 2011 and 2016, the employment rate for the broader population of people with disabilities increased from 33 percent to 36 percent (U.S. Census Bureau, American Community Survey 1-Year Estimates).¹⁸

¹⁸ American Community Survey estimates of the employment rate for people with disabilities are based on author calculations from Table B18120.

Exhibit 3-1. Trends in Annual Earnings in the First Six Years of BOND

Source: SSA administrative records from the MEF and MBR and statistics from the *Final Evaluation Report*.

Notes: Weights are used to ensure that the BOND subjects who met analysis criteria are representative of the national beneficiary population in the month of random assignment. Means are regression-adjusted for baseline characteristics. All dollar amounts are inflation-adjusted to 2016 dollars using the Consumer Price Index for Urban Wage Earnings and Clerical Workers (CPI-W).

Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a confirmatory standard of evidence (p-value adjusted by the multiple-comparisons procedure) and a two-tailed t-test with 9 degrees of freedom (resulting from a research design involving 10 study sites). For each year, the difference between C1 and T1 means was not statistically significant.

For 2016 there is no evidence of an impact on earnings. The point estimate for this impact—the 2016 adjusted difference between average earnings for the treatment and control group—is \$5, which is less than one-half of one percent of the mean value for C1 subjects (last pair of columns in Exhibit 3-1 and top row of Exhibit 3-2). The *p*-value for the test of the hypothesis of no impact is 0.868, is well above the *p* < 0.10 standard.

The estimate of the benefit offset's impact on 2016 earnings mirrors the impact findings for earnings from previous years. The earlier years' findings are evident in Exhibit 3-1: differences in the heights of each pair of treatment-control bars—the regression-adjusted annual earnings amounts for their respective subjects, inflation-adjusted to 2016 dollars—are very small.

Turning to SSDI benefits, we see that the level of average monthly SSDI benefits decreased by about 10 percent during the analysis period. The largest annual decrease occurred between 2015 and 2016—a three percent reduction. The decrease in the level of SSDI benefits paid over time reflects attrition from the SSDI rolls due to death, medical termination, or other reasons (see Exhibit I-1 of the *Final Evaluation Report*). The decrease may also be related to an increase in the proportion earning above BYA over the same period, which results in more benefit suspensions or terminations due to work (C1 subjects) or reductions under the benefit offset (T1 subjects).

Exhibit 3-2. Impact Estimates for 2016 Total Earnings and Total SSDI Benefits Paid

Outcome	T1 Mean	C1 Mean	Impact Estimate
Total earnings (January–December 2016)	\$1,716	\$1,711	\$5 ^a (\$30)
Total SSDI benefits paid (January–December 2016)	\$10,906	\$10,740	\$166 ^{####b} (\$25)

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

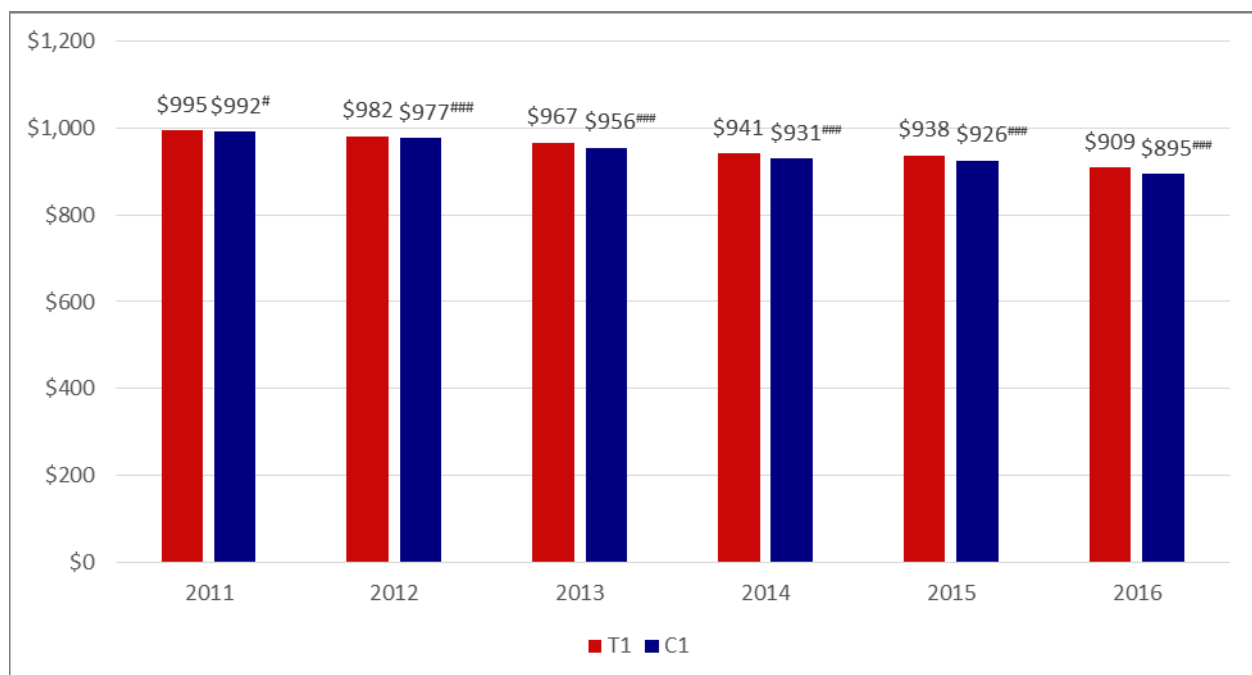
Notes: All earnings outcomes are based on earnings subject to Social Security taxes. Weights are used to ensure that the BOND subjects who met analysis criteria are representative of the national beneficiary population in the month of random assignment. Standard errors appear in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: T1 = 77,101; C1 = 891,429.

Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a confirmatory standard of evidence (*p*-value adjusted by the multiple-comparisons procedure) and a two-tailed t-test with 9 degrees of freedom (resulting from a research design involving 10 study sites).

^a The impact estimate for total earnings has a *p*-value after multiple-comparison adjustments of 0.868 and hence does not provide confirmatory evidence of an impact.

^b The impact estimate for total SSDI benefits paid has a *p*-value after multiple-comparison adjustments of less than 0.001 and hence provides confirmatory evidence of an impact.

Exhibit 3-3. Trends in Monthly SSDI Benefits Paid in the First Six Years of BOND

Source: SSA administrative records from the PHUS and MBR, and statistics from the *Final Evaluation Report*.

Notes: Weights are used to ensure that the BOND subjects who met analysis criteria are representative of the national beneficiary population in the month of random assignment. Means are regression-adjusted for baseline characteristics. All dollar amounts are inflation-adjusted to 2016 dollars using the Consumer Price Index for Urban Wage Earnings and Clerical Workers (CPI-W).

Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a confirmatory standard of evidence (*p*-value adjusted by the multiple-comparisons procedure) and a two-tailed t-test with 9 degrees of freedom (resulting from a research design involving 10 study sites).

There is strong evidence of a positive impact of the offset on SSDI benefits paid in 2016—the difference between the adjusted treatment and control means in the last pair of columns in Exhibit 3-3 (second row of Exhibit 3-2). The estimated impact on mean benefits paid in 2016 was \$166, equivalent to 1.5 percent of adjusted mean SSDI benefits paid to C1 subjects during 2016 (\$166/\$10,740).

The estimate of the benefit offset's impact on 2016 benefits is consistent with the evidence from previous years. Exhibit 3-3 reports the average monthly SSDI benefits paid from 2011 to 2016, inflation-adjusted to be equivalent to 2016 dollars. We present monthly benefit amounts rather than annual benefits because the 2011 estimate pertains to only eight months of the year. We see that the benefit offset had a positive and statistically significant impact on SSDI benefits paid in each year from 2011 through 2016.

Appendix C presents the estimated impact of the benefit offset on SSDI benefits due in 2016, which must be considered preliminary because the measure is based on adjustments made through January 2018 only. At this point, the benefits due impact estimate is not statistically significantly different than the benefits paid impact estimate.

3.1.2. Exploratory Impacts

A minority of Stage 1 subjects (about 15 percent) worked in 2016, and a much smaller portion (less than 3.5 percent) had earnings above BYA. The offset did not have a statistically significant impact on any of the four exploratory earnings outcomes.

The lack of statistically significant impacts on the percentage employed and the percentage with earnings above BYA in 2016 differs from the previous years' results. For 2015, the impacts on both of these outcomes are statistically significant, but in 2016 they are not (Exhibit B-1).¹⁹ For example, in 2015, 3.23 percent of T1 subjects had earnings above BYA compared to 2.97 percent of C1 subjects, implying a 0.27 percentage point impact—statistically significant at the 1-percent level. In 2016, the proportion earning above BYA rose for both groups, to 3.48 for T1 subjects and 3.36 for C1 subjects. Because the change was larger for the C1 subjects than for the T1 subjects, the 2016 impact estimate is smaller (0.12 percentage points) and below the threshold required for statistical significance.

The lack of statistically significant impacts on the proportions earning above two and three times BYA is also a change from the 2015 results. The change follows patterns that are symmetric to those observed for the percentage employed and the percentage with earnings above BYA. In each case the difference in the 2015 and 2016 impacts on the proportion earnings above two and three times BYA is not statistically significant (Exhibit B-1). However, the changes from 2015 to 2016 were sufficient to make estimates that were statistically significant in 2015 insignificant in 2016. In 2015, we found negative and statistically significant impacts on the proportion earning more than two times BYA and three times BYA.²⁰ Between 2015 and 2016, there was a smaller increase in the proportion of C1 subjects with earnings above two times BYA and above three times BYA than for T1 subjects.

¹⁹ The offset also had positive and statistically significant impacts on the percentage employed and the percentage with earnings above BYA in 2014 (Gubits et al. 2018b).

²⁰ The offset had a negative and statistically significant impact on the percentage with earnings above three times BYA in 2014. The impact on the percentage with earnings above two times BYA was not statistically significant in 2014.

The majority (84 percent) of the control group received an SSDI benefit in 2016 and the control group received SSDI benefits for an average of 9.8 months during the year. Benefit receipt was higher for treatment subjects by a statistically significant amount (Exhibit 3-4). The proportion of T1 subjects with an SSDI benefit payment is 1.2 percentage points higher, equivalent to 1.4 percent of the adjusted mean for C1 subjects. The mean number of months with SSDI payments is 0.13 months higher, equivalent to 1.3 percent of the adjusted mean for C1 subjects. This pattern of significant findings is consistent with the pattern from earlier years, and changes in estimated impacts across years are not statistically significant (Exhibit B-2). These findings presumably reflect 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 in earlier years, a minority of BOND subjects received an SSI payment in 2016 and there is no evidence of impacts on total SSI benefits paid, the proportion with at least one month with an SSI payment, or the number of months with SSI payments in 2016.

Exhibit 3-4. Impact Estimates for 2016 Employment and Benefit Receipt

Outcome	T1 Mean	C1 Mean	Impact Estimate
<i>Employment (January–December 2016)</i>			
Employment during year (%)	15.04	14.91	0.13 (0.16)
Earnings above BYA (%)	3.48	3.36	0.12 (0.09)
Earnings above 2x BYA (%)	1.51	1.60	-0.09 (0.05)
Earnings above 3x BYA (%)	0.83	0.85	-0.02 (0.04)
<i>Benefit Receipt (January–December 2016)</i>			
At least one month with an SSDI benefit paid (%)	85.01	83.81	1.20*** (0.17)
Number of months with SSDI payments	9.98	9.84	0.13*** (0.02)
At least one month with an SSI benefit paid (%)	14.28	14.36	-0.08 (0.13)
Total SSI benefits paid	\$365	\$368	-\$3 (\$7)
Number of months with SSI payments	1.60	1.61	-0.00 (0.02)

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

Notes: All earnings outcomes are based on a measure of earnings subject to Social Security taxes. Weights are used to ensure that the BOND subjects who met analysis criteria are representative of the national beneficiary population in the month of random assignment. Standard errors appear in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: T1 = 77,101; C1 = 891,429.

*/**/***/*** Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a two-tailed t-test with 9 degrees of freedom (resulting from a research design involving 10 study sites) and with no multiple-comparisons adjustment.

In summary, the absolute values of the 2016 estimates for impacts on employment and earnings above each of the three BYA thresholds are smaller than their corresponding 2015 values, and fewer are statistically significant—in each case the change in the point estimate of the impact from 2015 to 2016 is not statistically significant. The smaller estimated impacts in 2016 do not represent a large change in absolute or relative terms, despite the loss of statistical significance. The 2016 impacts on SSDI benefits are positive and statistically significant, consistent with 2015 results.

The loss of statistical significance and general dampening of the absolute value of estimated impacts of the offset on employment and earnings relative to the three BYA thresholds is consistent with the hypothesis that reversion to current law in 2016 affected estimated impacts, although by itself is not sufficient to conclude that such dampening occurred. Nearly 10 percent of T1 subjects reverted to current law in 2016, but many of them may not have been earning above BYA in 2015. To the extent feasible, the next section examines how their impacts changed from 2015 to 2016 relative to the changes in impacts for T1 subjects whose rules did not revert to current law.

3.2. Reversion to Current Law

Economic theory predicts that reversion to current law would cause some T1 subjects to reduce their earnings below BYA while causing others to increase their earnings to more than two or three times BYA. To explore whether there is evidence of such impacts in 2016—the first year in which the rules for a substantial share of T1 subjects reverted to current law—we estimated impacts for a new subgroup pair: the pre-BOND TWP completers and their complement—those who did not complete the TWP before BOND. As described in Section 2.1.3, the earnings rules for T1 pre-BOND TWP completers reverted to current law in May 2016. The large majority of T1 subjects who, according to administrative data, did not complete the TWP pre-BOND were subject to BOND rules for all of 2016. The exception is 0.8 percent of T1 subjects who reverted to current law between May and December of 2016. These subjects were included in the complement group to avoid using information from the post-random assignment period to define these two subgroups, and are such a small share of the complement subgroup that their inclusion will have negligible effects on the estimates for that subgroup.²¹

If reversion to current law affected earnings and employment outcomes in 2016, then we would expect it to be manifested most clearly in the impacts for the pre-BOND TWP completers, and especially on impacts for the percentage with 2016 earnings above each of the three BYA thresholds. Specifically, among the pre-BOND TWP completers, we would expect reversion to current law to dampen the absolute value of impacts in 2016 relative to what these impacts would have been had there been no reversion to current law. We use the change in impacts between 2015 and 2016 for the complementary subgroup as a benchmark to judge the change in impacts between 2015 and 2016 for the pre-BOND TWP completers. The power to detect the predicted differences across the two groups is limited by the relatively small sample size of the pre-BOND TWP completion subgroup (the T1 subjects in this subgroup represent 11.6 percent of the T1 sample).

²¹ About 0.14 percent of T1 subjects who did not complete the TWP pre-BOND reverted to current law in May 2016 (Exhibit A-1). This presumably occurred because SSA administrative data on TWP completion used for this analysis does not match the information SSA used to assign BPP dates and determine the timing of return to current law. See Appendix A.

We explore this hypothesis in two stages. First, we provide an orientation to these new subgroups by discussing their means and impacts in 2016. Next, we compare the change in impacts from 2015 to 2016 across the two subgroups to assess whether there is evidence of a change in behavior related to beneficiary reversion to current law in 2016. All impact estimates in this section are exploratory.²²

3.2.1. 2016 Earnings, Employment, and Benefit Outcomes for TWP Subgroups

Differences in outcomes between C1 subjects in the two subgroups are large and align with expectations. Those who completed the TWP pre-BOND (first three columns of Exhibit 3-5) have a relatively higher employment rate (42 percent) and relatively high earnings (more than \$6,000) in 2016. In contrast, those who did not complete the pre-BOND TWP (columns 4 through 6 of Exhibit 3-5) had a lower employment rate (11 percent) and lower average annual earnings (less than \$1,200). TWP completion at any time signals an ability to work and earn above the TWP threshold (\$810 in 2016), and it is not surprising that it continues to be predictive of employment and earnings several years later. SSDI benefit outcomes are also smaller for pre-BOND TWP completers among C1 subjects than for the balance of C1 subjects, presumably reflecting more frequent benefit suspension or termination due to work.

The offset had a negative and statistically significant effect on the proportion with earnings above twice BYA among pre-BOND TWP completers and no statistically significant effects on any other earnings and employment outcomes for either subgroup. For the pre-BOND TWP completers, the estimated impact on the proportion with 2016 earnings above twice BYA is -0.56 percentage points. This impact is 0.53 percentage points more negative than the impact for those who did not complete the TWP pre-BOND, a difference that is statistically significant at the 10 percent level.

The paucity of statistically significant findings for the employment and earnings above BYA outcomes may be due to low power—the result of the relatively small size of the pre-BOND TWP completer subgroup.

Among both subgroups, the offset had a positive, statistically significant effect on total SSDI benefits paid, with a larger impact for the pre-BOND TWP completers. Among those who completed the TWP pre-BOND, the offset caused an estimated average \$728 increase in total SSDI benefits paid in 2016 (8 percent of the corresponding C1 mean) compared to an \$88 increase in SSDI benefits paid (1 percent of the corresponding C1 mean) among those who did not complete the TWP pre-BOND. The difference in impacts, \$640, is statistically significant at the 1 percent level. There was also a statistically significant difference in impacts on number of months with SSDI payments, in the same direction.

²² These subgroups were not included in BOND's original evaluation plan (Bell et al. 2011), which did not anticipate an analysis of impacts after 2015.

Exhibit 3-5. Impact Estimates in 2016 for Subgroups Defined by TWP Completion in or Before April 2011

Outcome	Completed the TWP Pre-BOND			Did Not Complete TWP Pre-BOND			Estimated Difference in Impact (7)
	T1 Mean (1)	C1 Mean (2)	Impact Estimate (3)	T1 Mean (4)	C1 Mean (5)	Impact Estimate (6)	
Earnings and Employment Outcomes (January–December 2016)							
Total earnings	\$6,194	\$6,212	-\$19 (\$206)	\$1,128	\$1,120	\$8 (\$24)	\$26 (\$208)
Employment during year (%)	42.30	41.56	0.74 (0.62)	11.45	11.41	0.04 (0.18)	-0.70 (0.64)
Earnings above BYA (%)	12.56	12.17	0.39 (0.68)	2.29	2.21	0.08 (0.06)	-0.31 (0.68)
Earnings above 2x BYA (%)	5.67	6.23	-0.56* (0.30)	0.96	0.99	-0.03 (0.04)	0.53† (0.30)
Earnings above 3x BYA (%)	3.46	3.49	-0.03 (0.29)	0.48	0.51	-0.02 (0.03)	0.01 (0.29)
Benefit Outcomes (January–December 2016)							
Total SSDI benefits paid	\$10,360	\$9,632	\$728*** (\$101)	\$10,977	\$10,890	\$88*** (\$25)	-\$640††† (\$104)
Number of months with SSDI payments	9.94	9.40	0.55*** (0.07)	9.98	9.90	0.08*** (0.02)	-0.47††† (0.07)
Total SSI benefits paid	\$174	\$181	-\$6 (\$13)	\$390	\$392	-\$2 (\$8)	\$4 (\$15)
Number of months with SSI payments	1.08	1.09	-0.01 (0.05)	1.67	1.68	-0.00 (0.02)	0.01 (0.05)

Source: SSA administrative records from the MEF, PHUS, SSR, MBR, DCF, eWork, and DAF.

Notes: All earnings outcomes are based on earnings subject to Social Security taxes. Weights are used to ensure that the BOND subjects who met analysis criteria are representative of the national beneficiary population in the month of random assignment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Pre-BOND TWP Completion T1 = 7,418; Pre-BOND TWP Completion C1 = 111,973; No Pre-BOND TWP Completion T1 = 69,683; No Pre-BOND TWP Completion C1 = 779,456.

*/**/*** Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a two-tailed t-test with 9 degrees of freedom (resulting from a research design involving 10 study sites) and with no multiple comparisons adjustment.

†/††/††† Difference in impact estimates is significantly different from zero at the .10/.05/.01 levels, respectively, using an F-test.

3.2.2. Changes in Outcomes from 2015 to 2016 for TWP Subgroups

To support the assessment of whether 2016 impacts represent statistically significant departures from 2015 impacts in the directions expected because of reversion to current law in 2016, Exhibit 3-6 displays the earnings and employment results for 2016 along with the corresponding results for 2015 and changes from 2015 to 2016. We focus on impacts for the percent employed and percentages with earnings above BYA, two times BYA, and three times BYA among pre-BOND TWP completers because theory suggests that, if anything, reversion to current law will reduce the absolute values of previously observed impacts on these percentages—making the impact on the percentage with earnings above BYA less positive and the impacts on earnings above two and three times BYA less negative. We would expect any changes related to current law reversion to be evident only in this group, because only 0.8 percent of the other group reverted to current law in 2016, and some did so only late in the year.

We do not find evidence that reversion to current law reduced the impact from 2015 to 2016 on the percentage employed or the percentage with earnings above BYA.

Examining impacts on the percentage with earnings above BYA, the impact estimate declined for the pre-BOND TWP completers between 2015 and 2016, by -0.46—a statistically insignificant amount. We do not know how this impact estimate would have changed in the absence of reversion to current law. To better assess the possibility of reversion to current law affecting the 2016 impact estimate, we compare the change between years for this subgroup to the corresponding change for those who did not complete the TWP pre-BOND. The latter is also negative, and in this case statistically significant: -0.11. The difference, -0.35, is not statistically significant. Hence, the difference does not constitute evidence that reversion to current law reduced the impact on this outcome for the pre-BOND TWP completers. We also note the size of decline in this impact's point estimate's relative to the 2015 impact estimate was essentially the same for the two groups: 58 percent for the TWP completers and 54 percent for the complementary group.²³ This similarity may be just coincidental, but may point to some other factor as the cause of the decline in the impact estimates for both groups.

We also do not find clear evidence that reversion to current law caused some T1 subjects to increase their earnings to above two or three times BYA. We do see that the 2016 estimates of the impact on the proportion with earnings above two and three times BYA among pre-BOND TWP completers are each less negative than in 2015, but these differences are not statistically significant. In addition, the 2015 to 2016 estimated changes in impacts are not statistically significantly different from those observed for the group that did not complete the TWP pre-BOND.

To summarize, if reversion to current law affected earnings for some T1 subjects in 2016, we would expect to see a dampening of impacts on the percentage with 2016 earnings above each of the three BYA thresholds among pre-BOND TWP completers. The evidence indicates, however, that any such dampening was not large enough to be detected in annual outcomes for 2016. Although the point estimates for these three impacts changed in the direction anticipated from 2016 to 2015, none of the changes were statistically significant. This could be because the small size of the pre-BOND completers subgroup provides only limited statistical power.²⁴ Regardless of the limitations in power, however, qualitatively similar changes were observed for those who did not complete their TWP prior to BOND.

²³ There is no evidence to suggest that the reduction in impact on the proportion with earnings above BYA for those who did not complete the TWP pre-BOND (-0.11) is driven by beneficiaries who completed their TWP after BOND assignment in 2011 and in time to revert to current law by December 2016. Those who reverted to current law in 2016 comprise 0.8 percent of the subgroup that did not complete the TWP-pre BOND. If the point estimate of the change in impact for those in this subgroup who reverted to current law in 2016 was equal to the change in impact among the TWP completers (-0.46), it would account for about 3 percent of the point estimate for the group that did not complete the TWP pre-BOND [calculated as $0.008 * (-0.46) / (-0.11) = 0.033$].

²⁴ The small size of the pre-BOND completers subgroup provides only limited statistical power. The minimum detectable change in impact between 2015 and 2016 on the proportion with earnings above BYA is 1.79 percentage points, a much larger magnitude than the estimated 2015 impact of 0.85 percentage points. The minimum detected change in impacts is calculated as the multiplier of 2.716 times the estimated standard error for change in impact of 0.66. The multiplier 2.716 is the sum of critical values for 20 percent in one tail of the distribution (i.e., 80 percent power) and 5 percent in one tail of the distribution (i.e., a 10-percent t-test) using 9 degrees of freedom.

The full sample results show statistically significant impacts on the proportion earning above each of the three BYA thresholds in 2015 but not in 2016 (Exhibit B-1). The findings by pre-BOND TWP subgroups do not provide evidence that reversion to current law explains this pattern (Exhibit 3-6). There is some evidence of dampening among the subgroup that did not complete the TWP before BOND, of whom only a tiny share reverted to current law in 2016. As the subgroup that did not complete the TWP before BOND includes 88 percent of the Stage 1 sample, the changes in impacts from 2015 to 2016 for this group accounts for the bulk of the changes for the full sample.

Available evidence from the pre-BOND TWP subgroups suggests that the dampening of estimated impacts was caused by something other than reversion to current law. It is possible that the observed changes from 2015 to 2016 are simply due to random chance. The statistics, however, are also consistent with the possibility that a change in some external factor had positive effects on earnings outcomes for both groups. It is plausible, for instance, that improving job opportunities had a larger impact on the percentage of control subjects with earnings above BYA than on the corresponding percentage for treatment subjects because a larger share of treatment subjects capable of earning above BYA had achieved above BYA earnings in earlier years. The observed numerical increases in the proportions of subjects earning above BYA from 2015 to 2016, shown in Exhibit B-1, are consistent with that explanation: 0.40 percentage points for C1 subjects and 0.24 percentage points for T1 subjects. Although the evidence is consistent with the explanation posited, it is insufficient for us to be confident that the explanation is correct.

Exhibit 3-6. Impact Estimates for Change in Earnings and Employment Outcomes from 2015 to 2016 for Stage 1 Subgroups Defined by TWP Completion Pre-BOND

Outcome	Completed the TWP Pre-BOND			Did Not Complete TWP Pre-BOND			Estimated Difference in Impact (7)
	T1 Mean (1)	C1 Mean (2)	Impact Estimate (3)	T1 Mean (4)	C1 Mean (5)	Impact Estimate (6)	
Total Earnings							
2015	\$5,806	\$5,821	-\$15 (\$139)	\$956	\$941	\$15 (\$23)	\$31 (\$141)
2016	\$6,194	\$6,212	-\$19 (\$206)	\$1,128	\$1,120	\$8 (\$24)	\$26 (\$208)
Change from 2015 to 2016	\$388	\$391	-\$3 (\$159)	\$171	\$179	-\$8 (\$16)	-\$5 (\$160)
Employment during Year (%)							
2015	39.95	39.39	0.55 (0.57)	9.74	9.52	0.22 (0.13)	-0.34 (0.58)
2016	42.30	41.56	0.74 (0.62)	11.45	11.41	0.04 (0.18)	-0.70 (0.64)
Change from 2015 to 2016	2.35	2.16	0.19 (0.52)	1.71	1.89	-0.17 (0.19)	-0.36 (0.55)
Earnings Above BYA (%)							
2015	12.35	11.50	0.85* (0.43)	2.04	1.85	0.19** (0.06)	-0.66 (0.44)
2016	12.56	12.17	0.39 (0.68)	2.29	2.21	0.08 (0.06)	-0.31 (0.68)

Outcome	Completed the TWP Pre-BOND			Did Not Complete TWP Pre-BOND			Estimated Difference in Impact (7)
	T1 Mean (1)	C1 Mean (2)	Impact Estimate (3)	T1 Mean (4)	C1 Mean (5)	Impact Estimate (6)	
Change from 2015 to 2016	0.22	0.68	-0.46 (0.66)	0.25	0.36	-0.11* (0.05)	-0.35 (0.67)
Earnings Above 2x BYA (%)							
2015	5.25	5.93	-0.69** (0.26)	0.79	0.85	-0.06 (0.04)	0.63†† (0.26)
2016	5.67	6.23	-0.56* (0.30)	0.96	0.99	-0.03 (0.04)	0.53† (0.30)
Change from 2015 to 2016	0.43	0.30	0.13 (0.24)	0.17	0.15	0.03 (0.03)	-0.10 (0.24)
Earnings Above 3x BYA (%)							
2015	3.11	3.43	-0.33 (0.20)	0.40	0.45	-0.06* (0.03)	0.27 (0.20)
2016	3.46	3.49	-0.03 (0.29)	0.48	0.51	-0.02 (0.03)	0.01 (0.29)
Change from 2015 to 2016	0.35	0.06	0.29 (0.23)	0.09	0.05	0.03 (0.03)	-0.26 (0.23)

Source: SSA administrative records from the MEF, MBR, DCF, eWork, and DAF.

Notes: All earnings outcomes are based on a measure of earnings subject to Social Security taxes. Weights are used to ensure that the BOND subjects who met analysis criteria are representative of the national beneficiary population in the month of random assignment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics. All dollar amounts are inflation-adjusted to 2016 dollars using the Consumer Price Index for Urban Wage Earners and Clerical Workers (CPI-W).

Unweighted sample sizes: Pre-BOND TWP Completion T1 = 7,418; Pre-BOND TWP Completion C1 = 111,973; No Pre-BOND TWP Completion T1 = 69,683; No Pre-BOND TWP Completion C1 = 779,456.

*/**/** Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a two-tailed t-test with 9 degrees of freedom (resulting from a research design involving 10 study sites) and with no multiple comparisons adjustment.

†/††/††† Difference in impact estimates is significantly different from zero at the .10/.05/.01 levels, respectively, using an F-test.

3.3. Subgroups Based on Benefit Status at Baseline

Below we present the impact estimates for the subgroups defined by duration of SSDI benefit receipt (Exhibit 3-7) and SSI status (Exhibit 3-8). For reasons noted in Chapter 2, we consider all subgroup estimates as exploratory outcomes. For each pair of subgroups, we first describe adjusted outcome means for C1 subjects in the two subgroups; these values reflect population differences for the subgroups under current law.²⁵ We then describe impacts within each subgroup of the pair and discuss any evidence of differences in impacts across each pair of subgroups.

3.3.1. Duration Since Award

Consistent with findings from all previous BOND evaluation reports, there are differences in the *levels* of 2016 earnings and benefits paid for C1 subgroups by duration of prior SSDI receipt (Exhibit 3-7). Short-duration C1 subjects—defined as those on benefits for 36 months or less at baseline—have higher mean

²⁵ 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).

earnings (\$2,035 versus \$1,572) and are more likely to be employed (16.1 versus 14.4 percent) than long-duration C1 subjects. In addition, short-duration subjects have higher SSDI benefit payments than long-duration subjects (\$11,475 versus \$10,428). Total SSI payments are relatively lower for the short-duration group (\$315 versus \$391).

There is no evidence of different impacts across the benefit duration subgroups. The statistically significant impacts are quite similar across the two subgroups. Point estimates for impacts on SSDI benefits for short and long duration are \$107 and \$186, respectively. There are also positive impacts for months with SSDI benefits for both subgroups: 0.10 months (over a 12 month period) for short duration subjects and 0.15 months for long-duration subjects.

Exhibit 3-7. Impact Estimates for Subgroups Defined by Duration of SSDI Receipt

Outcome	Short Duration			Long Duration			Estimated Difference in Impact (7)
	T1 Mean (1)	C1 Mean (2)	Impact Estimate (3)	T1 Mean (4)	C1 Mean (5)	Impact Estimate (6)	
Earnings and Employment Outcomes (January–December 2016)							
Total earnings	\$2,019	\$2,035	-\$16 (\$55)	\$1,586	\$1,572	\$14 (\$39)	-\$30 (\$67)
Employment during year (%)	15.97	16.12	-0.15 (0.24)	14.63	14.39	0.25 (0.20)	-0.39 (0.31)
Earnings above BYA (%)	4.20	4.19	0.00 (0.13)	3.17	3.01	0.16 (0.13)	-0.16 (0.18)
Earnings above 2x BYA (%)	1.88	2.01	-0.13 (0.09)	1.35	1.42	-0.08 (0.06)	-0.05 (0.11)
Earnings above 3x BYA (%)	1.06	1.08	-0.03 (0.08)	0.73	0.75	-0.02 (0.05)	-0.00 (0.09)
Benefit Outcomes (January–December 2016)							
Total SSDI benefits paid	\$11,582	\$11,475	\$107** (\$36)	\$10,614	\$10,428	\$186*** (\$32)	-\$79 (\$48)
Number of months with SSDI payments	9.88	9.78	0.10*** (0.03)	10.02	9.87	0.15*** (0.03)	-0.04 (0.04)
Total SSI benefits paid	\$309	\$315	-\$5 (\$10)	\$389	\$391	-\$1 (\$9)	-\$4 (\$14)
Number of months with SSI payments	1.25	1.26	-0.02 (0.03)	1.76	1.76	0.00 (0.02)	-0.02 (0.03)

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

Notes: All earnings outcomes are based on a measure of earnings subject to Social Security taxes. Weights are used to ensure that the BOND subjects who met analysis criteria are representative of the national beneficiary population in the month of random assignment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Short Duration T1 = 38,663; Short Duration C1 = 209,756; Long Duration T1 = 38,438; Long Duration C1 = 681,673.

*/**/*** Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a two-tailed t-test with 9 degrees of freedom (resulting from a research design involving 10 study sites) and with no multiple comparisons adjustment.

†††/†††† Difference in impact estimates is significantly different from zero at the .10/.05/.01 levels, respectively, using an F-test.

3.3.2. SSI Benefit Status

Consistent with expectations and previous findings, there are also large differences in levels of 2016 earnings and benefit levels for C1 subgroups by SSI status (Exhibit 3-8). Relative to concurrent subjects, SSDI-only subjects have higher mean earnings (\$1,803 versus \$1,296) and higher mean SSDI benefit payments (\$11,898 versus \$5,541). This presumably reflects the fact that SSDI-only beneficiaries generally have more substantial earnings histories than concurrent beneficiaries. The percentage employed in 2016 was lower for SSDI-only (14.6 versus 16.2). Given that concurrent subjects are classified on the basis of SSI payments at the time of random assignment, it is not surprising that they had substantially more months with SSI payments in 2016 than did SSDI-only subjects (8.1 months over a 12 month period versus 0.2 months). The fact that some subjects in the SSDI-only group received SSI benefits after random assignment presumably reflects sufficiently large declines in their assets or income from other sources to satisfy the SSI means test.

Consistent with findings from previous years, we observe no evidence of differences in 2016 impacts by baseline SSI status. Within each subgroup, there are no impacts on earnings and employment outcomes. The offset has a statistically significant impact on SSDI benefits that is similar for both groups.

Exhibit 3-8. Impact Estimates for Subgroups Defined by Baseline SSI Status

Outcome	SSDI-Only			Concurrent			Estimated Difference in Impact (7)
	T1 Mean (1)	C1 Mean (2)	Impact Estimate (3)	T1 Mean (4)	C1 Mean (5)	Impact Estimate (6)	
Earnings and Employment Outcomes (January–December 2016)							
Total earnings	\$1,821	\$1,803	\$18 (\$35)	\$1,242	\$1,296	-\$54 (\$69)	\$72 (\$77)
Employment during year (%)	14.91	14.62	0.29 (0.17)	15.60	16.22	-0.62 (0.60)	0.91 (0.62)
Earnings above BYA (%)	3.65	3.49	0.16 (0.09)	2.72	2.79	-0.07 (0.23)	0.23 (0.24)
Earnings above 2x BYA (%)	1.70	1.77	-0.07 (0.06)	0.66	0.82	-0.16 (0.10)	0.09 (0.12)
Earnings above 3x BYA (%)	0.96	0.98	-0.02 (0.06)	0.23	0.28	-0.05 (0.06)	0.03 (0.09)
Benefit Outcomes (January–December 2016)							
Total SSDI benefits paid	\$12,066	\$11,898	\$168*** (\$28)	\$5,678	\$5,541	\$137** (\$47)	\$30 (\$55)
Number of months with SSDI payments	10.03	9.90	0.12*** (0.02)	9.75	9.58	0.18*** (0.05)	-0.06 (0.06)
Total SSI benefits paid	\$36	\$37	-\$0 (\$2)	\$1,847	\$1,860	-\$12 (\$37)	\$12 (\$37)
Number of months with SSI payments	0.17	0.17	-0.00 (<0.01)	8.07	8.08	-0.01 (0.08)	0.01 (0.08)

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

Notes: All earnings outcomes are based on a measure of earnings subject to Social Security taxes. Weights are used to ensure that the BOND subjects who met analysis criteria are representative of the national beneficiary population in the month of random assignment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: SSDI-only T1 = 64,696; SSDI-only C1 = 694,110; Concurrent T1 = 12,405; Concurrent C1 = 197,319.

*/**/*** Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a two-tailed t-test with 9 degrees of freedom (resulting from a research design involving 10 study sites) and with no multiple comparisons adjustment.

†††/†††† Difference in impact estimates is significantly different from zero at the .10/.05/.01 levels, respectively, using an F-test.

4. Discussion

The pattern of findings in 2016 for the two confirmatory outcomes (Exhibit 3-2) is similar to the pattern in each of the previous five years: we find no impact on earnings, but strong evidence of a positive impact on SSDI benefits paid in 2016. The dollar magnitude of the impact on SSDI benefits paid in 2016 is qualitatively similar to and not statistically different from the impact in 2015.

For the exploratory earnings and employment outcomes for the overall sample, we find no evidence of impacts on any of the employment and earnings outcomes in 2016. This is in contrast to the 2015 results, which found statistically significant positive impacts on employment and earnings above BYA and negative impacts on earnings above two and three-times BYA. However, the differences in the size of these impacts between 2015 and 2016 are not statistically significant. The exploratory benefit estimates are consistent with 2015 results. Notably, we found a positive and statistically significant impact of the offset on the number of months with SSDI payments in 2016, nearly identical to the impact in 2015.

The full sample results show statistically significant impacts on the proportion earning above each of the three BYA thresholds in 2015 but not in 2016 (Exhibit B-1). We did not find evidence that this change was explained by reversion to current law (Exhibit 3-6), although the statistical power for this analysis was low. The decline may just be due to idiosyncratic factors that affect the outcomes of individual beneficiaries. It is possible, however, that some other factor that had differential effects on treatment and control group outcomes contributed to the decline in these impact estimates. One possibility is the strengthening of the economy. That strengthening presumably explains why every earnings outcome for the full C1 group increased from 2015 to 2016 (Exhibit B-1). It could be that improving job opportunities had a larger impact on the percentage of control subjects earning above BYA than on treatment subjects because treatment subjects were more likely to have adjusted to earnings above BYA in earlier years. It is not possible, however, to differentiate this explanation from the possibility that changes in impacts observed are simply due to chance.

Regardless of the explanations for the different patterns of statistically significant impact estimates across years, the most important finding is that the 2016 results provide no statistically significant evidence of impacts on earnings and employment across all T1 beneficiaries. Without a sufficiently large positive impact on earnings—especially on the percentage of beneficiaries with earnings above BYA—the benefit offset will lead to higher SSDI benefits in a given calendar year relative to benefits without the offset. Indeed, we find strong evidence that the benefit offset increased SSDI benefits paid in 2016. This result holds when estimating impacts on SSDI benefits due in 2016 (i.e. benefits that SSA should have paid to the beneficiary based on eligibility and earnings in 2016). These benefit increases accrue to beneficiaries who would have engaged in SGA in absence of the offset.

References

- Bakker, B. F. 2009. *Trek alle register open! [Open up the registers!]*. Amsterdam, Netherlands: VU University.
- Bell, Stephen H., Daniel Gubits, David Stapleton, David Wittenburg, Michelle Derr, Arkadipta Ghosh, Sara Ansell, and David Greenberg. 2011. “BOND Implementation and Evaluation: Evaluation Analysis Plan.” Submitted to the Social Security Administration (contract deliverable 16.1 under Contract SS00-10-60011), Office of Program Development & Research. Cambridge, MA: Abt Associates, and Washington, DC: Mathematica Policy Research.
- Croake, Sarah, Priyanka Anand, Christopher Jones, Katherine Morrison, Cara Orfield, David Stapleton, Denise Hoffman, David Mann, Judy Geyer, Daniel Gubits, Stephen Bell, Andrew McGuirk, Tyler Rose, and David Wittenburg. 2017. “BOND Implementation and Evaluation: 2017 Stage 1 Interim Process, Participation, and Impact Report.” Submitted to the Social Security Administration (contract deliverable 24c2.2 under Contract SS00-10-60011), Office of Research, Demonstration, and Employment Support. Cambridge, MA: Abt Associates.
- Groen, Jeffrey A. 2012. “Sources of Error in Survey and Administrative Data: The Importance of Reporting Procedures.” *Journal of Official Statistics* 28(2): 173-198.
- Gubits, Daniel, Michelle Derr, Jillian Berk, Ann Person, David Stapleton, Denise Hoffman, Stephen Bell, Rachel Cook, and David Wittenburg. 2013. “BOND Implementation and Evaluation: Stage 2 Early Assessment Report.” Submitted to the Social Security Administration (contract deliverable 24a.2 under Contract SS00-10-60011), Office of Program Development & Research. Cambridge, MA: Abt Associates, and Washington, DC: Mathematica Policy Research.
- Gubits, Daniel, David Stapleton, Stephen Bell, Michelle Wood, Denise Hoffman, Sarah Croake, David R. Mann, Judy Geyer, David Greenberg, Austin Nichols, Andrew McGuirk, Meg Carroll, Utsav Kattell, and David Judkins. 2018a. “BOND Implementation and Evaluation: Final Evaluation Report, Volume 1.” Submitted to the Social Security Administration (contract deliverable 24e.2 under Contract SS00-10-60011), Office of Research, Demonstration, and Employment Support. Cambridge, MA: Abt Associates, and Washington, DC: Mathematica Policy Research.
- Gubits, Daniel, David Stapleton, Stephen Bell, Michelle Wood, Denise Hoffman, Sarah Croake, David R. Mann, Judy Geyer, David Greenberg, Austin Nichols, Andrew McGuirk, Meg Carroll, Utsav Kattell, and David Judkins. 2018b. “BOND Implementation and Evaluation: Final Evaluation Report, Volume 2 – Technical Appendices.” Submitted to the Social Security Administration (contract deliverable 24e.2 under Contract SS00-10-60011), Office of Research, Demonstration, and Employment Support. Cambridge, MA: Abt Associates, and Washington, DC: Mathematica Policy Research.
- Hoffman, Denise, Sarah Croake, David R. Mann, David Stapleton, Priyanka Anand, Chris Jones, Judy Geyer, Daniel Gubits, Stephen Bell, Andrew McGuirk, David Wittenburg, Debra Wright, Amang Sukasih, David Judkins, and Michael Sinclair. 2017. “BOND Implementation and Evaluation: 2016 Stage 1 Interim Process, Participation, and Impact Report.” Submitted to the Social Security Administration (contract deliverable 24c2.1 under Contract SS00-10-60011), Office of Research, Demonstration, and Employment Support. Cambridge, MA: Abt Associates, and Washington, DC: Mathematica Policy Research.

- Livermore, Gina, Allison Roche, and Sarah Prenovitz. 2009. "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, Office of Retirement and Disability Policy. Washington, DC: Mathematica Policy Research.
- Riley, Gerald R., and Kalman Rupp. 2012. "Expenditure Patterns Under the Four Major Public Cash Benefit and Health Insurance Programs for Working-Age Adults with Disabilities." *Journal of Disability Policy Studies* 25(2): 71-80.
- Schochet, Peter Z. 2008. "Technical Methods Report: Guidelines for Multiple Testing in Impact Evaluations." Submitted to the U.S. Department of Education, Institute of Education Science, NCEE 2008-4018. Princeton, NJ: Mathematica Policy Research.
- Stapleton, David C., Stephen H. Bell, David C. Wittenburg, Brian Sokol, and Debi McInnis. 2010. "BOND Implementation and Evaluation: BOND Final Design Report." Submitted to the Social Security Administration (contract deliverable 2.2 under Contract SS00-10-60011), Office of Program Development & Research. Cambridge, MA: Abt Associates, and Washington, DC: Mathematica Policy Research.
- Stapleton, David, David C. Wittenburg, Daniel Gubits, David Judkins, David Mann, and Andrew McGuirk. 2013. "BOND Implementation and Evaluation: First-Year Snapshot of Earnings and Benefit Impacts for Stage 1." Submitted to the Social Security Administration (contract deliverable 24c.1 under Contract SS00-10-60011), Office of Program Development & Research. Cambridge, MA: Abt Associates, and Washington, DC: Mathematica Policy Research.
- Stapleton, David, David Wittenburg, David R. Mann, Denise Hoffman, and Andrew McGuirk. 2014. "BOND Implementation and Evaluation: Second-Year Snapshot of Earnings and Benefit Impacts for Stage 1." Submitted to the Social Security Administration (contract deliverable 24c.2 under Contract SS00-10-60011), Office of Research, Demonstration, and Employment Support. Cambridge, MA: Abt Associates, and Washington, DC: Mathematica Policy Research.
- Westfall, Peter H., S. Stanley Young. 1993. *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment*. New York: John Wiley & Sons, Inc.
- Wittenburg, David, David Stapleton, Michelle Derr, Denise W. Hoffman, and David R. Mann. 2012. "BOND Implementation and Evaluation: BOND Stage 1 Early Assessment Report." Submitted to the Social Security Administration (contract deliverable 24.1 under Contract SS00-10-60011), Office of Program Development & Research. Cambridge, MA: Abt Associates, and Washington, DC: Mathematica Policy Research.
- Wittenburg, David, David R. Mann, David Stapleton, Daniel Gubits, David Judkins and Andrew McGuirk. 2015. "BOND Implementation and Evaluation: Third-Year Snapshot of Earnings and Benefit Impacts for Stage 1." Submitted to the Social Security Administration (contract deliverable 24.c.3 under Contract SS00-10-60011), Office of Research, Demonstration, and Employment Support. Cambridge, MA: Abt Associates, and Washington, DC: Mathematica Policy Research.

Appendix A. Additional Information on TWP Completion Subgroups

This appendix provides details about how we constructed the pre-BOND TWP completer subgroup. In Section A.1, we describe the definition and data sources used to classify treatment and control subjects in an identical manner. In Section A.2, we present statistics on the pre-BOND TWP completion status of treatment subjects, to verify that their status lined up with expectations about reversion to current law. We go on to consider the reasons for minor but unavoidable discrepancies.

A.1 Pre-BOND TWP Completion Subgroups

In this report, we define a subgroup pair that we did not specify in the original analysis plan (Bell et al. 2011) and have not considered previously in the BOND evaluation: pre-BOND TWP completers—beneficiaries who completed the TWP before BOND random assignment—and their complement. The first group includes Stage 1 subjects who completed a TWP in or by April 2011. The second group includes Stage 1 subjects who either completed a TWP after April 2011 or did not complete a TWP.

We identify the subgroups based on SSA administrative data. First, we identify beneficiaries with TWP completion dates in or before April 2011 recorded in either a March 2018 extract from the Disability Control File (DCF) or a January 2018 extract from SSA’s electronic work reporting (eWork) system. If a beneficiary has multiple TWP completion dates in or before April 2011, we sort the dates chronologically and select the date that is closest to, but before random assignment.

Second, because some of the TWP completion dates may correspond to previous entitlement periods (periods of benefit receipt that were terminated before the start of BOND), we only consider TWP completion dates that are after the start of the entitlement period during which the beneficiary was randomly assigned to BOND. To identify the start of the relevant entitlement period, we use administrative data from the MBR provided by SSA in October 2018. We identify the most recent entitlement date that is before random assignment and only consider TWP dates that occurred after that entitlement date. If the beneficiary was missing information on entitlement date in the MBR, we followed the same process if an entitlement date was available in the 2016 Disability Analysis File (DAF). We assign beneficiaries with TWP completion dates after the SSDI start date associated with BOND and before BOND random assignment to the first subgroup. If the beneficiary was missing information on entitlement date in both the MBR and DAF, we do not consider entitlement date and only compare TWP completion date to the BOND random assignment date.

A.2 Composition of the T1 Pre-BOND TWP Completion Subgroup

Exhibit A-1 presents statistics on the status of the two T1 subgroups as of the end of 2016 (Exhibit A-1). SSA had terminated the benefits of some members of both groups prior to May 2016 due to death or medical recovery, so there was no reason for their benefits to revert to current law in May 2016 or later. We retain this group in both samples, however, because our subgroup classification is based on TWP completion before random assignment regardless of outcomes after BOND random assignment, and it is possible, although perhaps unlikely, that random assignment could have resulted in benefit termination for one of these reasons. About 8 percent of pre-BOND TWP completers and 11 percent of all others did not revert to current law due to death or medical improvement before May 2016.

We present statistics on reversion to current law, as determined by the Implementation Team in 2016 (rows two through four in Exhibit A-1). The two TWP completion subgroups approximately correspond to beneficiaries who reverted to current law in May 2016 and those who did not, but there are some differences. The bulk of the pre-BOND TWP completers still on the SSDI rolls in May reverted to current law in that month, as expected, but a small minority did not; 5 (less than 0.01 percent of the weighted observations) reverted to current law later in 2016 and 224 (3.1 percent weighted) had not reverted to current law by the end of the year. Of the 69,838 subjects in the other group, 75 (0.1 percent weighted) reverted to current law in May, contrary to expectations. Another 559 (0.7 percent weighted) reverted to current law later in 2016, but this is not unexpected because some treatment subjects completed their TWP between May and December of 2016.

Exhibit A-1. Reversion to Current Law among TWP Completion Subgroup Pairs

	T1 Subjects who Completed the TWP Pre-BOND	T1 Subjects who Did Not Complete the TWP Pre-BOND
Died or medically terminated by May 2016	575 (7.6%)	8,243 (11.4%)
Reverted to current law in May 2016	6,459 (87.1%)	75 (0.1%)
Reverted to current law in June to December 2016	5 (0.0%)	559 (0.7%)
Did not revert to current law in 2016	224 (3.1%)	60,961 (88.1%)
Total	7,263 (100.0%)	69,838 (100.0%)

Source: BOND Operations Data System (BODS) and SSA administrative records from the DCF, MBR, eWork, and DAF.

Notes: Sample sizes are unweighted. Percentages are weighted to ensure that the BOND subjects who met analysis criteria are representative of the national beneficiary population in the month of random assignment. Mismatches between the classifications are shown in red.

There are several reasons why the constructed subgroups may not align with the Implementation Team information on reversion to current law:

- *Timing differences.* SSA determined reversion to current law in 2016 based on information available at that time. In contrast the research classifications are based on administrative data extracted two years later. If SSA made any retroactive updates in the intervening period, this could generate misalignment between the two classifications.
- *Matching error.* The BOND analysis sample includes beneficiaries entitled on multiple records. In such cases, implementation staff could identify the correct record via a case review of treatment subject data. Our algorithm, however, may select a record different than the one used to determine reversion to current law. This is because, unlike case reviews by SSA staff, our algorithm applies hard-coded rules to electronic records in an identical manner for all BOND subjects, not just treatment subjects.

- *Errors in SSA administrative data files.*²⁶ In some cases, there may be errors in one or multiple SSA administrative data source relative to the information that SSA staff had available for each case. All data sources may contain errors, and administrative data is no exception. Researchers have documented a variety of potential errors in administrative sources including processing errors and editing errors, among others (Groen 2012, Bakker 2009).

We defined the analysis groups based on TWP completion dates. We did not use information on actual reversion dates (that is BPP end dates) determined by the Implementation Team because control beneficiaries do not have BPP end dates for reversion to current law. While we could have used the Implementation Team data to classify treatment subjects, that would have violated the principle of identifying treatment and control subgroups in an identical manner. Such a violation would risk the creation of imbalances across the treatment and control groups with respect to unobservable characteristics that could affect outcomes. Our approach ensures that differences in outcomes for the two groups provide unbiased estimates of impacts and are not confounded by differences in unobservable characteristics.

²⁶ SSA staff explained that they correct data errors as soon as they are identified. Indeed, SSA staff conducted case reviews to better understand the source of misalignment between our research classification and the Implementation Team's administrative classifications. SSA staff identified two cases with incorrect administrative data records, which they immediately updated.

Appendix B. Impact Estimates for Changes in Outcomes from 2015 to 2016

Exhibit B-1. Impact Estimates for Change in Earnings and Employment Outcomes from 2015 to 2016 for All Stage 1 Subjects

Analysis Period	T1 Mean	C1 Mean	Impact Estimate
Total Earnings			
2015	\$1,520	\$1,508	\$12 (\$26)
2016	\$1,716	\$1,711	\$5 (\$30)
Change from 2015 to 2016	\$197	\$204	-\$7 (\$21)
Employment During Year (%)			
2015	13.25	12.99	0.26* (0.13)
2016	15.04	14.91	0.13 (0.16)
Change from 2015 to 2016	1.79	1.92	-0.13 (0.19)
Earnings Above BYA (%)			
2015	3.23	2.97	0.27*** (0.07)
2016	3.48	3.36	0.12 (0.09)
Change from 2015 to 2016	0.24	0.40	-0.15 (0.09)
Earnings Above 2x BYA (%)			
2015	1.31	1.44	-0.13** (0.05)
2016	1.51	1.60	-0.09 (0.05)
Change from 2015 to 2016	0.20	0.16	0.04 (0.04)
Earnings Above 3x BYA (%)			
2015	0.71	0.80	-0.09** (0.04)
2016	0.83	0.85	-0.02 (0.04)
Change from 2015 to 2016	0.12	0.05	0.06 (0.03)

Source: SSA administrative records from the MEF and MBR.

Notes: All earnings outcomes are based on a measure of earnings subject to Social Security taxes. Weights are used to ensure that the BOND subjects who met analysis criteria are representative of the national beneficiary population in the month of random assignment. Standard errors appear in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics. All dollar amounts are inflation-adjusted to 2016 dollars using the Consumer Price Index for Urban Wage Earnings and Clerical Workers (CPI-W).

Unweighted sample sizes: T1 = 77,101; C1 = 891,429.

*/**/** Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a two-tailed t-test with 9 degrees of freedom (resulting from a research design involving 10 study sites) and with no multiple-comparisons adjustment.

Exhibit B-2. Impact Estimates for Change in Benefit Outcomes from 2015 to 2016 for All Stage 1 Subjects

Analysis Period	T1 Mean	C1 Mean	Impact Estimate
Total SSDI Benefits Paid			
2015	\$11,257	\$11,115	\$142*** (\$24)
2016	\$10,906	\$10,744	\$162*** (\$25)
Change from 2015 to 2016	-\$351	-\$371	\$20 (\$13)
Number of Months with SSDI Payments			
2015	10.25	10.12	0.13*** (0.02)
2016	9.98	9.84	0.13*** (0.02)
Change from 2015 to 2016	-0.28	-0.28	0.00 (<0.01)
Total SSI Benefits Paid			
2015	\$385	\$389	-\$4 (\$7)
2016	\$365	\$368	-\$3 (\$7)
Change from 2015 to 2016	-\$19	-\$21	\$2 (\$3)
Number of Months with SSI Payments			
2015	1.68	1.69	-0.01 (0.01)
2016	1.60	1.61	-0.00 (0.02)
Change from 2015 to 2016	-0.07	-0.08	0.00 (<0.01)

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

Notes: All earnings outcomes are based on a measure of earnings subject to Social Security taxes. Weights are used to ensure that the BOND subjects who met analysis criteria are representative of the national beneficiary population in the month of random assignment. Standard errors appear in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics. All dollar amounts are inflation-adjusted to 2016 dollars using the Consumer Price Index for Urban Wage Earnings and Clerical Workers (CPI-W) to adjust 2015 dollars for inflation to be equivalent to 2016 dollars.

Unweighted sample sizes: T1 = 77,101; C1 = 891,429.

*/**/*** Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a two-tailed t-test with 9 degrees of freedom (resulting from a research design involving 10 study sites) and with no multiple-comparisons adjustment.

Exhibit B-3. Impact Estimates for Change in Benefit Outcomes from 2015 to 2016 for Stage 1 Subgroups Defined by TWP Completion Pre-BOND

Outcome	Completed the TWP Pre-BOND			Did Not Complete TWP Pre-BOND			Estimated Difference in Impact (7)
	T1 Mean (1)	C1 Mean (2)	Impact Estimate (3)	T1 Mean (4)	C1 Mean (5)	Impact Estimate (6)	
Total SSDI Benefits Paid							
2015	\$10,574	\$9,935	\$638*** (\$84)	\$11,347	\$11,270	\$77** (\$25)	-\$561*** (\$88)
2016	\$10,360	\$9,632	\$728*** (\$101)	\$10,977	\$10,890	\$88*** (\$25)	-\$640*** (\$104)
Change from 2015 to 2016	-\$214	-\$304	\$90 (\$71)	-\$369	-\$380	\$11 (\$12)	-\$79 (\$72)
Number of Months with SSDI Payments							
2015	10.21	9.65	0.56*** (0.07)	10.26	10.18	0.08*** (0.02)	-0.48*** (0.07)
2016	9.94	9.40	0.55*** (0.07)	9.98	9.90	0.08*** (0.02)	-0.47*** (0.07)
Change from 2015 to 2016	-0.26	-0.25	-0.01 (0.02)	-0.28	-0.28	<0.01 (<0.01)	0.02 (0.03)
Total SSI Benefits Paid							
2015	\$182	\$187	-\$6 (\$13)	\$411	\$416	-\$4 (\$8)	\$1 (\$15)
2016	\$174	\$181	-\$6 (\$13)	\$390	\$392	-\$2 (\$8)	\$4 (\$15)
Change from 2015 to 2016	-\$7	-\$7	-\$1 (\$5)	-\$21	-\$23	\$2 (\$3)	\$3 (\$5)
Number of Months with SSI Payments							
2015	1.14	1.14	<0.01 (0.05)	1.75	1.76	-0.01 (0.02)	-0.01 (0.05)
2016	1.08	1.09	-0.01 (0.05)	1.67	1.68	-0.00 (0.02)	0.01 (0.05)
Change from 2015 to 2016	-0.07	-0.05	-0.01 (0.01)	-0.08	-0.08	<0.01 (<0.01)	0.02 (0.01)

Source: SSA administrative records from the PHUS, SSR, MBR, DCF, eWork, and DAF.

Notes: All earnings outcomes are based on a measure of earnings subject to Social Security taxes. Weights are used to ensure that the BOND subjects who met analysis criteria are representative of the national beneficiary population in the month of random assignment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics. All dollar amounts are inflation-adjusted to 2016 dollars using the Consumer Price Index for Urban Wage Earners and Clerical Workers (CPI-W).

Unweighted sample sizes: Pre-BOND TWP Completion T1 = 7,418; Pre-BOND TWP Completion C1 = 111,973; No Pre-BOND TWP Completion T1 = 69,683; No Pre-BOND TWP Completion C1 = 779,456.

*/**/*** Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a two-tailed t-test with 9 degrees of freedom (resulting from a research design involving 10 study sites) and with no multiple comparisons adjustment.

†/††/††† Difference in impact estimates is significantly different from zero at the .10/.05/.01 levels, respectively, using an F-test.

Appendix C. Benefits Due Impact Estimates

In this appendix, we present benefits impact estimates using the benefits due measure and then compare them to the main impacts in this report, which are measured using benefit paid in 2016. The benefits due estimate must be considered preliminary because they reflect retroactive adjustments through January 2018 only. The impact estimates based on benefits due show positive impacts for both SSDI benefit outcomes (Exhibit C-1). There is strong evidence of a positive impact (\$134) on total SSDI benefits due. In addition, the findings show that T1 subjects typically receive SSDI benefits for a longer period (0.16 months over a 12 month period) than C1 subjects. There are no statistically significant impacts for either SSI outcome.

Exhibit C-1. Impact Estimates for 2016 Benefits Due

Outcome	T1 Mean	C1 Mean	Impact Estimate
Total SSDI benefits due	\$10,937	\$10,803	\$134*** (\$24)
Number of months with SSDI benefit due	9.98	9.82	0.16*** (0.02)
Total SSI benefits due	\$356	\$357	-\$1 (\$7)
Number of months with SSI benefit due	1.59	1.59	-0.00 (0.02)

Source: SSA administrative records from the MBR and SSR.

Notes: All earnings outcomes are based on a measure of earnings subject to Social Security taxes. Weights are used to ensure that the BOND subjects who met analysis criteria are representative of the national beneficiary population in the month of random assignment. Standard errors appear in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics. Dollar values are not adjusted for inflation.

Unweighted sample sizes: T1 = 77,101; C1 = 891,429.

*/**/** Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a two-tailed t-test with 9 degrees of freedom (resulting from a research design involving 10 study sites) and with no multiple-comparisons adjustment.

The group means and estimated impacts are substantively similar across the benefits due and benefits paid in measures (Exhibit C-2). There are no statistically significant differences between the benefits due and benefits paid impact estimates. The difference in total SSDI benefits between measures, \$28 is about one-quarter of one percent of the control mean.

Estimated impacts for benefits due for previous years presented in Gubits et al. (2018, Exhibit F-4) are larger in every year than the corresponding estimates for benefits paid, although not by a statistically significant amount. Hence, the fact that the point estimate for the impact on benefits due in 2016 is smaller than the estimate for benefits paid is somewhat of a departure from previous findings.

The change in the direction of this relationship—although not statistically significant in any year—may have occurred for several reasons. First, the underlying data used to create benefits due was extracted 17 months after the end of the analysis period used in Gubits et al. (2018), compared to 13 months for the 2016 estimate. We would expect to observe additional changes to benefits due with a later extraction date and these changes may be uneven across treatment and control groups. Second, fewer retroactive adjustments may have been required for the 2016 treatment subjects because by the end of 2016 the backlog of their pending work CDRs was much smaller than at the end of 2015. Third, 2016 was the first year in which some beneficiaries reverted to current law. SSA took efforts to ensure that the benefits of significant number of offset users in this group were adjusted timely, presumably reducing the likelihood

of retroactive adjustments. Further, retroactive adjustments made to the benefits of those who reverted to current law rules may have in some instances reduced their benefits due relative to their benefits paid. Finally, it is important to reiterate that the differences in all years are not statistically different. They all point to the same conclusion: the benefit offset had a positive impact on SSDI benefits.

Exhibit C-2. Impact Estimates for SSDI Benefits Due Compared to SSDI Benefits Paid

Outcome	Benefits Due			Benefits Paid ^a			Difference in Impact
	Treatment Mean	Control Mean	Impact Estimate	Treatment Mean	Control Mean	Impact Estimate	
Total SSDI benefits	\$10,937	\$10,803	\$134*** (\$24)	\$10,906	\$10,744	\$162*** (\$25)	-\$28 (\$35)
Number of months with SSDI benefits	9.98	9.82	0.16*** (0.02)	9.98	9.84	0.13*** (\$0.02)	0.03 (0.03)
Total SSI benefits	\$356	\$357	-\$1 (\$7)	\$365	\$368	-\$3 (\$7)	\$2 (\$9)
Number of months with SSI benefits	1.59	1.59	-0.00 (0.02)	1.60	1.61	-0.00 (0.02)	0.00 (0.02)

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

Sample Sizes: T1 = 77,101; C1 = 891,429.

Notes: Standard errors are in parentheses. Impact estimates are regression-adjusted for baseline characteristics. Weights are used to ensure that the BOND subjects who met analysis criteria are representative of the national beneficiary population in the month of random assignment.

*/**/** Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a two-tailed t-test with 9 degrees of freedom (resulting from a research design involving 10 study sites) and with no multiple-comparisons adjustment.

^a Total SSDI benefits paid presented in Exhibit 3-2 was designated as a confirmatory outcome and thus was adjusted for multiple comparisons. For this table, the hypothesis tests for all outcomes, including total SSDI benefits paid, were not adjusted for multiple comparisons.