

Note to readers: This January 16, 2018 version of the report replaces the February 27, 2017 version previously posted on the SSA website, correcting an error in the impact estimation.

Caution to readers: The estimates produced from IRS earnings and SSA benefit paid data in this report were later updated to include refinements to the analytic methodology and data. The specific variables affected are: Total earnings, Employment, Earnings above BYA, Earnings above 2XBYA, Earnings above 3XBYA, Total SSDI benefits paid, Number of months with SSDI payments, Total SSI benefits paid, and Number of months with SSI payments. The data and statistical methods used to produce these estimates have been updated over the course of the demonstration, making the published estimates in this report out of date. For the most up-to-date estimates, please refer to the Final Evaluation Report which will be available in late 2018.

BOND Implementation and Evaluation 2015 Stage 2 Interim Process, Participation, and Impact Report

Deliverable 24c2.3

January 16, 2018

Submitted to:

Social Security Administration

Attn: Ms. Joyanne Cobb

Office of Research, Demonstrations and
Employment Support (ORDES)

500 E Street, SW

9th Floor, RM #905

Washington, DC 20254

Contract No. SS00-10-60011

Prepared by:

Daniel Gubits

Judy Geyer

Denise Hoffman

Sarah Croake

Utsav Kattel

David Judkins

Stephen Bell

David Stapleton



Table of Contents

Acronyms Used in This Report..... vii

Terminology..... viii

Executive Summary ix

1. Introduction..... 1

 1.1. The BOND Policy Test 1

 1.2. The BOND Evaluation 4

 1.2.1. The Process Analysis..... 5

 1.2.2. The Participation Analysis 9

 1.2.3. The Impact Analysis..... 11

 1.3. The Current Report..... 12

 1.3.1. Data Sources for the Process Analysis 12

 1.3.2. Data Sources for the Participation Analysis 14

 1.3.3. Data Sources for the Impact Analysis 15

 1.4. Organization of the Report 15

2. BOND Study Sites and Disability Service Environment 17

 2.1. BOND Study Sites..... 17

 2.2. Geographic Characteristics..... 17

 2.3. Economic Indicators..... 17

 2.4. Number and Staffing Configuration of BOND Benefits Counseling Providers..... 18

 2.5. Number of BOND Stage 2 treatment subjects enrolled 19

 2.6. Availability and Use of Employment Services and Other Work-Focused, Disability-Related Resources 20

 2.7. Non-BOND SSDI Benefits Counseling Services 20

 2.8. Conclusion..... 21

3. BOND Benefits Counseling 24

 3.1. Design of BOND Counseling 24

 3.2. BOND Post-Entitlement Services 25

 3.3. Caseloads..... 26

 3.4. Analysis of Counseling Receipt 30

 3.5. Analysis of Beneficiary Survey Responses 37

 3.6. Conclusion..... 42

4. Knowledge of How Earnings Affect Calculation of Benefits and Future Benefit Eligibility and How Stage 2 Subjects Describe BOND 43

 4.1. Role of Understanding of the Benefit Offset in the BOND Logic Model..... 43

 4.2. How Stage 2 Subjects Think SGA-Level Earnings Would Affect Benefits Prior to Random Assignment 44

 4.3. How Stage 2 Subjects Think SGA-Level Earnings Would Affect Benefits at One Year After Random Assignment 46

 4.4. How Stage 2 Subjects Think SGA-Level Earnings Would Affect Future Benefit Eligibility One Year After Random Assignment 51

 4.5. How Stage 2 Treatment Subjects Describe BOND One Year After Random Assignment..... 54

 4.6. Summary 55

5. Using the Benefit Offset 56

5.1. Introduction 56

5.2. Implementation of the Benefit Adjustment Process 56

5.2.1. Eligibility for Benefit Adjustment..... 58

5.2.2. Documentation of Eligibility for Benefit Adjustment 58

5.2.3. Annual Earnings Estimate Submission 62

5.2.4. Benefit Adjustment 63

5.2.5. Delayed or Incorrect Benefit Adjustment and Potential Effects 65

5.3. Statistics on Offset Use and Benefit Adjustment 68

5.3.1. Pathway to Benefit Adjustment..... 68

5.3.2. Offset Use and Benefit Adjustment..... 74

5.4. Conclusion..... 77

6. Methodology and Context 79

6.1. Administrative Outcome Definitions and Theoretical Impacts 81

6.2. Administrative Features of the Offset That Influence Impacts 83

6.3. Impact Estimation Methodology 84

6.4. Final Analysis Sample Sizes 87

7. Impacts on Annual Earnings and SSDI Benefits Measured in Administrative Data 88

7.1. Employment, Earnings, Benefits, and Health Trends in the Control Group 89

7.2. Confirmatory Impacts..... 92

7.3. Exploratory Impacts for the Full Sample 94

7.3.1. Exploratory Impacts on Earnings-Related Outcomes 94

7.3.2. Exploratory Impacts on Benefit-Related Outcomes 96

7.4. Impact Variation by Beneficiary Background Characteristics 98

7.5. Summary of Results 100

8. Impacts on Additional Employment Outcomes and on Health-Related Outcomes 101

8.1. Impact Estimates for Survey-derived Employment Outcomes 101

8.2. Impact Estimates for Health Outcomes 106

8.3. Impact Estimates for Marital Status Outcomes 107

8.4. Summary 108

9. Conclusion..... 109

9.1. WIC and EWIC Services..... 109

9.2. Awareness of How Earnings Affect Calculation of Benefits 110

9.3. Use of the Benefit Offset..... 111

9.4. Delays with Benefit Adjustment and Improper Payments 111

9.5. Impact of the Offset and EWIC on Earnings..... 112

9.6. Impact of the Offset and EWIC on SSDI Benefits 112

9.7. Taking Stock..... 113

References..... 114

Appendix A. Impact Methodology 117

Appendix B. Subgroup Exhibits for 2013 Earnings and Benefit Impacts 136

Appendix C. 2011 & 2012 Earnings and Benefits Impacts..... 164

List of Exhibits

Exhibit 1-1.	Overview of BOND Random Assignment Process.....	3
Exhibit 1-2.	Previous Reports on BOND Participation, Process, and Impact Analyses for Stage 1 and Stage 2	5
Exhibit 1-3.	Implementation Findings through November 2012 from the Stage 2 Early Assessment Report	7
Exhibit 1-4.	Qualitative Data Collection – Fall 2014 Focus Groups.....	14
Exhibit 2-1.	BOND Site-Level Characteristics	22
Exhibit 2-2.	BOND Site-Level Employment Rates	23
Exhibit 3-1.	WIC and EWIC Caseloads Over Time.....	27
Exhibit 3-2.	WIC Caseload by Site in January 2014 and December 2014.....	27
Exhibit 3-3.	Engaged EWIC Cases in January 2014 and December 2014.....	29
Exhibit 3-4.	WIC and EWIC Caseloads per FTE.....	29
Exhibit 3-5.	Receipt of Benefits Counseling (through December 2014) by T21 and T22 Subjects	31
Exhibit 3-6.	Receipt of Benefits Counseling (through December 2014) by T21 and T22 Subjects, by Employment Status at Enrollment	32
Exhibit 3-7.	WIC Services for T21 Subjects (through December 2014).....	33
Exhibit 3-8.	EWIC Services for T22 Subjects (through December 2014)	33
Exhibit 3-9.	Receipt of EWIC Services by T22 Subjects by Calendar Year.....	34
Exhibit 3-10.	EWIC Vs. WIC Impact Estimates (T22 Vs. T21) on Counseling Receipt (through December 2014)	35
Exhibit 3-11.	EWIC vs WIC Impact Estimates (T22 vs T21) on Counseling Receipt (through December 2014) by Baseline Employment Status	36
Exhibit 3-12.	Estimated Impacts on Self-Reported Use of Training-Related Counseling of Stage 2 Volunteers: All Policy Comparisons	38
Exhibit 3-13.	Estimated Impacts on Self-Reported Service Locations, Referrals, and Personal Assistance Services of Stage 2 Volunteers: All Policy Comparisons	39
Exhibit 3-14.	Satisfaction with Counseling Services, Conditional on Service Receipt	41
Exhibit 4-1.	At Baseline: How Stage 2 Subjects Think Benefits Would Change as a Result of Earnings above SGA, Full Sample	46
Exhibit 4-2.	One-Year Follow-up: How Stage 2 Subjects Think Benefits Would Change as a Result of Earnings above SGA, Full Sample.....	47
Exhibit 4-3.	One-Year Follow-up: How Stage 2 Subjects Think Benefits Would Change as a Result of Earnings above SGA, By Baseline Employment Status	49
Exhibit 4-4.	One-Year Follow-up: How Stage 2 Subjects Think Future Benefit Eligibility Would Change as a Result of Earnings Above SGA, Full Sample and By Baseline Employment Status.....	52
Exhibit 4-5.	One-Year Follow-up: How Stage 2 Treatment Subjects Describe BOND	54
Exhibit 5-1.	Benefit Adjustment under Current Law and the BOND Benefit Offset.....	57

Exhibit 5-2. Percentage of Stage 2 Treatment Subjects with Documented Steps towards Benefit Adjustment (through December 31, 2014) 69

Exhibit 5-3. EWIC versus WIC Impact Estimates on Steps towards Benefit Adjustment (through December 31, 2014) 70

Exhibit 5-4. Stage 2 Treatment Beneficiary Characteristics by Steps towards Benefit Adjustment (through December 31, 2014)..... 71

Exhibit 5-5. Predictors of Benefit Adjustment for Stage 2 Treatment Subjects (through December 31, 2014) 73

Exhibit 5-6. Cumulative Stage 2 Treatment Subjects Eligible for an Adjustment Under the Offset and With an Adjustment, Based on Data Through December 31, 2015..... 76

Exhibit 6-1. Definitions of Confirmatory and Exploratory Administrative Outcomes and Predicted Signs of Impacts 83

Exhibit 6-2. Stage 2 Analysis Sample Composition 87

Exhibit 7-1. Trends in Employment and Earnings, Current Law Control Group (C2) 90

Exhibit 7-2. Earnings Trends for Employed Subjects, Current Law Control Group (C2) 90

Exhibit 7-3. Self-Reported Health in 2011 and 2013, Current Law Control Group (C2) 91

Exhibit 7-4. Trends in Annual SSDI Benefits Paid, Current Law Control Group 91

Exhibit 7-5. Estimated Impacts on 2013 Total Earnings and Total SSDI Benefits Paid of Stage 2 Volunteers: All Policy Comparisons..... 93

Exhibit 7-6. Estimated Impacts on 2013 Employment and Benefits Receipt for Stage 2 Volunteers: All Policy Comparisons 96

Exhibit 8-1. Estimated Impacts on Employment-Related Outcomes of Stage 2 Volunteers Over the Full Follow-Up Period or at the Time of Survey Interview: All Policy Comparisons..... 103

Exhibit 8-2. Estimated Impacts on Self-Reported Health Status of Stage 2 Volunteers: All Policy Comparisons..... 106

Exhibit 8-3. Estimated Impacts on Marital Status of Stage 2 Volunteers All Policy Comparisons..... 107

Exhibit A-1. W-2 Wage and Tax Statement..... 120

Exhibit A-2. Stage 2 Impact Estimates on Confirmatory Outcomes Illustrating the Multiple Comparison Adjustment on *p*-values 123

Exhibit A-3. Administrative-Data Covariates Included in Stage 2 Impact Regressions..... 125

Exhibit A-4. Baseline Survey Covariates Included in Stage 2 Impact Regressions..... 126

Exhibit B-1. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T21 Vs. C2) for Subgroups Defined by Duration of SSDI Receipt..... 136

Exhibit B-2. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. C2) for Subgroups Defined by Duration of SSDI Receipt..... 137

Exhibit B-3. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. T21) for Subgroups Defined by Duration of SSDI Receipt 138

Exhibit B-4. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 + T21 vs C2) for Subgroups Defined by Duration of SSDI Receipt 139

Exhibit B-5. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T21 Vs. C2) for Subgroups Defined by Employment at Baseline 140

Exhibit B-6. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. C2) for Subgroups Defined by Employment at Baseline 141

Exhibit B-7. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. T21) for Subgroups Defined by Employment at Baseline 142

Exhibit B-8. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 + T21 vs C2) for Subgroups Defined by Employment at Baseline 143

Exhibit B-9. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T21 Vs. C2) for Subgroups Defined by Access to Medicaid Buy-in Programs 144

Exhibit B-10. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. C2) for Subgroups Defined by Access to Medicaid Buy-in Programs 145

Exhibit B-11. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. T21) for Subgroups Defined by Access to Medicaid Buy-in Programs..... 146

Exhibit B-12. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 + T21 vs C2) for Subgroups Defined by Access to Medicaid Buy-in Programs..... 147

Exhibit B-13. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T21 Vs. C2) for Subgroups Defined by Age at Baseline 148

Exhibit B-14. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. C2) for Subgroups Defined by Age at Baseline 149

Exhibit B-15. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. T21) for Subgroups Defined by Age at Baseline 150

Exhibit B-16. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 + T21 vs C2) for Subgroups Defined by Age at Baseline 151

Exhibit B-17. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T21 Vs. C2) for Subgroups Defined by Primary Impairment of Major Affective Disorder..... 152

Exhibit B-18. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. C2) for Subgroups Defined by Primary Impairment of Major Affective Disorder..... 153

Exhibit B-19. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. T21) for Subgroups Defined by Primary Impairment of Major Affective Disorder..... 154

Exhibit B-20. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 + T21 vs C2) for Subgroups Defined by Primary Impairment of Major Affective Disorder..... 155

Exhibit B-21. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T21 Vs. C2) for Subgroups Defined by Primary Impairment of Back Disorder ... 156

Exhibit B-22. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. C2) for Subgroups Defined by Primary Impairment of Back Disorder ... 157

Exhibit B-23. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. T21) for Subgroups Defined by Primary Impairment of Back Disorder..... 158

Exhibit B-24. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 + T21 vs C2) for Subgroups Defined by Primary Impairment of Back Disorder..... 159

Exhibit B-25. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law T21 Vs. C2) for Subgroups Defined by Education at Baseline..... 160

Exhibit B-26. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. C2) for Subgroups Defined by Education at Baseline 161

Exhibit B-27. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. T21) for Subgroups Defined by Education at Baseline..... 162

Exhibit B-28. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 + T21 vs C2) for Subgroups Defined by Education at Baseline..... 163

Exhibit C-1. Estimated Impacts on 2012 Earnings and Benefits of Stage 2 Volunteers: Confirmatory Results, All Policy Comparisons 164

Exhibit C-2. Estimated Impacts on 2012 Earnings and Benefits of Stage 2 Volunteers: Exploratory Results, All Policy Comparisons 165

Exhibit C-3. Estimated Impacts on 2012 Outcomes of the Offset Plus WIC Compared to Current Law (T21 Vs. C2) for Subgroups Defined by Duration of SSDI Receipt..... 166

Exhibit C-4. Estimated Impacts on 2012 Outcomes of the Offset Plus EWIC Compared to Current Law (T22 Vs. C2) for Subgroups Defined by Duration of SSDI Receipt..... 167

Exhibit C-5. Estimated Impacts on 2012 Outcomes of the Offset Plus EWIC Compared to the Offset Plus WIC (T22 Vs. T21) for Subgroups Defined by Duration of SSDI Receipt 168

Exhibit C-6. Estimated Impacts on 2012 Outcomes of the Offset Plus WIC Compared to Current Law (T21 Vs. C2) for Subgroups Defined by Year of Study Enrollment 169

Exhibit C-7. Estimated Impacts on 2012 Outcomes of the Offset Plus EWIC Compared to Current Law (T22 Vs. C2) for Subgroups Defined by Year of Study Enrollment 170

Exhibit C-8. Estimated Impacts on 2012 Outcomes of the Offset Plus EWIC Compared to the Offset Plus WIC (T22 Vs. T21) for Subgroups Defined by Year of Study Enrollment 171

Exhibit C-9. Estimated Impacts on 2011 Earnings and Benefits of Stage 2 Volunteers: All Policy Comparisons..... 172

Acronyms Used in This Report

ADC	Adult Disability Cessation	I&R	Information and Referral
AEE	Annual Earnings Estimate	IRS	Internal Revenue Service
AIME	Average Indexed Monthly Earnings	IRWE	Impairment Related Work Expenses
BODS	BOND Operations Data System	ORDES	Office of Research, Demonstration and Employment Support
BOND	Benefit Offset National Demonstration	PII	Personal Identifying Information
BS&A	Benefits Summary and Analysis	SGA	Substantial Gainful Activity
BSAS	BOND Stand Alone System	SNAP	Supplemental Nutrition Assistance Program
BYA	BOND Yearly Amount	SSA	Social Security Administration
CDR	Continuing Disability Reviews	SSDI	Social Security Disability Insurance
CIL	Center for Independent Living	SSI	Supplemental Security Income
CWIC	Community Work Incentive Coordinators	SVRA	State Vocational Rehabilitation Agency
DAC	Disabled Adult Child	TSA	Transfer of Skills Analysis
DWB	Disabled Widow/Widowers Benefits	TTW	Ticket-to-Work
EN	Employment Network	TWP	Trial Work Period
EPE	Extended Period of Eligibility	WC	Workers' Compensation
ESP	Employment Service Plan	WIC	Work Incentive Counseling, Or Counselor
EWIC	Enhanced Work Incentives Counseling, Or Counselor	WIP	Work Incentives Plan
FTE	Full-Time Equivalent	WIPA	Work Incentives, Planning, and Assistance
GP	Grace Period	WISE	Work Incentives Seminar Events

Terminology

1. **Prospective BOND subjects:** Beneficiaries in the pool eligible for potential random assignment at Stage 1.
2. **Stage 2 solicitation pool:** SSDI-only beneficiaries to be recruited for Stage 2.
3. **Stage 2 volunteers:** Those subjects who volunteer for Stage 2.
4. **BOND subjects:** Beneficiaries assigned to any of the five BOND treatment or control groups, at either stage (see **Exhibit 2-3**). Terms for subjects in specific groups are as follows:
 - a. **Treatment subjects:** All subjects offered the use of the benefit offset, including:
 - i. **T1 subjects** or **Stage 1 treatment subjects:** Those offered the offset at Stage 1.
 - ii. **Stage 2 treatment subjects:** Those offered the offset at Stage 2, including:
 - (1) **T21 subjects** or **Stage 2 offset-only subjects:** Stage 2 volunteers offered the offset, but not offered enhanced work-incentives counseling.
 - (2) **T22 subjects** or **Stage 2 offset-EWIC subjects:** Stage 2 volunteers offered both the offset and enhanced work-incentives counseling.
 - b. **Control subjects:** Those whose benefits will continue to be determined by current law.
 - i. **C1 subjects** or **Stage 1 control subjects:** Those assigned to the Stage 1 control group.
 - ii. **C2 subjects** or **Stage 2 control subjects:** Stage 2 volunteers assigned to the Stage 2 control group.
5. **BOND users:** Those treatment subjects who take up a BOND treatment. These include:
 - a. **Offset users** – All treatment subjects who have their benefits reduced by the offset.
 - b. **EWIC users** – All treatment subjects who use EWIC services. They can only be subjects in the T22 group.
 - c. **WIC users** – All treatment subjects who use WIC services. They can be subjects in the T1 or T21 groups.

Executive Summary

As part of the Ticket to Work (TTW) and Work Incentives Improvement Act of 1999, Congress directed the Social Security Administration (SSA) to test alternative Social Security Disability Insurance (SSDI) work rules designed to increase the incentive for SSDI beneficiaries to work and reduce their reliance on benefits. In response, SSA has undertaken the Benefit Offset National Demonstration (BOND), a random assignment test of variants of SSDI program rules governing work and other supports. BOND incorporates a \$1 for \$2 benefit offset that allows beneficiaries to retain more of their monthly cash benefit while working.

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

BOND takes place in 10 large sites, each corresponding to the service area of one of 53 SSA Area Offices. The 10 sites are a random sample of the 53 candidate areas to ensure that the evaluation's findings are nationally representative. In total, the Stage 2 sample includes 12,744 beneficiaries in three assignment groups:

- 4,854 volunteers in the T21 group that receives the benefit offset and regular work incentives counseling (WIC);
- 3,041 volunteers in the T22 group that receives the benefit offset and enhanced work incentives counseling (EWIC); and
- 4,849 volunteers in the C2 group that continues to receive benefits and counseling according to current SSDI rules.

This report is the first of two Stage 2 *Interim Process, Participation, and Impact Reports*. This report documents results of the Stage 2 process and participation analyses through the fourth calendar year of implementation (2014) and documents impacts on earnings and benefit outcomes during the third calendar year of implementation (2013).

Summary of Key Findings

Process and Participation Analyses

The process and participation analyses examine benefits counseling offered in the demonstration, beneficiary understanding of current law and benefit offset rules, and use of the benefit offset. Overall, this analysis reveals that treatment group subjects received counseling as intended but experienced substantial lags in their benefit adjustments under the offset. The key findings are that:

- Stage 2 volunteers assigned to the offset plus EWIC (T22 subjects) received more counseling services than those randomly assigned to the offset plus WIC (T21 subjects).
- About half of the treatment subjects demonstrated an accurate understanding of how the benefit offset works. The understanding of how higher earnings affect SSDI is not substantially more

accurate for T22 subjects than for T21 subjects despite the expectation that EWIC would improve this outcome.

- By the end of 2014, the proportion of the offset-plus-EWIC group (T22) with at least one month of offset use (11.0 percent) was 2.2 percentage points greater than the corresponding proportion for the offset-plus-WIC group (T21) (8.8 percent).
- Several beneficiary characteristics are associated with benefit adjustment under the offset in or before December 2014, including young age, good or very good general health status, back or musculoskeletal disorders and injuries as primary impairments, baseline employment, high baseline average indexed monthly earnings, and residence in a county with a high employment rate among people with disabilities.
- Actual benefit adjustment often lags the eligibility for a benefit adjustment due to: failure of some offset-eligible beneficiaries to report their earnings; delays in the processing of beneficiaries' work continuing disability reviews (CDRs); and deficiencies in the BOND Stand Alone System (BSAS) that caused delays in automated reconciliation of beneficiary earnings and thus benefits adjustments.

Impact Analysis

Among the multiple outcomes we analyze in this report, two are of paramount interest: total earnings in 2013 and total SSDI benefits paid in 2013.

Confirmatory findings on 2013 Earnings and SSDI Benefits:

We do not detect effects of the benefit offset on either total earnings in 2013 or total SSDI benefits paid in 2013, the report's two confirmatory outcomes. In addition, we do not find an effect on earnings or benefits for the offset-plus-EWIC group compared to the offset-plus-WIC group.

We also estimated the impact of the offset on a large number of exploratory outcomes and subgroups. We classify these additional analyses as exploratory and therefore do not make a correction for having run multiple statistical tests of impact (as we did for the confirmatory outcomes).

Impact of the Offset and of EWIC on Earnings Behavior (exploratory findings):

We find exploratory evidence of an increase in the proportion of beneficiaries in both treatment groups with 2013 earnings above the BOND Yearly Amount (BYA), at which beneficiaries become eligible for the offset. We also found an increase in employment for the offset-plus-WIC group.

Impact of the Offset and of EWIC on SSDI Benefits (exploratory findings):

We find exploratory evidence of a positive effect on number of months of SSDI receipt in 2013 for the two offset groups compared to the current law control group. The relatively small share of beneficiaries induced by the offset to increase earnings above BYA (about 2 percent of all treatment subjects) compared to the much larger proportion of subjects with earnings above BYA under current law (about 9 percent of all control group subjects) suggests that the offset will have a positive effect on SSDI benefits paid *for* 2013, after all retroactive adjustments are made. However, data for this outcome are not yet available. The final BOND evaluation report will include an analysis of benefits paid *for* 2013.

1. Introduction

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 that are designed to increase the incentive for SSDI beneficiaries to work and to reduce their reliance on SSDI benefits. In response, SSA has undertaken the Benefit Offset National Demonstration (BOND), a random assignment test of variants of SSDI program rules governing work and other supports. SSA, in conjunction with several contractors led by Abt Associates, developed the infrastructure and supports required to implement BOND.

The BOND project includes two stages. Stage 1 is designed to examine how a national benefit offset would affect earnings and program outcomes for the entire SSDI population. Stage 2 is designed to learn about impacts for those most likely to use the offset (recruited and informed volunteers) and to determine the impacts of the addition of Enhanced Work Incentives Counseling (EWIC) to the offset.

This report, the first of two Stage 2 *Interim Process, Participation, and Impact Reports*, documents results of the Stage 2 process and participation analyses through the fourth calendar year of implementation (2014). This report also documents impacts on earnings and benefit outcomes during the third calendar year of implementation (2013).

The first Stage 2 *Snapshot Report* (Gubits et al., 2014) documented Stage 2 impacts on earnings and benefit outcomes during the first and second calendar years (2011 and 2012) of implementation. Future reports—*Interim Process, Participation, and Impact Report* in 2017, a second *Snapshot Report* in 2016 and the *Final Report* in 2017—will track Stage 2 impacts through 2015. A parallel series of reports is being produced for Stage 1, the first of which were released in 2013 and 2015.

This introductory chapter describes the benefit offset and Stage 2 of the demonstration (Section 1.1). The chapter also reviews the objectives of the BOND evaluation and the research questions addressed by the process, participation, and impact analyses (Section 1.2). Section 1.2 also summarizes primary findings to date on the implementation and impacts of BOND as documented in previous reports. The chapter concludes with a description of the data sources used in the current report (Section 1.3) and the organization of the remainder of the report (Section 1.4).

1.1. The BOND Policy Test

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, SSDI benefits are lost if a beneficiary's countable monthly earnings exceed the monthly Substantial Gainful Activity (SGA) amount after completing a nine-month Trial Work Period (TWP) and a three-month grace period (GP). In 2013, the year for which Stage 2 impacts are analyzed for this report, the SGA amount was \$1,040 per month for non-blind beneficiaries and \$1,740 per month for blind beneficiaries. The complete loss of

¹ Other benefits include Medicare for those on the rolls for at least 24 months. These benefits are extended for a lengthy period following suspension of SSDI benefits, but not indefinitely. Some SSDI beneficiaries also receive Supplemental Security Income (SSI), Medicaid, or other public or private benefits that can be reduced or eliminated as earnings increase.

benefits for sustained earnings in excess of the SGA amount is sometimes called the “cash cliff.” Economic theory suggests that the cash cliff discourages some beneficiaries from working at all and encourages those who work and could earn above the SGA level to keep their earnings below that level.

BOND replaces the cash cliff with a “ramp” (benefit offset), with the policy objective of encouraging beneficiaries who can work above the SGA level to increase their earnings and reduce their reliance on benefits. More specifically, the benefit offset is expected to increase the earnings of those who otherwise might not work at all, or would earn less than the SGA amount. If such individuals engage in SGA under the benefit offset, their benefits eventually will be reduced. Countering the possible reduction in SSDI benefit outlays are benefits paid under BOND to those who would have had earnings above the SGA amount in the absence of BOND. Thus, the direction of the net impact on mean earnings and benefits of all beneficiaries will depend on the size of the impacts for beneficiaries who would not engage in SGA under current law, relative to the size of the impacts for those who would. Those in the latter group lose their benefits entirely under current law, whereas, under the benefit offset, many will be eligible for a reduced SSDI benefit.

BOND also changes the administrative processes used to adjust benefits, and 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 2013, \$12,480 for non-blind and \$20,280 for blind Stage 2 treatment subjects)². The benefit offset reduces benefits by \$1 for every \$2 in countable annual earnings in excess of the BYA following the completion of the GP. It can also be helpful to beneficiaries who have variable monthly earnings. SSA continues to pay benefits monthly under BOND, but the monthly payment amount is based on expected annual earnings. In the following calendar year, SSA reconciles payments to actual countable earnings, based on information provided by the Internal Revenue Service (IRS), documentation provided by the beneficiary, or both.

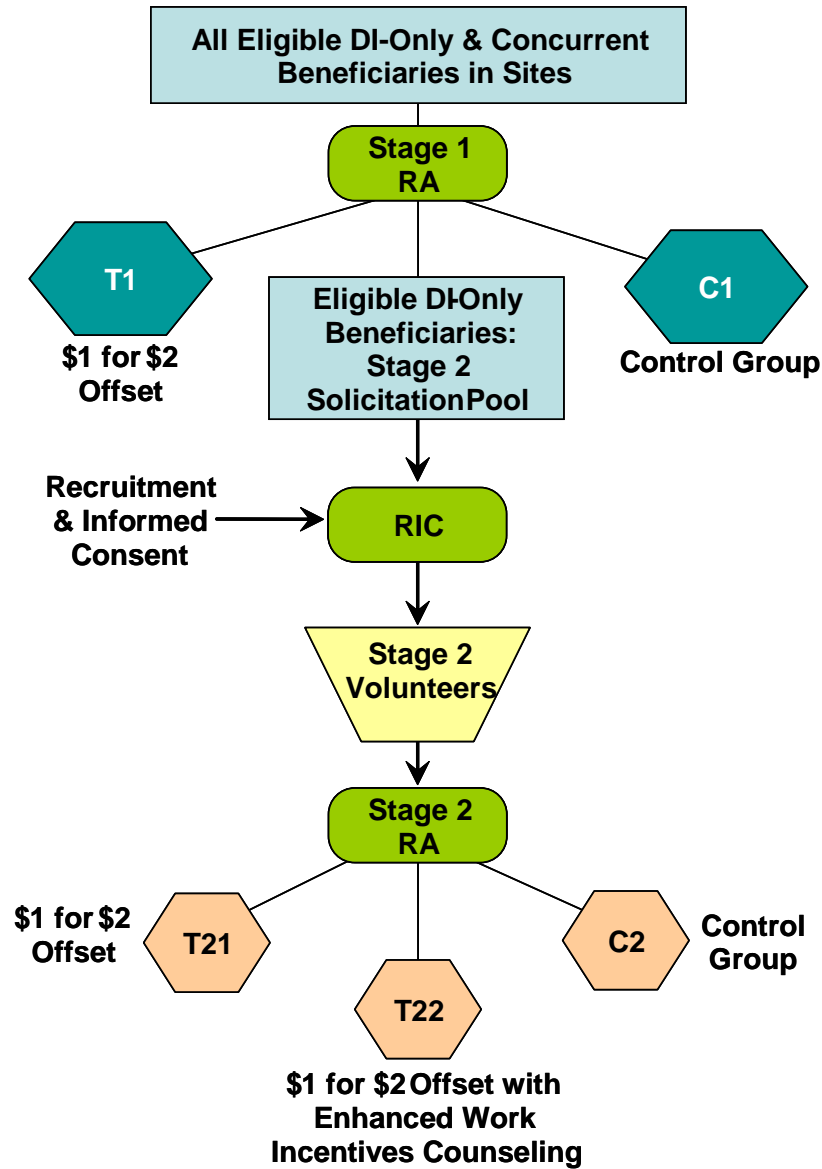
Treatment group beneficiaries eligible for the offset can use it for a 60-month participation period, which begins the month after random assignment for those who completed the TWP before that point or in the month after a given beneficiary’s TWP ends, provided that the TWP is completed by September 30, 2017. Those who do not complete the TWP by that date will lose their opportunity to use the offset. SSA will not permanently terminate benefits because of work during the participation period, even if benefits fall to zero because of earnings. SSA will apply current rules at the end of the participation period and will terminate the benefits of those engaged in SGA after the point when all remaining GP months have been used.

As noted earlier, BOND includes two stages—Stages 1 and 2—that test the benefit offset’s impact on the overall SSDI population and on those who have signaled interest in employment (see Exhibit 1-1). The goal of Stage 1 is to examine how a national benefit offset and changes to ancillary supports would affect earnings and program outcomes for the entire SSDI population. To test BOND, Abt Associates randomly selected ten large study sites to statistically represent the nation. In Stage 1, the demonstration randomly

² Stage 2 impacts documented in this report are analyzed through calendar year 2013. In 2014 the BYA was \$12,840 for non-blind beneficiaries and \$21,600 for blind beneficiaries.

Exhibit 1-1. Overview of BOND Random Assignment Process

BOND Sample Enrollment



DI = disability insurance; RA = random assignment; RIC = recruitment and informed consent.

assigned beneficiaries to a treatment group receiving the offset (T1 subjects) or to a control group continuing under standard rules (C1 subjects). By design, T1 and C1 subjects were to have access to counseling of roughly similar level of intensity: C1 subjects were to have access to counseling under an existing program—Work Incentives Planning and Assistance (WIPA)—whereas treatment subjects were to have access to similar counseling services, customized to the special rules that apply to their benefits—Work Incentives Counseling (WIC). The two groups should be identical except for the BOND intervention, so that any statistically significant differences in outcomes between T1 and C1 subjects can confidently be attributed to the intervention—the basic impact measurement strategy in a randomized experiment. The final Stage 1 analysis sample contains a total of 968,713 subjects, spread across T1 (77,115) and C1 (891,598).³

Stage 2 also uses an experimental design to learn about the impacts of the benefit offset on those beneficiaries most likely to use it—informed and recruited volunteers—and to determine the effects of the delivery of more intensive Enhanced Work Incentives Counseling (EWIC) services relative to current law and relative to WIC services. To achieve these goals, Stage 2 uses three-way random assignment into an offset-plus-WIC group (T21 subjects), an offset-plus-EWIC group (T22 subjects), and a current-law benefits group (C2 subjects). Concurrent beneficiaries—SSDI beneficiaries who also were receiving Supplemental Security Income (SSI) at the time of random assignment—were included in Stage 1 but excluded from Stage 2, because the interaction between SSI and SSDI diminishes the value of the SSDI benefit offset.⁴

Random assignment for Stage 2 occurred between March 1, 2011 and September 28, 2012, with 40 percent of volunteers enrolling in the study in 2011 and 60 percent of volunteers enrolling in 2012. In total, the Stage 2 sample includes 12,744 beneficiaries. The random assignment ratio for the T21, T22, and C2 assignment groups was 8:5:8; ultimately, 4,854 volunteers were assigned to the T21 group, 3,041 volunteers were assigned to the T22 group, and 4,849 volunteers were assigned to the C2 group. This report examines interim results of the Stage 2 process, participation, and impact analyses.

1.2. The BOND Evaluation

Abt Associates, in partnership with Mathematica Policy Research, is conducting a comprehensive evaluation of the BOND interventions, including studies of demonstration implementation, beneficiary participation, net impacts, and net social costs and benefits. The evaluation also will include cross-cutting analyses that combine findings from these four components. It is the expectation that, drawing on the various components of the evaluation, these studies will deepen our understanding of how the BOND interventions affected beneficiaries. Initial findings from the process and participation analysis have been published in previous reports as shown in Exhibit 1-2.

This section describes the research questions addressed by the process, participation, and impact analyses and summarizes key findings documented in previous reports.

³ See Stapleton et al. (2014) for details of the sample and initial impact estimates.

⁴ See Bell et al. (2011) for more details on the random assignment process and reasons for excluding concurrent beneficiaries from Stage 2 but not Stage 1.

The BOND Evaluation Team

Abt Associates, in partnership with 25 other organizations, is implementing and evaluating the Benefit Offset National Demonstration (BOND) under contract to the U.S. Social Security Administration. To ensure the objectivity of the evaluation, separate teams conduct the implementation (the “Implementation Team”) and evaluation (the “Evaluation Team”) components of the project. The current report reflects exclusively the views of the evaluation team, led by Evaluation Co-Directors Stephen Bell of Abt Associates and David Stapleton of Mathematica Policy Research. These 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) and Principal Investigator (Howard Rolston) have joint responsibility for coordinating the implementation and evaluation efforts, including, respectively, managing the day-to-day operations of the project and overseeing the effective and efficient implementation of the BOND design. Within this structure, full authority over and responsibility for the content of all evaluation reports rests with the Evaluation Co-Directors.

Exhibit 1-2. Previous Reports on BOND Participation, Process, and Impact Analyses for Stage 1 and Stage 2

Analysis	Stage 1 Reports	Stage 2 Reports
Participation and Process Analysis	<ul style="list-style-type: none"> Wittenburg et al. (2012) Derr et al. (2015) 	<ul style="list-style-type: none"> Gubits et al. (2013) Derr et al. (2015)
Impact Analysis	<ul style="list-style-type: none"> Stapleton et al. (2013) Stapleton et al. (2014) Stapleton et al. (2015) 	<ul style="list-style-type: none"> Gubits et al. (2014)

1.2.1. The Process Analysis

The overarching objective of the process analysis is to document the BOND intervention, creating a foundation for interpreting impacts and assessing the fidelity of the implementation of BOND. To that end, the process study is designed to evaluate the implementation of BOND within and across the study sites over time and to assess the fidelity of the implementation compared to the original design. It includes seven rounds of field work activities over the course of the demonstration and relies on multiple data sources, including feedback from beneficiaries. As described in the Evaluation Analysis Plan (Bell et al. 2011), the process analysis uses a mix of qualitative and quantitative data to address five broad research questions:⁵

1. How was the intervention implemented for Stage 1 and Stage 2? How did the implementation evolve over time?
2. Were the recruitment and enrollment processes for Stages 1 and 2 implemented as designed? If significant deviations occurred, why did they occur?

⁵ The research questions have been slightly modified from the BOND Evaluation Analysis Plan (Bell et al. 2011).

3. Were WIC and EWIC services implemented as designed? To what extent did EWIC services differ from WIC services?
4. Were the processes for reporting earnings, determining TWP completion, and making benefit adjustments for Stages 1 and 2 implemented as designed? How well did they perform?
5. What are the likely implications for demonstration outcomes? What are the lessons for national implementation of a benefit offset?

Findings to Date on the Implementation of BOND

The evaluation's early process analysis findings are summarized in the *Stage 2 Early Assessment Report* released in August 2013 (Gubits et al., 2013) and in *Process Study Report* released in February 2015 (Derr, et al., 2015). This section describes key findings from the process analysis documented in these reports.

The early assessment of Stage 2 implementation covered the period through November 2012, two months after participant enrollment ended. That study revealed that key processes were in place but that challenges remained for this portion of the demonstration, and emphasized the need for continued improvement. Exhibit 1-3 summarizes findings from the Stage 2 Early Assessment Report. Core activities during the start-up period included opening and closing the BOND site offices, enrollment, and a shift in responsibilities for preparing information for work Continuing Disability Reviews (CDR). In the initial demonstration design, the plan was for BOND field staff to complete work CDR preparation activities and submit needed information to SSA. This responsibility later shifted to SSA staff. In addition, the BOND implementation team and SSA made changes to a number of procedures and tools intended to improve the performance of demonstration processes. Another noteworthy event during this time was a change in how counseling services to BOND control subjects and all non-BOND beneficiaries were provided. WIPA funding ended June 30, 2012, with no indication that it would be reinstated. More than a year later, in August 2013, Congress re-appropriated WIPA funding. Changes to the WIPA program created some counseling service disruptions for control subjects, although SSA continued to provide some WIPA-like services through the Ticket to Work Help Line and Community Work Incentive Coordinators. In those sites that relied on staff to provide both BOND and WIPA services, these changes to the WIPA program created disruptions in BOND staffing.

In the Stage 2 Early Assessment Report, the BOND evaluation team concluded that the Stage 2 volunteer group was likely to successfully serve its purpose of testing impacts of the offset and of enhanced counseling as an add-on to the offset. One of the major accomplishments was that BOND slightly exceeded its overall enrollment target for Stage 2 (12,744 enrolled and are in the analysis sample versus 12,601 targeted enrollments) and more than half of the sites met their individual targets. In addition, the Stage 2 random assignment was successful in creating three well-matched study groups (T21, T22, and C2). Furthermore, as intended, EWIC services differed from WIC services on three primary fronts: EWIC staff (1) contacted beneficiaries proactively, (2) routinely followed up with beneficiaries and referral organizations, and (3) used a more systematic assessment process.

Exhibit 1-3. Implementation Findings through November 2012 from the Stage 2 Early Assessment Report

Topic	Findings
BOND Infrastructure	<ul style="list-style-type: none"> • BOND staff at site offices and WIC and EWIC providers did not fully understand some critical BOND procedures, including the BODS, development of work CDRs, and preparation of annual earnings estimates (AEEs). • The large amounts of information conveyed during training made it difficult for field staff to fully absorb all of the important points, and opportunities to practice new procedures before actually using them were limited. • Training and technical assistance resources, although helpful, were not wholly successful in improving field staff understanding. • At a number of junctures when issues arose in the operation of the demonstration, the Implementation Team and SSA responded by making changes to BOND procedures and tools.
Outreach and Enrollment	<ul style="list-style-type: none"> • BOND slightly exceeded its Stage 2 enrollment target and enrolled 12,744 volunteers from February 1, 2011, to September 28, 2012. • About half of study enrollments took place in the field (rather than in site offices), a rate which was higher and more costly than anticipated. • Although many of the features of recruitment were conducted uniformly across sites, there is some evidence that the efforts fell short of the desired level of uniformity.
Random Assignment	<ul style="list-style-type: none"> • Stage 2 random assignment was successful in creating three well-matched study groups (T21, T22, and C2).
WIC and EWIC Services	<ul style="list-style-type: none"> • The differences in caseloads of WIC and EWIC workers were more modest than expected, mainly due to lower-than-anticipated take-up of WIC services. • The main differences between EWIC and WIC services, as implemented, were that, compared to WIC staff, EWIC staff (1) contacted beneficiaries proactively, (2) followed up with beneficiaries and referral organizations, and (3) used a more systematic assessment process. • Consistent with the design, almost all (97 percent) of T22 subjects had some contact with an EWIC counselor, whereas less than a third (28 percent) of T21 subjects had contact with a WIC counselor.
Pathway to the Offset	<ul style="list-style-type: none"> • About 2 percent of Stage 2 treatment subjects were known to have used the offset by the end of 2012. • Benefit adjustment for offset-eligible beneficiaries was not completed on a timely basis due, in part, to delayed work CDRs and annual earnings estimates (AEEs). • Shifting work CDR responsibility from BOND field staff to SSA improved the process, but not the timeliness, of processing work CDRs. Continued delays reflect the large initial backlog, the effort required to complete CDRs, and the limited availability of qualified staff at SSA. • Problems with the BSAS used to adjust benefits delayed adjustment for many offset users not in contact with the demonstration by five months—above and beyond the delay inherent in this back-door adjustment process. • The lack of communication between the BOND field staff and SSA prevented the former from informing beneficiaries about the status of their cases, including potential improper payments.

Source: "Stage 2 Early Assessment Report," Gubits et al. 2013.

Despite these accomplishments, important challenges remained as of December 2012. Most notably, site-office staff and WIC and EWIC staff did not fully understand some critical BOND procedures, such as using the BOND Operations Data System (BODS), developing work CDRs, and preparing annual earnings estimates (AEEs). Based on feedback from staff interviews, it appears that the following factors contributed to this limited understanding of BOND: the newness of the intervention, the inherent complexity of the demonstration, and numerous changes in policies and procedures, and staff turnover.

Derr et al. (2015) reported on BOND implementation through the third calendar year (2013). Overall, the analysis showed that BOND implementation had gradually improved so that, as of fall 2013, the demonstration was, in large part, functioning as designed. The three primary findings for Stage 2 reported in Derr et al. 2015 are: (1) consistent with the BOND design, clear distinctions existed between WIC and EWIC services; (2) SSA had adjusted the benefits of 7.3 percent of Stage 2 subjects under the offset and the number using the offset was growing steadily; and (3) delays in the adjustment of benefits under the offset had been lengthy and, in many cases, led to improper payments for beneficiaries, but delays were becoming substantially shorter.

A central feature of the demonstration is testing whether counseling services more intensive than those available under the current law increase the size of the impacts of the benefit offset on earnings, benefits, and other outcomes. Essential to this test is a clear distinction between WIC services—those designed to be comparable to services available to all beneficiaries provided by CWICs (benefits counselors) under WIPA—and EWIC services designed to feature key enhancements compared to WIC. Derr et al. 2015 found that WIC and EWIC services differed substantially in both the quantity and nature of services provided as intended in the design. The analysis showed substantial differences in all sites, but differences in some sites were much larger than in others. The take-up rate of benefits counseling was lower than expected for WIC services (T1 and T21 subjects) and met expectations for EWIC services (T22 subjects). True to the design of BOND, nearly all subjects in the EWIC group received some benefits counseling, compared to about one-third of the WIC group. As intended, T22 subjects receive more intensive services than T21 subjects.

The percentage of treatment subjects who had used the offset is steadily growing. Based on benefit adjustments made through May 2014, 1.2 percent of Stage 2 treatment subjects had used the offset by the end of 2011. By the end of 2013, SSA had identified 7.3 percent of Stage 2 treatment subjects as offset users. Derr et al. 2015 concluded that offset use is likely to continue to rise as SSA retroactively adjusts the benefits of offset users through 2013 and more treatment subjects engage in SGA and qualify for the offset.

Derr et al. (2015) also found that since the initial implementation of BOND, delays with benefit adjustments under the offset had been common. Because such lags are common under current law, BOND control subjects also have presumably experienced substantial delays. Most of the causes of lags in benefit adjustment are the same under the offset and current law. Under both sets of rules, delays can occur because (1) beneficiaries do not report earnings in a timely manner (to SSA field offices or, in the case of the offset, to the demonstration); (2) backlogs occur in tasks such as processing work CDRs; or (3) beneficiaries are slow to respond to requests for information when the adjustment process is started. Three considerations suggest that the demonstration likely contributed to the lags for treatment subjects: (i) the need to inform treatment subjects about the offset, (ii) start-up problems in the post-entitlement processes, and (iii) delayed review of IRS data on earnings. The process analysis documented

improvements in the timeliness for benefit adjustment under BOND, the most notable of which is the decline in work CDR processing times. Some delays remained at the end of 2013 but expansion in resources available to perform work CDRs and the centralization of post-entitlement work were expected to further reduce delays.

1.2.2. The Participation Analysis

The participation analysis documents the engagement of prospective BOND subjects in work activity from before entry into the demonstration through the end of the demonstration. The evaluation team expected that the use of BOND demonstration services by both treatment groups, including their interest in volunteering to participate in these services in Stage 2, will vary across beneficiary subgroups (e.g., younger versus older workers). The participation analysis summarizes patterns of participation by subgroups and more broadly examines whether the BOND interventions influenced all types of work activity. Gubits et al. (2013) and Derr et al. (2015) reported on early results of the participation analysis from analysis of Stage 2 recruitment and participation in WIC and EWIC. In this report, the evaluation team presents findings about beneficiary engagement in work and, for the first time, beneficiary engagement with WIC and EWIC.

Recruitment of volunteers for Stage 2 is the first component of the participation analysis. Stage 2 subjects targeted for recruitment had to volunteer in order to participate in the demonstration. Stage 2 recruitment thus provides an important opportunity to examine beneficiary interest prior to the delivery of intervention services. This information provides insights about the characteristics of treatment and control subjects recruited in Stage 2 and, more broadly, the characteristics of beneficiaries who potentially could be targeted in future SSA demonstrations for return-to-work interventions based on their interest in BOND. This recruitment analysis addresses the following questions:

1. To what extent do subjects targeted for recruitment for Stage 2 volunteer?
2. What characteristics distinguish volunteers from non-volunteers?

The second component of the participation analysis examines beneficiary engagement in work and use of demonstration services. The findings will provide information on the engagement of prospective BOND subjects in work activities and, for the treatment groups, use of demonstration services, across both stages. The analysis for the second component will address the following questions:

3. To what extent do subjects in each treatment group work or use employment services and benefits counseling?
4. Who works, uses counseling services and other work incentives, and eventually uses the offset?
5. How did the demonstration affect the use of work-incentive counseling and the services delivered by counselors?
6. What characteristics distinguish offset users from others?
7. How do work and use of work incentives vary across demonstration groups?
8. How do work and use of work incentives change with time?

Gubits et al. (2013) examined the first two research questions about Stage 2 recruitment. The evaluation team expected that beneficiaries who had received SSDI for a short duration (defined as three years or less) would be more likely to work—and so be more responsive to the work incentives in BOND—than beneficiaries who had received SSDI for a longer time. Therefore, SSA especially sought information in Stage 2 for beneficiaries who had received SSDI for a short duration. To do this, the evaluation team set a goal of having at least 50 percent of volunteers be short-duration recipients. Because only 32 percent of SSDI-only beneficiaries overall fall into this subpopulation, this goal was accomplished by oversampling short-duration beneficiaries into the Stage 2 outreach waves. Short-duration beneficiaries make up 64 percent of all Stage 2 subjects.⁶

The analysis of Stage 2 recruitment showed that beneficiaries who volunteered to enroll in Stage 2 differed from non-volunteers in several key characteristics. Women volunteered for the demonstration at a higher rate than men and volunteers tended to be younger than non-volunteers, with a mean age of 47.6 years vs. 49.2 years for non-volunteers. Volunteers were generally representative of the solicitation pool in terms of primary impairment, although it is notable that people with mental disorders were slightly more likely than other beneficiaries to volunteer. Beneficiaries with short-duration SSDI receipt (defined here as having received benefits for 36 months or less) volunteered at higher rates than those with longer SSDI receipt, resulting in a mean duration among volunteers of 53.4 months compared to 73.2 months among non-volunteers. While monthly SSDI benefit amounts were similar between volunteers and non-volunteers, disabled adult children were less likely to volunteer as were those who had a representative payee.

Gubits et al. (2013) reported on use of BOND counseling through December 2012. The analysis showed much higher participation in EWIC compared to WIC. Of T22 subjects, 97 percent had received some EWIC services by the end of December 2012. In contrast, 31 percent of T21s enrolled in 2011 had received WIC services, and 26 percent of T21s enrolled in 2012 had participated in WIC. Derr et al. (2015) analyzed participation in WIC and EWIC through the end of 2013 and found similar contrasts. Derr et al. (2015) also compared WIC participation by baseline employment status and found that a higher proportion (45 percent) of T21s who were working at baseline had received WIC compared to either T21s who were looking for work at baseline (37 percent had received WIC), or T21s who were not working and not looking for work at baseline (24 percent had received WIC). The current report presents additional findings on beneficiary engagement in work and use of demonstration services.

⁶ This percentage reflects two factors. First, of the beneficiaries solicited to volunteer for Stage 2, 53 percent came from the short-duration subpopulation, oversampling by a factor of 1.68. Short-duration beneficiaries were also more likely to volunteer once solicited: 6.4 percent did so compared to 4.2 percent of long-duration beneficiaries.

1.2.3. The Impact Analysis

The impact analysis for Stage 2 includes three pairwise comparisons which serve to address three research questions:

Research Question	Addressed by Comparison of
A. What is the <i>impact of the benefit offset on SSDI-only beneficiaries who volunteer for BOND</i> , compared to current law?	T21 to C2
B. What is the <i>impact of the benefit offset plus enhanced work incentives counseling</i> on SSDI-only beneficiaries who volunteer for BOND, compared to current law?	T22 to C2
C. What is the <i>incremental effect of enhanced work incentives counseling when added to the benefit offset</i> (i.e., EWIC vs. WIC), for SSDI-only beneficiaries who volunteer for BOND?	T22 to T21

In addition to the benefit offset, WIC and EWIC, and the change to an annual accounting period, some administrative differences may have influenced impact estimates for the first two research questions in the early years of the BOND project (including 2012). The administrative procedures established to provide T21 and T22 subjects with information and to implement benefit adjustments under the offset likely affected the speed with which retroactive payment adjustments were made and improper past payments recovered. Because of how they are measured, these adjustments are especially important for the estimated impacts on benefits paid. By necessity, the impact estimates in this document focus on benefits paid *in* 2013 rather than impacts on benefits paid *for* 2013 after all retroactive benefit adjustments and repayments of improper payments have been completed. The latter are not observed in the data available for this report and might be quite different. Future reports will include estimates of the impact of BOND on benefits paid *for* the years in the evaluation period.

Finally, for the T21 to C2 comparison, T21 subjects have access to counseling services that are tailored to the benefit offset but are otherwise intended to be comparable to counseling services available to all beneficiaries under current law and hence offered to C1 subjects. Although not intended, it is possible that in the implementation of the counseling services offered to C2 and T21 subjects differ in ways that have an impact on earnings and benefits above and beyond the impact of the offset itself.

Gubits et al. (2014) reported findings from the analysis of impacts through the second calendar year of Stage 2 implementation (2012), which is the first year in which the offset and its work incentives counseling components could have affected outcomes for all Stage 2 subjects. Due to methodological issues that were unresolved at the time of report release, Gubits et al. (2014) provided estimates that were generalizable to the 10 study sites only, but not to the entire nation. As is discussed in Appendix A of this report, these methodological issues have been resolved. Appendix C of this report presents Stage 2 impacts on earnings and benefit outcomes in 2012 that are representative of those who would have volunteered for Stage 2 in the national beneficiary population. As a primary goal of the BOND evaluation is to provide nationally-representative estimates, the estimates presented in Appendix C supersede those presented in Gubits et al. (2014). Using the improved, national-representative methodology, those impact estimates show that:

- There is no confirmatory evidence that the offset rules combined with standard work incentives counseling (WIC) increased either mean earnings or SSDI benefits paid in 2012, compared to current law earnings rules and similar counseling services. There is suggestive evidence (from a non-confirmatory test) that the offset-plus-WIC intervention increased the proportions of beneficiaries with earnings above the BOND Yearly Amount (BYA).
- There is no confirmatory evidence that the offset rules combined with enhanced work incentives counseling (EWIC) increased either mean earnings and or SSDI benefits paid in 2012, compared to current law earnings rules and standard counseling services.
- EWIC, when combined with the offset, did not have any detectable effects on 2012 earnings and benefit outcomes beyond those produced by the offset-plus-WIC intervention.

1.3. The Current Report

The current report documents findings from the Stage 2 process, participation, and impact analyses. This section describes the data sources used in the report.

1.3.1. Data Sources for the Process Analysis

The process analysis involves seven rounds of data collection activities over the course of the demonstration and relies on multiple sources. This report uses information collected in the first six rounds of process study data collection covering BOND implementation through 2014. The process analysis team has collected data from beneficiary focus groups conducted during site visits to the BOND sites, focus group discussions with WIC and EWIC providers, and interviews with the BOND implementation team at SSA, Abt Associates, Mathematica Policy Research, and implementation partners. The process study team also used administrative data from BODS on the delivery of BOND services and beneficiary status and survey data from the Stage 2 baseline survey to characterize BOND beneficiaries. This section provides an overview of the qualitative data collection activities.

During the fifth round of process analysis data collection in 2013, the process analysis team conducted focus groups with beneficiaries in conjunction with site visits to each BOND site. In each site, the team conducted one focus group with WIC users (T1 and T21 subjects) and one group with EWIC users (T22 subjects), for a total of 10 WIC and 10 EWIC groups. As is standard with this mode of data collection, the participants were not intended to be representative of all BOND treatment group subjects who used counseling services; instead, they were selected in a manner that seemed likely to efficiently yield useful information about participants' demonstration experiences from their perspective. For convenience, the team held the focus groups in the area of each site with the highest concentration of beneficiaries. The process study team obtained a list of beneficiaries who had used some BOND services and lived within 20 miles of the focus group location. Beneficiaries were contacted by telephone; those who agreed to participate received a follow-up letter with the location, date, and time of the session. Beneficiaries who attended the group received \$25 in cash for their time. Participants answered semi-structured questions on a few key topics to encourage interaction and discussion.

In the fall of 2014, the study team conducted the sixth round of qualitative data collection. This sixth round of data collection consisted of two main activities: 1) telephone/online focus groups with WIC and EWIC supervisors and staff during October 2014, and 2) telephone interviews with key members of the BOND Implementation Team and SSA BOND Operations team in January 2015. Data collection focused

on documenting the changes made to BOND since fall 2013 data collection and staff perceptions of those changes. Data collection was also part of an ongoing effort to capture the successes, challenges, and lessons in implementing BOND and how they might influence the impact of the demonstration.

Telephone focus groups. The process analysis team conducted 10 telephone/online focus groups with 51 WIC and EWIC supervisors and staff (Exhibit 1-4)⁷. The team organized the groups to collect data separately from sites in which post-entitlement responsibilities had been centralized versus those in which they were not centralized, from supervisors versus staff, and from those involved in WIC services versus EWIC services. To identify potential participants the process analysis team asked the BOND Implementation Team to recommend WIC and EWIC supervisors and staff who represented each staff and supervisor role in each site. To recruit participants the team sent an email invitation two to four weeks ahead of the scheduled time, followed by a telephone call to non-responders, and an email reminder one to two days ahead of the meeting.

During each 90-minute focus group, trained facilitators led the telephone discussion using protocols and ran an online polling feature to capture answers to multiple choice questions. The focus group topics included: staff or supervisor education and experience; context and service environment; BOND organizational and staffing infrastructure; roles and responsibilities of WIC and EWIC staff; unengaged T22 subjects (EWIC groups only); perception of centralization (centralized sites only), perception of post-entitlement responsibilities (non-centralized sites only); improper payments; influence of BOND on beneficiary behavior; and successes and challenges. During the focus groups, the facilitators invited (but did not require) participants to respond to an on-line, multiple-choice poll. The process evaluation team analyzed the multiple choice poll responses across all focus groups and identified themes within each focus group.

Telephone interviews with SSA and BOND Implementation Team. In January 2015, the Evaluation Team conducted 5 telephone interviews with 10 key members of the BOND Implementation and Operations Teams. We selected team members most familiar with BOND processes, changes to processes, and the reasons for those changes, and completed interviews with all identified individuals.

During each 60-minute telephone interview, interviewers used a protocol tailored to the role of each respondent. The interviews focused on clarifying the information discussed during the focus groups and identifying key changes to implementation. Across all five interviews, the topics of discussion included: the BOND service environment; WIC and EWIC services; WIC and EWIC organizational performance; unengaged T22 subjects; work CDRs; centralization of post-entitlement work; AEE and EOYR mailers; and benefit adjustments and improper payments. After each interview, the Evaluation team identified the information and clarifications most relevant to understanding the impact of the demonstration.

⁷ The 18 WIC non-supervisory staff who participated in focus groups represented 39 percent of the total 46 staff at all sites and the 14 EWIC non-supervisory staff who participated represented 49 percent of the total 29 staff at all sites.

Exhibit 1-4. Qualitative Data Collection – Fall 2014 Focus Groups

	Focus Group	Sites Represented	Number of Focus Group Participants	Live Poll Respondents
Supervisors				
1	WIC Centralized Group (supervisors)	AZ, CO, DC, NNE, SFL, WI	8	7
2	WIC Non-Centralized (supervisors)	GD, WNY	2	2
3	EWIC Centralized (supervisors, pilot)	AZ, CO, GD, GH, SFL	5	5
4	EWIC Non-Centralized (supervisors)	AL, DC, WI, WNY	4	3
Staff				
5	WIC Centralized Group 1 (staff, pilot)	AZ, GH, NNE	3	3
6	WIC Centralized Group 2 (staff)	DC, NNE, SFL, WI	10	10
7	WIC Non-Centralized (staff)	AL, GD, WNY	5	5
8	EWIC Centralized Group 1 (staff)	AZ, CO	3	3
9	EWIC Centralized Group 2 (staff)	NNE, SFL	5	5
10	EWIC Non-Centralized (staff)	AL, DC, WI, WNY	6	6
Total	10 groups*	All BOND sites	51	49

*In addition, we held five calls with individuals who were unable to attend the scheduled focus groups, where the absence left a gap in representation for a particular site and role. Specifically, these calls were with: Colorado WIC staff, Northern New England WIC supervisor, Western New York WIC staff, Greater Houston WIC supervisor, Greater Houston EWIC staff. As of November 2014 BOND had a total of 46 WIC counselors and 29 EWIC counselors.

1.3.2. Data Sources for the Participation Analysis

The participation analysis relies on operations data and information from two beneficiary surveys.

BOND Operations Data System (BODS) and Beneficiary Tracking System (BTS) include data documenting outreach and recruitment, random assignment, participation in the BOND counseling interventions (WIC and EWIC); and information provided by SSA administrative systems on the number of TWP months completed, and whether SSA has determined that disability has ceased because of work above SGA. These variables are used to identify BOND subjects who may be eligible for the benefit offset.

BOND baseline survey data were collected during Stage 2 enrollment and prior to random assignment. The baseline survey provides information on background characteristics of beneficiaries that are not measured well in the administrative records, for virtually all volunteers who participated in the Stage 2 experiment. The baseline survey variables include recent employment history, current income, contextual information on demographic and family status, and an assessment of the subject's current understanding of SSDI work incentives and attitudes toward the demonstration. The evaluation team uses these data to form subgroups for separate analysis, to provide covariates in the impact analyses, and to help adjust for non-response to the follow-up surveys.

The BOND interim survey was collected approximately one year after Stage 2 random assignment (median 13 months). The interim survey was conducted by phone with field follow up and collected information on service receipt in the first year after Stage 2 random assignment to establish the

treatment/control differential in service receipt and to measure the extent to which Stage 2 treatment group members participated in employment-related services. The interim survey also collected information on employment and earnings at the time of the survey, receipt of education and training, health and functional status, awareness of BOND or SSA work incentives, and demographic information. The study team attempted to complete the interim survey with all Stage 2 subjects and achieved an 84 percent response rate.

1.3.3. Data Sources for the Impact Analysis

For the impact analysis documented in this report, administrative data for calculating earnings and benefit impacts were available through calendar year 2013. Earnings are measured from the SSA Master Earnings File (MEF), which contains longitudinal information on wages and self-employment income reported to the IRS. SSA considers the MEF records almost 100 percent complete for calendar year 2013 when SSA extracted them for this report in February 2015.⁸ Benefit outcomes are measured from SSA's Payment History Update System (PHUS) for SSDI.

1.4. Organization of the Report

The remainder of the report is structured in seven chapters. Chapter 2 updates results from the 2013 process analysis about the BOND study sites and disability service environment.

Using data from BTS and the Stage 2 interim survey, Chapter 3 describes benefits counseling in BOND and updates process and participation analysis presented in previous reports. The analysis explores the content of benefits counseling and the extent to which Stage 2 treatment subjects received benefits counseling.

Chapter 4 presents findings for the participation analysis. The chapter analyzes data from the Stage 2 interim survey to assess awareness of BOND and understanding of how earnings affect SSDI benefits among the Stage 2 subjects.

Chapter 5 updates the process and participation analysis regarding use of the benefit offset among Stage 2 subjects since the 2013 process analysis. The chapter also analyzes impacts on offset use for T22 subjects compared to T21 subjects and steps taken by beneficiaries toward use of the benefit offset.

Chapters 6 presents the methodology used in the impact analysis. Chapters 7 and 8 present findings from the Stage 2 impact analysis. Chapter 7 focuses on impacts on 2013 earnings and SSDI benefits measured from SSA administrative data. Chapter 8 presents impact estimates on additional employment, health, and marital status outcomes measured in the survey.

Chapter 9 provides a summary of key findings and conclusions.

⁸ Because the data are collected by the IRS and are therefore subject to IRS access rules, SSA staff have direct access to MEF data, but contractors do not. Consequently, qualified SSA staff accessed the data, submitted programs developed by the BOND team to estimate impacts, reviewed output to ensure that it complied with privacy requirements, and then transmitted the output to the evaluation team. The MEF earnings data are updated annually. The 2013 earnings data for this report were extracted in February 2015.

Three appendices provide additional technical details. Appendix A provides a full description of the impact estimation methodology and Appendix B provides impact estimates for all of the subgroup analyses we conducted. Appendix C provides updated 2011 and 2012 earnings and benefits impacts that are generalizable to the national population of would-be volunteers.

2. BOND Study Sites and Disability Service Environment

2.1. BOND Study Sites

BOND was designed to produce nationally representative estimates of the impact of the benefit offset. Towards that goal, the 10 randomly selected BOND sites reflect national variation in their environments. Although the evaluation does not estimate site-specific impacts, understanding site-level variation in the environment, and how site environments change during BOND, provides useful context for interpreting study findings and will inform any future implementation of a national program.

The sites differ in six salient ways: (1) geographic characteristics; (2) economic indicators; (3) number and staffing configuration of BOND benefits counseling providers; (4) number of BOND Stage 2 treatment subjects enrolled; (5) availability of employment services and other work-focused, disability-related resources; and (6) presence of non-BOND SSDI benefits counseling services. These dimensions are summarized in Exhibit 2-1, and discussed in greater detail in previous reports (Derr et al. 2015, Gubits et al. 2013). The following sections provide a summary of this information and, where relevant, describe changes and new observations.

2.2. Geographic Characteristics

Sites vary in terms of the number of states and communities included in their catchment areas, population density, and dispersion of SSDI beneficiaries. Service delivery is more complex in sites where providers must understand and navigate multiple sets of state and community policies and resources, and tailor service delivery accordingly. For example, the four-state Northern New England site relies on four different state VR agencies to provide services to beneficiaries. In addition, benefits counseling staff stated that beneficiaries in rural areas may face challenges regarding access to jobs and employment support services.

2.3. Economic Indicators

In two ways, the relative strength of the local economic environment may affect beneficiaries' opportunities to engage in SGA, a necessary step toward using the BOND offset. First, if there are few job openings, it may be difficult for individuals with disabilities to find employment. Evidence suggests that, while it is typically more difficult to find a job during periods of high unemployment, individuals with disabilities have an even more difficult time finding jobs than those without a disability (Livermore et al. 2012). Second, in a weak economy, declines in state revenues result in less funding to support services for people with disabilities.

The unemployment rate is the conventional indicator of the strength of the local economy. However, the employment rate provides a more accurate picture of opportunities for work among people with disabilities because it reflects job activity among all potential workers. In contrast, the unemployment rate omits discouraged workers who have stopped looking for work, and many people with disabilities are in this group. As a result, the employment rate tends to fluctuate more than the unemployment rate over the business cycle, and that is especially true for people with disabilities.

Leading up to the start of BOND enrollment in 2011, the national employment rate among people without disabilities had fallen from 75.0 percent in 2007 (before the 2008 recession) to 72.8 percent in 2011, a 1.9

percent decline.⁹ For people with disabilities, the national employment rate had fallen from 36.2 percent to 32.6 percent, a substantially larger relative decline: 9.9 percent. These changes were mirrored in the mean rates across the 15 states represented in the 10 BOND sites: from 76.3 percent for people without disabilities in 2007 to 74.5 percent in 2011, a 2.4 percent decline, and from 38.1 percent to 34.9 percent for those with disabilities, an 8.3 percent decline.

From 2011 to 2013, the period of analysis for this report, the national employment rate increased for people with and without disabilities: to 74.2 percent for people without disabilities, a 1.9 percent increase, and to 33.9 percent for people with disabilities, a 4.0 percent increase (Exhibit 2-2). The average rates in the states represented in the BOND sites also increased, by similar amounts: to 75.6 percent for people without disabilities, a 1.5 percent increase, and to 35.6 percent for people with disabilities, a 4.4 percent increase.

During this period, the change in state-level employment rates among people with disabilities varied across the 15 states in the BOND sites. The employment rate for people with disabilities fell during this period in two states: by 8.0 percent and 0.6 percent in Vermont and Maine, respectively, both in the Northern New England site. The remaining 13 states all experienced increases in the employment rate for people with disabilities, with the highest percentage increases in the two Northern New England States (New Hampshire, 13.6 percent, and Massachusetts, 10.1 percent) and the District of Columbia (13.0 percent).¹⁰

Similarly, the 2013 state-level employment rates for people with disabilities also varied across the 15 states included in the BOND sites. In 2013, six of the 10 BOND sites included at least one state with an employment rate for people with disabilities lower than the national average employment rate for people with disabilities. Wyoming, part of the Colorado/Wyoming site, had the highest employment rate among people with disabilities, at 50.7 percent, while the lowest employment rate was 27.1 percent in Alabama.

2.4. Number and Staffing Configuration of BOND Benefits Counseling Providers

To deliver BOND WIC and EWIC services, the BOND Implementation Team contracted with local providers already engaged in disability service delivery. Cross-site variation in available providers and geographic coverage areas across sites led to cross-site variation in BOND provider arrangements. As detailed in Exhibit 2-1, the arrangements varied with respect to the number of providers in the sites, the type of provider organization (non-profit agency, state vocational rehabilitation (VR) agency, educational institution, other state agency, advocacy organization, or provider association) and the providers' staffing models (dispersed, in which staff allocate a portion of their time to BOND, versus consolidated, in which most staff devote all of their time to BOND).

Cross-site variation is reflected in providers' flexibility to accommodate reductions in FTEs over the course of the demonstration, the need for coordination and oversight, counselor knowledge of local systems, accessibility of services to beneficiaries, and currency of skills and trainings.

⁹ See notes to Exhibit 2-2 for sources.

¹⁰ Some variation in state-level estimates is expected due to sampling error.

- Providers' staffing arrangements and overall size affected their flexibility to respond to planned demonstration reductions in FTEs. Because they had more options when reassigning staff to non-BOND work when their hours on BOND were reduced or eliminated, larger providers, such as state VR agencies, were better positioned to accommodate reductions in FTEs. Similarly, sites with dispersed staffing structures had greater flexibility to accommodate changes because multiple staff shared part-time BOND counseling roles along with work from other funding sources.
- The number of providers and their staffing arrangements affected the proximity and content of services to beneficiaries. Sites that covered larger geographic areas, especially across multiple states, were more likely to have multiple providers or dispersed staffing structures. These arrangements placed counselors closer to beneficiaries across the site and retained counselors with knowledge of local resources.
- Sites with more providers and dispersed staffing structures required greater coordination and oversight relative to sites with fewer providers or more consolidated staffing structures.
- Provider and staffing configurations also affected counselors' ability to stay current on skills and related training. Staff in sites with fewer providers and more consolidated staffing structures found it easier to consult with their on-site colleagues for support, complete trainings, build expertise and otherwise stay up-to-date on BOND policies and procedures. These factors in turn may have affected the quality of post-entitlement work,¹¹ such as calculating Annual Earnings Estimates. A review by the BOND Implementation Team found that, relative to WIC providers with a consolidated staffing model, WIC providers with a dispersed staffing model made more errors in BOND post-entitlement work.

In addition, two recent changes affected BOND providers. The first change was a reduction in FTEs for counseling staff. The Implementation Team had planned this reduction in anticipation of smaller workloads for WIC and EWIC staff as the demonstration proceeded. Second, to improve the quality of post-entitlement work, in December 2013, the BOND Implementation Team shifted the majority of post-entitlement work to a centralized team. Centralization was implemented in Arizona/Southeastern California (EWIC and WIC), Colorado/Wyoming (EWIC and WIC), DC Metro (WIC), Greater Detroit (EWIC), Greater Houston (EWIC and WIC), Northern New England (EWIC and WIC), South Florida (EWIC and WIC), and Wisconsin (WIC) in December 2013, and in Alabama (WIC) in January 2015. The implications of these changes for WIC and EWIC services are discussed in Chapter 3.

2.5. Number of BOND Stage 2 treatment subjects enrolled

Across the sites, the number of BOND subjects in the Stage 2 treatment and control groups varies. As previously reported (Derr et al., 2015), the magnitude of this variation is approximately proportional to the number of BOND-eligible beneficiaries when enrollment began in May 2011. The highest enrollment in the Stage 2 treatment groups is 1,064 subjects in South Florida; the lowest are 639 and 641 in the DC metro and Colorado/Wyoming sites, respectively (see Exhibit 2-1 for the number of Stage 2 BOND subjects by site).

¹¹ Post-entitlement work is discussed in detail in Chapter 3. It refers to activities required to facilitate the BOND benefit adjustment process.

2.6. Availability and Use of Employment Services and Other Work-Focused, Disability-Related Resources

To engage in SGA and use the offset, some beneficiaries require referrals to employment service providers. These service providers include state VR agencies and other providers acting as employment networks (ENs) under SSA's Ticket to Work program. WIC and EWIC counselors can refer BOND subjects to such providers, just as work-incentive counselors serving control subjects might do for SSDI beneficiaries under current law. For example, a counselor might refer a beneficiary in need of career counseling or assistive technology. However, WIC and EWIC counselors have reported that these resources have not been consistently available to BOND treatment or control subjects, due to waiting lists at VR agencies, the number of local ENs, and variations in the quality of services, all of which may impact the timing and extent of offset use (Section 2.6 of Derr et al. 2015).¹² These factors are all external to BOND.

In recent focus groups, WIC and EWIC staff and their supervisors were asked to comment on the availability and quality of employment support services. In an online poll,¹³ nearly half of supervisors (47 percent) reported a decrease in non-BOND funding of disability-related services in the past year. Most staff (81 percent) reported that they usually refer beneficiaries to VR agencies, but 68 percent also said they refer beneficiaries to other ENs on occasion. In discussion, participants explained that other ENs were not a primary referral source due to limited local presence, and because few beneficiaries were ready to work full time—a goal many ENs require of the clients they serve. Overall, 43 percent reported that it was difficult or very difficult to get beneficiaries connected with employment services, while 35 percent described service access as very or somewhat easy. When considering beneficiaries who had received services, 38 percent said they met beneficiaries' needs well or very well, while 35 percent reported that services did not meet the needs of BOND beneficiaries (they did so "poorly" or "very poorly").

2.7. Non-BOND SSDI Benefits Counseling Services

The WIPA program provides benefits counseling for SSDI beneficiaries more broadly, including the BOND control group. SSA suspended funding for WIPA when authority for the program ended in June 2012, but most BOND sites maintained some level of counseling services for control subjects until SSA was able to reinstate the program in August 2013 (see Section 1.2.1 and Derr et al. 2015). The lapse of WIPA funding did not affect funding for WIC and EWIC services to treatment subjects, but did lead to WIC and EWIC staffing changes in six of the ten BOND sites (see Exhibit 2-3 in Derr et al. 2015), because WIC and EWIC services were being provided by WIPA grantees that had to reconfigure their staffing after the loss of WIPA funding.

According to a member of the BOND Implementation Team, the later reinstatement of WIPA may have helped provider organizations adjust to reductions in BOND full-time equivalent (FTE) positions, which

¹² See Honeycutt and Stapleton (2013) for more information on wait times for SSDI beneficiaries at state VR agencies and evidence that long VR wait times for beneficiaries have a negative impact on their employment and benefit outcomes.

¹³ Participants were invited (but not required) to respond to an on-line, multiple-choice poll during the focus group. For the question directed only to supervisors, 17 of 19 responded. For the remaining questions described, the number of WIC and EWIC counselors and supervisors who responded to the poll ranged from 37 to 43, of 51 possible respondents.

took place around the same time. WIPA funding allowed provider organizations to pay for staff time that was no longer reserved for BOND-related duties.

2.8. Conclusion

Variation in the implementation of BOND across sites has been present throughout the BOND demonstration, and reflects the underlying diversity of the site environments. This same diversity would cause variation in the implementation of a national program.

In addition, several contextual changes and cross-site themes may help inform the results of the BOND impact evaluation. First, employment rates have improved since the start of the demonstration. As a result, beneficiaries' opportunities to use the offset may be increasing over time. Second, the availability and quality of employment support services for BOND beneficiaries remain inconsistent across sites, and pose a significant challenge to work and use of the offset for some beneficiaries. There is no indication that either the economic environment or availability and quality of service referrals was different for treatment subjects than for control subjects, although the 14-month interruption in funding for WIPA likely had some impact on the availability of counseling services (including referrals made by counselors) to some control subjects during that period.

Exhibit 2-1. BOND Site-Level Characteristics

Sites	Geographic Characteristics			Number of Stage 2 BOND Subjects ¹			BOND Benefits Counseling Providers			Centralized Post-Entitlement Process
	Number of States	Population Density ²	Geographically Dispersed ³	T21	T22	C2	Number	Types	Dispersed Staffing	
Alabama	Single	94 (AL)	X	500	314	499	1	• Nonprofit		X (WIC)
Arizona/SE California	Multiple (1 full, 1 partial)	56 (AZ) 239 (CA)		625	386	622	2*	• Nonprofit		X
Colorado/Wyoming	Multiple (2)	49 (CO) 5.8 (WY)	X	395	246	391	2	• Nonprofit • State VR	X (WIC)	X
DC Metro	Multiple (1 full, 3 partial)	9,856 (DC) 203 (VA) 595 (MD) 77 (WV)		393	246	393	3*	• Nonprofit • Other ⁴		X (WIC)
Greater Detroit	Partial	175 (MI)		444	279	444	1	• Nonprofit		X (EWIC)
Greater Houston	Partial	96 (TX)		421	262	419	1	• Nonprofit	X (WIC)	X
Northern New England	Multiple (3 full, 1 partial)	147 (NH) 43 (ME) 839 (MA) 68 (VT)	X	462	291	463	5	• Nonprofit • State VR • University • Med Center	X (ME, WIC; VT WIC)	X
South Florida	Partial	96 (FL)		654	410	654	1	• Nonprofit		X
Western New York	Partial	411 (NY)		463	292	464	5*	• Nonprofit • Advocacy organization	X (WIC)	
Wisconsin	Partial	105 (WI)	X	579	363	580	8	• Nonprofit • State Health Agency • University • For-profit	X (EWIC, WIC)	X (WIC)
U.S. Average	N/A	87	N/A	4,936	3,089	4,929	N/A	N/A	N/A	N/A

Source: Based on staff interviews from BOND site visits.

* Indicates sites that rely on Virginia Commonwealth University to provide telephonic EWIC services to some or all T22s, depending on the site.

¹ The target sample for the Stage 2 volunteer pool was 12,600 subjects, with 4,800 T21 subjects, 3,000 T22 subjects, and 4,800 C2 subjects.

² Population density indicates number of individuals per square mile of land, 2010.

³ Geographic dispersion defined as 20 percent of the SSDI population living outside of the Metropolitan Statistical Area (MSA). Based on findings from “Social Security Administration \$1 for \$2 Benefit Offset Demonstration: Site Visit Report” (September 2008).

⁴ Association of disability service providers.

N/A = Not applicable.

Exhibit 2-2. BOND Site-Level Employment Rates

Sites	State for Employment Rate	Employment Rate for People without Disabilities, ages 18-64 (%)			Employment Rate for People with Disabilities, ages 18-64 (%)		
		2011	2013	Percent Change	2011	2013	Percent Change
Alabama	Alabama	70.2	70.5	0.4	26.5	27.1	2.3
Arizona/SE California	Arizona	69.9	71.3	2.0	32.8	33.6	2.4
	California	69.5	71.1	2.3	31.4	32.7	4.1
Colorado/Wyoming	Colorado	76.3	77.3	1.3	41.4	42.3	2.2
	Wyoming	80.1	79.4	-0.9	47.8	50.7	6.1
DC Metro	District of Columbia	71.5	73.5	2.8	30.0	33.9	13.0
Greater Detroit	Michigan	70.2	73.4	4.6	28.9	29.9	3.5
Greater Houston	Texas	73.5	74.7	1.6	36.9	38.7	4.9
Northern New England	Maine	78.1	78.8	0.9	31.4	31.2	-0.6
	Massachusetts	76.9	77.9	1.3	31.7	34.9	10.1
	New Hampshire	79.5	80.3	1.0	36.8	41.8	13.6
	Vermont	80.0	79.6	-0.5	36.2	33.3	-8.0
South Florida	Florida	70.6	72.2	2.3	29.2	30.5	4.5
Western New York	New York	72.1	73.3	1.7	31.3	32.2	2.9
Wisconsin	Wisconsin	78.7	80.1	1.8	38.7	40.9	5.7
Average in Site States	--	74.5	75.6	1.5	34.1	35.6	4.4
U.S.	--	72.8	74.2	1.9	32.6	33.9	4.0

Source: 2013 data are from the 2014 Annual Disability Statistics Compendium, based on data from U.S. Census Bureau, 2013 American Community Survey, American FactFinder, Table B18120; <<http://factfinder2.census.gov>>; (accessed 22 September 2014). 2011 data are from the 2012 Annual Disability Statistics Compendium, based on data from U.S. Census Bureau, 2011 American Community Survey, American FactFinder, Table B18120; <<http://factfinder2.census.gov>>; (accessed 24 September 2012).

3. BOND Benefits Counseling

Benefits counseling is a key component of BOND. The intention of the counseling developed for BOND is to enable beneficiaries to understand and take advantage of the offset. BOND includes two types of counseling: (1) basic work incentives counseling (WIC), which is by design comparable to the counseling available under current law; and (2) enhanced work incentives counseling (EWIC). WIC staff serve beneficiaries assigned to treatment in Stage 1 (T1 subjects) and Stage 2 volunteers assigned to the T21 treatment group. EWIC staff serve only Stage 2 volunteers assigned to the T22 treatment group.

3.1. Design of BOND Counseling

All SSDI beneficiaries are eligible to receive benefits counseling from a Work Incentives, Planning, and Assistance (WIPA) project. SSA distributes available funds to 103 WIPA projects that provide counseling to SSDI beneficiaries on using work incentives to increase employment and earnings. WIPA projects also refer beneficiaries to employment support programs, such as state vocational rehabilitation agencies (SVRAs) or employment networks (ENs). The primary objective of WIPA is to equip beneficiaries to make the best use of work incentives to increase their employment.

The goal of WIC is to provide subjects in the offset-only treatment groups (T1 and T21) with counseling services that have the same overall content and intensity as WIPA agencies' information and advice to beneficiaries under the status quo. The only intended change, relative to the status quo, is that WIC providers advise subjects about the effects of earnings on benefits under the offset. The intent of the demonstration is to measure the impacts of the benefit offset when implemented with this minimal, necessary change in the nature of WIC.

Stage 2 of BOND includes a test of the effect of offering EWIC on the size of the impacts from introduction of the benefit offset. The T22 group receives EWIC, which is intended to differ from the regular WIC described above. Counselors that provide EWIC to T22 subjects do not provide WIC to other treatment subjects. EWIC services include the WIC services described above. The more intensive components of EWIC services include counselor outreach to routinely contact the beneficiary, the development of a detailed employment support plan based on assessments of vocational skills and interests, and assistance in helping beneficiaries obtain the resources and support they need to find employment, as well as the ongoing support they need to keep it. The primary difference between the services is that EWIC staff take a proactive approach to contacting beneficiaries on an on-going basis to inform them about demonstration services. EWIC staff were instructed to contact all T22 beneficiaries within two weeks of random assignment and contact them thereafter at least once per month over the course of BOND. The requirements for EWIC contact were modified in early 2014 after all T22 subjects had received at least 18 months of monthly contact. From that date forward, all engaged T22 subjects must receive at least quarterly contact from EWIC staff and only those likely to go into the offset must receive monthly contact. In contrast, WIC staff provide the same type of information, though only to beneficiaries who contact them.¹⁴ WIC staff do not conduct the assessments of vocational skills and interests or the employment support plan that are part of the EWIC design.

¹⁴ For additional details on the design of EWIC and WIC, the reader is referred to Section 5.1 of the Stage 2 Early Assessment Report (Gubits et al. 2013) and Section 5.2 of the BOND Final Design Report (Stapleton et al 2010).

In addition to benefits counseling, the WIC and EWIC staff also assist beneficiaries with reporting work to BOND, completing 820/821 CDR forms, and occasionally appealing the outcome of work CDR decisions. Staff at a limited number of sites also provide BOND post-entitlement services, defined as collecting information from the beneficiary to develop an annual earnings estimate (AEE), collecting and reviewing documentation for non-countable income (which SSA deducts from earnings to calculate benefits), and assisting beneficiaries to submit this information to SSA. Section 3.2. describes post-entitlement services.

3.2. BOND Post-Entitlement Services

In BOND, post-entitlement services are defined as activities conducted to collect estimates and documentation of beneficiaries' earnings and to submit them to SSA via the BOND data system.¹⁵ In December 2013, the centralized staff from the BOND Implementation Team assumed responsibility for providing post-entitlement services in several BOND sites selected for their improvement potential. Prior to centralization, WIC and EWIC counselors played this role. The completion of post-entitlement work has direct effects on monthly benefit amounts established under the offset. Many WIC and EWIC counselors found it difficult to perform post-entitlement work, citing challenges in mastering and staying abreast of procedural changes related to this work. After centralization, WIC and EWIC counselors continued to provide all benefits counseling services for beneficiaries (the counseling components have not changed) and to provide answers to basic questions pertaining to benefits. Staff difficulty in mastering post-entitlement work was a significant factor in the decision to centralize.

In focus group discussions, counselors at sites that have centralized post-entitlement work strongly favored the centralization, because it freed them to have more time for benefits counseling, but also noted some positive aspects to conducting the work themselves. Some counselors acknowledged that the experience gained by conducting post-entitlement work in the past has helped them to be better counselors. While counselors prefer to spend more time on counseling, in some cases centralization appears to have led beneficiaries to have confusion about the separation of tasks between role of the centralized staff and their counselor. Perhaps as a result, counselors report a loss of the beneficiaries' trust. In addition, counselors said they sometimes "feel blind" to the activities of the centralized staff, for example the counselors do not know the timeframes for processing, or when letters are sent out.

The Implementation Team has taken steps to promote communications in centralized sites. WICs and EWICs have real-time access to individual BTS records and can see activities completed by each team. The Implementation Team also developed and distributed a manual describing the responsibilities of each team member. Prior to and during initial implementation of the centralization, the Implementation Team held calls with each sites' WICs, EWICs, and centralized team member to clarify roles and to discuss specific cases.

¹⁵ Post-entitlement work involves helping the beneficiary calculate an annual earnings estimate (AEE). Exhibit 6-2 in the Stage 2 Early Assessment Report contains a worksheet used to calculate the annual earnings estimate (AEE). Post-entitlement work also involves documenting and substantiating evidence of non-countable income that should be deducted from earnings to calculate benefits. Non-countable income is used to appeal overpayment decisions from automated reconciliation for past years and is submitted ahead of automated reconciliation for SSA to take into account when processing the automated reconciliation.

Staff at non-centralized sites report that they spend the majority of their time focused on post-entitlement work, but that benefits counseling does not get overlooked because they incorporate it into the post-entitlement work. During focus group polling, approximately two-thirds of non-centralized staff favored maintaining control of post-entitlement work because it helps them understand post-entitlement processes and facilitates close relationships with beneficiaries.

Section 5.2 of this report offers further discussion of centralization and its impact on the accuracy of beneficiaries' annual earnings estimates (AEEs).

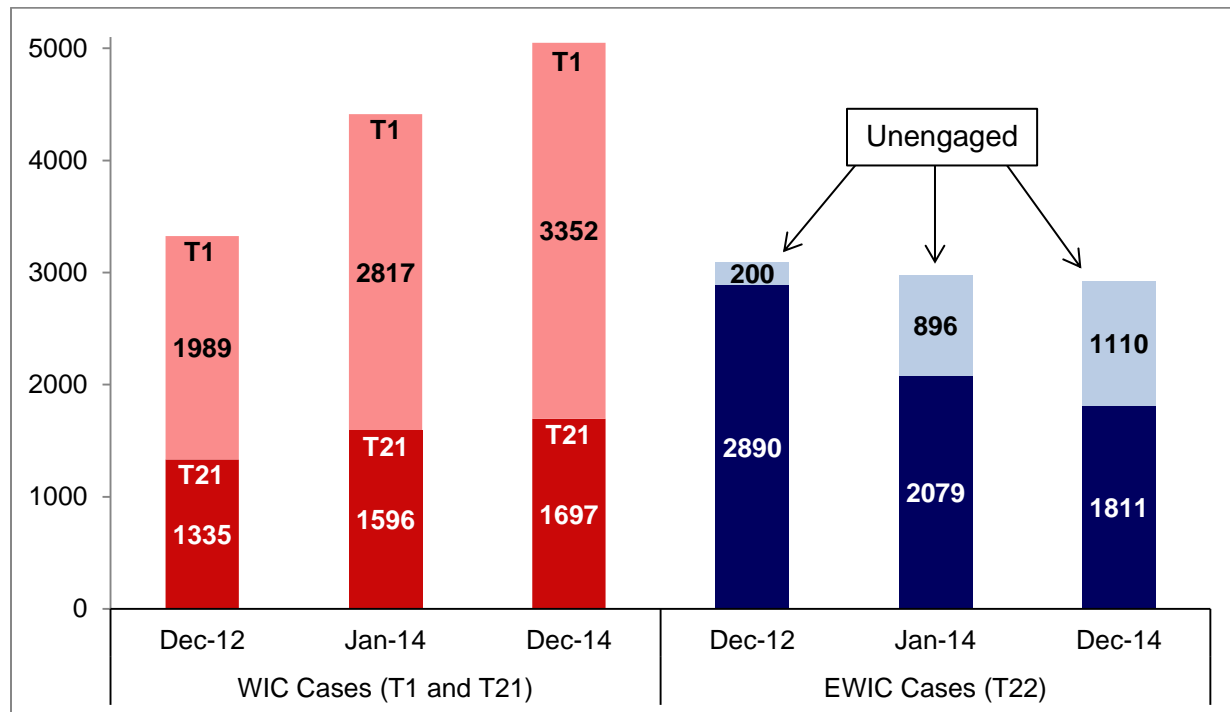
3.3. Caseloads

The design of WIC and EWIC services has important implications for the potential evolution of the caseload over the course of the demonstration. The EWIC caseload was determined by the number of subjects randomly assigned to T22; as a result, the maximum EWIC caseload was fixed as of September 2012. In contrast, the WIC caseload is determined by the motivation of treatment subjects to seek services, and can continue to increase because T1 and T21 subjects may choose to reach out to WIC staff for the first time at any point up until September 2017.

The number of WIC cases (T1 and T21 beneficiaries who have used services at some point) has grown, while the number of EWIC cases has not. This difference is by design because WIC beneficiaries (T1 and T21) have to seek information and referral (I&R) and/or counseling in order to enter the caseload. In contrast, the intention for T22 subjects is that they would be actively contacted by counselors. Exhibit 3-1 displays three snapshots of the total WIC and EWIC caseloads between December 2012 and December 2014. The WIC caseload is defined as the number of T1 and T21 beneficiaries for whom counselors have recorded a case note or outreach attempt at some point. The WIC caseload increased by about 14 percent, whereas the EWIC caseload has been relatively steady.

The cumulative count of beneficiaries in the WIC caseload has increased at every BOND site in 2014. Exhibit 3-2 lists the change in the total WIC caseload by site, ranging from a low of 11 percent in Greater Detroit to 26 percent in Colorado/Wyoming. The change in the T21-specific WIC caseload by site ranges from a low of 3.7 percent in Alabama to 18.8 percent in Colorado/Wyoming. Although the change in the T21 caseload has been generally smaller than the change in the T1 caseload, 27 percent of T21 beneficiaries were active in WIC by the end of December 2012 whereas 2.6 percent of T1 beneficiaries were active in WIC by the end of the December 2012.

Exhibit 3-1. WIC and EWIC Caseloads Over Time



Source: BTS data as of December 31, 2014.

Note: The number of WIC cases is the number of cases for which a counselor made a contact attempt or a case note.

Exhibit 3-2. WIC Caseload by Site in January 2014 and December 2014

	WIC Caseload (T1 and T21)			WIC Caseload (T21 only)		
	Jan-14	Dec-14	Change in Caseload (percentage)	Jan-14	Dec-14	Change in Caseload (percentage)
Alabama	357	420	17.6	109	113	3.7
Arizona/Southeastern California	522	590	13.0	202	213	5.4
Colorado/Wyoming	192	242	26.0	64	76	18.8
DC Metro	333	373	12.0	100	105	5.0
Greater Detroit	471	524	11.3	205	213	3.9
Greater Houston	343	408	19.0	118	126	6.8
Northern New England	568	645	13.6	166	180	8.4
South Florida	631	713	13.0	260	271	4.2
Western New York	425	485	14.1	170	186	9.4
Wisconsin	571	649	13.7	202	214	5.9
TOTAL	4,413	5,049	14.4	1596	1697	6.3

Source: BTS data as of January 16, 2014 (as reported in Derr et al. 2015) and December 31, 2014.

Note: The number of WIC cases is the number of cases for which a counselor made a contact attempt or a case note.

As expected and shown in Exhibit 3-1, the EWIC caseload did not increase after Stage 2 enrollment ended in September 2012. The EWIC caseload declined only slightly (1.8 percent), as cases would be removed only if a subject withdrew from BOND or became deceased. Another way to analyze the caseload is to look at the number of engaged T22 subjects. For EWIC staff, the number of engaged beneficiaries is the most accurate representation of the number of active clients, and all service benchmarks are defined for this group. The number of engaged T22 subjects declined from January 2014 to December 2014 by 12.9 percent. A counselor can designate a beneficiary as unengaged if the beneficiary is incarcerated, asks not to be contacted, is not responsive to repeated contact attempts, or is not interested in employment at this time. The unengaged T22 subjects receive two letters per year reminding them that they are in BOND and providing them with contact information of EWIC staff. If an unengaged subject is found to need a work CDR or AEE, either a WIC or EWIC or centralized post-entitlement team member will work with the beneficiary to obtain the AEE, despite the “unengaged” status. So even if a beneficiary is designated as unengaged from EWIC he or she will receive assistance to submit an AEE if one is needed. Moreover, “unengaged” is not a permanent status: unengaged subjects can return to engaged status whenever they request to do so by reaching out to an EWIC counselor.

As illustrated in Exhibit 3-1, the share of T22 subjects classified as unengaged increased from 30 percent in January 2014 to 38 percent in December 2014.¹⁶ Exhibit 3-3 displays the number of engaged EWIC cases at each BOND site. Although the share of engaged subjects decreased in all sites, changes were particularly large in some sites. Arizona/Southern California, for example, decreased its number of engaged subjects by 43.6 percent. The Implementation Team believes that sites have dissimilar unengagement rates regardless of guidance due to differences in agency philosophy. Some EWIC providers are reluctant to categorize beneficiaries as unengaged even if they are unresponsive to contacts for several months. Other providers believe that if a beneficiary is unresponsive to contacts for an extensive period, the unresponsiveness indicates unwillingness to engage with EWIC.

The ratio of subjects requiring service to the number of counselors can affect the quality and intensity of benefits counseling. If an EWIC counselor is overstretched, he/she might find it difficult to provide the enhanced services as designed. Conversely, if a WIC counselor has a very small caseload, a counselor’s inclination might be to provide services beyond the scope of WIC benefits counseling. To estimate this counseling “burden”, we examined the number of T1, T21, and T22 cases for which a counselor made a case note or an outreach attempt any time during the 2014 calendar year. This measure is an imperfect estimate of counseling burden because it does not capture the intensity of each contact, but data on counseling time spent with each beneficiary is not available. On average, the number of EWIC cases per FTE was very similar to the number of WIC cases per FTE in December 2014. As shown in Exhibit 3-4, WIC staff served an average of 122 beneficiaries per FTE, and EWIC staff served 109 beneficiaries per FTE. The number of EWIC cases-per-counselor was intended to be lower than WIC cases-per-counselor because EWIC staff are expected to spend more time with individual clients. The number of WIC cases-per-counselor was larger than EWIC cases-per-counselor in six of the ten sites (Exhibit 3-4).

¹⁶ The share of T22 subjects classified as unengaged increased from 30 percent in January 2014 ($896/2975 = 0.30$) to 38 percent in December 2014 ($1110/2921 = 0.38$).

Exhibit 3-3. Engaged EWIC Cases in January 2014 and December 2014.

	Engaged EWIC Cases		Change in Caseload (percentage)
	Jan-14	Dec-14	
Alabama	139	104	-25.2
Arizona/Southeastern California	195	110	-43.6
Colorado/Wyoming	210	192	-8.6
DC Metro	197	187	-5.1
Greater Detroit	220	216	-1.8
Greater Houston	155	151	-2.6
Northern New England	221	203	-8.1
South Florida	286	250	-12.6
Western New York	231	221	-4.3
Wisconsin	225	177	-21.3
TOTAL	2,079	1,811	-12.9

Source: BTS data as of January 16, 2014 (as reported in Derr et al. 2015) and December 31, 2014.

Exhibit 3-4. WIC and EWIC Caseloads per FTE

	WIC Cases per FTE 2014	EWIC Cases per FTE 2014	Ratio of EWIC-Cases-per- FTE to WIC-Cases-per- FTE
Alabama	73	97	1.33
Arizona/Southeastern California	146	108	0.73
Colorado/Wyoming	94	116	1.24
DC Metro	86	104	1.21
Greater Detroit	139	118	0.85
Greater Houston	140	125	0.89
Northern New England	128	81	0.63
South Florida	128	119	0.93
Western New York	129	140	1.08
Wisconsin	157	85	0.54
AVERAGE	122	109	0.94

Source: BTS data as of December 31, 2014.

Note: Staff FTE allocations are for 2014. T1, T21, and T22 subjects are classified as a “case” if a counselor recorded a case note or contact attempt for that subject at any point in 2014.

As reported in Gubits et al. (2013) and Derr et al. (2015), significant variations remain in caseload size across sites. As of December 2014, the size of WIC caseloads ranged from 73 to 157 subjects per FTE, whereas the number of engaged EWIC cases per FTE ranged from 81 to 140 subjects. The ratio of EWIC to WIC caseloads varied significantly across sites, from Alabama, where the EWIC caseload was 33 percent larger than the WIC caseload, to Wisconsin, where the EWIC caseload was about half as large as the WIC caseload. Some of this variation reflects variation in the application of site-specific criteria for classifying a T22 subject as unengaged. Site visits also revealed significant variation in caseload size within a site for each type of counselor, primarily due to differences in counselors' tenure and geographic location.

The comparison of WIC and EWIC caseloads is complicated by several factors. By design, all EWIC staff are required to initiate monthly contact with all T22 subjects assigned to them. In contrast, WIC caseloads are composed of T21 and T1 subjects who proactively contact WIC counselors. Moreover, EWIC counselors, by design, engage in different types of counseling activities with T22 subjects than WIC counselors engage with T21 and T1 subjects. The next section analyzes the impact of random assignment to T21 or T22 on various forms of counseling receipt.

3.4. Analysis of Counseling Receipt

The BOND Beneficiary Tracking System (BTS) contains records on the dates and services delivered to T21 and T22 subjects. This section compares counseling receipt of T21 and T22 subjects, but not C2 subjects because administrative data on WIPA participants is tracked separately (outside of BOND). The methodology for this analysis is discussed in Chapter 6 and Appendix A of this report. In brief, we conduct multivariate regression analysis in order to control for beneficiaries' baseline characteristics and apply sample weights to produce estimates that are nationally representative for beneficiaries who would volunteer for the offset.

There are clear differences in the amount of counseling received by T21 and T22 subjects, as intended by BOND's design (Stapleton et al 2010). Because T21 and T22 subjects were comparable at baseline in their demographics, benefit histories, and employment experience, the observed differences in counseling can be attributed to differences in the WIC and EWIC models, including differences in outreach by counseling staff, caseload sizes, and service delivery instructions.

Exhibit 3-5 displays the percentage of subjects in each Stage 2 treatment group who received any benefits counseling, including those who only received I&R and those who received counseling beyond I&R. By the end of 2014, 96 percent of T22 beneficiaries had received benefits counseling since study enrollment, and 62 percent had received benefits counseling in 2014. By contrast, 37 percent of T21 beneficiaries received benefits counseling since study enrollment, and 9 percent received benefits counseling in 2014. The large difference in the proportion of subjects receiving counseling is consistent with design, due to the EWIC mandate to conduct outreach to all T22 subjects. The estimated impacts of T22 compared to T21 on counseling receipt are large and statistically significant in every time period. In the most recent calendar year, 2014, T22 subjects were 53 percentage points more likely to receive benefits counseling than T21 beneficiaries. Since study enrollment, T22 subjects were 59 percentage points more likely to receive benefits counseling than T21 subjects.

Exhibit 3-5. Receipt of Benefits Counseling (through December 2014) by T21 and T22 Subjects

Time Period	Percentage of Subjects who received any benefits counseling in time period		
	Offset and EWIC (T22)	Offset and WIC (T21)	Estimated Impact on Counseling Receipt (T22 vs T21)
Enrollment - December 2014(Full time Period)	95.5	36.9	58.6*** (3.3)
Random assignment - December 2012	89.6	28.2	61.4*** (3.6)
January 2013 - December 2013	79.6	14.2	65.4*** (4.3)
January 2014 - December 2014	61.5	8.8	52.8*** (6.5)

Source: Analysis of BTS records, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 and Appendix A for details about the methodology. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: T21 = 4,936 and T22 = 3,089.

*/**/** 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.

The estimates in Exhibit 3-6 show strong evidence that assignment to T22 led to a substantial increase in counseling receipt for all subgroups defined by employment status at baseline. The estimated impact of assignment to T22 on use of counseling is largest for subjects not working and not looking for work at baseline. For those not working and not looking for work at baseline, assignment to T22 increased the proportion of participants receiving benefits counseling by 66 percentage points, compared to an increase of 46 percentage points for those who were working at baseline.

Exhibit 3-6. Receipt of Benefits Counseling (through December 2014) by T21 and T22 Subjects, by Employment Status at Enrollment

Time Period	Percent of Subjects who received any benefits counselling from enrollment to Dec. 2014		
	Offset and EWIC (T22)	Offset and WIC (T21)	Estimated Impact of EWIC instead of WIC Given Offset (T22 vs T21)
All Subjects	95.5	36.9	58.6*** (3.3)
Not Working and Not Looking for Work	94.2	28.5	65.6*** (3.7)
Not Working but Looking for Work	96.7	40.5	56.2*** (3.6)
Working at baseline	95.8	50.2	45.7*** (3.4)

Source: Analysis of BTS records, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 and Appendix A for details about the methodology. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: T21 = 4,936, 2,589, 1,079, and 1,229 for all subjects, subjects not working and not looking for work, subjects not working but looking for work, and subjects working at baseline respectively.

T22 = 3,089, 1,645, 701, and 720 for all subjects, subjects not working and not looking for work, subjects not working but looking for work, and subjects working at baseline respectively.

*/**/** 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.

Exhibits 3-7 and 3-8 show that among those who received benefits counseling, the intensity of benefits counseling was somewhat greater for T22 subjects than for T21 subjects. Of the 36 percent of T21 subjects who received benefits counseling through December 2014, 86 percent received more extensive benefits counseling than I&R (Exhibit 3-7). Of the 95 percent of T22 subjects who received benefits counseling through December 2014, 99 percent received more extensive benefits counseling than I&R (Columns 2 and 3, Exhibit 3-8). The outcome measure in column (3) in Exhibits 3-7 and 3-8 is defined the same manner for T21 and T22 subjects. Column (4) of Exhibit 3-8 also examines the use of EWIC-specific tools or assessments provided exclusively to T22 subjects. Almost all T22 subjects who received benefits counseling beyond an I&R used one or more of the specific tools or assessments that were provided exclusively to T22 subjects through EWIC.

Exhibit 3-7. WIC Services for T21 Subjects (through December 2014)

BOND Site	Percent of T21s receiving WIC (1)	Of those Receiving WIC, Percent Only Information and Referral (2)	Of those Receiving WIC, Percent with Additional WIC Services (3)
Alabama	23.3	16.2	83.8
Arizona/SE California	35.3	19.0	81.0
Colorado/Wyoming	24.3	4.2	95.8
DC Metro	28.5	28.8	71.2
Greater Detroit	49.1	9.7	90.3
Greater Houston	34.0	2.8	97.2
Northern New England	40.0	11.4	88.6
South Florida	43.0	18.1	81.9
Western New York	42.4	22.8	77.2
Wisconsin	37.3	7.9	92.1
T21 Total	36.1	14.3	85.7

Source: Analysis of BTS records. Sample size = 4,936

Exhibit 3-8. EWIC Services for T22 Subjects (through December 2014)

BOND Site	Percent of T22s receiving EWIC (1)	Of those Receiving EWIC, Percent Only Information and Referral (2)	Of those Receiving EWIC, Percent with Additional EWIC Services (3)	Of those Receiving EWIC, Percent with Use of One or More EWIC-Specific Tools or Assessments (4)
Alabama	95.5	0.7	99.3	99.0
Arizona/SE California	91.9	0.3	99.7	98.3
Colorado/Wyoming	100.0	0.0	100.0	100.0
DC Metro	96.3	0.0	100.0	100.0
Greater Detroit	97.8	0.7	99.3	98.2
Greater Houston	99.2	0.0	100.0	100.0
Northern New England	96.6	2.5	97.5	97.5
South Florida	87.2	1.1	98.9	96.4
Western New York	98.6	0.3	99.7	99.3
Wisconsin	94.8	1.5	98.5	98.0
T21 Total	95.2	0.7	99.3	98.5

Source: Analysis of BTS records. Sample size = 3,089

As shown in Exhibit 3-9, counselors delivered the majority of EWIC services between participant enrollment and December 2012. Exhibit 3-9 shows that each year has seen a substantial decrease in the level of each EWIC counseling activity. This decrease is consistent with program design. Assessments, employment support plans (ESP), and a Benefits Summary and Analysis (BS&A), are one-time activities intended to occur soon after contact is established with the beneficiary. For example, the BS&A summarizes current benefits and provide case-specific information on how the offset and other work incentives would affect the beneficiary's SSDI and other possible benefits, such as Supplemental Nutrition Assistance Program (SNAP) benefits and health care coverage. Exhibit 3-9 shows that 34 percent of T22 subjects completed a BS&A by the end of 2012 and 3.5 percent of T22 subjects completed a BS&A in 2014. There are similar declines in the completion of detailed employment support plans (ESP) based on assessments of vocational skills and interests, pre-employment skills training among those with documented need, services coordination among those with documented need, and skills assessments. EWIC is intended to have several key ongoing components, including service coordination and referrals; these later activities constitute the majority of substantive contact with engaged T22 subjects in 2014.

Exhibit 3-9. Receipt of EWIC Services by T22 Subjects by Calendar Year

	Percent of T22 subjects who received service		
	Enrollment - Dec 2012	Jan. - Dec. 2013	Jan. - Dec. 2014
Any Contact	99.6	94.4	76.2
Barriers and needs assessment	77.7	25.2	11.2
Skills assessment	67.8	28.1	9.5
ESP	60.9	22.4	3.3
Service coordination among those with documented need	81.0	67.9	43.6
Pre-employment skills training among those with documented need	49.2	28.7	9.6
I&R assessment	82.6	10.9	3.3
Baseline assessment	75.4	10.3	0.6
BS&A	34.1	19.3	3.5
WIP	33.5	17.4	2.1
Referral	46.5	27.0	12.1

Source: Analysis of BTS records. Sample size 3089

Exhibits 3-10 and 3-11 show strong evidence that assignment to T22 resulted in higher rates of service receipt than did assignment to T21, and this holds true for all subgroups defined by employment status at baseline. Almost all of the T22 subjects received some counseling, and more than half of all T22 subjects received a BS&A or referral (Exhibit 3-10). By contrast, 37 percent of T21 subjects received some counseling, and 15 percent received a BS&A or referral (Exhibit 3-10). For all subjects, assignment to T22 rather than T21 more than doubles the percent the participants who receive a BS&A or referral for all subjects (Exhibit 3-10).

Exhibit 3-10. EWIC Vs. WIC Impact Estimates (T22 Vs. T21) on Counseling Receipt (through December 2014)

	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome with Offset and WIC (T21) (2)	Estimated Impact on Counseling Receipt (T22 vs T21) (3)
All Subjects			
Percentage who received any counseling	95.5	36.9	58.6*** (3.3)
Percentage who received counseling beyond I&R	89.9	36.0	53.9*** (2.9)
Percentage who received a BS&A	51.8	15.0	36.8*** (5.0)
Percentage who received a referral	62.3	12.4	49.9*** (3.5)

Source: Analysis of BTS records, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics. Unweighted sample sizes: T21 = 4,936 and T22 = 3,089.

*/**/** 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.

For T22s, counseling receipt is somewhat similar for all subgroups defined by baseline employment status. For T21 subjects, beneficiaries with greater workforce attachment were more likely to have received I&R and counseling beyond I&R than beneficiaries who were not working and not looking for work at baseline. This relationship between counseling and baseline employment status for T21 subjects is expected because T21 subjects who are working or looking for work are more likely to be affected by the offset and thus seek counseling. For both T21 and T22 subjects, beneficiaries with greater workforce attachment were more likely to have received a BS&A than their respective counterparts with less attachment (Exhibit 3-11). This relationship between BS&A receipt and baseline employment status was expected because a BS&A is more relevant for beneficiaries who are employed or looking for work.

Exhibit 3-11. EWIC vs WIC Impact Estimates (T22 vs T21) on Counseling Receipt (through December 2014) by Baseline Employment Status

	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome with Offset and WIC (T21) (2)	Estimated Impact on Counseling Receipt (T22 vs T21) (3)
Working for Pay at Baseline			
Percentage who received any counseling	95.8	50.2	45.7*** (3.4)
Percentage who received counseling beyond I&R	93.2	48.4	44.7*** (3.2)
Percentage who received a BS&A	66.5	24.0	42.5*** (4.8)
Percentage who received a referral	56.4	14.9	41.6*** (4.5)
Looking for Work at Baseline			
Percentage who received any counseling	96.7	40.5	56.2*** (3.6)
Percentage who received counseling beyond I&R	91.8	39.5	52.3*** (3.3)
Percentage who received a BS&A	55.0	15.1	39.9*** (5.0)
Percentage who received a referral	66.7	14.1	52.6*** (5.3)
Neither Working Nor Looking for Work at Baseline			
Percentage who received any counseling	94.2	28.5	65.6*** (3.7)
Percentage who received counseling beyond I&R	86.9	28.1	58.8*** (3.0)
Percentage who received a BS&A	43.2	10.3	32.9*** (5.6)
Percentage who received a referral	63.0	10.4	52.6*** (3.3)

Source: Analysis of BTS records, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: T21 = 1,229, 1,079, and 2,589 for subjects working for pay at baseline, subjects looking for work at baseline, and subjects neither working nor looking for work at baseline respectively.

T22 = 720, 701, and 1,645 for subjects working for pay at baseline, subjects looking for work at baseline, and subjects neither working nor looking for work at baseline respectively.

*/**/** 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.

3.5. Analysis of Beneficiary Survey Responses

This section presents the experimental impact of treatment on counseling receipt as reported by the beneficiaries in the Stage 2 interim survey, which was asked of all Stage 2 subjects roughly one year after their enrollment in the study.¹⁷ For beneficiaries who received counseling, we also report service locations, referrals, and personal assistance services received, as well as satisfaction with counseling services. In this section, we estimate impact estimates comparing T22 and T21 subjects, as well as comparing each of those groups with the C2 subjects.

“Counseling receipt” as measured in administrative data is a different construct than “counseling receipt” as measured in survey data. In the administrative data, counselors record every interaction they have with beneficiaries, including one-way outreach activities. These records do not reflect whether or not the recipient successfully understood and utilized the information gained from the counseling activity. In contrast, the survey data asked beneficiaries questions such as “Have you talked to someone about how work and earnings affect your Social Security benefits?”, “Have you received a work or job assessment?”, “Have you received help to find a job?”, etc. Beneficiary responses to these questions may be mediated by imperfect communication or understanding, the salience of the counseling content received, or recall error. Self-reported rates of counseling receipt are expected to be lower than rates computed using administrative data, but they are complementary to the analysis of counseling receipt because they are influenced by the salience of counseling events to the beneficiaries.

Although self-reported rates of counseling receipt are lower than administrative reports of counseling receipt, the impact estimates are consistent with what we found using the administrative data: the impacts of T22 in comparison to the current-law control group are substantial and the impact of T22 in comparison to T21 is also large and significant. Exhibit 3-12 displays the impact of T22 and T21 on counseling receipt as reported by the beneficiaries. Compared to T21 subjects, T22 subjects had higher rates of benefit counseling service receipt, use of vocational rehabilitation services, use of work or job assessments, receipt of help in finding a job, enrollment in school or classes, and receipt of advice on modifying a job or workplace (Exhibit 3-12).¹⁸ We conclude that EWIC induced more beneficiaries to get help towards seeking employment and increasing earnings, as compared to WIC.

T21 beneficiaries had a statistically significant higher rate of benefit counseling service receipt than C2 beneficiaries, as expected because of the need to understand the offset. There is also evidence that T21 beneficiaries had a slightly higher rate of job search help than C2 beneficiaries, but the impacts of T21 compared to C2 were not significant on other types of counseling, suggesting that there was no large

¹⁷ The response rate was 84 percent and the median timing of response was 13 months after study enrollment.

¹⁸ The survey asked respondents about services received from any counselor, rather than specifically asking the T21s about WIC, T22 about EWIC, and C2 about WIPA. The section begins by asking the respondent “Since [date of random assignment] have [you] talked to someone about how work and earnings affect [your] Social Security benefits and assistance from other programs? This is sometimes referred to as benefits counseling, and could be done by a benefits counselor, work incentives counselor, or someone from [PROGRAM NAME]?” Next, it asks a series of questions about “how satisfied [you are] with [your] benefits counseling experience”; next, it asks a series of questions about “different types of services or support that [you] may have received to improve [your] ability to work independently.”

effect of the offset-plus-WIC on T21 subjects' receipt of counseling aimed specifically at finding work and increasing their earnings.

Exhibit 3-12. Estimated Impacts on Self-Reported Use of Training-Related Counseling of Stage 2 Volunteers: All Policy Comparisons

	Average Outcome with Offset and WIC (T21) (1)	Average Outcome with Offset and EWIC (T22) (2)	Average Outcome under Current Law (C2) (3)	Estimated Impact of Offset + WIC vs. Current Law (T21 vs. C2) (4)	Estimated Impact of Offset + EWIC vs. Current Law (T22 vs. C2) (5)	Estimated Impact of EWIC Instead of WIC Given Offset (T22 vs. T21) (6)
Use of Vocational Rehabilitation Service (%)	5.2	7.1	4.8	0.4 (0.6)	2.3*** (0.7)	1.9** (0.7)
Use of Benefit Counseling Services (%)	20.6	50.3	14.8	5.8*** (1.5)	35.5*** (3.4)	29.7*** (2.4)
Use of employment supports since study enrollment:						
Work or job assessment (%)	11.8	24.8	10.5	1.3 (0.9)	14.3*** (2.2)	13.0*** (2.0)
Help to find job (%)	12.9	20.5	11.1	1.8* (0.9)	9.4*** (1.1)	7.6*** (1.1)
Training to learn new job or skill (%)	8.3	8.3	8.1	0.1 (0.7)	0.2 (0.8)	0.0 (0.8)
Enrolled in school or classes (%)	7.2	9.0	6.7	0.5 (0.7)	2.3** (0.7)	1.9** (0.7)
Advice about modifying job or workplace (%)	7.1	12.3	6.3	0.7 (0.7)	5.9*** (1.1)	5.2*** (0.9)
On-the-job training, coaching or support services (%)	11.3	12.4	10.9	0.3 (0.8)	1.5 (1.0)	1.1 (1.0)
Personal care assistance (%)	6.9	7.7	5.7	1.1* (0.6)	1.9* (0.9)	0.8 (0.9)
Transportation assistance (%)	9.4	10.3	10.7	-1.3 (0.8)	-0.4 (0.9)	0.9 (1.0)
Help in keeping a job (%)	4.1	4.9	3.9	0.2 (0.7)	1.0 (0.7)	0.8 (0.7)
Any kind of assistive device (%)	9.4	9.5	10.2	-0.8 (0.8)	-0.7 (0.9)	0.1 (0.9)
Other (%)	10.1	9.8	8.8	1.4 (0.8)	1.0 (0.9)	-0.4 (1.0)

Source: Analysis of BOND Stage 2 12-Month Interim Survey, with covariate data and sample weight construction using the BOND Stage 2 Baseline Survey and baseline SSA administrative data.

Notes: Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: T21 = 4,217 , T22 = 2,641 , C2 = 4,021.

*/**/** 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.

Exhibit 3-13 describes the place and referral sources for learning a new job or skill, the place and referral sources for job training, and types of personal assistance received. Out of 18 policy comparisons on the location of service receipt, the analysis finds one significant difference: T22 subjects were more likely to report having received training to learn a new job or skill at a vocational rehabilitation agency, relative to T21 subjects. There is strong evidence that T22 subjects were more likely to be referred by a benefits counselor to training programs than T21 subjects and C2 subjects, which is consistent with the design plan that EWIC would offer more intensive counseling than WIC and current law.

Exhibit 3-13. Estimated Impacts on Self-Reported Service Locations, Referrals, and Personal Assistance Services of Stage 2 Volunteers: All Policy Comparisons

	Average Outcome with Offset and WIC (T21) (1)	Average Outcome with Offset and EWIC (T22) (2)	Average Outcome under Current Law (C2) (3)	Estimated Impact of Offset + WIC vs. Current Law (T21 vs. C2) (4)	Estimated Impact of Offset + EWIC vs Current Law (T22 vs. C2) (5)	Estimated Impact of EWIC Instead of WIC Given Offset (T22 vs. T21) (6)
Percent training to learn new job or skill	8.3	8.3	8.1	0.1 (0.7)	0.2 (0.8)	0.0 (0.8)
Percent training to learn a new job or skill, by location of training:						
A vocational rehabilitation agency	2.5	3.4	2.5	-0.0 (0.4)	0.9 (0.5)	0.9* (0.5)
A welfare agency	0.2	0.2	0.2	-0.1 (0.1)	-0.1 (0.1)	0.0 (0.1)
A mental health agency	0.8	0.9	0.5	0.3 (0.2)	0.4 (0.3)	0.1 (0.3)
A state agency	1.6	2.0	2.0	-0.3 (0.4)	0.0 (0.4)	0.4 (0.4)
A workforce center or unemployment office	1.1	1.4	1.0	0.2 (0.2)	0.4 (0.3)	0.3 (0.3)
An employer	3.0	2.8	3.0	0.1 (0.6)	-0.1 (0.6)	-0.2 (0.5)
Percent training to learn a new job or skill, by type of person who referred:						
Parent/guardian	0.1	0.0	0.0	0.0 (0.0)	-0.0 (0.0)	-0.0 (0.0)
Spouse/partner	0.0	0.1	0.1	-0.0 (0.1)	-0.0 (0.1)	0.0 (0.1)
Friend	0.8	0.8	0.7	0.0 (0.3)	0.1 (0.3)	0.1 (0.3)
Job Coach	0.3	0.3	0.3	0.0 (0.2)	0.0 (0.2)	-0.0 (0.2)
Employer/supervisor	0.6	0.5	0.3	0.3* (0.1)	0.2 (0.2)	-0.1 (0.2)
Other relative	0.4	0.2	0.1	0.3 (0.2)	0.0 (0.1)	-0.3 (0.2)
Benefit specialist	0.4	1.5	0.4	0.0 (0.2)	1.1*** (0.3)	1.1*** (0.3)
Medical Provider	0.3	0.3	0.5	-0.2 (0.2)	-0.2 (0.2)	0.1 (0.2)
Other	1.2	1.0	1.1	0.1 (0.3)	-0.1 (0.3)	-0.2 (0.3)
Percent who received job training, job coaching, or support services	11.3	12.4	10.9	0.3 (0.8)	1.5 (1.0)	1.1 (1.0)

	Average Outcome with Offset and WIC (T21) (1)	Average Outcome with Offset and EWIC (T22) (2)	Average Outcome under Current Law (C2) (3)	Estimated Impact of Offset + WIC vs. Current Law (T21 vs. C2) (4)	Estimated Impact of Offset + EWIC vs Current Law (T22 vs. C2) (5)	Estimated Impact of EWIC Instead of WIC Given Offset (T22 vs. T21) (6)
Percent received job training, job coaching, or support services, by location type:						
A vocational rehabilitation agency	3.8	4.7	3.3	0.5 (0.5)	1.5* (0.7)	0.9 (0.7)
A welfare agency	0.3	0.5	0.6	-0.3 (0.2)	-0.1 (0.2)	0.2 (0.2)
A mental health agency	1.3	1.6	1.1	0.2 (0.3)	0.5 (0.4)	0.3 (0.4)
A state agency	2.2	3.3	2.5	-0.3 (0.5)	0.8 (0.6)	1.0* (0.5)
A workforce center or unemployment office	2.2	2.5	1.9	0.3 (0.4)	0.6 (0.5)	0.3 (0.5)
An employer	6.0	5.8	5.4	0.6 (0.8)	0.5 (0.7)	-0.1 (0.7)
Percent received job training, job coaching, or support services, by type of person who referred:						
Parent/guardian	0.2	0.2	0.3	-0.1 (0.2)	-0.1 (0.2)	-0.0 (0.2)
Spouse/partner	0.1	0.0	0.0	0.1 (0.0)	-0.0*** (0.0)	-0.1 (0.0)
Friend	1.0	0.9	0.9	0.1 (0.3)	0.1 (0.3)	-0.1 (0.3)
Job Coach	0.5	0.7	0.4	0.1 (0.2)	0.3 (0.3)	0.2 (0.2)
Employer/supervisor	1.0	1.1	1.1	-0.0 (0.2)	-0.0 (0.3)	0.0 (0.3)
Other relative	0.4	0.4	0.4	0.1 (0.2)	0.1 (0.2)	-0.0 (0.2)
Benefit specialist	0.8	2.2	0.8	-0.0 (0.2)	1.4** (0.5)	1.4*** (0.4)
Medical Provider	0.5	0.7	0.5	-0.1 (0.2)	0.2 (0.2)	0.3 (0.2)
Other	1.4	1.7	1.5	-0.1 (0.3)	0.2 (0.4)	0.2 (0.4)
Percent who received personal assistance service	6.1	6.9	6.5	-0.4 (0.6)	0.3 (0.7)	0.7 (0.7)
Percent who received personal assistance service, by assistance type:						
Job coach	3.5	3.8	4.2	-0.7 (0.6)	-0.4 (0.6)	0.3 (0.5)
Sign language interpreter	0.5	0.4	0.3	0.2 (0.2)	0.2 (0.2)	-0.1 (0.2)
Reader/interpreter for the blind	0.5	0.9	0.4	0.1 (0.2)	0.5* (0.3)	0.4 (0.3)
Personal care attendant/Personal assistant	2.0	1.7	1.5	0.5 (0.4)	0.2 (0.4)	-0.2 (0.4)

Source: Analysis of BOND Stage 2 12-Month Interim Survey, Stage 2 Baseline Survey and baseline SSA administrative data.

Notes: Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: T21 = 4,163, T22 = 2,595, C2 = 3,961.

*/**/** 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.

The survey also collected data on counseling service satisfaction. Naturally a response to these questions on service satisfaction is conditional on having received counseling services. Exhibit 3-14 displays how satisfied Stage 2 volunteers were with various aspects of counseling services. The data presented in this exhibit only reflect the ratings given by subjects who received counseling services and thus may not accurately predict how subjects who did not seek counseling services would rate the service if they had received it. Analysis conditional on counseling receipt is used for informational and contextual purposes only (as opposed to experimental impact analysis) because counseling receipt was not randomly assigned.

Overall satisfaction is high, with at least 80 percent in all experimental arms reporting that aspects of the service were “good,” “very good,” or “excellent.” The satisfaction indicators as measured for EWICs (column 1) are higher numerically than the measures for WIC and WIPA (columns 2 and 3).

Exhibit 3-14. Satisfaction with Counseling Services, Conditional on Service Receipt

	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome with Offset and WIC (T21) (2)	Average Outcome with WIPA (C2) (3)
Subjects who received benefits counseling			
Percentage who rate the overall service provided by a benefits or work incentive counselor as “good”, “very good”, or “excellent”	87.4	84.4	82.0
Percentage satisfied with how soon a benefits counselor or work incentives counselor was available	93.0	87.0	87.3
Percentage satisfied with the time it took a counselor to reach participant by phone after participant tried to reach a counselor	90.4	83.9	83.3
Percentage satisfied with the time it took a counselor to return a call after participant left a message	92.4	86.0	83.0
Percentage who rate the benefits or work incentives counselor’s courtesy as “good” or “excellent”.	94.9	92.2	90.9
Percentage who rate the length of time spent with a benefits or work incentives counselor as “good” or “excellent”.	89.0	84.9	83.0
Percentage who agree that benefits or work incentives counselor clearly explained how earnings would affect the participant’s cash benefits, medical insurance, and other types of assistance.	90.5	88.1	83.1
Percentage who agree that they after talking with the counselor, they knew what they or the counselor was supposed to do next.	87.6	84.0	80.4
Percentage who agree that written materials given to them by the counselor about the participant’s personal situation and benefits clearly told the participant what s/he needed to know.	86.4	83.4	80.1
Percentage who agree that pamphlets and booklets received from the benefits or work incentives counselor helped the participant to understand how work and earnings affect benefits.	83.9	82.2	79.0

Source: Analysis of BOND Stage 2 12-Month Interim Survey.

Notes: Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment.

Unweighted sample sizes: T21 = 859, T22 = 1,273, C2 = 573.

3.6. Conclusion

Evidence presented in this chapter reinforces earlier findings that T22 subjects are receiving more counseling service than T21 subjects, as intended by study design. Based on demonstration administrative data, by December 2014, 96 percent of T22 had received benefits counseling compared to 37 percent of T21 subjects.

Results based on counseling questions in the interim survey data reinforce the findings from the administrative data and add nuance. They show that, from the beneficiaries' perspective, T22 had a much larger impact on use of counseling services in comparison to C2 than T21—a 35 percentage point increase versus a 6 percentage point increase. Further, while there is no evidence that T21 had an effect on counseling activities specifically focused on employment search or increased earnings, there is such evidence for T22 relative to C2 and to T21.

As anticipated in planning, the percent of Stage 2 beneficiaries receiving counseling declined in 2014 in comparison to the year prior, and this decline was proportionally larger for beneficiaries assigned to WIC services as compared to beneficiaries assigned to EWIC services. The caseloads for EWIC and WIC counselors decreased slightly at all sites, and beneficiaries receiving services report high levels of satisfaction with counseling services.

Beyond counseling receipt, this chapter presents evidence that assignment to T22 relative to T21 increased beneficiary activities that are presumably intended to lead to employment or higher earnings, including use of vocational rehabilitation service, a work or job assessment, receipt of assistance in finding a job, enrollment in school or classes, and receipt of advice about modifying a job or work place. The next chapter examines whether the impacts on increased counseling service receipt for T22 subjects as compared to T21 subjects coincide with impacts on subjects' understanding of the offset.

4. Knowledge of How Earnings Affect Calculation of Benefits and Future Benefit Eligibility and How Stage 2 Subjects Describe BOND

This chapter presents information about how Stage 2 subjects describe BOND and their knowledge of how their benefits and benefit eligibility are affected by working for pay. This information is based on responses to the 12-month Interim Survey of Stage 2 beneficiaries.¹⁹ The chapter begins with a review of the importance of understanding the offset in the BOND logic model.

4.1. Role of Understanding of the Benefit Offset in the BOND Logic Model

The BOND Final Design Report (Stapleton et al. 2010) states two objectives of the demonstration:

- Establish and test a cash benefit system with better work incentives, to improve financial returns of working.
- Develop and test work incentive counseling systems to improve beneficiary understanding of work incentives and ability to use them.

The BOND logic model posits that the offer of the benefit offset (i.e., a “better work incentive” than exists in current law) may eventually lead to increased earnings, reduced benefits, and other positive long-term impacts. Implicit in the logic of BOND is that beneficiaries need to understand the benefit offset offer in order to change their behavior in response to this new work incentive. While there is much to understand about how the offset works (see Chapter 5 for an explanation of the process of using the offset), it seems reasonable that, at a minimum, treatment subjects need to grasp how they gain from the offer—that they will not lose all their SSDI benefits when earning above the SGA level—in order for them to make a change in their employment behavior in response to the offer.

For a beneficiary, full understanding of the offset involves understanding how his or her combined income from benefits and earnings will vary at different levels of earnings. This is a complex relationship, relatively difficult to absorb from either a graphical or narrative presentation. At the design stage of BOND, there was some concern on the part of SSA and the design team that study subjects would have difficulty achieving this full understanding, reflecting reports about such difficulty in the Benefit Offset Pilot Demonstration. It was thought that additional work incentives counseling might be necessary to facilitate the use of the offset offer. Therefore, the enhanced work incentives counseling (EWIC) component was incorporated into the Stage 2 test of the offset. Stapleton et al. (2010) states that, among a number of purposes, “EWIC is expected to increase the impact of the offset by improving beneficiary understanding of how higher earnings will affect SSDI and other benefits.”

The implementation team explained the benefit offset offer to Stage 2 subjects prior to enrollment in order to create interest in volunteering for the study and to allow individuals to provide informed consent. Presumably, those that volunteered had some understanding which spurred them to enroll. After

¹⁹ The overall response rate to the interim survey was 84 percent. The response rates for the T21 and T22 groups (85.4 percent for T21 and 85.5 percent for T22) were somewhat higher than for the C2 group (81.6 percent).

enrollment, work incentives counselors were available to explain the offset and other service provided by BOND to treatment subjects. For control subjects, WIPA services and other services in the community were available to explain the current law rules about how earnings affect benefits.²⁰

The Stage 2 Baseline and Interim Surveys asked subjects about how their benefits and benefit eligibility are affected by earning above the SGA limit. Their interim survey responses allow us to address the important questions of “How well do study subjects understand the basic details of the offset offer 12 months after random assignment?” and “Does EWIC improve understanding of the offset?” The following sections describe the survey questions and the responses of the Stage 2 subjects.

4.2. How Stage 2 Subjects Think SGA-Level Earnings Would Affect Benefits Prior to Random Assignment

In the baseline and interim surveys, all Stage 2 subjects were asked about how their monthly disability cash benefits would change if they were to earn more than the SGA limit after the Trial Work Period (TWP). The evaluation team expected that baseline responses to these questions would be no different across assignment groups because the baseline survey was administered prior to random assignment. In contrast, the evaluation team expected that the 12-month follow-up responses of the treatment group members would be sharply different from those of the control group members because they are subject to different rules (i.e., offset rules versus current-law rules). It was also expected that the greater amount of counseling in EWIC would lead to more accurate responses for T22 beneficiaries compared to T21 beneficiaries.

The questions on the survey were:

Introduction: *“Under the current rules of the Social Security Disability Insurance program, disability beneficiaries are allowed to earn up to \$1010 per month without a change to your benefits. This limit is called the level of Substantial Gainful Activity or SGA and Social Security increases this limit each year to adjust for inflation. When disability beneficiaries go to work while receiving disability benefits, SSA ignores the cap of \$1,010 for up to 9 months, no matter how much a beneficiary earns from work. (The SGA for a blind beneficiary is \$1690.)”*

“We’d like to know which of the following things you think would happen to your monthly disability cash benefits if you were to work and earn more than the SGA limit of \$1010 per month after those initial months have passed. ([IF NEEDED:] The SGA for a blind beneficiary is \$1690.) Thinking about the amount of your disability cash benefits, if you earned more than \$1010 after those initial months...”

²⁰ WIPA services were unavailable during the period from July 2012 to July 2013, a substantial portion of the first year after random assignment for many C2 subjects. During this period, call center staff at the Ticket to Work Help Line provided basic information and referral services. In addition, from April 2013 to April 2014, 20 community work incentive coordinators provided Employment Success Advisor (ESA) services to beneficiaries who were currently employed, had a job offer pending, or were actively seeking employment. The ESA services were similar to those available through the WIPA program. During the 13 months that ESA services were available, 969 beneficiaries throughout the nation (both BOND and non-BOND beneficiaries) were served by the ESA program.

- Do you think your benefits would stay the same?
- Do you think you would lose your monthly benefits completely?
- Do you think your benefits would be reduced but that you would be able to keep receiving some of your monthly disability benefits?
- How do you think those benefits would be reduced?

From their responses to these questions, the evaluation team classified subjects as having one of these five perceptions.

- Benefits would stay the same
- Benefits would be reduced but not to \$0
- Benefits would be reduced to \$0
- Benefits would neither stay same, nor be reduced
- Don't know whether benefits would change

If subjects stated that they thought that their benefits would be reduced but not to \$0, they were further classified by their perception of the reduction amount:

- By full amount of benefits (equivalent to “reduced to \$0”)²¹
- By half, \$1 reduction for \$2 in earnings
- By other amount
- Don't know how much reduction

Exhibit 4-1 shows the responses of all Stage 2 subjects at baseline (i.e., prior to randomization). When the Stage 2 subjects gave these responses they did not know to which group they would be assigned and thus did not know whether the offset would be available to them. Hence, we find that responses are roughly equivalent across the three assignment groups, as expected.²² A little more than a fifth of subjects respond that their benefits would not be reduced to \$0 if they worked above the SGA-level after the TWP. The evaluation team does not interpret these baseline responses as the untainted perception of current law rules. Instead, it seems reasonable that the recruitment process—during which the benefit offset was explained to potential volunteers—could have affected the responses to the survey questions to some degree. The survey questions did not state that the subjects should respond according to current law rules,

²¹ This category is for the few subjects who gave the inconsistent answers that they did not think they would lose their benefits completely but that they thought their benefits would be reduced by the full amount of their benefit.

²² A chi-squared test of the responses in the top panel has a p-value of .837, indicating that we cannot reject the null hypothesis of equivalence across the three groups. A chi-squared test of the responses in the bottom panel has a p-value of less than .0001, indicating that we reject the null hypothesis of equivalence across the three groups. Since differences across columns are not numerically large, we interpret this p-value as a chance imbalance between groups.

and it may have been confusing to some whether the questions were about current law rules or about benefit offset rules. In addition, like all survey responses, these baseline responses may suffer from some amount of measurement error. That is, the responses might not reflect the true level of understanding of current law rules.

Exhibit 4-1. At Baseline: How Stage 2 Subjects Think Benefits Would Change as a Result of Earnings above SGA, Full Sample

	Offset Rules		Current-Law Rules
	T21 (1)	T22 (2)	C2 (2)
All Stage 2 Subjects			
If earnings are above SGA-level beyond TWP months, percent of subjects who think:			
Benefits would stay the same	2.1	1.8	1.9
Benefits would be reduced but not to \$0	23.0	21.3	21.6
Benefits would be reduced to \$0	68.2	70.2	70.4
Benefits would neither stay same, nor be reduced	1.9	2.1	1.6
Don't know whether benefits would change	4.7	4.7	4.6
Percent of subjects who think reduction amount would be^a:			
By full amount of benefit (equivalent to "reduced to \$0")	0.6	0.3	0.6
By half, \$1 reduction for \$2 in earnings	17.1	17.1	16.9
By other amount	3.4	2.7	3.0
Don't know how much reduction	1.9	1.2	1.1

Source: Analysis of Stage 2 12-Month Survey, baseline SSA administrative data, and the Stage 2 Baseline Survey.

Notes: Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of volunteers for offset participation in the nation.

^a The four rows of this panel sum to the percent who think benefits would be reduced but not to \$0.

Unweighted sample sizes: T21 = 4,884, T22 = 3,051, C2 = 4,865.

4.3. How Stage 2 Subjects Think SGA-Level Earnings Would Affect Benefits at One Year After Random Assignment

In contrast to the baseline results, the evaluation team expected that at follow-up the responses of treatment subjects in the T21 and T22 groups would more often be consistent with the offset rules. The expected responses for treatment subjects who understand the offset are that "Benefits would be reduced but not to \$0," and that benefits would be reduced "by half, \$1 reduction for \$2 in earnings." The evaluation team further expected that if there was a difference in the understanding of what would happen to benefits after the TWP, that the offset-plus-EWIC subjects of T22 would be better informed than the offset-plus-WIC subjects of T21 because of the greater contact with counseling staff for those in T22. Finally, it was expected that C2 subjects would provide the response consistent with current law, that "benefits would be reduced to \$0."

Exhibit 4-2 shows the responses of the survey respondents. We find that there are some differences in response pattern between treatment and control subjects, but that the differences are not very large. About half of the treatment subjects gave the correct response that benefits would be reduced but not to \$0, but

another 40 percent of treatment subjects gave the incorrect response that benefits would be reduced to \$0. Likewise, about half of control subjects gave the correct response that benefits would be reduced to \$0, while another 40 percent gave the incorrect response that benefits would be reduced but not to \$0. In contrast to expectations, there was little substantive difference in the expected response of T21 and T22 subjects that benefits would be reduced but not to \$0.

Exhibit 4-2. One-Year Follow-up: How Stage 2 Subjects Think Benefits Would Change as a Result of Earnings above SGA, Full Sample

	Offset Rules		Current-Law Rules	Estimated Impact of EWIC Instead of WIC (T22 vs. T21) (4)
	T21 (1)	T22 (2)	C2 (2)	
All Stage 2 Subjects(N = 6,859)				
If earnings above SGA-level beyond TWP months, percent of subjects who think:				
Benefits would stay the same	3.4	2.4	2.7	-0.95 (0.68)
Benefits would be reduced but not to \$0	50.1	52.4	39.6	2.24 (1.46)
Benefits would be reduced to \$0	41.3	40.7	52.2	-0.68 (1.44)
Benefits would neither stay same, nor be reduced	1.3	1.0	1.2	-0.30 (0.33)
Don't know whether benefits would change	3.9	3.6	4.2	-0.31 (0.54)
Of those who think reduced but not to \$0, percent of subjects who think reduction amount would be:^a				
By full amount of benefit (equivalent to "reduced to \$0")	0.3	0.4	0.6	0.02 (0.16)
By half, \$1 reduction for \$2 in earnings	40.4	44.0	27.9	3.63** (1.45)
By other amount	5.8	5.3	7.2	-0.48 (0.88)
Don't know how much reduction	3.6	2.6	3.9	-0.92* (0.48)

Source: BOND Stage 2 12-Month Interim Survey, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: Differences in responses between T21 and T22 beneficiaries are tested for statistical significance because they are both subject to the offset rules and so their responses should be the same. C2 beneficiaries are subject to current-law rules and so should have different responses from the treatment beneficiaries.

Weights for sample selection and survey non-response are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Impact estimates are regression-adjusted for baseline characteristics.

^a The four rows of this panel sum to the percent who think benefits would be reduced but not to \$0.

Unweighted sample sizes: T21 = 4,174, T22 = 2,620, C2 = 3,975.

*/**/** Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a two-tailed t-test (with no multiple comparisons adjustment).

The second panel shows the percentages of the assignment groups with different responses about the benefit reduction amount. The wording of the survey questions leaves some ambiguity about what would be the correct response for those beneficiaries subject to the offset rules. Among the T21 group, 46 percent gave one of the two possible correct responses (that benefits would be reduced by half or by some other amount²³). The proportion is slightly higher among the T22 group, at 49 percent. Assignment to EWIC increased the proportion of those who correctly said that the reduction amount would be \$1 for every \$2 in earnings, and reduced the proportion of those who thought benefits would be reduced but not to \$0 and did not know how much the reduction would be. We find that 28 percent of the C2 group thought that their benefits would be reduced by half, implying that they mistakenly believed they were eligible for the offset.

How might a little more than a quarter of C2 beneficiaries have given a response consistent with offset rules rather than the current-law rules they are subject to? As at baseline, it seems plausible that these C2 beneficiaries were influenced by their contact with the demonstration, and might have given different responses had they never directly encountered BOND. At the time of their recruitment into the study, these C2 beneficiaries were provided an explanation of how the \$1 for \$2 offset works. The survey mentioned the benefit offset twice prior to these questions about benefit rules. The information provided during study recruitment and the mentions of the offset in the survey may have created confusion about what the benefit rules are for those C2 beneficiaries who gave incorrect responses.

Overall, Exhibit 4-2 suggests that about half of the treatment beneficiaries did not demonstrate an understanding of what would happen to benefits if they earned above the SGA-level after the TWP at the time of the survey. Only about half of each assignment group gave correct responses for how their benefits would be calculated and a little more than a quarter of C2 subjects mistakenly thought their benefits would be calculated under the offset rules. There is some evidence that EWIC improves understanding of how higher earnings affect benefits relative to WIC. However, this difference (3.6 percentage points) is small relative to the 30 percentage point gap in receipt of benefit counseling shown in Chapter 3 (Exhibit 3-12).

Exhibit 4-3 examines whether understanding of how benefits would be affected by high earnings differs according to baseline employment. The same outcomes as above are presented separately for the three subgroups of (a) working for pay at baseline, (b) looking for work at baseline, and (c) not working or looking for work at baseline. We expect that those working at baseline would be most cognizant of the offset rules given that they are most likely to have these rules make a material difference in their total income in the near term. Likewise, we expect that those not working at baseline would be least knowledgeable about the rules because they are less likely to be affected by the rules in the near term.

²³ The wording of the survey question said that benefits would be reduced “by half the amount of your benefits, that is a \$1 reduction in benefits for every \$2 you earn for work.” This is technically incorrect, as offset rules reduce benefits \$1 for every \$2 a beneficiary earns *above the BYA amount*, rather than for the entire earnings amount. The evaluation team expected that relatively few subjects would change their answer to “by some other amount” because of this finely-grained detail.

Perhaps the most notable feature of the results in Exhibit 4-3 for the T21 and T22 groups is the small amount of substantive variation across these subgroups. In addition, those working at baseline are not the most cognizant of the offset rules.

Exhibit 4-3. One-Year Follow-up: How Stage 2 Subjects Think Benefits Would Change as a Result of Earnings above SGA, By Baseline Employment Status

	Offset Rules		Current-Law Rules	Estimated Impact of EWIC Instead of WIC (T22 vs. T21) (4)
	T21 (1)	T22 (2)	C2 (2)	
Working for Pay at Baseline (N = 1,724)				
If earnings above SGA-level beyond TWP months, percent of subjects who think:				
Benefits would stay the same	3.4	2.0	2.4	-1.39 (0.93)
Benefits would be reduced but not to \$0	44.6	50.3	29.0	5.67 (3.17)
Benefits would be reduced to \$0	47.0	43.4	63.0	-3.69 (3.37)
Benefits would neither stay same, nor be reduced	0.9	0.4	1.4	-0.54 (0.38)
Don't know whether benefits would change	4.0	3.9	4.2	-0.06 (1.63)
Of those who think reduced but not to \$0, percent of subjects who think reduction amount would be: ^a				
By full amount of benefit (equivalent to "reduced to \$0")	0.1	0.4	0.3	0.27 (0.16)
By half, \$1 reduction for \$2 in earnings	36.8	43.9	21.3	7.11* (3.63)
By other amount	5.1	4.1	4.3	-0.98 (0.85)
Don't know how much reduction	2.6	1.9	3.1	-0.72 (0.95)
Looking for Work at Baseline (N = 1,484)				
If earnings above SGA-level beyond TWP months, percent of subjects who think:				
Benefits would stay the same	1.8	1.8	2.7	-0.04 (1.02)
Benefits would be reduced but not to \$0	52.3	52.3	39.4	-0.04 (1.88)
Benefits would be reduced to \$0	40.6	41.9	51.9	1.30 (2.29)
Benefits would neither stay same, nor be reduced	1.2	0.9	1.6	-0.31 (0.60)
Don't know whether benefits would change	4.1	3.2	4.4	-0.91 (0.69)
Of those who think reduced but not to \$0, percent of subjects who think reduction amount would be:				
By full amount of benefit (equivalent to "reduced to \$0")	0.6	0.2	0.5	-0.45 (0.28)
By half, \$1 reduction for \$2 in earnings	42.3	44.3	29.3	1.98 (2.30)

	Offset Rules		Current-Law Rules	Estimated Impact of EWIC Instead of WIC (T22 vs. T21) (4)
	T21 (1)	T22 (2)	C2 (2)	
By other amount	5.6	5.7	6.0	0.08 (1.56)
Don't know how much reduction	3.8	2.1	3.6	-1.66 (1.02)
Not Working or Looking for Work at Baseline (N = 3,606)				
If earnings above SGA-level beyond TWP months, percent of subjects who think:				
Benefits would stay the same	3.9	2.9	2.9	-1.02 (1.14)
Benefits would be reduced but not to \$0	52.2	53.3	45.0	1.15 (1.75)
Benefits would be reduced to \$0	38.7	38.9	46.9	0.24 (1.32)
Benefits would neither stay same, nor be reduced	1.5	1.3	1.0	-0.21 (0.29)
Don't know whether benefits would change	3.8	3.6	4.1	-0.16 (0.68)
Of those who think reduced but not to \$0, percent of subjects who think reduction amount would be: ^a				
By full amount of benefit (equivalent to "reduced to \$0")	0.3	0.4	0.8	0.10 (0.16)
By half, \$1 reduction for \$2 in earnings	41.5	43.8	30.5	2.27 (1.83)
By other amount	6.4	5.9	9.3	-0.50 (1.39)
Don't know how much reduction	4.0	3.2	4.4	-0.71 (0.64)

Source: BOND Stage 2 12-Month Interim Survey, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: Differences in responses between T21 and T22 beneficiaries are tested for statistical significance because they are both subject to the offset rules and so their responses should be the same. C2 beneficiaries are subject to current-law rules and so should have different responses from the treatment beneficiaries.

Weights for sample selection and survey non-response are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Impact estimates are regression-adjusted for baseline characteristics.

^a The four rows of this panel sum to the percent who think benefits would be reduced but not to \$0.

Unweighted sample sizes: Working for pay at baseline: T21 = 1,067, T22 = 628, C2 = 986. Looking for work at baseline: T21 = 887, T22 = 583, C2 = 861. Not working or looking for work at baseline: T21 = 2,190, T22 = 1,393, C2 = 2,107.

*/**/** Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a two-tailed t-test (with no multiple comparisons adjustment).

For the C2 group, we expect better awareness of current-law rules among those who will be affected by these rules. Therefore, we expect a positive correlation between labor market activity and the correct

response that benefits would be reduced to \$0. This is, in fact, what we find. Those control subjects who are working for pay at baseline are most likely to think that benefits will be completely suspended as a result of earning over the SGA-level after the TWP. Of the C2 subjects who were working for pay at baseline, 63 percent gave correct responses that their benefits would be cut to \$0. This proportion declines to 52 percent for those who were looking for work at baseline and to 47 percent for those neither working nor looking for work at baseline. Although those working at baseline are the most knowledgeable on average among the three subgroups, a little more than a fifth of this subgroup gave a response consistent with offset rules, rather than current-law rules.

Additional Analysis

Given the relatively low awareness of offset rules among those most likely to be affected by these rules in the near term, we examined treatment subjects among three other subgroups where we expected better understanding of offset rules:

- *those with a bachelor's degree or higher at baseline*, representing 17 percent of the sample;
- *those who had worked since random assignment*, representing 41 percent of the sample; and
- *those who were working 30 hours or more per week at the time of the survey*, representing 6.5 percent of the sample.

In the first two subgroups, the correct response of “benefits would be reduced but not to \$0” was given by roughly the same proportion of subjects as in the full T21 and T22 samples. In the third subgroup, which by virtue of its high level of employment has the most reason to understand how SGA-level earnings affect benefits, the correct response was actually more rare. Only 33 percent of the T21 subjects working 30 hours or more per week gave the correct response. EWIC did make a significant difference for this subgroup, though, as 44 percent of T22 subjects gave the correct response.

4.4. How Stage 2 Subjects Think SGA-Level Earnings Would Affect Future Benefit Eligibility One Year After Random Assignment

All Stage 2 subjects were asked as part of the Stage 2 Interim Survey about how eligibility for disability benefits would change if they earned above SGA after the TWP. Unlike our expectations for the responses to how benefits would change, we do not expect differing responses for treatment and control subjects. While the offset rules imply indefinite eligibility for benefits, the rules themselves only apply for the five-year BOND participation period. Since BOND itself is finite, the correct response for treatment subjects would be that they would remain eligible for benefits for some time, but eventually they would have to re-apply for benefits. And this is also true for the C2 subjects who are under current law.

The survey questions were:

- Do you think you would remain eligible for disability benefits in the future, no matter how much you earn from work? That is, you would never have to re-apply for benefits?
- Do you think you would remain eligible for disability benefits for a while, but eventually you would no longer be eligible to receive benefits? That is, do you think eventually you would have to re-apply for benefits?

From their responses to these questions, the evaluation team classified subjects as having one of these perceptions:

- Would remain eligible for benefits indefinitely (never re-apply)
- Would remain eligible for a while (eventually would have to re-apply)
- Would neither remain eligible indefinitely nor for a while
- Don't know about future eligibility

Exhibit 4-4 shows the response for each treatment group for the full sample and by baseline employment status. We find that 70 percent of the treatment subjects and 74 percent of the C2 subjects give the expected answer that they would remain eligible for a while but eventually would have to re-apply.

Exhibit 4-4. One-Year Follow-up: How Stage 2 Subjects Think Future Benefit Eligibility Would Change as a Result of Earnings Above SGA, Full Sample and By Baseline Employment Status

	Offset Rules		Current-Law Rules	Estimated Impact of EWIC Instead of WIC (T22 vs. T21) (4)
	T21 (1)	T22 (2)	C2 (2)	
All Stage 2 Subjects (N = 6,859)				
If earnings above SGA-level beyond TWP months, percent of subjects who think:				
Would remain eligible for benefits indefinitely (never re-apply)	11.1	11.8	8.6	0.73 (1.01)
Would remain eligible for awhile (eventually would have to re-apply)	69.7	70.0	73.9	0.29 (1.34)
Would neither remain eligible indefinitely nor for awhile	5.2	5.0	4.8	-0.14 (0.67)
Don't know about future eligibility	14.1	13.2	12.7	-0.88 (1.00)
Working for Pay at Baseline (N = 1,724)				
If earnings above SGA-level beyond TWP months, percent of subjects who think:				
Would remain eligible for benefits indefinitely (never re-apply)	13.2	13.0	9.1	-0.19 (0.92)
Would remain eligible for awhile (eventually would have to re-apply)	67.1	68.5	73.5	1.38 (1.28)
Would neither remain eligible indefinitely nor for awhile	5.8	3.7	5.7	-2.01 (1.31)
Don't know about future eligibility	13.9	14.7	11.6	0.82 (1.61)
Looking for Work at Baseline (N = 1,484)				
If earnings above SGA-level beyond TWP months, percent of subjects who think:				
Would remain eligible for benefits indefinitely (never re-apply)	10.5	12.0	9.5	1.44 (3.26)
Would remain eligible for awhile (eventually would have to re-apply)	70.9	69.6	72.5	-1.27 (3.55)

	Offset Rules		Current-Law Rules	Estimated Impact of EWIC Instead of WIC (T22 vs. T21) (4)
	T21 (1)	T22 (2)	C2 (2)	
Would neither remain eligible indefinitely nor for awhile	4.4	4.5	4.4	0.10 (1.74)
Don't know about future eligibility	14.2	14.0	13.6	-0.26 (3.04)
Not Working or Looking for Work at Baseline (N = 3,606)				
If earnings above SGA-level beyond TWP months, percent of subjects who think:				
Would remain eligible for benefits indefinitely (never re-apply)	10.3	11.0	8.0	0.67 (1.06)
Would remain eligible for awhile (eventually would have to re-apply)	70.3	70.8	74.7	0.54 (1.23)
Would neither remain eligible indefinitely nor for awhile	5.2	5.9	4.5	0.67 (1.18)
Don't know about future eligibility	14.1	12.3	12.9	-1.88** (0.68)

Source: BOND Stage 2 12-Month Interim Survey, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: Weights for sample selection and survey non-response are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: All subjects: T21 = 1,067, T22 = 628, C2 = 986. Looking for work at baseline: T21 = 887, T22 = 583, C2 = 861. Not working or looking for work at baseline: T21 = 2,190, T22 = 1,393, C2 = 2,107.

Impact estimates are regression-adjusted for baseline characteristics.

*/**/** Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a two-tailed t-test (with no multiple comparisons adjustment).

Another 11-12 percent of treatment subjects think that they would remain eligible for benefits indefinitely. It is possible that some treatment subjects answered the questions with the offset rules in mind without taking into account the finite nature of the demonstration. It is also possible that some or all of these responses are not considering offset rules and are misperceptions of current-law rules.

The subgroup results shown in the bottom three panels of Exhibit 4-4 reveal little variation by baseline employment status. We might expect that more of those working for pay at baseline would have a correct perception of future eligibility. However, those working for pay at baseline have about the same distribution of perceptions of future eligibility as those looking for work and those not working at baseline.

Overall, while we see more correct responses about future eligibility than about how benefits change with earnings above the SGA-level, there are still a nontrivial number of beneficiaries who are uncertain about future eligibility or who have incorrect perceptions of the rules. About a quarter of C2 subjects and between 18 to 30 percent of treatment subjects are uncertain or have incorrect perceptions of future eligibility.

4.5. How Stage 2 Treatment Subjects Describe BOND One Year After Random Assignment

As part the Stage 2 Interim Survey, Stage 2 treatment subjects were asked how they would describe “the BOND program” to a friend or relative. The evaluation team created response codes and applied up to five codes to each open-ended response. Except for the 1 percent of survey responses that were completed by proxies, all Stage 2 subjects were asked the question.

Exhibit 4-5 shows how T21, T22, and C2 subjects described the “BOND program.” The percentages in each of the first two columns sum to more than 100 percent because the description provided by each respondent could receive up to five codes. The most frequent response, provided by about 45 percent of treatment subjects and 32 percent of control subjects, was “offset program/allows people to work/make more money/not lose benefits.” The next two most frequent responses, each provided by roughly a quarter of treatment subjects, were “help to find employment/return to work/job counseling” and “good/helpful/would recommend.” The last row shows that 10 percent of T22 subjects could not describe BOND, compared to 13 percent of T21 subjects and 18 percent of C2 subjects.

Exhibit 4-5. One-Year Follow-up: How Stage 2 Treatment Subjects Describe BOND

Outcome	Percent of T21 subjects who gave response (1)	Percent of T22 subjects who gave response (2)	Percent of C2 subjects who gave response (3)	Estimated Impact of Offset and WIC vs. Current Law (T21 vs. C2) (4)	Estimated Impact of Offset and EWIC vs. Current Law (T22 vs. C2) (5)	Estimated Impact of EWIC instead of WIC Given Offset (T22 vs. T21) (6)
How Stage 2 Subjects Would Describe the BOND Program to a Friend or Relative						
Good/helpful/ would recommend	22.9	25.3	19.1	3.8*** (1.1)	6.2*** (1.2)	2.4* (1.3)
Not helpful/poor/dislike program	4.4	3.3	5.2	-0.8 (0.6)	-1.9** (0.8)	-1.1 (0.7)
Promotes higher self-esteem/independence/better quality of life	2.4	2.5	1.8	0.6 (0.4)	0.6 (0.4)	0.1 (0.5)
Incentive program/encourages people to work	2.1	2.2	1.0	1.0** (0.3)	1.2*** (0.4)	0.1 (0.5)
Offset program/ allows people to work/ make more money/not lose benefits	45.8	43.5	31.6	14.2*** (1.3)	11.9*** (1.6)	-2.3 (1.4)
Help to find employment/return to work/job counseling	22.8	28.2	20.3	2.4** (1.1)	7.9*** (1.3)	5.4*** (1.3)
Help with job training/education	3.1	5.1	2.9	0.3 (0.5)	2.3** (1.0)	2.0** (0.8)
Determine improvements/services needed/effectiveness of services (for people to return to work)	3.1	3.7	6.9	-3.8*** (0.8)	-3.2*** (0.7)	0.6 (0.6)
Different levels of assistance/different groups/lottery/randomly chosen	3.4	2.5	10.0	-6.6*** (1.0)	-7.5*** (1.0)	-0.9 (0.5)
Don't understand program/confusing/ complicated	1.0	0.8	1.2	-0.2 (0.3)	-0.4 (0.3)	-0.2 (0.3)
Don't know/don't remember/nothing/ refused	12.6	10.3	18.1	-5.5*** (0.9)	-7.9*** (1.1)	-2.4** (0.9)

Source: BOND Stage 2 12-Month Interim Survey, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: Respondents were asked how they would describe BOND to a friend or relative. The open ended responses were coded into the categories shown in the exhibit. The maximum number of codes that any response contained was five.

Weights for sample selection and survey non-response are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: T21 = 4,174, T22 = 2,620, C2 = 3,975

Impact estimates are regression-adjusted for baseline characteristics.

*/**/** Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a two-tailed t-test (with no multiple comparisons adjustment).

4.6. Summary

The results presented in this chapter show that about half of the treatment subjects provided a 12-month survey response consistent with how the benefit offset works. Also, the understanding of T22 subjects is only slightly more accurate than that of T21 subjects. In addition, the understanding of the offset does not appear to be greater for subjects who were working when they enrolled in the study, as might be expected. The 12-month survey responses for the control group subjects do not provide evidence of widespread understanding of program rules. Only about half of the control group subjects gave a response consistent with an accurate understanding of the current-law rules that apply to them. In fact, a little more than a quarter of control subjects gave a response consistent with the offset rules, perhaps reflecting the information they heard about the offset during the enrollment process. Although these control group subjects gave responses consistent with offset rules, it is unknown whether their behavior is influenced by a belief that the offset is available to them. (Such behavior would serve to diminish the measured impact of the offset offer.) What seems clear, however, is that almost half of control subjects do not have a firm grasp of the current-law rules on earnings that apply to them, indicating that confusion about rules is relatively common among SSDI beneficiaries.

Roughly three-quarters of treatment and control subjects have a correct perception of how earnings would affect their future benefits eligibility.

Taken together, these results indicate that by 12 months after study enrollment about half of the treatment subjects do not demonstrate an understanding of a presumably crucial prerequisite for a behavioral response to the offset—understanding the benefit offset offer. In addition, while one of the key purposes of EWIC was to improve beneficiary understanding of the benefit offset offer, these results suggest that as of 12 months after enrollment, understanding had improved by only a small amount.

5. Using the Benefit Offset

5.1. Introduction

Use of the benefit offset and timely and accurate benefit adjustment under the offset rules rely on the implementation of a number of complex processes. Since the start of the demonstration, the BOND Implementation Team and SSA staff have streamlined and improved several of these processes. However, issues with other processes essential to administering the benefit offset have continued and SSA staff have uncovered additional challenges.

Despite these challenges, Stage 2 beneficiaries are reaching significant milestones in moderate numbers. By the end of 2014, SSA had adjusted benefits under the offset for nearly one in ten Stage 2 treatment subjects. In addition, another five percent of Stage 2 treatment subjects had reached significant milestones along the pathway to the offset.

This chapter describes the structure and implementation of the benefit offset and statistics related to the use of the benefit offset. Section 5.2 discusses the pathway to the offset and necessary steps for benefit adjustment under the offset rules. The chapter compares and contrasts these steps to current law requirements that apply to BOND control subjects. This will aid in understanding impact estimates that compare treatment and control subjects in subsequent chapters. We then discuss key implementation findings, including difficulties in administering the offset, solutions, and progress to date. We base this discussion on qualitative interviews conducted in late 2014 and early 2015 with: WIC and EWIC administrators and counselors; a technical assistance provider; the BOND lead for post-entitlement work (described in Section 5.2.3); the liaison between the BOND sites and the BOND Implementation Team; the Implementation Team director and deputy director; and the ORDES BOND work unit within SSA. In most cases and unless specifically noted, these findings apply to both Stage 1 and Stage 2 subjects. Finally, in Section 5.3, we describe the frequency of and trends in offset use and subsequent benefit adjustment among Stage 2 treatment subjects through the end of 2014.

5.2. Implementation of the Benefit Adjustment Process

This section describes the four primary milestones in the SSA administrative process on the path toward the first adjustment of benefits under the offset rules: (1) eligibility for benefit adjustment, (2) documentation of eligibility for benefit adjustment, (3) AEE submission, and (4) benefit adjustment. Exhibit 5-1 shows that the first two milestones are the same for treatment subjects and control subjects en route to benefit suspension due to engagement in SGA. Beneficiaries in both groups must demonstrate prolonged engagement in SGA in order to be eligible for benefit adjustment under the offset (treatment subjects) or benefit suspension (control subjects). SSA documents engagement in SGA via a work CDR. The last two milestones, completion of an AEE and benefit adjustment, are unique to treatment subjects. Because it takes considerable time to receive and process the required information, SSA often applies the first benefit adjustment retroactively, back to the month of eligibility for a reduction (treatment subjects) or suspension (control subjects).²⁴

²⁴ In some cases, treatment beneficiaries may be due money from SSA (known as an underpayment) at the time of their first benefit adjustment under the offset. In these cases, the first benefit adjustment for treatment beneficiaries may be an increase rather than a reduction. This has most commonly happened because SSA had

Exhibit 5-1. Benefit Adjustment under Current Law and the BOND Benefit Offset

Here we provide a brief overview of benefit adjustment under current law and procedures developed to administer the benefit offset tested in BOND. A detailed review of current-law SSA rules and work incentives is provided in Chapter 2 of Stapleton et al. (2010). A detailed description of the BOND innovation is available in Sections 6.1 and 6.2 of Gubits et al. (2013) and Section 5.2 of Derr et al. (2015).

Establishing and Documenting Eligibility for Benefit Adjustment under Current Law and the Benefit Offset

Current-law SSDI rules and procedures govern BOND treatment and control subjects as they work and proceed to become eligible for benefit adjustment:

- During the ***Trial Work Period (TWP)***, beneficiaries are entitled to attempt work without affecting benefits. In 2014, a TWP month was any month in which an SSDI beneficiary had earnings of at least \$770 or worked at least 80 self-employed hours. The TWP consists of nine such months in a rolling 60-month window.
- SSA conducts a ***Work Continuing Disability Review (Work CDR)*** to confirm beneficiaries' continued eligibility for benefit receipt. In SSA's terminology, disability "ceases" for beneficiaries who engage in SGA after completing the TWP.
- During the ***Grace Period (GP)***, SSA pays benefits at their full amount for the SGA cessation month and for two additional months of SGA.

Benefit Adjustment

If beneficiaries continue to engage in SGA after the GP, rules regarding work and the associated processes differ between BOND treatment and control subjects.

Current Law (BOND Control Subjects)

During the three-year ***Extended Period of Eligibility (EPE)***, SSA suspends benefits in any month in which a beneficiary engages in SGA. Engagement in SGA after that results in benefit termination.

Benefit Offset (BOND Treatment Subjects)

During the five-year ***BOND Participation Period***, SSA adjusts treatment group members' benefits on the basis of annual (calendar year) earnings.

- Beneficiaries submit an ***Annual Earnings Estimate (AEE)*** as a basis for SSA to adjust benefits for the current year.
- SSA uses the ***BOND Stand Alone System (BSAS)*** to adjust benefits in accordance with an AEE. SSA pays the reduced level of benefits in equal monthly amounts throughout the year.

Some beneficiaries with earnings fail to report earnings and do not submit an AEE in the year in which they are first eligible for an offset adjustment. SSA identifies these beneficiaries through a review of IRS earnings data after the end of the year and, after providing the beneficiary with an opportunity to submit additional information, adjusts benefits retroactively through a process called ***reconciliation***. SSA similarly reconciles the past benefits of those who submitted an AEE but had earnings that were substantially different from those estimated.

suspended benefits prior to the beneficiaries' enrollment in BOND and assignment to treatment, and benefit suspension continued for subsequent month until SSA made a retroactive adjustment under the offset.

5.2.1. Eligibility for Benefit Adjustment

Beneficiaries must engage in SGA for a sustained period to become eligible for the benefit offset. Two factors related to the demonstration may affect sustained engagement in SGA: the availability of employment services and beneficiaries' understanding of the work incentives. Many beneficiaries seeking to return to SGA-level employment need employment-related services to facilitate their efforts. Apart from counseling, employment services are not a component of BOND; treatment subjects presumably have no more or less access to such services than control subjects. BOND WIC and EWIC counselors may refer treatment subjects to existing service providers, as WIPA counselors may for control subjects. However, through 2014, treatment subjects (and presumably control subjects) have not always been able to obtain employment services when needed for reasons external to BOND (see the discussion in Section 2.6). Beneficiary understanding of the work incentives is critical to the impact of the benefit offset on beneficiary behavior. If treatment subjects do not understand that, in contrast to current law, the benefit offset allows them to keep some of their SSDI benefits if they engage in SGA after the TWP and GP, they will be less likely to take advantage of the opportunity than were they better informed.

5.2.2. Documentation of Eligibility for Benefit Adjustment

For many beneficiaries who became eligible for adjustment in the first three years of the demonstration, impediments to completing documentation required to establish offset eligibility via the work CDR process delayed the first adjustment of benefits under the offset. There are three steps in the process: (1) SSA or BOND staff identify those in need of a work CDR, (2) beneficiaries, often with the help of SSA or BOND staff, compile accurate information on work history, and (3) SSA verifies the information and completes the work CDR. These processes have not always functioned as intended. Furthermore, several processes differ across treatment and control subjects in ways that might lead to systematic differences in the timeliness of work CDR completion.

As described in this section and in previous reports (Gubits et al. (2013), Derr et al. (2015)), these processes involve various SSA and BOND staff. To clarify the discussion, it is helpful to recognize that SSA staff includes staff in several different components of SSA. First, the ORDES BOND work unit provides operational support for BOND. This involves collecting work CDR documentation from beneficiaries, distributing some work CDRs to other SSA components for processing, and directly processing the balance of work CDRs. Second, staff at local SSA field offices also are available to assist BOND treatment subjects with activities such as collecting work reports, assisting with work CDR paperwork, and processing select work CDRs. Third, staff at SSA processing centers also process work CDRs for some BOND treatment subjects.²⁵ In addition, BOND staff—specifically WIC and EWIC staff and members of the BOND Implementation Team—are available to collect beneficiary work reports and ask the ORDES BOND work unit to initiate a work CDR. As described below and in previous reports, these responsibilities have evolved over the course of the demonstration.

²⁵ As described in Derr et al. (2015), staff at SSA processing centers and local SSA field offices began processing work CDRs in August 2013 and January 2014, respectively, as part of a larger effort to decrease the backlog of work CDRs at that time.

Identifying Beneficiaries in Need of a Work CDR

The process for identifying beneficiaries in need of a work CDR differs depending on whether beneficiaries report their earnings. SSA requires all SSDI beneficiaries to report earnings. BOND treatment subjects and Stage 2 control subjects may report earnings to either BOND staff or SSA. According to BTS data as of March 2015, despite similar employment rates,²⁶ less than 14 percent of Stage 2 control subjects had ever submitted work reports to the demonstration. This compares to 39 percent of Stage 2 treatment subjects. The difference may occur because control subjects are more likely to report their earnings to SSA rather than to BOND staff or because they are less likely to report earnings to either entity. These reporting differences are salient because of process differences for beneficiaries who do and do not report earnings and also because of varying practices across BOND and SSA in initiating work CDRs among beneficiaries in the former group.

When beneficiaries submit work reports to BOND staff, those staff are responsible for reviewing available data to determine if the BOND beneficiaries are in need of work CDRs. If the BOND beneficiaries are in need of work CDRs, BOND staff follow the BOND guidelines for the timing of CDR initiation and, if appropriate, submit requests to SSA's ORDES BOND work unit to initiate work CDRs. Since the start of BOND, WIC and EWIC staff have shared responsibility for identifying beneficiaries in need of a work CDR among those who report earnings with the site office staff who enrolled Stage 2 subjects (until September 2012) and BOND Implementation Team staff.

According to members of the BOND Implementation Team and as indicated in Gubits et al. (2013), because of the complexity of the task and competing demands on their time, staff initially struggled to identify which beneficiaries required a work CDR. In response, in March 2013, the BOND Implementation Team began monthly reviews of BTS data to identify beneficiaries in need of a work CDR. In contrast, SSA staff screen work reports as they are submitted.

During the first three years of the demonstration, structural differences occurred in the timing of work CDR initiation, depending on which entity received work reports. Outside of BOND, SSA's guidance to SSA staff throughout the demonstration period has been to initiate a work CDR whenever a beneficiary reports new work at any level. SSA field office staff and processing center staff can initiate this process. Prior to Spring 2014, guidance issued to BOND staff was to request a work CDR if the beneficiary was known to have earnings over the SGA amount and was likely to have completed seven or more TWP months. In an effort to reduce overpayments, which grow as delays in CDR completion lengthen, in Spring 2014 SSA changed the guidance issued to BOND staff to conform to the general SSA guidance for current-law beneficiaries.

As noted above, despite working at similar rates, treatment subjects were more likely than control subjects to ever have submitted a work report to BOND staff (39 percent compared to 14 percent). It is therefore plausible that control subjects submitted work reports to SSA staff (rather than BOND staff) at a higher rate than treatment subjects. If this was indeed the case, then the difference in guidance issued to SSA staff and BOND staff may have led to accelerated work CDR processing for control subjects relative to treatment subjects prior to Spring 2014. However, data are not currently available to measure any potential differences in the timing of work CDR initiation between the two groups.

²⁶ According to 2013 administrative data, Stage 2 treatment and control subjects worked at similar rates: between 33 and 36 percent were employed at any point during the year (see Exhibit 7-6).

Treatment and control subjects who do not report their earnings will face delays in the work CDR process compared to beneficiaries who report their earnings to BOND or SSA. Under both current law (SSA 2011) and under BOND (Derr et al. 2015), many beneficiaries fail to report work. In early 2015, several members of the BOND Implementation Team reported that this behavior appears to continue. For beneficiaries who do not report work, the start of the work CDR process cannot occur until SSA receives information on earnings from another source, typically IRS earnings records. SSA reviews IRS earnings records three times a year, typically starting in late spring or summer of the following calendar year.²⁷ SSA receives earnings information sooner if beneficiaries report their earnings in real time.

Developing Work History

When either SSA or BOND staff identify the need for a work CDR, ORDES BOND work unit staff send the beneficiary administrative forms and a request to document past work activity. In many cases, beneficiaries need assistance in completing these forms. Either BOND staff or SSA field office staff are available to provide assistance, if requested.

According to both BOND and ORDES staff, the process of developing work history generally works well (Derr et al. 2015). However, there are some exceptions. In some cases, beneficiaries may not provide information in a timely fashion (Derr et al. 2015). SSA field office staff are required to assist beneficiaries with work CDR paperwork, including beneficiaries participating in BOND. ORDES and Implementation Team staff reported some instances in which SSA field office staff did not assist BOND treatment subjects with work CDR paperwork, because of their treatment group status. ORDES BOND work unit staff reported that, towards the end of the period covered in this report, they had begun to take corrective action to address this issue.

SSA Processing of Work CDRs

SSA has been delayed in processing work CDRs for treatment subjects since the start of the demonstration. To some extent, BOND inherited work CDR delays because some BOND subjects were already overdue for work CDRs before they enrolled in the demonstration. In fiscal year 2010—before the demonstration began enrolling subjects—it took SSA an average of 124 days to process work CDRs (SSA 2011). It appears that long processing times persist for BOND treatment subjects for reasons explained below. WIC and EWIC staff cited delays in work CDR processing as a key barrier to timely benefit adjustment.

SSA has dedicated additional resources to process treatment group work CDRs over time, but ORDES BOND work unit staff reported that they remain insufficiently staffed to process the CDRs on a timely basis. As described in Derr et al. (2015), staffing changes within the work unit and ongoing assistance from both SSA processing centers and field offices have all helped to reduce the work CDR backlog. Under the current arrangement, for work CDRs initiated by their staff, the work unit processes cases expected to result in cessation internally and sends work CDRs for treatment subjects not expected to have a disability cessation to SSA processing centers or field offices for processing. SSA processing centers and field offices also initiate and process work CDRs. When SSA processing centers and field offices process work CDRs, they transfer any BOND treatment cases that result in cessation to the work unit for final processing. At the start of 2015, the work unit queue—which includes Stage 1 and Stage 2

²⁷ The timing of SSA review of earnings information was misstated in Footnote 57 of Derr et al. (2015).

treatment subjects—had 617 cases in need of full work CDR processing and 520 cases with cessation dates in need of final processing. Three work unit staff are responsible for processing the CDRs, in addition to other responsibilities. The work unit staff reported that this workload is much higher than the typical workload for the SSA field office staff who process work CDRs outside of the demonstration.

The addition of SSA field offices to assist in BOND work CDR processing has also led to coordination challenges. According to WIC and EWIC staff, the process and the multiple points of contact are confusing to some beneficiaries. Some counselors also expressed frustration that they are unsure where work CDRs are being developed. The same staff noted that they are sometimes unable to ask questions or get status updates, most frequently in cases for which SSA field offices and processing centers are processing the work CDR. In addition, according to two members of the Implementation Team, the ORDES BOND work unit faces coordination challenges in sharing beneficiary work reports stored in BTS with SSA processing center and field office staff, who do not have access to that information. Beneficiary work reports are typically more recent relative to the earnings reports based primarily on IRS data housed at SSA. The Implementation Team staff explained that WIC and EWIC counselors often let ORDES BOND work unit staff know that information is available on BTS, and the work unit then makes this information available to the SSA staff processing the work CDR.

The effect of additional resources on processing times is unclear, but recent statistics suggest that treatment group work CDRs are still subject to long processing times. BTS data on work CDRs are too preliminary to facilitate a comparison of processing times across years.²⁸ According to a snapshot from SSA's eWork system, in February 2015, 56 percent of the BOND treatment group work CDR cases were more than 270 days (nine months) old. That duration prevents timely benefit adjustment under the offset. At the same point in time, according to the same source, 1 percent of all work CDR cases being processed nationwide were more than 270 days old. The implication is that, once SSA initiates a work CDR, likely offset users in the BOND treatment group are subject to longer work CDR processing times than members of the BOND control group.

Beneficiaries may encounter delays at several points in the work CDR process. Regardless of the source, the process does not facilitate timely benefit adjustment. Of the 1,247 cessation dates recorded in BTS for treatment subjects between March 2014 and May 2015, SSA assigned 4 percent of cessation dates within three months after cessation. Three months is a notable milestone because, including the month of cessation, this is the period for which benefits are protected from adjustment due to engagement in SGA as part of the GP. The implication is that adjustment of the benefits of almost all of those who engaged in SGA after disability cessation has not been timely, as is discussed in Section 5.2.5.

²⁸ We analyzed processing times for 1,212 work CDRs requested by BOND staff in 2012 using data from two different points in time and found varying results. Data from March 3, 2015 reveal that, at that time, 558 work CDRs resulted in cessation and the average processing time was 281 days. The balance may include work CDRs still being processed, cases SSA determined to have a continuance (rather than a cessation), or cases for which SSA deemed a work CDR was unnecessary. These statistics update results presented in the BOND Process Study Report (Derr et al. 2015), which were based on data from April 9, 2014 and found that 478 work CDRs requested by BOND staff in 2012 resulted in cessation and had an average processing time of 227 days.

5.2.3. Annual Earnings Estimate Submission

The third milestone on the pathway to the offset, completion of an AEE, is unique to BOND treatment subjects and necessary for accurate benefit adjustment. During the EPE, and after beneficiaries use the three grace period months, under current law SSA suspends the benefits of those it determines to be engaging in SGA. Under the BOND rules after the GP months SSA instead adjusts benefits based on annual earnings, necessitating a directive to adjust benefits throughout the year: an AEE.

Accurate and timely AEEs are a necessary step for proper benefit adjustment and to help beneficiaries minimize over and underpayments. When SSA completes a work CDR and identifies a treatment subject as engaging in SGA and offset eligible, SSA suspends benefits until the beneficiary submits an AEE.²⁹ In subsequent years, BOND staff advise beneficiaries to submit an AEE for each calendar year and strive to collect AEEs before the start of a new calendar year. If a beneficiary does not submit a new AEE, SSA will adjust benefits in accordance with the most recent AEE submitted.

The processes for identifying beneficiaries in need of an AEE and completing AEEs have evolved over the course of the demonstration. BOND field staff initially had difficulty identifying beneficiaries in need of an AEE. However, since the Implementation Team began monthly reviews of BTS data in early 2013, this process has run smoothly (Derr et al. 2015). Beneficiaries who submit an AEE for the first time do so with the help of a WIC, EWIC, or BOND Implementation Team member, in person or over the phone. Demonstration staff collect subsequent-year AEEs by mail. BOND Implementation Team staff began mailings to collect 2013 AEEs in late 2012, but responses were lower than expected. To increase response rates for 2014 and 2015 AEEs, the Implementation Team mailed postcards in advance of the 2014 and 2015 mailers. BOND Implementation Team staff also called beneficiaries who were unresponsive to mailed requests for AEEs. The AEE mail collection efforts for 2014 and 2015 yielded responses from 76 and 69 percent of contacted beneficiaries, respectively.

As discussed in Section 3.2, early in the demonstration, staff struggled to accurately complete AEEs along with other post-entitlement work.³⁰ Some staff became proficient with AEEs; for other staff, problems persisted through fall 2013. To address this issue, in December 2013 the Implementation Team centralized post-entitlement work in the majority of the 10 BOND sites (Derr et al., 2015).³¹ In centralized post-entitlement sites, members of the BOND Implementation Team, known as the Centralized Post-Entitlement (CPE) Team, complete AEEs.

Centralizing post-entitlement work has led to several improvements, most notably in the quality of AEEs. Members of the BOND Implementation Team explained that post-entitlement work is nuanced, technical, and complex. Many of the CPE team members now responsible for the work have experience working at

²⁹ We do not know the number of beneficiaries for whom SSA suspended benefits because of an outstanding AEE.

³⁰ Post-entitlement work includes preparing AEEs and collecting documentation of non-countable income used for reconciliation.

³¹ The demonstration centralized post-entitlement work in Arizona/Southeastern California (EWIC and WIC), Colorado/Wyoming (EWIC and WIC), DC Metro (WIC), Greater Detroit (EWIC), Greater Houston (EWIC and WIC), Northern New England (EWIC and WIC), South Florida (EWIC and WIC), and Wisconsin (WIC) in December 2013, and in Alabama (WIC) in January 2015.

SSA and, in large part as a result of that experience, have the necessary skills to complete the work. Furthermore, while some WIC and EWIC staff had expressed disinterest in this work, this sentiment appears to be rare among the CPE team. In fall 2014, a large majority (86 percent) of WIC and EWIC focus group participants in centralized sites favored centralization of AEEs and other post-entitlement work. Furthermore, BOND staff perceived that beneficiaries received higher quality support as a result and submitted fewer records with errors. Indeed, in late 2012 and early 2013, about 30 percent of submitted AEEs contained errors (Derr et al. 2015). In 2014, only about 1 percent of initial AEEs contained errors.³² Although comparable statistics across years were not available, BOND staff perceived that the timeliness of AEE submission and post-entitlement work had improved, in part due to the streamlined process. ORDES BOND work unit staff agreed that the centralized process was more streamlined and noted that they received fewer trivial questions from staff responsible for post-entitlement work. In addition, work unit staff observed that improper payments resulting from inaccurate AEEs had declined.

Although centralization of AEEs and other post-entitlement work has yielded mostly positive results, centralization has also created a few new challenges. First, WIC, EWIC, and CPE team staff indicated that some beneficiaries are confused about the roles of various staff and to whom to direct questions. Indeed, in sites in which the Implementation Team centralized post-entitlement work, beneficiaries may interact with their designated counselor as well as a CPE team member. Implementation Team leadership has instructed WIC and EWIC staff to answer post-entitlement questions and then refer beneficiaries to a CPE team member for further assistance.

Second, in centralized sites, some WIC and EWIC staff reported that they are inadequately informed about the status of the post-entitlement work, and therefore poorly positioned to answer beneficiary questions or anticipate delays or other issues. Each centralized site has a designated CPE team member to handle post-entitlement work, so staff know to whom to direct inquiries, but some counseling staff described being less informed than when they had direct responsibility for the work.

5.2.4. Benefit Adjustment

Since the beginning of the demonstration, SSA has encountered challenges with the process for making benefit adjustments. SSA uses the BOND Stand Alone System (BSAS), a computer program that interfaces with SSA's data systems, to adjust SSDI benefits for treatment subjects according to BOND rules after completion of an AEE. SSA also uses BSAS to conduct automated reconciliation with IRS data for beneficiaries with an established disability cessation. BSAS functions as intended for the former purpose, but functionality has been problematic for the latter.

SSA typically uses BSAS to make contemporaneous adjustments in accordance with AEEs in two instances. The first is to adjust the benefits of beneficiaries who enter the offset for the first time during the current calendar year. These are typically beneficiaries who interact with BOND or SSA staff to report earnings and, when requested, complete administrative paperwork. Beneficiaries who enter the

³² BTS statistics on AEE errors exclude 2014 AEEs submitted by mail and those submitted by the two most senior CPE staff who assist beneficiaries with AEEs. The statistics presented in the body of this report are for AEEs that are likely the more complicated records being prepared (e.g. initial AEEs, including partial-year AEEs), according to Implementation Team staff. Accordingly, the error rate for all AEEs may be even lower than 1 percent.

offset proactively are referred to as front-door offset entrants (see Section 5.2.1 of Derr et al. (2015)). The second is to adjust the benefits of beneficiaries who have previously used the offset and submit revised AEEs or AEEs for a new calendar year.

During automated reconciliation, SSA uses BSAS to reconcile benefits paid—what the agency paid the beneficiary during the year—to benefits due as determined by IRS earnings reports, net of any non-countable income reported by the beneficiary and approved by SSA. Automated reconciliation affects both beneficiaries for whom SSA adjusted benefits under the offset in accordance with an AEE and beneficiaries with disability cessation dates for whom SSA had not adjusted benefits. The latter consists of beneficiaries who did not report earnings to the demonstration. These beneficiaries are referred to as back-door offset entrants (see Section 5.2.1 of Derr et al. (2015)). SSA identifies back-door entrants as offset-eligible through reviews of IRS records that occur after each calendar year. Once SSA identifies beneficiaries with unreported substantial earnings, conducts a work CDR, and assigns the beneficiary a cessation date, SSA uses automated reconciliation to adjust benefits for the calendar year immediately prior and manual reconciliation to adjust benefits in years for which they have already run automated reconciliation, as applicable. SSA used automated reconciliation to adjust the benefits of 45 percent of Stage 2 beneficiaries whose first month of eligibility for the offset was in 2011 and 26 percent of those whose first eligibility month was in 2012, based on May 2014 data (Derr et al. 2015).³³

The benefit adjustment process appears to be working smoothly in adjusting benefits in accordance with AEEs. Even at the start of the demonstration, delays were no longer than three days, and a BSAS correction in December 2012 successfully addressed the underlying issues causing those delays (Gubits et al. 2013). According to ORDES and BOND staff, BSAS currently functions well for adjustments based on AEEs.

For retroactive adjustments made through automated reconciliation, BSAS does not appear to be working as well. We have previously reported that issues with BSAS functionality were responsible for long delays in the automated reconciliation processes for 2011 and 2012 benefits, and they have continued for the 2013 benefit adjustment process. As previously noted in Derr et al. (2015), automated reconciliation for a given year is scheduled for August of the following year, but SSA delayed the 2011 automated reconciliation by five or six months (conducted in January and February 2013) and 2012 automated reconciliation by one or two months (conducted in four batches in September and October 2013). SSA began 2013 automated reconciliation in late April 2015—an eight month delay. The direct result of such delays is an extended wait for benefit adjustment, resulting in an extended period of time over which beneficiaries may accumulate improper payments and perhaps not understand how their earnings are affecting their benefits.³⁴ The delay in 2013 automated reconciliation did not allow time for inclusion of related statistics in this report.

³³ Based on May 2014 data, 42 of 94 Stage 2 beneficiaries whose first offset month was in 2011 entered the offset through automated reconciliation as did 91 out of 350 with a first offset month in 2012. It is possible that SSA will complete work CDRs and assign cessation dates to additional beneficiaries with unreported earnings who were eligible for the offset in 2011 and 2012. If so SSA will adjust benefits for these beneficiaries using manual processes, to account for earnings in 2011 and 2012.

³⁴ Until automated reconciliation, some beneficiaries may be unaware that they are using the offset or that work activity affected and affects their past and current benefits. For example, a delay in 2012 EOYR from August

SSA delayed the 2013 automated reconciliation when ORDES BOND work unit staff discovered that BSAS had incorrectly calculated over and underpayments for some beneficiaries during the previously run 2011 and 2012 automated reconciliations. The work unit manually checks and confirms all underpayments and discovered the problem when reviewing cases with underpayments. According to BOND Implementation Team leadership, the contractor responsible for developing and maintaining BSAS developed a solution that fixed the underlying problem for all beneficiaries except concurrent beneficiaries (concurrently receiving SSDI and SSI benefits) and those dually entitled to SSDI benefits. The work unit must manually process these cases.

The ORDES BOND work unit staff will correct the BSAS errors, but at the expense of delays in other BOND processes. Work unit staff are currently working to identify cases they believe were processed incorrectly and were unsure of the pervasiveness of the issue. Although the work unit reviewed and confirmed underpayments, overpayments were not subject to the same scrutiny and beneficiaries may have been informed of inaccurate overpayment amounts. Once identified, work unit staff will correct the errors, but they reported that this process will take a long time. As with delays in work CDR processing at the work unit, staff attribute delays in identifying and correcting cases of inaccurate overpayments to insufficient resources relative to their workload. Furthermore, the work unit staff time spent manually processing the cases that could not be automated via BSAS and resolving the previous years' errors could otherwise be spent processing work CDRs.

5.2.5. Delayed or Incorrect Benefit Adjustment and Potential Effects

A number of challenges have hindered timely and proper benefit adjustment in BOND. Though many initial implementation challenges have been addressed, some problems remain. In this section, we consider the potential consequences of improper benefit adjustments for beneficiaries' perceptions and behaviors.

Improper Payments

Delays and other problems with benefit adjustment may lead to improper payments—when SSA pays beneficiaries more or less than they were entitled to, a difference that SSA later reconciles. In addition to improper payments due to delays in beneficiary reporting, work CDR processing, and automated reconciliation, improper payments can arise from several factors including revised AEEs, incorrect AEEs, and errors in BSAS.³⁵ Improper payments fall into three categories:

- *Underpayments* occur when beneficiaries receive less in benefits than they were entitled.³⁶ When SSA recognizes the underpayment, the agency issues beneficiaries a lump-sum check.

2013 to October 2013 could result in additional accumulation of improper payments between August and October 2013. If a working beneficiary was aware of his status in August 2013, he could submit an AEE for the remainder of the year and avoid accumulating additional incorrect payments for that year.

³⁵ See Derr et al. (2015) Section 5.5.2 for a more detailed discussion of the genesis of payment errors.

³⁶ We do not distinguish underpayments that accrue within a calendar year from those that occurred in a previous calendar year because, from the beneficiary's point of view, it is difficult to distinguish between the two. SSA handles both in the same manner: reconciling the difference by issuing the beneficiary a lump-sum check.

- *Overpayments* occur when beneficiaries receive more benefits than they were entitled in a previous calendar year and SSA does not recognize the discrepancy until after the end of the year. When SSA identifies the overpayment, it requires beneficiaries to repay the amount owed, either by check or through withheld future benefits. Beneficiaries have the right to appeal an overpayment and SSA may agree to set up a repayment plan to mitigate financial hardship.
- *Incorrect overpayments* occur when beneficiaries receive more in benefits than they were entitled in the current calendar year and SSA recognizes the discrepancy before the end of the year. In these cases, to recoup the payment, benefit checks are withheld immediately until the payment is recovered or until the end of the calendar year, whichever comes first.³⁷

Control subjects are also subject to improper payments, but the prevalence and size of improper payments to control subjects may differ from those of treatment subjects.³⁸ Because of the offset, treatment subjects encounter smaller overpayments and underpayments than do beneficiaries subject to current law, holding duration of the period and the size of the full benefit constant.

Differences in the prevalence and size of improper payments by BOND group may also arise because of the difference in the accounting periods used for the treatment and control beneficiaries. Control beneficiaries are subject to current law, under which SSA uses a monthly accounting period, while SSA adjusts the benefits of treatment subjects based on an annual accounting period. Because SSA continues to pay treatment subjects' benefits monthly, each month's benefits are based on average monthly earnings over the entire year. Beneficiaries may have difficulty estimating annual earnings with accuracy, which leads to improper payments. Because the accounting period for determining monthly benefits is annual, changes in earnings as reflected in revised AEEs submitted after January (or after the first offset month, if later) affect benefits paid in previous months within the same calendar year, so the result is incorrect payments. Beneficiaries with inaccurate AEEs that are not corrected via revised AEEs will encounter over or underpayments. According to ORDES BOND work unit staff, the latter payment errors are likely to be small and were anticipated to occur in BOND. Refer to the BOND Process Study Report (Derr et al. 2015) for a detailed discussion of improper payments.

Beneficiary Experiences with Improper Payments

According to BOND staff, early in the demonstration benefit adjustment was often coupled with improper payments, as SSA and BOND staff were establishing processes, beneficiaries were learning about their reporting requirements, and SSA staff were contending with a large backlog of work CDR cases (Gubits et al. 2013, Derr et al. 2015). Indeed, in 2013, WIC and EWIC staff in seven of the ten BOND sites said that all or nearly all of their clients in the offset encountered an improper payment (Derr et al. 2015).

³⁷ At the end of each calendar year, incorrect overpayments are eligible to be reclassified as overpayments. According to ORDES staff, SSA will withhold benefits until a beneficiary with an incorrect overpayment either submits a new AEE after the start of a new calendar year or until SSA runs automated reconciliation for the previous year, whichever comes first. In addition, a beneficiary can request ORDES to process the overpayment earlier by asking for a beneficiary initiated reconciliation (BIR).

³⁸ We are currently exploring the use of SSA administrative data to quantify the prevalence and size of work-related overpayments and underpayments for BOND treatment and control subjects.

According to Implementation Team staff, because beneficiaries failed to report earnings or because of delays in work CDR processing, improper payments remained common among beneficiaries with first offset adjustments in 2014. The result in either case is usually a retroactive cessation date, which SSA often assigns many months after the beneficiary was first eligible for the offset. Beneficiaries with retroactive cessation dates who continue to engage in SGA typically face improper payments. WIC and EWIC staff reported that it seemed almost inevitable that offset users would have an improper payment—especially an incorrect overpayment—during their first year of the offset. Several Implementation Team members responsible for post-entitlement work agreed that it was rare for a beneficiary to enter the offset and not have an incorrect overpayment.

Following entry into the offset, improper payments are generally less frequent and smaller in size, according to WIC and EWIC staff. The potential size of improper payments is largest in the year the beneficiary is first notified of benefit adjustment. This is because the adjustment can encompass many years of improper payments, including months before enrollment in BOND. In subsequent years, adjustments generally encompass only one calendar year. In addition, improvements in the accuracy of AEEs have helped to facilitate proper adjustment, based on reports from both ORDES and Implementation Team staff.

Relative to previous years, in 2014, BOND staff were better able to help treatment subjects anticipate improper payments and their effects. In 2012, for example, WIC and EWIC staff struggled to identify and notify beneficiaries of pending improper payments (Gubits et al. 2013). In 2013, WIC and EWIC staff reported that they were generally able to alert beneficiaries to the potential for an improper payment (Derr et al. 2015). However, during 2014 interviews, Implementation Team staff noted that improper payments were a lingering source of confusion for some WIC and EWIC staff. Additional training and centralization of post-entitlement work helped improve this process. In addition, guidance that staff should discuss improper payments whenever they collect an AEE appears to have helped beneficiaries. According to WIC and EWIC staff, some beneficiaries are able to plan their expenditures to minimize financial difficulties when they are aware of pending improper payments (Derr et al. 2015).

Past reports have documented reported effects of overpayments and improper overpayments on beneficiaries including negative perceptions of the demonstration and its staff (Gubits et al. 2013) and financial hardship (Derr et al. 2015). During interviews in 2014, all participating WIC and EWIC staff indicated that the beneficiaries they serve generally have negative reactions to overpayments and incorrect overpayments. About one in three staff observed that beneficiaries generally subsequently recover from the initial negative reaction. However, for some beneficiaries, the negative reaction may persist and affect beneficiary behavior. According to a poll of WIC and EWIC staff, 43 percent had served a beneficiary who reduced earnings in response to an overpayment or incorrect overpayment. Implementation Team staff heard anecdotal reports of similar responses. This response is rarely in the beneficiary's financial interest because under the offset rules, beneficiaries' total income (benefits plus earnings) are reduced when earnings decline. We have no other information on the prevalence of this type of behavioral response. Of course, it is likely that beneficiaries subject to current law exhibit similar responses to improper payments, but in their case it might be in their financial interest to do so.

BOND staff suggested that negative reactions to overpayments and improper overpayments may be fueled in part by communication from SSA. WIC and EWIC staff reported that beneficiaries often find notices about improper payments to be confusing. Implementation Team staff added that these notices are

not only confusing but also alarming, although the same is true for overpayment notices for all SSDI beneficiaries. WIC and EWIC staff do not receive copies of the notices and instead often ask beneficiaries to send them a copy or read them the notice. If the beneficiary does not share the notice, WIC and EWIC staff counsel beneficiaries based on incomplete information. The same challenge applies for beneficiaries and WIPA-funded counselors under current law. SSA field office staff are responsible for assisting all SSDI beneficiaries with overpayments.

Underpayments also accrue to BOND subjects, but were generally not considered problematic by beneficiaries. The perception from ORDES staff was that as of early 2015, after the first benefit adjustment under the offset, underpayments are at least as likely as overpayments. WIC and EWIC staff reported that they generally encourage beneficiaries to submit AEEs that are slightly higher than their actual estimated earnings for the year, so that the beneficiaries will be less likely to encounter overpayments following automated reconciliation. The ORDES observation about the prevalence of underpayments is consistent with this guidance. In future rounds of data collection, we will gather additional information about beneficiaries' preferences related to AEEs and under versus overpayments.

5.3. Statistics on Offset Use and Benefit Adjustment

In this section, we present statistics on the progression of Stage 2 treatment subjects toward completing the milestones en route to an adjustment under the offset through December 2014. The analysis also considers beneficiary characteristics associated with benefit adjustment.

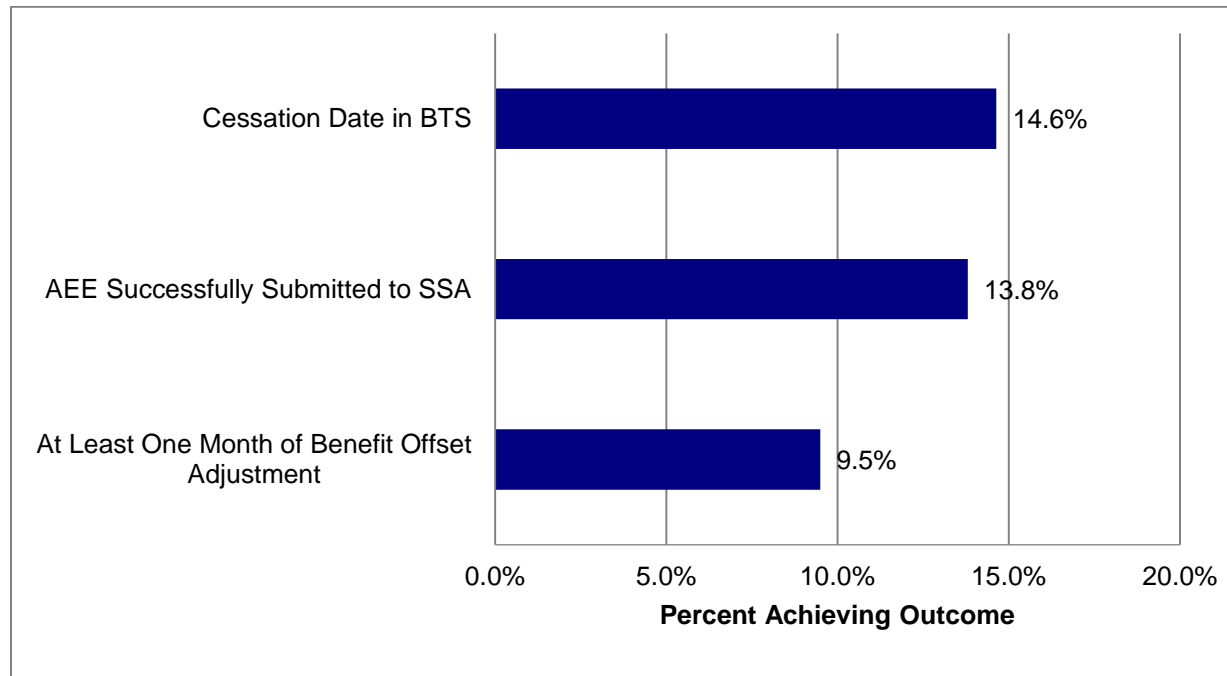
5.3.1. Pathway to Benefit Adjustment

By the end of 2014, SSA had adjusted the benefits of nearly one in every ten Stage 2 treatment subject. Another five percent of beneficiaries had reached significant milestones on the pathway to the benefit adjustment (Exhibit 5-2). According to data from December 2014, SSA had determined that 14.6 percent of Stage 2 treatment subjects had reached disability cessation due to SGA. Almost all of those subjects, or 13.8 percent of all Stage 2 treatment subjects, had successfully submitted an AEE to SSA. Finally, SSA had adjusted the benefits of 9.5 percent of Stage 2 treatment subjects.

Several factors may explain the differences between the number with cessation dates, the number who had submitted AEEs, and the number whose benefits had been adjusted. First, it takes time to reach a beneficiary and collect an AEE. Thus, beneficiaries who were contacted for an AEE, but have not completed it may help explain the small gap between those with a cessation date and those who submitted an AEE. Second, administrative delays could generate minor delays in the submission of an AEE or benefit adjustment. However, such delays are generally rare in the adjustment process after the assignment of a cessation date. Third, beneficiaries may reduce earnings below the SGA amount following the cessation month. BOND Implementation Team leadership believed this explanation was the most likely and cited a recent BTS statistic that 42 percent of beneficiaries in need of an AEE (e.g. beneficiaries with established cessation dates) earned below SGA in 2014, according to estimated earnings amounts as indicated on 2014 AEEs. We do not know the extent to which this reflects worsening medical problems or other changes in circumstances that prevent or discourage engagement in SGA versus the beneficiary's response to receiving personalized information about how their earnings will affect their benefits through an AEE. Furthermore, Exhibit 4-2 shows that 93 percent of Stage 2 treatment subjects were aware that, beyond the TWP, earnings above BYA result in some type of benefit reduction

(although 41 percent incorrectly thought that benefits would be entirely suspended). General knowledge of the potential for benefit reduction may impact beneficiary behavior.

Exhibit 5-2. Percentage of Stage 2 Treatment Subjects with Documented Steps towards Benefit Adjustment (through December 31, 2014)



Source: Analysis of BTS records, April 2015.

EWIC services were intended to increase beneficiary efforts to use the offset. An analysis of the impact of eligibility for EWIC services (that is, assignment to T22) on completing the three milestones on the pathway to benefit adjustment confirms that they have had such an effect. Exhibit 5-3 show estimates of impacts based on data through December 31, 2014. The exhibit shows that eligibility for EWIC services led to a 2.3 percentage point increase (15 percent) in the proportion of beneficiaries with a cessation date compared to eligibility for WIC services. Eligibility for EWIC services also increased the proportion of beneficiaries for whom an AEE was successfully submitted relative to WIC. Finally, eligibility for EWIC services had a 2.2 percentage point (24 percent) impact on the likelihood of benefit adjustment relative to WIC services. Together, the results suggest that the availability of enhanced benefits counseling accelerated actual adjustment of benefits. Although these results might reflect increased engagement in SGA, they might also reflect acceleration of processing times for those engaged in SGA. Chapter 7 presents estimates of the impacts of EWIC services on earnings above BYA and other earnings outcomes through 2013.

Exhibit 5-3. EWIC versus WIC Impact Estimates on Steps towards Benefit Adjustment (through December 31, 2014)

Steps Toward Benefit Adjustment	Average Outcome with Offset and EWIC (T22)	Average Outcome with Offset and WIC (T21)	Estimated Impact of EWIC Instead of WIC Given Offset (T22 vs. T21)
Cessation Date in BTS	17.7	15.4	2.3* (1.1)
AEE Successfully Submitted to SSA	16.9	14.1	2.8** (0.9)
At least one month of Benefit Offset Adjustment	11.2	9.0	2.2** (0.9)

Source: Analysis of BTS records from April 2015, with covariates from the baseline administrative SSA records, and the Stage 2 baseline survey used impact analysis regression equations.

Notes: Weights are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: T22 = 3,089; T21 = 4,936.

*/**/** Impact estimate is significantly different from zero at the 10, 5, and 1 percent levels, respectively, using a two-tailed t-test with 9 degrees of freedom.

Progression along the pathway to benefit adjustment varies substantially with beneficiary characteristics. Exhibit 5-4 compares baseline characteristics of Stage 2 treatment subjects (T21 and T22 combined) for three groups based on achievement of offset milestones through December 2014. The groups are: (1) those who did not have a benefit offset adjustment (“non-users”) and did not have a disability cessation date, (2) non-users with a cessation date, and (3) those who had at least one month of benefit adjustment. Compared to both non-user groups combined, at the time of enrollment, beneficiaries with an offset adjustment were more likely to be in one of the three youngest age groups (20-29, 30-39, and 40-44 years), have primary impairments of neoplasms, be the direct payee of benefits, work for pay, have a postsecondary degree or higher, and report good or very good health. Those with an offset adjustment also resided in counties with lower than average unemployment rates. These findings are consistent with a comparison of beneficiaries who had benefits adjusted through April 2013 to non-offset users in Gubits et al. (2013).

With regard to several characteristics, beneficiaries with an offset adjustment were more similar to non-users without a cessation date than to non-users with cessation dates. First, non-users without a cessation date and those with an offset adjustment both received SSDI for an average of 52 months relative to about 92 months for non-users with a cessation date. Second, non-users with a cessation date were more than 10 percentage points more likely to have a primary impairment of mental disorders than beneficiaries in the other two groups. Finally, non-users with a cessation date had the lowest AIME and monthly SSDI benefit amounts of the three groups. These results are surprising because we would expect the two groups of beneficiaries with a cessation date to be most similar. It is unknown what share of non-users with cessation dates will eventually have benefits adjusted and whether non-users with cessation dates are more likely to have benefit adjustments than non-users without cessation dates.

Exhibit 5-4. Stage 2 Treatment Beneficiary Characteristics by Steps towards Benefit Adjustment (through December 31, 2014)

Baseline Characteristic	Non-offset user, no cessation date (1)	Non-offset user, with cessation date (2)	Benefit offset adjustment by 12/31/14 (3)	P-value
Number of Beneficiaries	6,850	413	762	
Gender				
Male	49.5%	47.0%	46.9%	0.328
Female	50.6%	53.0%	53.2%	
Age				
20–29 years	5.5%	6.3%	9.2%	0.000
30–39 years	13.7%	18.6%	19.3%	
40–44 years	10.9%	14.8%	14.8%	
45–49 years	15.9%	17.7%	16.9%	
50–54 years	24.6%	20.8%	18.9%	
Over age 55	29.5%	21.8%	20.9%	
Mean age (years)	47.9	46.0	45.0	0.001
Primary Impairment				
Neoplasms	3.8%	4.6%	7.6%	0.000
Mental Disorders	30.4%	41.9%	31.8%	
Back or Other Musculoskeletal	26.5%	18.9%	22.3%	
Nervous System Disorders	6.6%	5.6%	7.4%	
Circulatory System Disorders	6.8%	3.6%	4.9%	
Genitourinary System Disorders	2.8%	3.2%	2.5%	
Injuries	4.3%	3.9%	6.2%	
Respiratory	2.4%	1.7%	2.0%	
Severe Visual Impairments	2.9%	1.5%	2.4%	
Digestive system	2.0%	2.2%	3.0%	
Other impairments	11.6%	13.1%	10.1%	
Length of SSDI Receipt				
Short duration (36 months or less)	65.8%	32.2%	60.9%	0.000
Number of Months Received SSDI	51.6	91.7	52.1	0.000
Benefit Amount and Status				
Monthly SSDI Benefits (\$)	\$1,082	\$1,069	\$1,174	0.000
AIME (May 2011) (\$)	\$1,867	\$1,740	\$2,195	0.000
Disabled adult child (DAC)	3.5%	2.4%	2.1%	0.029
Disabled widow beneficiary (DWB)	1.2%	0.7%	0.7%	0.273
Dually-entitled disabled adult child	1.1%	1.5%	1.1%	0.761
Dually-entitled disabled widow beneficiary	0.6%	0.7%	0.4%	0.571
Payee is other than self	8.1%	6.8%	5.5%	0.008

Baseline Characteristic	Non-offset user, no cessation date (1)	Non-offset user, with cessation date (2)	Benefit offset adjustment by 12/31/14 (3)	P-value
Local characteristics				
County unemployment rate (April 2011)	8.7%	8.4%	8.3%	0.004
County employment rate for people with a disability (2010)	34.3%	35.3%	35.5%	0.064
Employment status at baseline				
Working for pay	18.6%	40.1%	68.7%	0.000
Looking for work	22.8%	26.1%	17.2%	0.003
Education				
2- or 3-year postsecondary degree	14.6%	17.4%	17.7%	0.059
Bachelor's degree or higher	15.6%	20.2%	25.1%	0.025
Health status at baseline				
Health is fair or poor	66.0%	60.6%	51.6%	0.025

Sources: Analysis of the BTS and BOND Stage 2 Baseline Survey.

Note: p-values shown are from statistical tests of differences in percentages across the three groups. Groups of mutually-exclusive characteristics were tested for differences with chi-squared tests. Single characteristics not part of a mutually-exclusive group were tested for differences by F-tests. The Chi-squared statistic from omnibus statistical test of difference between groups across all characteristics is 1,362.14, with a p-value of 0.000.

In a regression analysis of Stage 2 treatment subjects, many beneficiary characteristics are predictive of benefit adjustment in or before December 2014 (Exhibit 5-5). Readers should interpret these results with caution because in a regression with this number of predictors we would expect a few coefficients to be statistically significant by chance—that is, significant even though the corresponding actual coefficient for the underlying population is zero.

Age was a statistically significant predictor of benefit adjustment. For example, beneficiaries ages 20-29 were 8.9 percentage points more likely to use the offset relative to beneficiaries ages 55 and older, holding other characteristics constant. Similarly, primary impairments of neoplasms, back or musculoskeletal disorders, injuries, respiratory conditions, and digestive system conditions were all associated with a statistically significant increase in the likelihood of offset use relative to beneficiaries with other impairments, and beneficiaries who reported fair or poor health were 3.1 percentage points less likely to have a benefit adjustment than those in good or very good health. Beneficiary attainment of a post-secondary degree (but no bachelor's degree) was associated with a 1.1 percentage point increase in offset use relative to beneficiaries with education less than a postsecondary degree. Disabled adult child beneficiaries and beneficiaries with representative payees were also less likely to have a benefit adjustment than beneficiaries without those characteristics.

Employment and past earnings were also predictive of benefit adjustment, other things constant. Baseline employment was associated with a 23.0 percentage point increased likelihood of benefit adjustment relative to those not looking for work and not working. Holding other characteristics constant, those looking for work were 4.0 percentage points more likely to have a benefit adjustment relative to those not looking for work and not working. Previous earnings was also a strong indicator of future benefit

adjustment. An increase in AIME by \$1,000 was associated with a 1.9 percentage point increase in the likelihood of benefit adjustment. In addition, increases in the county employment rate for people with disabilities as measured by the American Community Survey were associated with a small but statistically significant increase in benefit adjustment.

Exhibit 5-5. Predictors of Benefit Adjustment for Stage 2 Treatment Subjects (through December 31, 2014)

Predictor of Benefit Offset Adjustment	Coefficient (1)	Standard Error (2)	P-value
Gender			
Male	-0.86	0.66	0.223
Age			
20–29 years	8.93	2.32	0.004***
30–39 years	6.29	1.27	0.001***
40–44 years	5.76	1.32	0.002***
45–49 years	3.61	1.18	0.014**
50–54 years	1.46	0.55	0.025**
Over age 55	0.00	--	--
Primary Impairment			
Neoplasms	6.49	3.07	0.064*
Mental Disorders	0.53	1.66	0.758
Back or Other Musculoskeletal	2.37	0.92	0.029**
Nervous System Disorders	0.66	1.60	0.691
Circulatory System Disorders	1.76	1.62	0.307
Genitourinary System Disorders	-1.57	1.41	0.292
Injuries	4.50	1.70	0.027**
Respiratory	3.60	1.70	0.064*
Severe Visual Impairments	-1.96	2.06	0.368
Digestive system	5.85	2.45	0.040**
Other impairments	0.00	--	--
Length of SSDI Receipt			
Short duration (36 months or less)	-1.18	1.16	0.336
Months Received SSDI	-0.07	0.11	0.501
Benefit Amount and Status			
Monthly SSDI Benefits (\$1,000)	-0.19	0.11	0.120
AIME (May 2011) (\$1,000)	1.91	0.20	0.000***
Disabled adult child (DAC)	-6.66	2.24	0.015**
Disabled widow beneficiary (DWB)	2.05	2.21	0.377
Payee is other than self	-4.95	1.29	0.004***
Local characteristics			
County unemployment rate (April 2011)	0.09	0.10	0.365
County employment rate for people with a disability (2010)	0.04	0.01	0.015**

Predictor of Benefit Offset Adjustment	Coefficient (1)	Standard Error (2)	P-value
Employment status at baseline			
Working for pay	22.99	1.91	0.000***
Looking for work	4.03	0.52	0.000***
Not working for pay or looking for work	0.00	--	--
Education			
2- or 3-year postsecondary degree	-1.06	0.56	0.089*
Bachelor's degree or higher	1.34	1.75	0.463
Less than a 2-year postsecondary degree	0.00	--	--
Health status at baseline			
Health is fair or poor	-3.06	0.75	0.003***

Source: Analysis of BTS records, baseline administrative SSA records, and Stage 2 Baseline Survey data.

Notes: The model was estimated using a clustered linear regression model without weights. The outcome variable is an indicator of whether the recipient has at least one month of Offset use through December 2014. Adjusted R-Squared: 0.1342. Model F-statistic is 20.66, p-value <.0001.

Sample size: 7,967.

***/**/* indicate statistical significance at the 0.01/0.05/0.10 levels.

5.3.2. Offset Use and Benefit Adjustment

In the previous section, we presented details about factors related to offset use and pre-offset milestones for beneficiaries who had used the offset in one or more months by December 2014. In this section, we examine the timing of benefit adjustment compared to eligibility for benefit adjustment. Exhibit 5-6 presents statistics on first month of offset eligibility and the first month SSA adjusted benefits, based on SSA administrative data and BTS records:³⁹

- The upper red line in the exhibit, *cumulative percentage of Stage 2 treatment subjects eligible for an adjustment under the offset*, shows the cumulative percentage of treatment subjects eligible for offset use in at least one month (i.e. earned above BYA during the participation period), based on administrative records through December 31, 2015. For instance, the value of 1.8 percent in January 2012 means that, as of December 2015, SSA had identified 1.8 percent of Stage 2

³⁹ This report discusses BOND implementation through 2014. This analysis is an exception. As of the revision of this document, data about the number of benefit adjustments SSA has made, and about the number of beneficiaries known to have begun a period of offset use, are available through December 2015. Exhibit 5-6 provides the BOND team's most up-to-date information on offset use and benefit adjustment through December 2014 and beyond. Differences between this exhibit and similar exhibits for this period in other reports (Derr et al. 2015) are due to retroactive benefit adjustments.

The data on cumulative eligible offset users are based on a monthly extract from SSA's Master Beneficiary Record (MBR) as well as calculation and verification of first offset month by SSA staff. The data on cumulative percentage with adjustments in 2011 and 2012 are from manual updates made by SSA staff to BTS and were verified by SSA staff. The data on cumulative adjustments in 2013 through 2015 are from a combination of BTS, BSAS, and MBR data.

treatment subjects as eligible for the offset in at least one month for the period from April 2011 to January 2012. The cumulative percentage known to be eligible at different points over the April 2011 to December 2015 interval will continue to increase, until SSA completes all retroactive adjustments for this period. This is especially true in the later months of the interval.

- The lower blue line in the exhibit, *cumulative percentage of Stage 2 treatment subjects with an adjustment*, provides information on the months in which initial benefit adjustments under the offset were made—usually later than the first month of offset qualification. It represents the percentage of beneficiaries whose benefits actually had been adjusted under the offset as of the month indicated based on administrative data through December 2015. As indicators of actual adjustments made, the points plotted on this line will not change as SSA processes additional adjustments retroactively.

The main reason the two series differ is that SSA's initial adjustment of benefits for a beneficiary generally occurs some amount of time after the first month of eligibility. Future automated reconciliations will raise the upper red line, especially for the late 2014 through 2015 interval, and hence introduce a larger gap between the red and blue lines over this interval.⁴⁰

An example is helpful in distinguishing between the two cumulative series. If a beneficiary had a cessation date and sufficient earnings to be first eligible for the offset in May 2012, but SSA adjusted his or her benefits in September 2013 (retroactively), the beneficiary would be included in the upper line starting in May 2012 and in the lower line starting in September 2013. In contrast, if the same beneficiary had entered through the front door and SSA had first adjusted his or her benefits under the offset contemporaneously with the first month of offset use (May 2012), the beneficiary would be included in both lines starting in that month. The value of the first series is slightly higher than the value of the second series in December 2014 because the first series reflects retroactive adjustments made in the first four months of 2015 while the second includes only cases for which adjustments were made by December 2014.

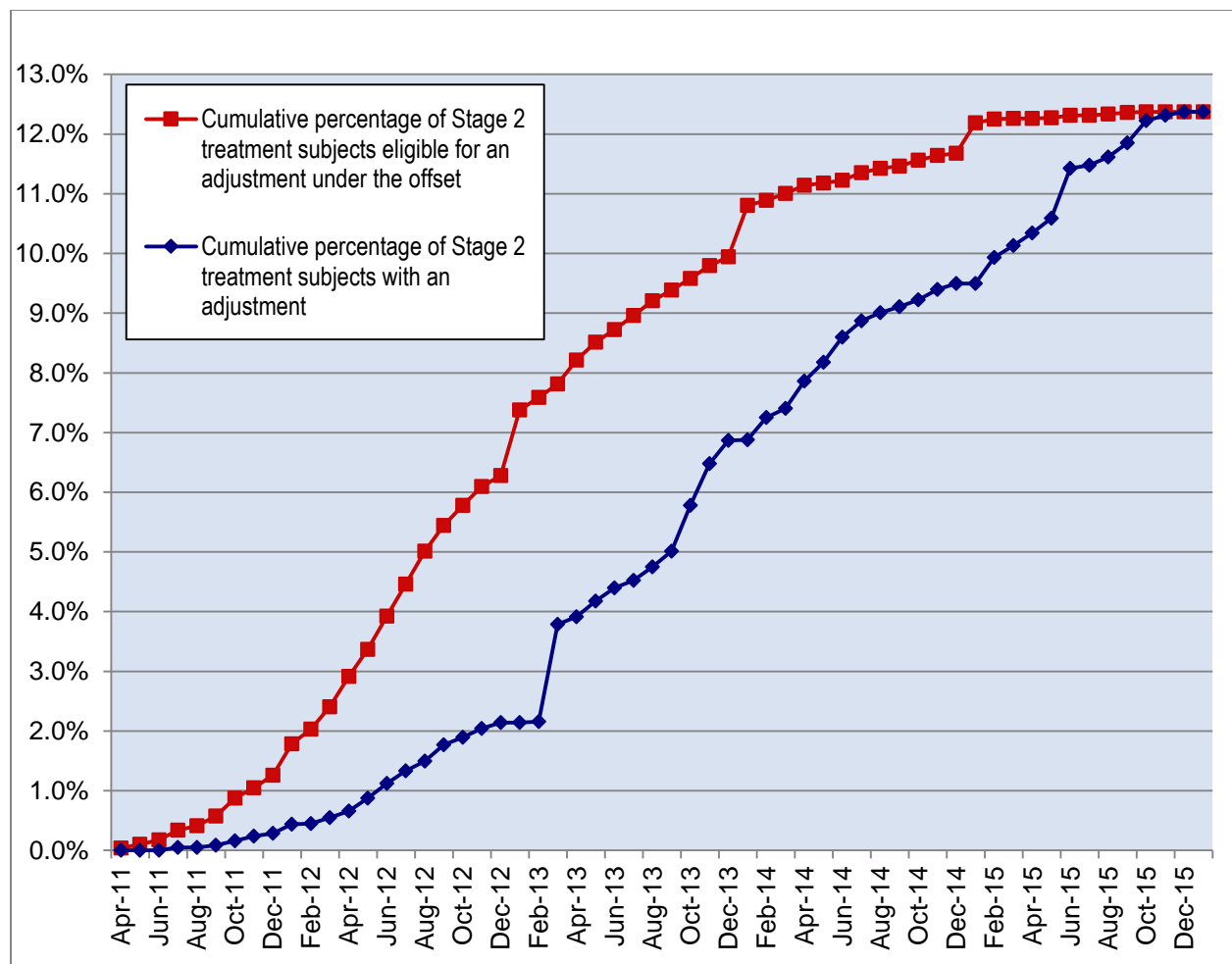
Based on data through the end of December 2015, the number of Stage 2 treatment subjects eligible for the offset grew throughout the demonstration period. The data reveal that 5.5 percent of Stage 2 treatment subjects (437 beneficiaries) were eligible for the offset for the first time before September 2012 (Exhibit 5-6)—the last month of the enrollment period. The gradual increase before that month reflects the fact that Stage 2 enrollment spanned a period of about 18 months, and treatment subjects were not eligible for the offset until at least one month after they enrolled (later if they had not already completed the GP).⁴¹

⁴⁰ The value of the both series are equal in December 2015 because they are from data in which all those known to have used the offset by December 2015 had their benefits adjusted by December 2015. SSA continues to retroactively identify additional T1 offset users in 2015 or earlier, so the cumulative percentage of T1 subjects who used the offset during the period (including in December 2015) will increase, but the percentage of initial adjustments during the period will not change, by construction.

⁴¹ The increase from December to January in all years will always be slightly larger than other month-to-month increases, due to cases where the offset eligibility is determined from IRS records rather than earnings summaries submitted by the beneficiary or where a beneficiary did not earn above BYA in the first year of offset eligibility but did so in a subsequent year. For example, consider a beneficiary who met the eligibility criteria for offset use, except for earnings above BYA, in 2013. If that beneficiary earned above BYA in 2014,

The percentage of beneficiaries eligible for the offset increased to 12.2 percent (978 beneficiaries) of Stage 2 treatment subjects by December 31st, 2014.⁴² This increase reflects the fact that there is a continuous flow of working beneficiaries entering into the BOND participation period and, upon completion of the GP, becoming eligible to use the offset.

Exhibit 5-6. Cumulative Stage 2 Treatment Subjects Eligible for an Adjustment Under the Offset and With an Adjustment, Based on Data Through December 31, 2015



Source: Monthly extracts from SSA’s MBR.

Note: Although the balance of this report covers the Stage 2 implementation period through December 2014, this exhibit incorporates most recent data on offset use are available through December 2015.

because of the annual accounting period, regardless of the month in which his work began, SSA would consider the beneficiary as eligible for adjustment under the offset starting in January 2014.

⁴² Statistics in early sections indicate that SSA had *adjusted the benefits* of 9.5 percent of Stage 2 treatment subjects by December 2014.

The number of Stage 2 treatment subjects with an adjustment also grew throughout the demonstration period, but with a different trajectory. A comparison of the lower line in Exhibit 5-6 to the corresponding upper line shows that SSA was delayed in making the first adjustments for many Stage 2 offset users. For example, in February 2013, 609 beneficiaries (7.6 percent) were eligible for the offset but SSA had only adjusted the benefits of 173 beneficiaries (2.2 percent). In the next month, March 2013, the number eligible for the offset grew marginally to 627 beneficiaries (7.8 percent) while the number of beneficiaries with an adjustment jumped to 304 (3.8 percent). The percentage of new users entering through automated reconciliation has declined since 2011. For Stage 2, 45 percent of new users for 2011 entered via the 2011 automatic reconciliation that occurred in January and February of 2013, compared to 26 percent for 2012 (conducted in four batches in September and October 2013), respectively (Derr et al. 2015).

5.4. Conclusion

Benefit adjustment for Stage 2 treatment subjects has steadily increased from 431 beneficiaries (5.4 percent) at the end of Stage 2 enrollment in September 2012 to 762 beneficiaries (9.5 percent) at the end of 2014. Stage 2 assesses the impact of the benefit offset on those most likely to use it and the Stage 2 sample includes recruited and informed volunteers. Within that sample, several beneficiary characteristics were predictive of benefit adjustment in or before December 2014. Young age, good or very good general health status, primary impairments including back or musculoskeletal disorders and injuries, baseline employment, high AIME, and residence in a county with a high employment rate among people with disabilities were all associated with an increased likelihood of benefit adjustment, holding other characteristics constant. In addition, eligibility for EWIC services had a 2.2 percentage point impact on the likelihood of benefit adjustment relative to eligibility for WIC services.

Because adjustments for a large number of beneficiaries are still in process, we expect the percentage with a benefit offset adjustment to increase in the future, above the 9.5 percent with benefit adjustments through December 2014. Indeed, SSA adjusted the benefits of an additional 0.1 percent of beneficiaries who were eligible for the offset in or before December 2014 as of December 31st, 2015. Beneficiaries who become eligible for the offset in or after 2015 will likely experience shorter delays than their predecessors because of recent process improvements—a change in the guidance on when to initiate work CDRs, a more effective method for identifying beneficiaries in need of a work CDR, and improvements in the quality of AEEs.

The rate of benefit adjustment is contingent upon beneficiaries' first establishing offset eligibility. By design, beneficiaries must work at a substantial level for more than 12 months (the nine-month TWP months and three-month GP) before they are eligible for offset use. Of course, the establishment of offset eligibility depends on working beneficiaries' ability and willingness to engage in sustained SGA-level work. Even for SSDI beneficiaries who are able and willing to earn above SGA, it may still take time before a beneficiary is able to use the offset for a variety of reasons. First, those interested, but unprepared, may require time to obtain counseling or employment-related services, address a variety of personal issues, or find an SGA-level job. Second, those who, in response to BOND, promptly initiated SGA-level work may need time to work towards total annual earnings that exceed BYA.

Once offset eligibility is established, benefit adjustment may be delayed for various reasons. We have identified three main sources of delays in proper benefit adjustment. First, the failure of some offset-eligible beneficiaries to report their earnings delays the start of the benefit adjustment process. Second, lags in the processing of work CDRs, once SSA recognizes the need for a work CDR, delay the

determination of when eligibility began and provides a longer time frame over which the beneficiary may accrue improper payments. These lags are the result of (1) a tiered workflow for processing work CDRs between SSA field offices or processing centers and the ORDES BOND work unit and (2) high work CDR caseload responsibility at the work unit. Many (44 percent) BOND treatment subjects in the ORDES work CDR queue encountered CDR processing times longer than 270 days. This delay alone prevents timely benefit adjustment under the offset. Finally, BSAS deficiencies have caused substantial delays in automated reconciliation, thereby delaying benefit adjustment for some first-time offset users.

The qualitative data suggest that offset use frequently leads to the occurrence of improper payments. There are a number of reasons for this. First, the retroactive nature of most initial benefit adjustments means that improper payments are often made during the first months of offset use. Second, ongoing offset users incur improper payments of some amount whenever their actual earnings differ from the amount they estimated at the beginning of the year. Finally, BSAS errors in benefit calculation have resulted in some payment errors to offset users.

As discussed in Section 4.1, the BOND logic model posits that beneficiaries need to understand the benefit offset in order to change their behavior in response to the incentive. Delays in delivery of the incentive via the adjustment process and the occurrence of improper payments are noteworthy because they may weaken beneficiary understanding of how the offset works.

6. Methodology and Context

With this chapter, we shift from the process analysis results presented in Chapters 3 through 5, to presentation of impact analysis results in Chapters 7 and 8. As an introduction to that presentation of impact analysis results, this chapter presents an overview of the data sources and methodology for the impact analysis. The discussion here is similar, but not identical, to the discussion in the Stage 2 Snapshot Report (Gubits et al 2014).

The goals for the Stage 2 evaluation are to learn about the impacts of the benefit offset for those most likely to use it (recruited and informed volunteers) and to determine the extent to which significant enhancements to the basic BOND-focused work incentives counseling affect offset utilization and impacts. For practical reasons, the design restricted the beneficiaries in Stage 2 to those most likely to use the offset. Specifically, attainment of the Stage 2 objectives requires more intensive data collection and more complex service delivery than is required for Stage 1. Restricting Stage 2 eligibility to those most likely to use the benefit offset reduces the sample sizes required for Stage 2 groups from tens of thousands to thousands.

Two aspects of this strategy for selecting the sample ensured that Stage 2 subjects would be likely to use the offset. First, Stage 2 sample does not include concurrent beneficiaries. The interaction between SSI and SSDI substantially diminishes the value of the SSDI offset to concurrent beneficiaries, so the design team expected that relatively few SSI beneficiaries would use the SSDI offset. Second, in contrast to the Stage 1 sample (which is randomly selected from all eligible SSDI beneficiaries), the Stage 2 sample is composed of self-selected volunteers. The Implementation Team recruited randomly selected eligible SSDI-only beneficiaries who responded to a solicitation to participate in BOND.

A key assumption of the design is that interest in using the offset led to the decision to volunteer for the study, and that this interest means that Stage 2 subjects will be more likely to use the offset than the average Stage 1 subject. Consistent with this assumption, as of May 1, 2015, 2.3 percent of the T1 group had at least one month of offset use, compared to 10.6 percent of the combined Stage 2 treatment groups.

This report presents estimates of the impact of the two Stage 2 treatment groups on both administrative outcomes and self-reported outcomes. Administrative data for calculating earnings and benefit impacts were available through calendar year 2013. These administrative data sources and their uses are listed below:

- 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 2013 when SSA extracted them for this report.
- Benefit outcomes are measured from SSA's Payment History Update System (PHUS) for SSDI and the Supplemental Security Record (SSR), for SSI.⁴³ The benefit records were complete.

⁴³ Because the data are collected by the IRS and are therefore subject to IRS access rules, SSA staff have direct access to MEF data, but contractors do not. Consequently, qualified SSA staff accessed the data, submitted programs developed by the BOND Evaluation Team to estimate impacts, reviewed output to ensure that it

- We supplement the administrative data with self-reported outcomes measured using the BOND 12-Month Interim Survey. The response rate of the 12-Month Interim Survey was 84 percent and the median timing of response was 13 months after study enrollment.
- We supplement the analysis with self-reported outcomes measured using the BOND Baseline Survey. Study participants completed the Baseline Survey just before they were randomized to one of the treatment arms. The response rate of the Baseline Survey was 99.3 percent.⁴⁴ In this report, we use the Baseline Survey to describe participants' pre-BOND employment status, earnings, health status, and understanding of SSDI benefits. We also use baseline survey data for covariates in the impact regressions (Chapter 6 and Appendix A).

We initially specified the methodology for the impact analysis in Bell et al. (2011). We later refined for the *First-Year Stage 1 Snapshot Report* (Stapleton et al. 2013). However, in the preparation of the 2014 Stage 2 Snapshot Report, a methodological issue arose having to do with the estimation of standard errors appropriate for statistical inference at the national level. As a result the 2014 Stage 2 Snapshot Report presented results representative of only the 10 BOND sites, rather than results representative at the national level. In the second half of 2014, the BOND Evaluation Team developed and tested an approach to estimate standard errors appropriate for statistical inference at the national level. We expect that this approach will be used in all future BOND reports. It was used for the first time in the Third Year Stage 1 Snapshot Report (Wittenburg et al. 2015) and used for the first time for Stage 2 impacts in this report. In addition to the changes related to producing nationally generalizable impact estimates, we also made a correction to the “within-site” component of the analytic weight—a change with no substantive consequence for the findings. Appendix A.1 summarizes these methodological changes. Appendix C of this report presents the previously reported First and Second Year impacts and standard errors for Stage 2 administrative outcomes re-estimated, this time using the new approach.

We initially specified the administrative outcomes for the impact analysis in Bell et al. (2011) and every evaluation report has used these specifications. This report uses the same measures, but clarifies that, in all reports, the administrative earnings measure includes only “Social Security earnings.” Social Security earnings are earnings that are taxable for Social Security purposes.⁴⁵ About 6 percent of the U.S. workforce holds jobs not covered by Social Security taxes. Furthermore, Social Security earnings are capped at a maximum taxable amount, \$113,700 for 2013. In 2013, 0.01 percent (one one-hundredth of a percent) of all Stage 2 subjects had earnings at or above \$113,700. Beneficiaries who are earning at or above that amount are unlikely to have a behavioral response to the offset.

Non-covered jobs constitute a larger omission. It is not feasible for this evaluation to obtain a more comprehensive measure of earnings from administrative data. As a result, reported estimates of earnings, employment, and the proportion with earnings above BYA have a small downward bias. The impact of the offset would be downwardly biased if some who are encouraged to work choose jobs not covered by

complied with privacy requirements, and then transmitted the output to the evaluation team. The MEF earnings data are updated annually. The 2013 earnings data for this report were extracted in February 2015.

⁴⁴ All 12,954 BOND subjects enrolled in Stage 2 completed a baseline survey. However, for 85 cases (less than one percent of the Stage 2 sample), the baseline data were corrupted on the laptops and could not be recovered. For these 85 cases, baseline data are not available.

⁴⁵ Appendix A.2 describes the earnings data sources in more detail.

Social Security. On the other hand, the estimate could have a slight upward bias due to the fact that the offset may induce some people with under the table earnings to report them. Measures of weekly earnings and employment taken from survey data should not be subject to the same source of bias (though they are subject to other biases; in particular, recall bias and non-response bias).

The subsections of this chapter review the outcome definitions, anticipated impacts, estimation methodology, and analysis sample. Additional details on the estimation methodology are available in Appendix A.

6.1. Administrative Outcome Definitions and Theoretical Impacts

The nine administrative outcomes for this report include two confirmatory outcomes (total earnings and total SSDI benefits paid *in* 2013) and seven exploratory outcomes (related to employment and benefits). The exploratory earnings outcomes include indicators for earnings in excess of each of three annual earnings thresholds defined by multiples of BYA (one, two, and three times BYA) and an indicator for employment during 2013 (defined as any earnings in 2013). The exploratory benefit outcomes include number of months with SSDI payments, total SSI benefits paid, and number of months with SSI payments⁴⁶—each in 2013. In the discussion that follows, we consider the expected direction of benefit offset impacts on these outcomes, abstracting from administrative factors that could themselves influence the impacts. We then turn to a discussion of administrative factors and their potential influence on impacts.

Although the goal of BOND is to test whether eliminating the SGA cash cliff and replacing it with the \$1 for \$2 offset ramp would increase return to work and earnings, and reduce beneficiary's reliance on SSDI benefits (Bell et al. 2011), the theoretical direction of impacts of the benefit offset on mean earnings and benefits is ambiguous. As described in detail in Bell et al. (2011), this ambiguity arises because the incentives created by the benefit offset vary with what the beneficiary's earnings would be under current law. T21 and T22 subjects who would have had no earnings or earnings below BYA under current law are expected, on average, to have higher earnings and lower SSDI benefits under the benefit offset. Conversely, some T21 and T22 subjects who would have had earnings above BYA under current law are expected to have lower mean earnings and higher mean SSDI benefits under the benefit offset.⁴⁷

The impact estimates from a random assignment study (such as those reported below) combine those in both groups. In order to estimate a positive impact on the mean earnings, the positive impact for those whose earnings would be less than BYA under current law would have to be larger than the expected negative impact for those who would earn more than BYA under current law.

⁴⁶ Although eligibility criteria for Stage 2 required that beneficiaries not be receiving SSI benefits at the time the study team determined eligibility (in the first six months of 2011), Stage 2 subjects could potentially become SSI recipients (for example, after spending down their assets enough to meet the resource test). Therefore, SSI benefits are included as an outcome variable.

⁴⁷ Empirically, there is evidence that some high-earning beneficiaries will reduce their earnings, but not reduce employment. Weathers and Hemmeter (2011) found evidence of a reduction in earnings by beneficiaries earning above SGA before random assignment in the Benefit Offset Pilot Demonstration.

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

Unlike for earnings and benefits paid (just discussed), theory does predict the signs of the impacts for five other outcomes, all of them exploratory. Theory predicts positive impacts on (i) employment, on the (ii) percentage of beneficiaries with earnings above BYA, and on (iii) months with SSDI payments. Theory also predicts negative impacts on (iv) SSI benefits and (v) months with SSI payments. These predictions can be verified by comparing the proportion of BOND subjects working below and above BYA in T21 and T22 versus the proportion in C2. As indicated earlier, for those who would have earnings below BYA under current law, theory predicts that the offset will increase both the percentage employed and the percentage of beneficiaries with earnings above BYA. Those who would have earnings above BYA under current law will have a stronger incentive to keep their earnings above BYA under the offset than they do under current law—even though some might work and earn less under the offset.

Theory also predicts that the impact on SSI benefits paid will be negative. The offset might have an impact on SSI payments to T21 and T22 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.⁴⁸

For the two remaining exploratory outcomes—earnings above two times BYA and earnings above three times BYA—it is not possible to predict the direction of impacts. For those treatment beneficiaries whose earnings would be less than BYA under current law, the offset is expected to have a positive average earnings effect, perhaps increasing the proportions with earnings above two times BYA and above three times BYA. Conversely, for those who would have had earnings above BYA under current law, the offset is expected to have a negative average earnings effect, perhaps decreasing the proportions with earnings above two times BYA and above three times BYA. Since the magnitudes of these opposing expected effects are not predicted by theory, it not possible to predict the overall direction of impact for any earnings threshold well above the BYA level. Exhibit 6-1 lists the administrative outcomes, provides a definition of each outcome, and indicates the predicted sign of the impact (positive, negative, or ambiguous).

⁴⁸ See Riley and Rupp (2012).

Exhibit 6-1. Definitions of Confirmatory and Exploratory Administrative Outcomes and Predicted Signs of Impacts

	Definition	Predicted Sign
Confirmatory Outcomes		
Total earnings in 2013	2013 Social Security earnings	?
Total SSDI benefits paid in 2013	Sum of SSDI benefit payments from January through December 2013; for SSDI workers, this includes benefits for dependent spouses and minor children, but not for DAC ^a ; for DAC and DWB, it includes only benefits payable to the DAC or DWB	?
Exploratory Outcomes		
Earnings Outcomes (January–December 2012)^b		
Employment in 2013	Indicator for any 2013 Social Security earnings	+
Earnings above BYA	Indicator for 2013 Social Security earnings greater than or equal to \$12,480 (non-blind subjects) or \$20,880 (blind subjects)	+
Earnings above 2 × BYA	Indicator for 2013 Social Security earnings greater than or equal to \$24,960 (non-blind subjects) or \$41,760 (blind subjects)	?
Earnings above 3 × BYA	Indicator for 2013 Social Security earnings greater than or equal to \$37,440 (non-blind subjects) or \$62,640 (blind subjects)	?
Benefit Outcomes (January–December 2012)		
Number of months with SSDI payments	Number of months with SSDI benefits paid above zero	+
Total SSI benefits paid	Sum of SSI benefit payment amounts from January through December 2013	–
Number of months with SSI payments	Number of months with SSI benefits paid above zero	–

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

^a For details on family benefits, see <https://www.socialsecurity.gov/pubs/EN-05-10024.pdf>; accessed May 27, 2014.

^b Earnings relative to BYA is based on earnings reported in the MEF.

6.2. Administrative Features of the Offset That Influence Impacts

The previous discussion abstracts from the administrative features of the benefit offset that the study team designed and implemented to facilitate use of the offset by T21 and T22 beneficiaries. As described in Bell et al. (2011), because these processes are necessarily different from current law processes, they are part of the T21 and T22 interventions being tested under BOND.

In the first years of BOND, the administrative factors most likely to affect outcomes concern the administrative processes leading to the adjustment of benefits—the special processes implemented for T21 and T22 subjects and the current processes that apply to C2 subjects. For T21 and T22 subjects, that process started shortly after their enrollment date, when they learned their random assignment status. Some of those eligible to use the offset informed the demonstration of their work activity as recommended and their benefits were eventually adjusted via an administrative process set up for that purpose. Others eligible to use the offset early did not contact the demonstration, however. Instead, SSA discovered their high earnings in its annual review of earnings reported to the IRS, and then initiated the process to adjust their benefits.

The benefits measures for this report are based on benefits paid *in* 2013, rather than benefits paid *for* 2013, which includes all future retroactive adjustments for 2013 benefits. These two measures will diverge according to the dollar value of retroactive adjustments made for 2013 benefits. Although this

dollar value is not yet known, we know that there must be retroactive adjustments of some dollar amount for the treatment subjects who did not pro-actively inform SSA of earnings above BYA. The BOND administrative data as of May 2015 show that 30 percent of T21 subjects and 24 percent of T22 subjects eligible to use the offset during 2013 did not have a benefit adjustment until after 2013. This implies that some adjustments to benefits paid to T21 and T22 subjects for 2013 are not reflected in benefits paid *in* 2013, and that there will be at least some discrepancy between benefits paid *in* 2013 and benefits paid *for* 2013.

The direction and size of the impacts of this administrative factor depend on how the processes for the T21 and T22 groups compare to the corresponding processes for C2 subjects. The most striking difference is that T21 and T22 subjects had to be notified about a change in the earnings rules before the benefit adjustment process could start, whereas C2 subjects were subject to rules that had been in place for many years. Also, T21 and T22 administrative processes had not been previously implemented on a large scale, resulting in start-up delays⁴⁹, whereas the C2 processes have been in place for many years.

One other administrative factor seems likely to have a positive impact on benefits paid *for* 2013, and possibly on benefits paid *in* 2013, but an ambiguous impact on 2013 earnings: the change from monthly to annual accounting. As discussed in Chapter 5 of this report, the justification for annual accounting was to simplify administration of the offset and to simulate the expected future accounting procedure should the benefit offset become national policy. While not the purpose of this change, the move to an annual accounting period is expected to help beneficiaries with highly variable earnings (for example, seasonal workers) to a significant degree. Under monthly accounting, earnings above SGA in any month reduce benefits for that month, but under annual accounting the benefit reduction for those same earnings might be smaller or zero because of earnings below the SGA amount in other months of the same year. Holding earnings constant, this administrative change is expected to increase the benefits paid to some beneficiaries; any increase in earnings due to this factor will reduce benefits (and correspondingly, any decrease in earnings will increase benefits). The theoretical sign of the impact of this administrative change on earnings is ambiguous.

6.3. Impact Estimation Methodology

SSA included Stage 2 in the demonstration in order to provide information about the impact of the benefit offset on beneficiaries who volunteer for the study and about the impact of EWIC vs. WIC. Given the self-selected nature of the Stage 2 sample, the impacts from Stage 2 do not generalize to the national SSDI caseload or to any easily identifiable subpopulation. Instead, the Stage 2 impacts presented in this report generalize to those who would have volunteered in the nation had they been solicited.

To estimate impacts, we compare mean outcomes for the T21, T22, and C2 groups to each other. For outcomes derived from administrative data, the sample means are weighted for differences in site-selection probabilities and in sampling rates into the solicitation pool across sampling strata. For outcomes derived from 12-month interim survey data, in addition to the factors for which the administrative data are weighted, the sample means are weighted for propensity to respond to the survey, in order to address the possibility of non-response bias. For both these types of outcomes, the means are

⁴⁹ This issue is described in Gubits et al. (2013) and Derr et al. (2015).

adjusted for the effects of small random differences in baseline characteristics.⁵⁰ The adjustments for differences in baseline characteristics also serve to reduce the standard errors. For each specific outcome, we test the null hypothesis of no impact. Each individual test uses a specified level of significance. For example, a 10 percent significance level means that if the null hypothesis is true, there is only a 10 percent chance that the test will mistakenly reject it.

The impact estimates are “intent to treat” estimates. For example, the benefit offset impacts capture the mean impact of the applicability of the benefit offset rules to the earnings of all T21 subjects, whether or not those subjects work and use the offset. Likewise, the benefit offset plus EWIC impacts capture the impact on all T22 subjects, whether or not they work. Hence, the impact estimates reflect “no impacts” for those treatment subjects who would not have any earnings under current law or the offset.

The Stage 2 impact analysis has a total of six confirmatory hypothesis tests: tests of impacts on the two confirmatory outcomes in each of the three pairwise comparisons. We group the four tests in the T21 vs. C2 and T22 vs. C2 comparisons together because they both involve impacts of the benefit offset. We perform a multiple comparison procedure on these four tests together to adjust the *p*-values of the tests. We perform a separate multiple comparison procedure to adjust the *p*-values of the two confirmatory tests in the T22 vs. T21 comparison. These adjustments are necessary because we are performing multiple hypothesis tests, which makes the probability of at least one Type I error (rejecting a true null hypothesis) larger than the significance level for the individual tests. To compensate for this effect, we adjust the test statistics for the confirmatory tests so that the probability of rejecting the null hypothesis of no impact within the “family” of tests (i.e., either within the four tests of T21 vs. C2 and T22 vs. C2 or within the two tests of T22 vs. T21) is equal to the specified significance level if the null hypothesis of no impact on any outcome in the tested group is true.⁵¹

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

We estimate impacts on administrative outcomes for the full Stage 2 assignment groups and for seven pairs of subgroups. We treat all subgroup analyses, including the tests of earnings and SSDI benefits paid, as exploratory. The first subgroup pair is defined by duration of SSDI benefit receipt at the point of solicitation into the demonstration.⁵² The duration subgroups are of interest because prior research and

⁵⁰ See Appendix A for a full description of the estimation model and the construction of analysis weights.

⁵¹ Our approach adjusts the *p*-values for the confirmatory outcomes using the Westfall-Young stepdown method. Details of the *p*-value adjustments for tests of impacts on the confirmatory outcomes appear in Appendix B. See Schochet (2009) for further discussion of the multiple comparisons problem.

⁵² We measure the duration of SSDI receipt from the outreach release date rather than from the date of random assignment in order to prevent endogenous selection into the duration subgroups. Some beneficiaries may have responded faster to outreach than others and the speed of their response may be correlated with their earnings and benefit outcomes. A short-duration beneficiary who took a long time to respond to outreach before enrolling in the study may have crossed the threshold into the long-duration definition (37 months or more of SSDI receipt) if duration is measured from random assignment. In order to rule out the possibility of subjects

program rules suggest that subjects who have been on the rolls for a *short duration* (defined here as three years or less) may 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 have their benefits adjusted, because they will have completed fewer TWP and GP months at the outset of the demonstration in comparison to long-duration subjects.

The second subgroup pair is defined by whether the participant lived in a state with a Medicaid Buy-In program.⁵³ Most states now have a Medicaid Buy-In program for persons with disabilities, who may otherwise be concerned that they will lose their Medicaid coverage if they enter or return to the workforce. Commercial or employer-based health insurance might not provide coverage for services and supports that enable people with disabilities to work and live independently. Therefore, theory would predict that study subjects with access to a Medicaid Buy-In program would be more likely to seek employment than study subjects without access to a Medicaid Buy-In program.

Two subgroup pairs are defined by specific disabilities: a primary impairment of Major Affective Disorder at baseline, and a primary impairment of Back Disorder at baseline. The incidence of Major Affective Disorder as a primary impairment has grown significantly in recent years, and there has been some expectation that we might see stronger treatment effects for subjects with this primary impairment. Similar to Major Affective Disorder, the incidence of Back Disorder as a primary impairment has increased in recent years, generating suspicion that we might see stronger treatment effects for subjects with this disorder than subjects with a different primary disability.

The remaining three subgroup pairs are defined by (i) employment status at baseline, (ii) age at baseline, and (iii) education at baseline. We expect that subjects who are employed, younger, or have higher education levels are more likely to use the offset because they face higher opportunity costs of not working. For example, those who are already working at baseline may be able to increase earnings to take advantage of the offset more readily than beneficiaries not already working. Beneficiaries with higher education may have more employment options than those with lower levels of education, in part because higher education may give more options for changing fields – for example, from construction to information systems. They may also be better able to understand the offset rules and, therefore, be more likely to change their behavior in response. Beneficiaries who are younger may also face more economic opportunity by changing fields through job training or other means than older beneficiaries, because they have more years before retirement to gain earnings and invest in a new career.

determining their subgroup membership after exposure to the study (which occurred when subjects first received solicitations to enroll), we measure duration from outreach release date.

⁵³ Some BOND sites serve participants from more than one state, and in some cases, state Medicaid Buy-In availability changed between 2011 and 2012 during the recruitment and randomization period. The majority of BOND participants at the Alabama (100%), Colorado/Wyoming (88%), and South Florida (99%) sites did not have access to Medicaid Buy-In at baseline; some in the DC Metropolitan Area (24%) also did not have access to Medicaid Buy-In at baseline. All others had access to Medicaid Buy-In (including Arizona/Southeast California, Greater Detroit, Greater Houston, Northern New England, Western New York, and Wisconsin).

6.4. Final Analysis Sample Sizes

Exhibit 6-2 presents the sizes for the overall sample and the subgroups. The final Stage 2 analysis sample contains a total of 12,744 subjects, spread across T21 (4,854), T22 (3,041), and C2 (4,849).

As would be expected if random assignment was properly implemented, the differences in baseline characteristics between assignment groups are small and appear to be due to chance (Gubits et al. 2013). An omnibus test for differences across all the characteristics shows that there is not a statistically significant difference between groups (Gubits et al. 2013). Baseline equivalence bolsters the case that any study findings of statistically significant impacts represent real impacts of the interventions, rather than systematic preexisting differences between the three groups or their environments.

Exhibit 6-2. Stage 2 Analysis Sample Composition

Random Assignment Group	Full Sample	Year of Stage 2 Enrollment	
		2011	2012
Stage 2 Sample Unweighted Counts			
T21	4,854	1,948	2,906
T22	3,041	1,212	1,829
C2	4,849	1,941	2,908
Stage 2 Sample Weighted Percentages			
T21	100%	42.6%	57.4%
T22	100%	42.2%	57.8%
C2	100%	42.9%	57.1%

Source: BOND Operations Data System (BODS).

Notes: The total sample size (T21 + T22 + C2) is 12,744. The Stage 2 analysis sample excludes 210 beneficiaries who are related to other BOND subjects (e.g., a primary and a DAC or two DACs with the same primary) to avoid contamination effects that might arise from the fact that almost all such beneficiaries (204 of the 210) were assigned to different BOND groups (see Appendix B for details on this adjustment). Because only six of these beneficiaries would have been able to be retained, it was not feasible to replicate the approach used for the Stage 1 analysis (where we were able to include pairs in which both members were assigned to the same group and revise the weights so that impact estimates reflect impacts for all beneficiary pairs with at least one member in Stage 1 (Stapleton et al. 2013)).

Weights are used to account for differing probabilities of selection into the Solicitation Pool by site and duration of SSDI receipt. The weighted Stage 2 sample size is 276,342 (the estimated number of Stage 2-eligible beneficiaries in the nation who would have volunteered had all Stage 2-eligible beneficiaries been offered the opportunity to enroll in the study).

This exhibit shows 1 additional T21 subject and 1 fewer C2 subject than Exhibit 2-2 in Gubits et al. 2014. The random assignment status of one Stage 2 subject was recorded as T21 and C2 in different subcomponents of BODS. We identified this discrepancy in March of 2015 and corrected it by placing the subject in the T21 group.

7. Impacts on Annual Earnings and SSDI Benefits Measured in Administrative Data

For the Stage 2 sample of volunteers, this chapter presents estimates of the impact of BOND on administrative earnings and benefits outcomes in 2013. Those randomly assigned to one of the treatment groups became subject to the offset work incentives as they enrolled between March and December 2011 (40 percent of T21 and T22 subjects) or during the first nine months of 2012 (60 percent). Hence, the duration of subjects' participation in Stage 2 ranges from 3 to 22 months at the start of 2013 and from 15 to 34 months at the end of the year. Later reports will examine impacts in 2014 and beyond when all treatment group subjects have been under the BOND benefit payment rules and work incentive counseling for longer periods of time.

For 2013 earnings and benefits, this chapter reports three policy comparisons:

- The impact of the benefit offset with standard work incentives counseling (WIC) compared to current law (T21 vs. C2).
- The impact of the benefit offset and enhanced work incentives counseling (EWIC) compared to current law (T22 vs. C2).
- The incremental impact of enhancing work incentives counseling (moving from WIC to EWIC) given the benefit offset (T22 vs. T21).

We refer the reader to Chapter 6 of this report for definitions of outcome variables, theories about possible impacts, administrative features of the offset that may influence impacts, and the impact estimation methodology we used to generate the results we present in the current chapter.

We begin in Section 7.1 with a discussion of employment, earnings, benefits, and health trends in the current law control group in order to set the context for interpreting the impacts of the offset (with WIC or EWIC). Section 7.2 then discusses the 2013 confirmatory results (i.e., impacts on SSDI benefits and earnings) for the full Stage 2 sample. Exploratory evidence regarding other employment- and benefit-related outcomes appears in Section 7.3. Later sections of the chapter look for variations in impact for beneficiaries with differing background characteristics and summarize the chapter's findings.

In the discussions of impact estimates in Section 7.2 through 7.4, we use particular language to signify differing levels of confidence that a non-zero impact has occurred. When the null hypothesis of no effect can be rejected with 99-percent confidence (i.e., with 0.01 statistical significance) we state that the estimate *provides strong evidence* that BOND had an effect on the tested outcome. When the null hypothesis of no effect can be rejected with 95-percent confidence (i.e., with 0.05 statistical significance) but not 99-percent confidence we state simply that the estimate *provides evidence* that BOND had an effect on the tested outcome. Finally, when the null hypothesis of no effect can be rejected with 90-percent confidence (i.e., with 0.10 statistical significance) but not 95-percent confidence we state that the estimate *provides some evidence* that BOND had an effect on the tested outcome.

We make a distinction between hypothesis testing pre-specified as confirmatory and hypothesis testing of a more exploratory nature. Statistically significant findings from confirmatory analyses meet a higher

standard of evidence—one that minimizes the possibility of “false positive” findings (i.e., apparent impacts where the truth is zero impact). In contrast, statistically significant findings from *exploratory* hypothesis tests offer suggestive evidence of other impacts that BOND may have achieved. Confirmatory tests are confined to results on total earnings and SSDI benefits received in 2013, generalized to represent a national population of volunteers to use the offset.

Under the high standard set for confirmatory evidence, we do not find evidence of an impact on 2013 earnings and SSDI benefits paid in 2013. Some of the exploratory impact estimates are statistically significant, however. While there is no evidence of an impact on confirmatory outcomes, there is some suggestive evidence of impacts on several exploratory outcomes: positive impacts on employment, proportion with earnings above BYA, and number of months of SSDI receipt. Future Stage 2 impact analysis reports will help us learn more about the impacts of the offset on future earnings, future benefits, and benefits paid *for* 2013 for Stage 2 volunteers, i.e. for SSDI recipients most inclined to use the offset.

In contrast, the comparison of EWIC versus WIC yields consistent results across confirmatory and exploratory analyses. At the confirmatory level of evidence, we do not find evidence of a differential impact of offset-plus-EWIC as compared to the offset-plus-WIC. In addition, a pattern of effects of EWIC compared to WIC does not emerge in the exploratory analyses.

7.1. Employment, Earnings, Benefits, and Health Trends in the Control Group

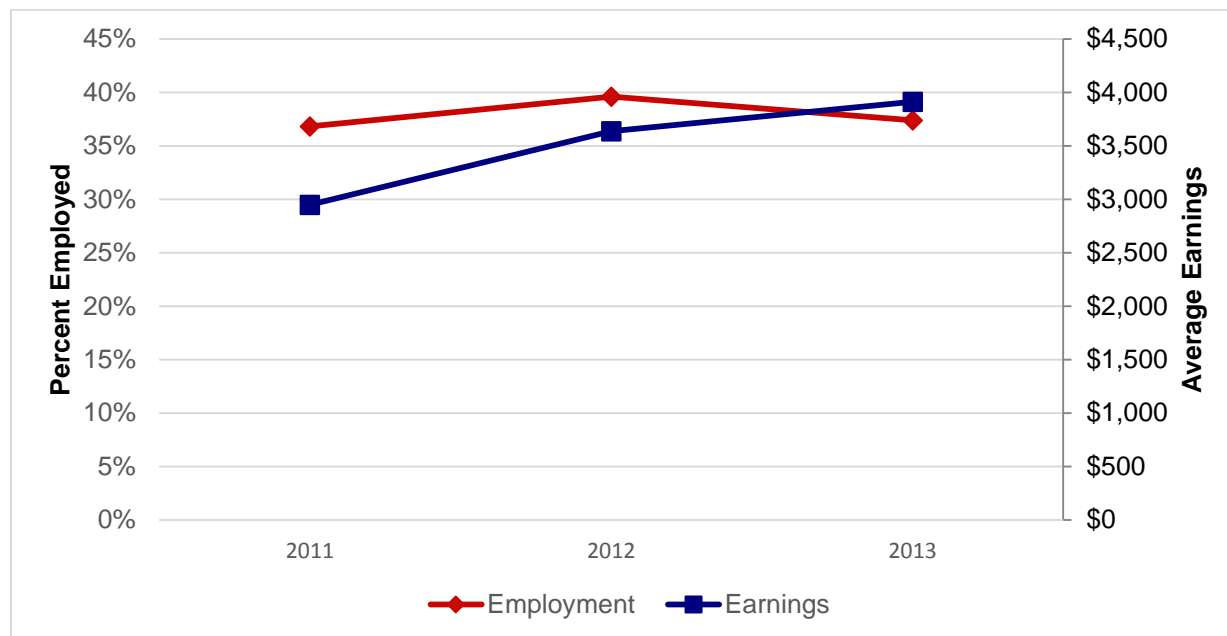
To provide a broader context for the findings, we briefly report trends in employment, earnings, health, and SSDI benefits received in the control group, spanning from 2011 through 2013. As depicted in Exhibit 7-1, 37 percent of subjects in the current law control group were employed at some point in 2011, with rates for 2012 and 2013 at 40 and 37 percent respectively—a statistically significant ($p=.006$) variation. Average earnings rose significantly over time from \$2,951 in 2011 to \$3,638 in 2012 and \$3,913 in 2013. The rise is consistent with SSDI beneficiaries volunteering for BOND at a time when they expected their earnings to rise (and thus to be able to benefit from the offset). As shown in Exhibit 7-2, earnings for employed subjects in the control group rose steadily between 2011 and 2013 from \$8,014 in 2011 to \$9,182 in 2012 to \$10,460 in 2013.

All subjects are aging throughout the study and thus we expect a slight decline in health for current law control group members. Exhibit 7-3 reports statistically significant differences in self-reported health for this sample. The percent of subjects reporting a health status of “good” or “excellent” remained roughly the same for the two years, but the share rating their health as “poor” increased from 16 to 22 percent and this difference is statistically significant ($p<.001$).

As shown in Exhibit 7-4, average SSDI benefits declined over the same period, likely driven by beneficiaries completing their TWPs or leaving the rolls through medical improvement, and to a small extent by death or incarceration.

Collectively, these trends provide context for interpreting estimated impacts of the demonstration. Average earnings and benefits are on a slight downward trend, likely driven by aging and declining health. However, among those employed, average earnings increased, suggesting that—like employed subjects in the control group—some members of the two treatment groups will be able to achieve sustained earnings above SGA and thus to benefit from the offset.

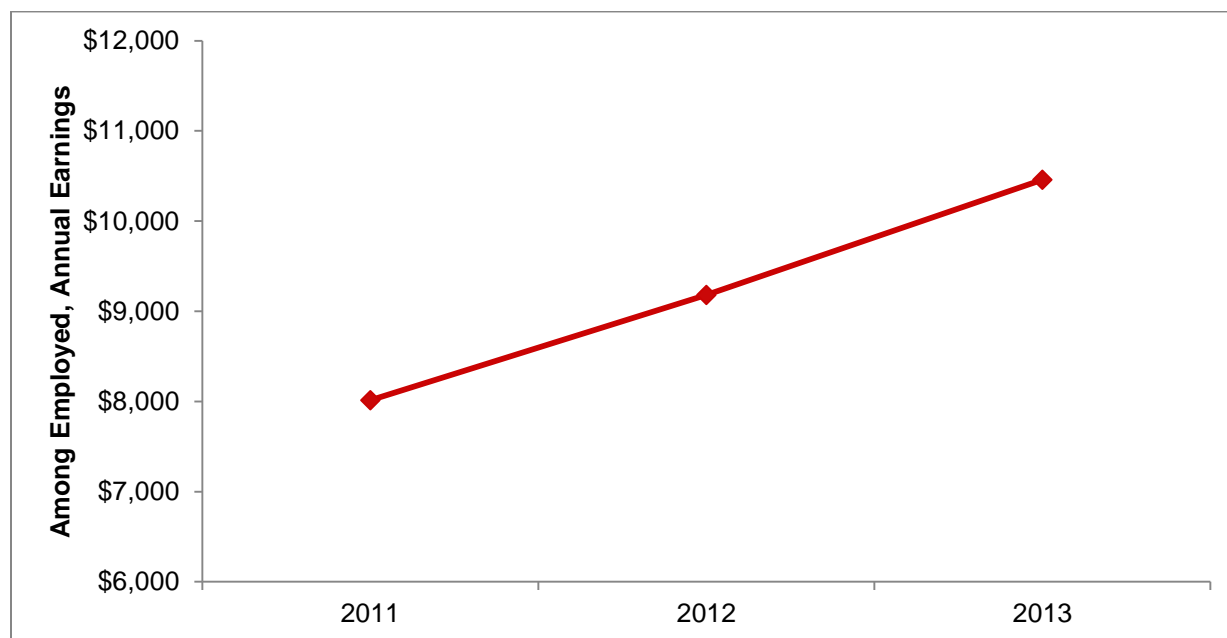
Exhibit 7-1. Trends in Employment and Earnings, Current Law Control Group (C2)



Source: Exhibits 7-5, 7-6, C-1, C-2, and C-9 of this report.

Note: Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of volunteers for offset participation in the nation.

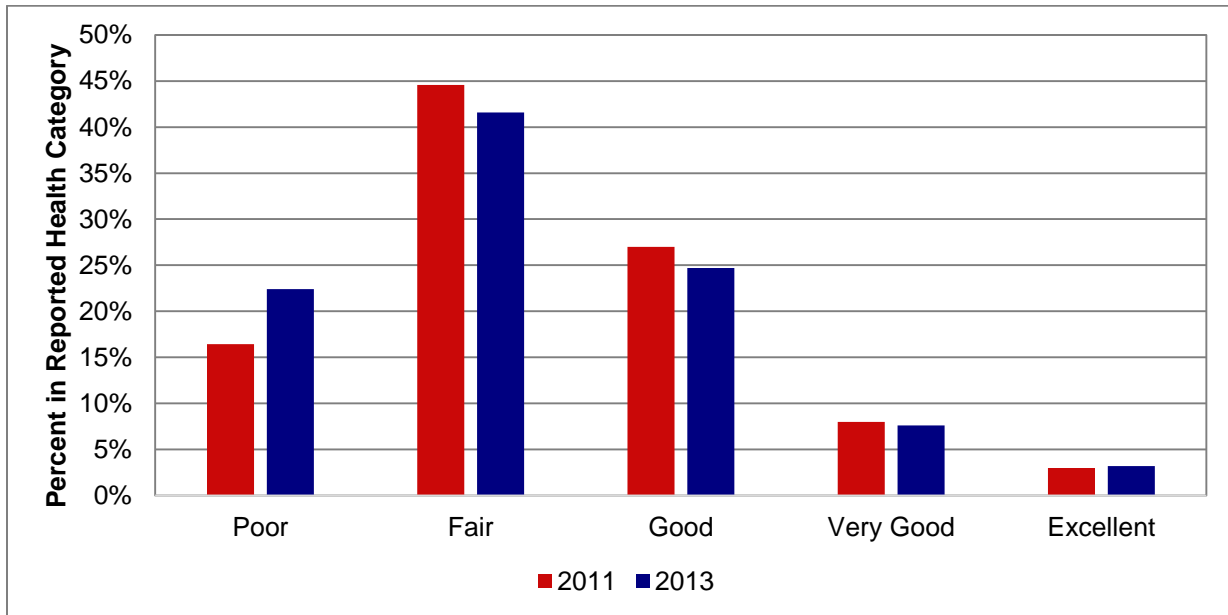
Exhibit 7-2. Earnings Trends for Employed Subjects, Current Law Control Group (C2)



Source: Exhibits 7-5, 7-6, C-1, C-2, and C-9 of this report.

Notes: Average C2 earnings are divided by the share of C2 subjects employed to obtain the average annual earnings among those who are employed. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of volunteers for offset participation in the nation.

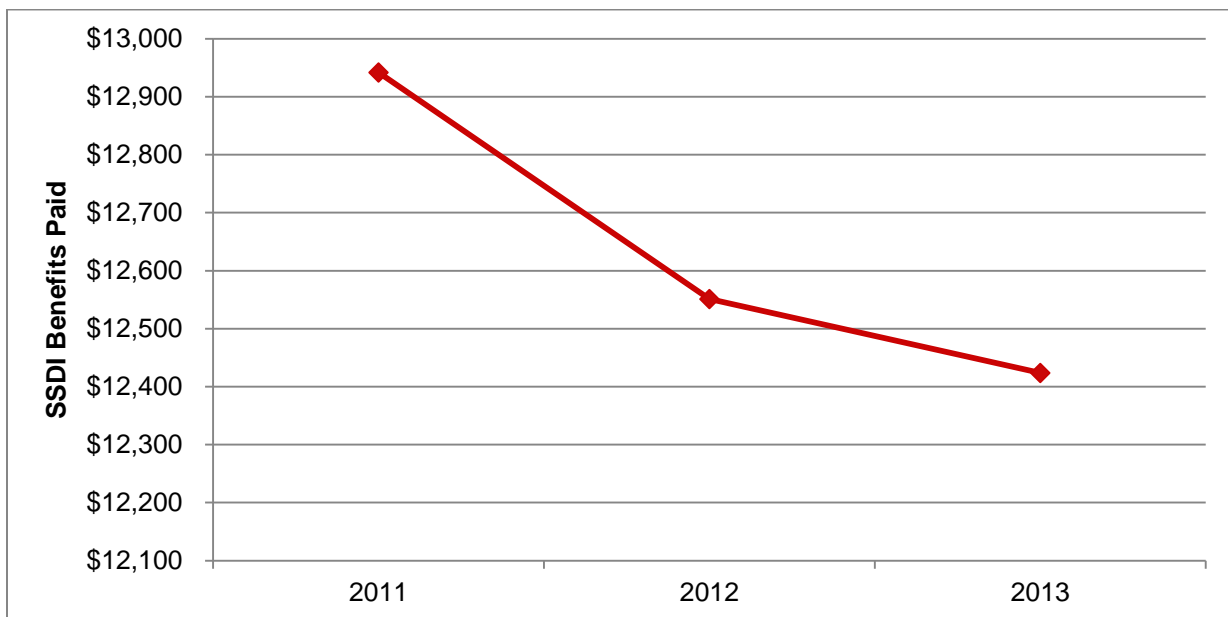
Exhibit 7-3. Self-Reported Health in 2011 and 2013, Current Law Control Group (C2)



Source: Stage 2 Baseline Survey (2011 – 2012) and Interim Follow-up Survey (13 months later).

Note: The unweighted baseline sample size is 4849. The unweighted interim follow-up sample size is 3961. Weights reflecting sample selection and survey nonresponse are used to ensure that the BOND subjects who met analysis criteria are representative of volunteers for offset participation in the nation.

Exhibit 7-4. Trends in Annual SSDI Benefits Paid, Current Law Control Group



Source: Exhibits 7-5, C-1, and C-2 of this report.

Note: Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of volunteers for offset participation in the nation.

7.2. Confirmatory Impacts

Among the many outcomes analyzed in the BOND evaluation, two outcomes are of paramount interest. These we examine for confirmatory evidence that one or both of the Stage 2 BOND interventions compared to current law are having an impact on beneficiaries:⁵⁴

1. Total earnings in the most recent available year (2013 in this report)
2. Total SSDI benefits in the most recent available year (2013 in this report)

The Stage 2 impact analysis has a total of six confirmatory hypothesis tests: tests of impacts on these two confirmatory outcomes in each of the three pairwise comparisons. We group the four tests in the T21 vs. C2 and T22 vs. C2 comparisons together because they both involve impacts of the benefit offset. We perform a multiple comparison procedure on these four tests together to adjust the p-values of the tests. We perform a separate multiple comparison procedure to adjust the p-values of the two confirmatory tests in the T22 vs. T21 comparison.

For total earnings received from January through December 2013, we do not find statistically significant effects on either treatment group, relative to current law (Exhibit 7-5). Estimated impacts on mean earnings (first row of the exhibit) are \$385 for the offset-plus-WIC compared to current law, and \$366 for the offset-plus-EWIC compared to current law. The adjusted p-values for these two impact estimates equal 0.135 and 0.259, respectively, and so do not meet the study's established standard for statistical significance ($p < 0.10$).⁵⁵ These findings are consistent with the failure to find evidence of an effect on 2012 earnings for the offset combined with either WIC or EWIC (see Appendix C).⁵⁶ The size of the

⁵⁴ These two outcomes were identified in the *BOND Evaluation Analysis Plan* (Bell, et al 2011) for confirmatory analysis, prior to the research team having access to outcome data for study subjects. Pre-specifying outcomes for confirmatory analysis prior to having access to outcome data is standard evaluation practice. It makes transparent that researchers have selected the study's confirmatory outcomes based on hypotheses developed prior to looking at the data, rather than based on the estimates of impact for many different outcomes. See the discussion of confirmatory outcomes in Chapter 6, Section 6.1, of this report.

In later reports, impacts on earnings and SSDI benefits in subsequent years (always the most recent available year) will become the confirmatory outcomes, supplanting the confirmatory impact estimates published in the current report. The practice of supplanting previous confirmatory impact estimates with the most recent available estimates reflects the supremacy of long-term impacts in determining the interventions' impacts on earnings and benefits.

⁵⁵ Before adjustment for multiple comparisons, the unadjusted p-value for the offset plus WIC compared to current law is 0.060 and the unadjusted p-value for the offset plus EWIC compared to current law is 0.229.

⁵⁶ The estimates published in the *First and Second Year Snapshot of Earnings and Benefits Impacts for Stage 2* report provided some confirmatory evidence that the benefit offset plus WIC and the benefit offset plus EWIC increased earnings in comparison to current law. However, the statistical tests in that report were for impacts on beneficiaries in only the 10 BOND sites, not the entire nation (Gubits et al. 2014). The 2013 tests presented here as well as those for 2011 and 2012 in Appendix C concern impacts for the national population, which involves more statistical uncertainty than findings in the earlier reports. The new findings show that the impact estimates for 2012, as well as for 2013, are not large enough to confirm nationwide impacts. See Appendix A for a discussion of the issues underlying the methodological differences between Gubits et al. (2014) and this Stage 2 *Interim Report*.

impact estimates for 2013 is similar to the size of the impact estimates on 2012 earnings: estimates for both years and both policy comparisons range between 8 and 10 percent of average earnings under current law (with rounding)—impacts not large enough to be statistically significant.

Similar to the 2012 findings, we find no evidence of an incremental effect of EWIC compared to WIC. The point estimate of impact on earnings for this comparison is negative, small (-\$19), and not statistically significant.

Exhibit 7-5. Estimated Impacts on 2013 Total Earnings and Total SSDI Benefits Paid of Stage 2 Volunteers: All Policy Comparisons

Outcome	Average Outcome with Offset and WIC (T21) (1)	Average Outcome with Offset and EWIC (T22) (2)	Average Outcome under Current Law (C2) (3)	Estimated Impact of Offset + WIC vs Current Law (T21 vs. C2) (4)	Estimated Impact of Offset + EWIC vs Current Law (T22 vs. C2) (5)	Estimated Impact of EWIC instead of WIC Given Offset (T22 vs. T21) (6)
Total earnings (January–December 2013)	\$4,298	\$4,279	\$3,913	\$385 ^a (\$179)	\$366 ^a (\$284)	-\$19 (\$198)
Total SSDI benefits paid (January–December 2013)	\$12,781	\$12,798	\$12,424	\$357 ^b (\$131)	\$374 ^b (\$135)	\$16 (\$170)

Source: Analysis of SSA administrative records (from the MEF, BODS, MBR, and SSR), with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: T21 = 4,854 , T22 = 3,041 , C2 = 4,849

###,### 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.

^a The impact estimates for total earnings for T21 vs. C2 and for T22 vs. C2 had p-values after multiple comparison adjustments of 0.135 and 0.259, respectively. Hence, the data do not provide confirmatory evidence of an impact in either case

^b The impact estimate for total SSDI benefits paid for T21 vs. C2 and for T22 vs. C2 both had p-values after multiple comparison adjustments of 0.104 Hence, the data do not provide confirmatory evidence of an impact.

We also do not find statistically significant effects on total SSDI benefits paid between January and December 2013, for either treatment group, relative to current law (second row of Exhibit 7-5). Estimated impacts on benefits are \$357 annually for the comparison of the offset plus WIC to current law and \$374 for the comparison of the offset plus EWIC to current law—about 3 percent of average benefits under current law (\$12,424). The adjusted p-value of both of these estimates is 0.104.⁵⁷

⁵⁷ Before adjustment for multiple comparisons, the unadjusted p-value for the offset plus WIC compared to current law is 0.023 and the unadjusted p-value for the offset plus EWIC compared to current law is 0.022.

These 2013 findings are consistent with results that also failed to provide evidence that the offset combined with WIC or combined with EWIC had an effect on 2012 benefits, relative to current law (see Appendix C).⁵⁸

Similar to the 2012 findings, we find no evidence of an incremental effect of EWIC compared to WIC. The point estimate of impact on SSDI benefits for this comparison is small (\$16) and not statistically significant.

7.3. Exploratory Impacts for the Full Sample

The previous section reported results for confirmatory outcomes, finding no evidence of impact on average earnings or SSDI benefits paid. This section considers potential impacts on other earnings- and benefit-related outcomes—outcomes tested for effects on an exploratory rather than confirmatory basis. It is possible for impacts on the exploratory outcomes to occur in the full sample even when the effects on the confirmatory outcomes are zero. Hence, we look for these ancillary impacts even in the absence of statistically significant effects on average earnings and SSDI benefits paid.

Seven other outcomes related to earnings and benefit amounts are being tracked in administrative data: any employment during the year and in various dollar ranges relative to BYA, number of months of SSDI receipt over a year, and total dollars and number of months of payments from the Supplemental Security Income (SSI) program. We report impacts on these measures for 2013 in this section. Consistent with the *BOND Evaluation Analysis Plan* (Bell et al. 2011), we consider these analyses to be exploratory and therefore do not make any correction for multiple comparisons. As a result, any statistically significant findings are suggestive of where further effects of the benefit offset plus EWIC or WIC may have taken place. Even if the intervention had no impact on any of the measures examined here, we would expect some of the impact estimates to be statistically significant by chance alone due to the fact that we conduct many hypothesis tests in this section.

7.3.1. Exploratory Impacts on Earnings-Related Outcomes

As stated in Chapter 6, the offset is predicted to have two countervailing effects on earnings: a *positive* effect on average earnings for those who would not engage in SGA under current law (i.e. without the offset) and a *negative* effect on average earnings for those who would earn above the SGA level under current law. The net result of these two changes can be an earnings impact in either direction or no earnings effect at all. Only for the first group can an employment gain accompany the earnings gain, moving employment upward for those who do not work at all under current law but who work under the offset. Both earnings and employment increase for these beneficiaries. In contrast, none of the

⁵⁸ The estimates published in the *First and Second Year Snapshot of Earnings and Benefits Impacts for Stage 2* report provided some confirmatory evidence that the benefit offset plus WIC (but not the benefit offset plus EWIC) increased SSDI benefits in comparison to current law. However, the statistical tests in that report were for impacts on beneficiaries in only the 10 BOND sites, not the entire nation (Gubits et al. 2014). The 2013 tests presented here as well as those for 2011 and 2012 in Appendix C concern impacts for the national population, which involves more statistical uncertainty than findings in the earlier reports. See Appendix A for a discussion of the issues underlying the methodological differences between Gubits et al. 2014 and this Stage 2 Interim Report.

beneficiaries predicted to have downward effects on earnings will experience a negative employment effect. These beneficiaries are predicted to work fewer hours while continuing to work.

Similarly, the predicted increase in the proportion with earnings above BYA is generated by an increase in earnings in response to the offset from some of those who would not engage in SGA under current law. In contrast, none of those who would earn above the SGA level under current law are predicted to lower their earnings below BYA. These beneficiaries are predicted to reduce their hours but to continue earning above BYA.

Movements in and out of other earnings categories—above twice BYA and above three times BYA—are also possible without net changes in earnings. Hence, for all of the exploratory outcomes of interest here impacts could occur even absent impacts on the two confirmatory outcomes.

Exhibit 7-6 provides exploratory evidence that the offset plus WIC increased the proportion of sample members employed (i.e., those with any earnings during 2013) and the proportion earning above BYA that year. It also presents exploratory evidence that the offset plus EWIC increased the proportion of sample members with earnings above BYA in 2013. These findings conform to an unambiguous prediction of theory that by removing the benefit cliff at earnings above BYA, the offset will increase employment and the proportion of beneficiaries with earnings above BYA. In the current law control group, 37 percent of beneficiaries had some employment in 2013 and 9 percent had earnings above the BYA. The offset plus WIC increased the proportion employed by 3 percentage points (an 8 percent increase, after rounding) and the proportion with earnings above the BYA by 2 percentage points (a 22 percent increase, after rounding). The offset plus EWIC did not yield a statistically significant increase in the total employment rate, compared to the current law control group, but it did increase the proportion with earnings above BYA by 2 percentage points (a 25 percent increase, after rounding).

That these employment effects are taking place without confirmatory evidence of impact on average earnings could be due to multiple factors. One possible explanation is that average earnings in the treatment groups may have increased, but not enough to be statistically significant. It is also possible that even as a greater proportion of subjects chose to earn above BYA, average earnings *within* one or more of the earnings ranges far above BYA may have declined (for illustration, a person who would earn 2.9 times BYA without the offset might choose to earn 2.0 times BYA if offered the offset). This possibility is consistent with theory, which predicts that subjects who under current law would choose to earn between BYA and the amount where benefits would be reduced to \$0 under the offset (i.e., the end of the offset “ramp,” on average about three times BYA) will *decrease* their earnings if the offset is available (thereby obtaining more leisure time at the same or greater total income). Finally, the estimated impact on the proportion with earnings above BYA is small (2.01 percentage points for the offset plus WIC, and 2.28 percentage points for the offset plus EWIC), and modest differences in mean earnings within this small proportion of the study sample are hard to detect because they are averaged with earnings for the rest of study sample.

Exhibit 7-6. Estimated Impacts on 2013 Employment and Benefits Receipt for Stage 2 Volunteers: All Policy Comparisons

Outcome	Average Outcome with Offset and WIC (T21) (1)	Average Outcome with Offset and EWIC (T22) (2)	Average Outcome under Current Law (C2) (3)	Estimated Impact of Offset + WIC vs Current Law (T21 vs. C2) (4)	Estimated Impact of Offset + EWIC vs Current Law (T22 vs. C2) (5)	Estimated Impact of EWIC instead of WIC Given Offset(T22 vs. T21) (6)
Employment (January–December 2013)						
Employment during year (%)	40.16	38.78	37.41	2.75** (0.96)	1.37 (1.07)	-1.38 (1.09)
Earnings above BYA (%)	11.11	11.38	9.10	2.01*** (0.62)	2.28** (0.73)	0.27 (0.73)
Earnings above 2x BYA (%)	4.09	3.75	3.57	0.51 (0.41)	0.18 (0.51)	-0.34 (0.46)
Earnings above 3x BYA (%)	1.55	1.67	1.59	-0.04 (0.27)	0.08 (0.35)	0.12 (0.30)
Benefit Receipt (January–December 2013)						
Number of months with SSDI payments	11.26	11.28	11.01	0.25*** (0.07)	0.26*** (0.08)	0.01 (0.07)
Total SSI benefits paid	\$42	\$48	\$37	\$6 (\$9)	\$11 (\$12)	\$6 (\$14)
Number of months with SSI payments	0.20	0.24	0.19	0.01 (0.03)	0.05 (0.04)	0.04 (0.05)

Source: Analysis of SSA administrative records (from the MEF, BODS, MBR, and SSR), with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: T21 = 4,854, T22 = 3,041, C2 = 4,849

*/**/** 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 (and with no multiple comparisons adjustment).

There is no evidence that EWIC does more (or less) than WIC to increase employment or the proportion with earnings above BYA in calendar year 2013.

7.3.2. Exploratory Impacts on Benefit-Related Outcomes

This section considers exploratory impacts on benefit-related outcomes. It is possible that impacts could occur for some of the benefit-related exploratory outcomes even in a scenario with no impact on total benefits paid. As stated in Chapter 6, the offset is predicted to have two countervailing effects on benefits: a *positive* effect on benefits for those who would engage in SGA and go into cessation under current law but who receive partial benefits under the offset, and a *negative* effect on benefits for those who increase

their earnings from something below the SGA level to above BYA because of the offset's incentive and thereby move from full benefits to partial benefits. The number of months with benefits will change only for the first group, increasing along with the increase in total benefits paid for that subpopulation. In contrast, beneficiaries who move to earnings above BYA because of the offset are unlikely to experience a *reduced number of months* receiving SSDI benefits. That would generally require earning more than three times BYA and receiving \$0 in benefits—an option available to them under current law but not chosen. On net, then, the number of months of SSDI benefits paid to the treatment groups can go up because of the offset even with no net effects on average SSDI benefits paid.

Effects on months with SSI payments and SSI payment amount from the offset can occur for T21 and T22 subjects only if they go onto SSI through asset spend-down following study enrollment. As noted in Chapter 6, the offset could raise earnings enough to slow entry onto SSI through asset spend-down for some beneficiaries or lower earnings enough to hasten spend-down and SSI entry by others. Cancelling effects on average earnings do not necessarily imply equal shares crossing this line between SSI ineligibility and eligibility, making impacts on these two SSI-related outcomes possible even in the absence of any true impact on average earnings (or on average SSDI benefits paid). Hence, for all of the exploratory outcomes of interest here, impacts could occur even absent impacts on the two confirmatory outcomes.

There is exploratory evidence that the offset—both with WIC and with EWIC—increased the number of months of SSDI receipt by a quarter of a month (roughly one week, or 0.25 months). This increase is a 2 percent increase over the average number of months of SSDI receipt for C2 subjects. This finding is expected because treatment subjects whose earnings would have been above BYA under current law receive partial benefits under the offset but would have had their benefits suspended under current law. There is no evidence of any impact of the offset-plus-WIC or the offset-plus-EWIC on SSI benefits received or number of months of SSI receipt in 2013.

There is no evidence that EWIC does more than WIC to affect the number of months of SSDI receipt, number of months of SSI receipt, or amount of SSI benefits.

7.4. Impact Variation by Beneficiary Background Characteristics

We also explored whether the Stage 2 treatments affected earnings and benefits differently for beneficiaries with different background characteristics. Seven potential impact “moderators” of this sort were discussed in Section 6.3 and are listed here for reference:

• Short-duration beneficiaries (those receiving benefits 36 months or less when entering BOND)	Vs.	• Longer-duration beneficiaries (those receiving benefits 37 months or more when entering BOND)
• Beneficiaries who are employed at baseline	Vs.	• Beneficiaries who are not employed at baseline
• Beneficiaries with access to Medicaid buy-in programs	Vs.	• Beneficiaries without access to Medicaid buy-in programs
• Younger beneficiaries (less than age 50)	Vs.	• Older beneficiaries (age 50 and older)
• Beneficiaries with primary impairment of major affective disorder	Vs.	• Beneficiaries with all other primary impairments
• Beneficiaries with primary impairment of back disorder	Vs.	• Beneficiaries with all other primary impairments
• Beneficiaries with an associate’s or bachelor’s degree at baseline	Vs.	• Beneficiaries with no postsecondary degree at baseline

Appendix B contains impact estimates for all of the subgroups defined by these moderating characteristics. All moderator analyses of impact variation are exploratory, and the significance tests reported here and in Appendix B are not adjusted for multiple comparisons (unlike the full-sample confirmatory impact estimates, which are adjusted for multiple comparisons). We examine how impacts vary with the moderators for each of four policy contrasts:

- The offset plus WIC compared to current law (T21 versus C2);
- The offset plus EWIC compared to current law (T22 versus C2);
- The offset with either type of work incentives counseling compared to current law (T22 combined with T21 versus C2); and
- The offset plus EWIC compared to the offset plus WIC (T22 versus T21).

The focus of the analysis is on whether impacts differ according to the moderating characteristic (e.g., *across* subgroups defined by duration of SSDI receipt) rather than whether impacts are detected within any particular subgroup defined by the moderators (e.g., impacts on short-duration beneficiaries). If the impacts do not differ in a statistically significant manner, our practice is to focus on the full sample impact estimates rather than any subgroup-specific impact estimate. This practice is often adopted because full sample impact estimates are more precise (i.e., have smaller standard errors) and cannot confidently be improved upon as information about particular subpopulations (Bloom and Michalopoulos 2013). Columns 7 in the exhibits of Appendix B display the estimate of the difference in impact for a subgroup and its counterpart and the corresponding standard error of the difference.

We examined impact moderation for all nine administrative outcomes discussed earlier in this chapter. With nine outcomes, seven moderators, and four policy comparisons, there are 252 tests of differences in impacts that could potentially produce evidence of moderation to highlight in this chapter. As shown in Appendix B, our analysis produced 32 statistically significant ($p < 0.10$) differences in impact magnitude between subgroups at the $p < 0.01$ significance level, absent adjustment for multiple comparisons, out of the 252 tests performed. If all of the hypothesis tests were independent and there were no true impact moderation there would be a 9.6 percent chance of 32 or more statistically significant moderator findings occurring at the $p < 0.10$ level.⁵⁹ In fact, the hypothesis tests are not independent, so the chance of 32 or more significant moderator findings is higher than 9.6 percent. Given the strong possibility that the great majority of these results are due to chance sampling variation, we present in this text only the subset of moderator tests that produce the strongest evidence: statistically significant impact variations across subgroups at the $p < 0.01$ level.

Seven differences in impact are statistically significant at the $p < 0.01$ level. Five of them stem from the comparison of subjects employed versus not employed at baseline. Specifically:

- Employment at baseline moderates the impact of T22 + T21 as compared to C2 on proportion with earnings above BYA (Exhibit B-8). A larger impact occurs for subjects employed at baseline than for subjects not employed at baseline.
- Employment at baseline moderates the impact of T22 + T21 as compared to C2 on total SSDI benefits paid (Exhibit B-8). A larger impact occurs for subjects employed at baseline than for subjects not employed at baseline.
- Employment at baseline moderates the impact of (a) T21 as compared to C2, (b) T22 as compared to C2, and (c) T22 + T21 as compared to C2 on number of months of SSDI payments (Exhibits B-5, B-6, and B-8, respectively). A larger impact occurs for subjects employed at baseline than for subjects not employed at baseline.

These findings are consistent with the hypothesis that subjects already working at baseline have a greater ability to earn—or interest in earning—at a level that takes advantage of the offset than do beneficiaries not already working at baseline.

The two additional differences in impact that are statistically significant at the $p < 0.01$ level concern the impact of T22 as compared to T21 for the subgroup defined by having a primary impairment of back disorder (Exhibit B-23). EWIC provided to T22 subjects as compared to WIC for T21 subjects created a more *negative* impact on earnings and the proportion with earnings greater than 3xBYA for subjects with a primary impairment of back disorder at baseline than for subjects with any other type of primary impairment at baseline. No hypotheses suggest themselves to account for why EWIC, compared to WIC, would have a particularly dampening effect on earnings for beneficiaries with back disorders.

⁵⁹ This calculation is one minus the probability that fewer than 32 hypothesis tests are statistically significant. $1 -$

$$\sum_{k=0}^{32-1} \binom{252}{k} (0.1)^k (0.9)^{252-k} = 0.096$$

Although not the main focus of the moderator analysis, about a quarter of the subgroup-specific impact estimates presented in Appendix B are statistically significant at the $p < .10$ level. Most of these statistically significant results reflect impacts that are apparent in the full-sample exploratory results.⁶⁰ These impacts are for the T21 versus C2 comparison on earnings; T22 versus C2 and T22 versus T21 comparisons on employment; the T21 versus C2 and T22 versus C2 comparisons on proportion with earnings above BYA; the T21 versus C2 and T22 versus C2 comparisons on total SSDI benefits paid; and the T21 versus C2 and T22 versus C2 comparisons on number of months with SSDI payments. Some subgroup-specific impact estimates will be statistically significant if the sample is well-powered for those outcome variables and there are also statistically significant full-sample exploratory impact findings.⁶¹

7.5. Summary of Results

We do not find confirmatory evidence that the offset plus WIC had an impact on total earnings in 2013 nor on total SSDI benefits received in 2013. Similarly, we found no confirmatory evidence that the offset plus EWIC changed these outcomes.

Some exploratory findings showed impacts:

- Consistent with theory, both the offset plus WIC and the offset plus EWIC increased the proportion of sample members with earnings above BYA compared to current law, with the offset plus WIC also increasing employment in 2013;
- The offset did more to increase employment in 2013 when provided in conjunction with WIC than when provided in conjunction with EWIC, an unexpected finding that is not consistent with findings from the 2012 analysis nor the logic model for EWIC;
- Both the offset plus WIC and the offset plus EWIC increased the average number of months of benefit receipt in 2013 compared to current law.

Similar to the 2012 findings, we find no evidence in 2013 that EWIC enhanced the impact of the offset compared to WIC.

As a complement to these findings from administrative data, the next chapter reports impact findings on employment, earnings, and other outcomes measured in the 12-month interim survey of Stage 2 beneficiaries.

⁶⁰ Of the 504 subgroup-specific impacts presented in Appendix B, 125 (25 percent) are statistically significant at the $p < .10$ level. Of these 125, 116 (or 93 percent) concern outcomes and policy comparisons for which statistically significant impacts are observed in the full sample exploratory results.

⁶¹ Given the large number of impacts examined, it is also to be expected that some subgroup-specific impact estimates will be statistically significant when the corresponding exploratory full-sample estimates are not significant. We are unable to determine whether these subgroup-specific estimates represent noteworthy new information or are simply due to chance.

8. Impacts on Additional Employment Outcomes and on Health-Related Outcomes

This chapter presents impact results for employment, health, and marital status outcomes derived from the 12-month interim survey of Stage 2 BOND subjects. Like Chapter 7, this chapter presents impact estimates for the three pairwise comparisons for Stage 2 volunteers: (i) offset-plus-WIC (T21) versus current law (C2), (ii) offset-plus-EWIC (T22) versus current law (C2), and (iii) EWIC (T22) versus WIC (T21). All statistical tests in this chapter are exploratory.⁶²

One advantage of survey data is that they allow for outcomes to be defined over time periods relative to the time of demonstration enrollment and random assignment (rather than periods that are tied to calendar years as is required when using SSA earnings data). A second advantage of survey data is that they enable analysis of a broader set of outcomes than the more limited administrative data provide.

One disadvantage of the survey data (compared to SSA earnings and disability benefits data) is that they may suffer from response errors. However, the survey data examined in this chapter are almost all point-in-time measures and therefore not subject to recall bias (often a major source of response errors). A second disadvantage of the survey data is that they do not contain complete information for the entire sample. However, the overall response rate to the Stage 2 interim survey was high, 84 percent. The response rates for the T21 and T22 groups (85.4 percent and 85.5 percent) were somewhat higher than for the C2 group (81.6 percent). This means that it is possible that survey-based impact estimates differ from what they would have been had survey data been available for all study subjects. To address the possible presence of this type of non-response bias, all estimates presented in this chapter are weighted to adjust for survey non-response.⁶³

8.1. Impact Estimates for Survey-derived Employment Outcomes

Exhibit 8-1 shows impact estimates for eight employment-related outcomes that are based on interim survey data:

1. *Any employment since random assignment*: equals 1 if the subject has performed any work for pay during the time period between random assignment and survey response (median time period is 13 months) and 0 otherwise. As discussed in Chapter 6, the predicted sign of impact on employment for the treatment groups compared to current law is positive. Therefore, we expect positive effects for this outcome in the comparisons with current law.

⁶² That is, they are not adjusted for multiple comparisons and carry a high risk of “false positives.” Hence, they cannot be considered as providing confirmatory evidence about the benefit offset even when shown as statistically significant; rather, they suggest additional possible BOND impacts beyond those that emerged from exploratory analyses in Chapter 7 above.

⁶³ The non-response weights cause survey respondents whose baseline characteristics are most similar to survey non-respondents to have the greater influence on analytic results, offsetting (in measurable, but not in unmeasurable, ways) the absence of the non-respondents from the analysis sample.

2. *Employed at time of response to the interim survey*: equals 1 if the subject is working for pay at the time of the survey and 0 otherwise. We expect positive effects for this outcome in the comparisons with current law.
3. *Hours per week working in current (main) job*: equals the typical hours spent working for pay at the time of the survey at the current job (for those working at only one job) or at the current main job (for those working at more than one job); the outcome equals 0 for those who are not working at the time of the survey. The predicted sign of impact on this outcome is ambiguous because while we expect more subjects to have non-zero hours of work, some subjects who under current law would earn above the BYA-level may reduce their work hours due to the offset, once able to supplement earnings with partial SSDI benefits.
4. *Weekly earnings in current (main) job*: equals self-reported weekly earnings in one's current or main job at the time of the survey; the outcome equals 0 for those who are not working at the time of the survey. As discussed in Chapter 6, the predicted sign of impact on earnings for the treatment groups compared to current law is ambiguous since the offset may induce earnings increases for some beneficiaries (depending on the range in which they would earn under current law, relative to the BYA) and earnings reductions for others.
5. *Current weekly earnings above weekly equivalent of BYA*: equals 1 if current weekly earnings from one's only or main job are above the weekly equivalent of BYA assuming work is spread evenly across the calendar year (defined as $7/365 \times \text{BYA}$) and 0 otherwise. As discussed in Chapter 6, the predicted sign of impact on the percentage with earnings above BYA for the treatment groups compared to current law is positive. Therefore, we expect positive effects for this outcome in the comparisons with current law.
6. *Current weekly earnings above weekly equivalent of $2 \times \text{BYA}$ (%)*: defined similarly to above. It is not possible to predict the direction of impacts on the percentage with earnings above $2 \times \text{BYA}$. This is because we do not know the relative magnitudes of the two effects predicted by theory. Some of those who are induced to work above BYA because of the offset may choose to earn above the $2 \times \text{BYA}$ level. On the other hand, some of those who would under current law earn above the $2 \times \text{BYA}$ level may choose to reduce their earnings somewhat to below that level when their earnings are supplemented by partial SSDI benefits under the offset.
7. *Current weekly earnings above weekly equivalent of $3 \times \text{BYA}$ (%)*: defined similarly to above. It is not possible to predict the direction of impacts on the percentage with earnings well above BYA, for the same reasons indicated in the point above.
8. *Currently doing volunteer work*: equals 1 if doing any volunteer work for an organization at the time of the survey and 0 otherwise. There is no theoretical prediction for how the offset will affect this outcome compared to current law. It is possible that some subjects use volunteer work to investigate their own readiness for work, which would imply a positive effect on volunteering, at least temporarily. On the other hand, if people work more for pay in order to reap the income benefits of the offset they may volunteer less (due to less available time and/or a reduced need for obtaining satisfaction through unpaid work).

We do not have theoretical predictions for the effect of EWIC compared to WIC for any of these outcomes. At the start of the demonstration, we expected that EWIC would increase the impact of the offset by improving beneficiary understanding of the offset. It is possible, however, that improved

understanding of the potential negative effect of higher earnings on non-SSDI benefits (such as SNAP and TANF) might discourage some subjects from pursuing increased employment.

Exhibit 8-1 provides measures of impact for the seven outcomes from the three policy comparisons of interest. Columns (1) to (3) show the regression-adjusted means for each outcome in the three random assignment groups. Columns (4) to (6) provide impact findings for the three policy comparisons: offset-plus-WIC compared to current law (T21 versus C2), offset-plus-EWIC compared to current law (T22 versus C2), and offset-plus-EWIC compared to offset-plus-WIC (T22 versus T21). All impact estimates presented in the chapter are exploratory and not adjusted for multiple comparisons. In the discussion of Exhibit 8-1 and of results later in the chapter, we focus attention on impact estimates that are statistically significant. We note in the discussion whether certain statistically significant impacts are consistent or inconsistent with related estimates of effects on the administrative earnings and benefits outcomes examined in Chapter 7.

Exhibit 8-1. Estimated Impacts on Employment-Related Outcomes of Stage 2 Volunteers Over the Full Follow-Up Period or at the Time of Survey Interview: All Policy Comparisons

Outcome	Average Outcome with Offset and WIC (T21) (1)	Average Outcome with Offset and EWIC (T22) (2)	Average Outcome under Current Law (C2) (3)	Estimated Impact of Offset + WIC vs Current Law (T21 vs. C2) (4)	Estimated Impact of Offset + EWIC vs Current Law (T22 vs. C2) (5)	Estimated Impact of EWIC instead of WIC Given Offset (T22 vs. T21) (6)
Any employment since random assignment (%)	40.6	42.2	38.4	2.2 (1.2)	3.8** (1.2)	1.6 (1.5)
Employed at time of 12-month survey (%)	28.5	29.3	26.5	2.0* (0.9)	2.8 (1.6)	0.8 (1.3)
Hours per week working in current (main) job	6.3	6.8	5.5	0.8** (.3)	1.3*** (.4)	0.5 (.3)
Weekly earnings in current (main) job	\$75	\$80	\$64	\$11* (\$5)	\$16** (\$7)	\$5 (\$5)
Current weekly earnings above weekly equivalent of BYA (%)	10.7	12.1	7.9	2.7** (1.1)	4.2*** (0.8)	1.4 (0.8)
Current weekly earnings above 2 x weekly equivalent of BYA (%)	3.8	3.4	3.0	0.8 (0.5)	0.4 (0.6)	-0.4 (0.5)
Current weekly earnings above 3 x weekly equivalent of BYA (%)	1.4	1.7	1.3	0.1 (0.3)	0.4 (0.3)	0.3 (0.3)
Currently doing volunteer work (%)	15.5	17.6	16.0	-0.5 (0.9)	1.6 (1.0)	2.1 (1.4)

Source: BOND Stage 2 12-Month Interim Survey, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: Weights for sample selection and survey non-response are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: T21 = 4,150, T22 = 2,602, and C2 = 3,961

Impact estimates are regression-adjusted for baseline characteristics.

*/**/** Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a two-tailed t-test (with no multiple comparisons adjustment).

The first row of Exhibit 8-1 shows results for any employment since random assignment. Thirty-eight percent of the current law group (C2 sample) had performed some work for pay since enrolling in the study when surveyed at a median duration of 13-months. We see exploratory evidence that the offset plus EWIC increased the share that ever worked by an estimated 3.8 percentage points compared to the current law group. This result contrasts with the lack of an effect on employment in calendar years 2011, 2012, or 2013 seen in the administrative data for this comparison.

In part, the difference may be due to differences in measurement for self-reported employment and administrative data. Several factors might explain the discrepant results. These factors include:

- *Difference in reference period:* While the administrative employment outcome is defined as any earnings during calendar year 2013, the first survey employment outcome is defined as any employment over about 13 months (the interval varies by sample member and spans months in 2011, 2012, and 2013). The second survey employment outcome is for a single point in time—the time of survey response, which might have been either in 2012 or 2013.
- *Differences between Social Security and total earnings:* Survey responses can include earnings that are not subject Social Security taxes, whereas the administrative data include only earnings that are subject to Social Security taxes. In addition, survey responses may omit some employment due to recall errors that do not afflict the IRS-based calendar year employment indicator.
- *Difference in samples/non-response bias:* While the administrative data provide information on all study subjects, the survey data contain information only for survey respondents. Ideally survey weights would completely adjust for such differential survey non-response. In practice, non-response weights can only correct for differential non-response correlated with observables. In as much as non-response is correlated with unobservables (as seems at least to some extent likely), non-response bias may explain some of the inconsistency in estimated impacts.

Looking further at the second row of Exhibit 8-1, we find some exploratory evidence that the offset plus WIC increased the share of subjects working at the time of the survey compared to current law, by 2.0 percentage points. This finding is consistent with exploratory findings regarding positive employment effects in calendar year 2013 reported in Chapter 7. No effects on current employment are found for the other two policy comparisons.

The third row of Exhibit 8-1 shows a positive effect of the offset, combined with either counseling approach, on hours worked per week. Inclusive of those not working at the point of interview (who are coded at 0 hours of work per week) we see a gain of about 1 hour in both offset groups on top of the 5.5 hours worked per week on average under current law.⁶⁴ Average weekly earnings (with non-workers

⁶⁴ When only looking at those who are working at the time of the survey, the hours worked per week is about 22 hours and the hourly wage is about \$13 per hour. Because the benefit offset has an effect on the share of sample members working at the time of the survey, the comparisons across groups of hours worked per week and hourly wage *only for those working* do not represent the impacts of the offset for a fixed group of beneficiaries.

coded as \$0 per week) are also pushed upward in both offset groups, when compared to current law. The increase is \$11 with WIC and by \$16 with EWIC. These represent gains of 17 and 25 percent, respectively, slightly larger in percentage terms than corresponding changes in the share currently employed or in average hours worked per week.⁶⁵ These results contrast with calendar year effects on earnings during 2013 (see Chapter 7, Exhibit 7-5), for which the results implied that neither intervention generated any confirmatory evidence of impact. The differences in measurement or time frame for the two sets of results noted above may explain why we find no effect on 2013 earnings, but do find effects on current weekly earnings at the time of the survey.⁶⁶ There is no evidence that EWIC had an effect either on number of weekly work hours or on mean weekly earnings compared to WIC.

The fifth row of Exhibit 8-1 shows positive effects on the proportion of subjects with current weekly earnings above the weekly equivalent of BYA for the two treatment groups compared to the current law control group. These point-in-time effects are consistent with those found in Chapter 7 for all of calendar year 2013.

None of the policy comparisons produced evidence of effects on weekly earnings above two times the weekly equivalent of BYA or earnings above three times the weekly equivalent of BYA. Nor is any evidence discovered that the policies have effects on the rate of volunteer work at the point of interview (last row of the exhibit).

There is no evidence that EWIC had a different effect on employment-related outcomes than WIC.

Therefore, we focus on the comparisons of hours worked per week and current earnings for *all sample members*, which are estimates for a fixed group of beneficiaries.

⁶⁵ The increase in average hours is consistent with the increase in current weekly earnings under the reasonable assumption that the offset did not decrease hourly wages for those who would have worked under current law. To see why, envision the increase in average hours worked as a combination of more treatment subjects working (raising their weekly hours from 0 to some larger number) and treatment subjects who would have worked even under current law working more hours per week. Regardless of the hourly wage rate received, “new” treatment group workers increase average earnings for the sample as a whole (because their counterparts under current law had \$0 earnings). Add the assumption that “existing” workers do not earn less per hour and it must follow that current weekly earnings for the treatment group as a whole in each policy comparison has its average earnings pushed upward by the increase in average hours.

⁶⁶ We note that the confirmatory analysis uses a higher standard of evidence (i.e., statistical significance at the $p < 0.10$ level after multiple comparisons adjustment) than the exploratory analysis (which uses statistical significance at the $p < 0.10$ level with no multiple comparisons adjustment). This difference in standard of evidence, however, does not explain the difference in results discussed here, as the administrative results on earnings during 2013 shown in Exhibit 7-5 are not statistically significant even when not adjusted for multiple comparisons.

8.2. Impact Estimates for Health Outcomes

The 12-month interim survey included a question that asked Stage 2 subjects to rate their own general health as excellent, very good, good, fair, or poor. We expect effects on health outcomes to be mediated by changes in employment behavior, and we found exploratory evidence of positive impacts on employment (Exhibits 7-6 and 8-1). Theory does not predict whether the offset will be beneficial, deleterious, or inconsequential to the health of those induced to change their employment behavior. More work could improve functioning or heighten attention to medical needs, thereby improving self-perceived health.⁶⁷ Alternatively, expanded work effort could cause health to deteriorate through added physical exertion and/or less time to attend to health care needs.

Exhibit 8-2 shows the distribution of study subjects among these health categories and how that distribution might have been changed by the offset, or by EWIC instead of WIC. Under current law, 11 percent of subjects rated their own health as excellent or very good at the time of the follow-up interview. A quarter rated their health as good, 42 percent as fair, and 23 percent as poor. The only statistically significant change in the shares attributable to the BOND interventions (of 15 potential shifts across three policy comparisons) is a 2.5 percentage-point reduction in the share reporting their health as “poor” for offset-plus-EWIC subjects compared to current law subjects.

Exhibit 8-2. Estimated Impacts on Self-Reported Health Status of Stage 2 Volunteers: All Policy Comparisons

Outcome	Average Outcome with Offset and WIC (T21) (1)	Average Outcome with Offset and EWIC (T22) (2)	Average Outcome under Current Law (C2) (3)	Estimated Impact of Offset + WIC vs Current Law (T21 vs. C2) (4)	Estimated Impact of Offset + EWIC vs Current Law (T22 vs. C2) (5)	Estimated Impact of EWIC instead of WIC Given Offset (T22 vs. T21) (6)
Health reported as excellent (%)	3.0	3.3	3.2	-0.2 (0.5)	0.1 (0.6)	0.3 (0.6)
Health reported as very good (%)	8.3	8.4	7.7	0.6 (0.7)	0.7 (0.9)	0.1 (0.8)
Health reported as good (%)	25.6	25.2	24.8	0.7 (1.2)	0.4 (1.2)	-0.3 (1.2)
Health reported as fair (%)	41.8	43.0	41.8	-0.0 (1.3)	1.2 (1.4)	1.3 (1.4)
Health reported as poor (%)	21.3	20.0	22.5	-1.2 (1.0)	-2.5** (1.1)	-1.3 (1.1)

Source: BOND Stage 2 12-Month Interim Survey, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: Weights for sample selection and survey non-response are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: T21 = 4,150, T22 = 2,602, and C2 = 3,961

Impact estimates are regression-adjusted for baseline characteristics.

*/**/** Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a two-tailed t-test (with no multiple comparisons adjustment).

⁶⁷ Given that all SSDI beneficiaries qualify for government-funded medical insurance (Medicare), we would not expect added employment to increase access to health care services and thereby improve self-perceived health.

8.3. Impact Estimates for Marital Status Outcomes

The 12-month interim survey asked Stage 2 subjects for their current marital status: married, widowed, divorced, separated, or never married. To the extent that the offset allows some SSDI beneficiaries to achieve higher incomes by combining above-BYA earnings and partial SSDI benefits, it may affect marriage and marital stability through economic channels.⁶⁸ Exhibit 8-3 shows the distribution of study subjects across the marital statuses, and how that distribution might have been changed by the BOND interventions.

As the exhibit shows, about a third of current law subjects were married at the time of the survey interview, 27 percent were divorced, and 31 percent had never been married. About 4 percent were widowed and 7 percent separated. The only statistically significant change in the shares attributable to the BOND interventions (of 15 potential shifts across three policy comparisons) is a 1.5 percentage-point increase in the share separated for offset-plus-EWIC subjects compared to offset-plus-WIC subjects. Absent any significant impact on earnings or SSDI benefits for this comparison—the two possible economic routes for BOND to affect marital outcomes—we have no theoretical basis to expect an impact of this sort.

Exhibit 8-3. Estimated Impacts on Marital Status of Stage 2 Volunteers All Policy Comparisons

Outcome	Average Outcome with Offset and WIC (T21) (1)	Average Outcome with Offset and EWIC (T22) (2)	Average Outcome under Current Law (C2) (3)	Estimated Impact of Offset + WIC vs Current Law (T21 vs. C2) (4)	Estimated Impact of Offset + EWIC vs Current Law (T22 vs. C2) (5)	Estimated Impact of EWIC instead of WIC Given Offset (T22 vs. T21) (6)
Married (%)	31.2	30.4	31.0	0.1 (0.6)	-0.6 (0.8)	-0.7 (0.7)
Widowed (%)	4.5	4.1	4.3	0.2 (0.3)	-0.2 (0.4)	-0.4 (0.4)
Divorced (%)	27.8	27.2	27.0	0.8 (0.6)	0.3 (0.6)	-0.5 (0.6)
Separated (%)	5.7	7.2	6.6	-0.9 (0.6)	0.6 (0.6)	1.5* (0.7)
Never married (%)	30.9	31.1	31.2	-0.3 (0.4)	-0.1 (0.5)	0.1 (0.5)

Source: BOND Stage 2 12-Month Interim Survey, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: Weights for sample selection and survey non-response are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: T21 = 4,150 ,T22 = 2,602 ,and C2 = 3,961

Impact estimates are regression-adjusted for baseline characteristics.

*/**/** Impact estimate is significantly different from zero at the .10/.05/.01 levels, respectively, using a two-tailed t-test (with no multiple comparisons adjustment).

⁶⁸ The Seattle-Denver Income Maintenance Experiment tested the policy of a negative income tax which provided workers with higher disposable income. Hannan, Tuma, and Groeneveld (1977) found a positive effect of the negative income tax on marital dissolution, a result which received widespread attention from policy researchers. For technical reasons, these findings are controversial and later analyses (see Cain and Wissoker 1990, and Hannan and Tuma 1990) led to more nuanced interpretation of the findings.

8.4. Summary

This chapter presents exploratory evidence consistent with the economic theory underlying BOND and its intended policy effects. The statistically significant findings also provide exploratory evidence that the offset has had positive effects on two outcomes for which the predicted sign of the impact is ambiguous: earnings and weekly hours worked at the time of the survey.

As predicted by theory, we find a positive effect of the offset plus WIC compared to current law on employment at the time of the survey interview and a positive effect of the offset plus EWIC compared to current law on any employment since random assignment. Also as predicted by theory, we find positive effects of the offset plus WIC and the offset plus EWIC on the proportion of beneficiaries with current earnings above the weekly equivalent of BYA. The magnitudes of these effects are about 2 to 4 percentage points. The positive effect on employment for the offset-plus-WIC group and the increase in the proportion earning above the weekly equivalent of BYA for both offset groups are consistent with the effects found in Chapter 7 for calendar year 2013.

In addition to effects predicted by economic theory, we also find positive effects on hours of work per week (including 0 hours for non-workers) and on weekly earnings in the current job (including \$0 for non-workers) for the two offset groups compared to current law. These impacts are larger in percentage terms (i.e., relative to current law levels in the control group) than the employment effects, translating into \$572 per year for the offset-plus-WIC group (15 percent of current law earnings) and \$884 per year for the offset-plus-EWIC group (21 percent of current law earnings).

One year after random assignment, there is no evidence that intensive counseling affects employment-related, health, and marriage outcomes differently than WIC counseling.

9. Conclusion

This chapter summarizes and discusses the findings of this first interim report on Stage 2 of the Benefit Offset National Demonstration (BOND). The purpose of Stage 2 is to learn more about the impacts on those beneficiaries most likely to use the offset (recruited and informed volunteers) and to determine the extent to which enhancements to counseling services affect impacts. Readers should keep in mind the selected nature of the sample when considering the results of this report. The results generalize to all those Stage 2-eligible (SSDI-only and ages 18-59) beneficiaries in the nation who would have volunteered had they been given the opportunity. The results, however, do not generalize to the entire SSDI population.

The process and participation analyses in this report cover the period from the beginning of Stage 2 enrollment in March 2011 through December 2014, with particular focus on calendar year 2014.⁶⁹ The impact analysis examines administrative earnings and benefits data from calendar year 2013 and survey data collected (in 2012 and 2013) about a year after each Stage 2 subject enrolled in the study. Impacts are measured as differences in outcomes between three sets of beneficiaries who volunteered to receive the offset: a random subset assigned to receive the offset plus work incentives counseling (WIC; the “T21” treatment group); a second random subset assigned to receive the offset plus enhanced work incentives counseling (EWIC; the “T22” treatment group); and a final random subset to whom standard SSDI benefit rules were applied without the offset (the “C2” current law control group).

Among the many outcomes we analyze in this report, two are of paramount interest: total earnings in 2013 and total SSDI benefits paid in 2013. All other analyses are exploratory and therefore we do not make a correction for having run multiple statistical tests of impact (as we did for the confirmatory outcomes). Even if BOND truly had no impact, one would expect some of the exploratory impact estimates to be statistically significant by chance alone. Thus, even nominally (i.e. without corrections for multiple comparisons) significant exploratory findings should be interpreted with care.

Stage 2 is implemented in 10 sites that, collectively, are nationally representative. Over the course of the demonstration, factors external to BOND in these sites, such as employment opportunities, employment support systems, and counseling services for beneficiaries subject to current law (including BOND control subjects), have varied in ways that reflect variation in the same factors across all areas in the country.

9.1. WIC and EWIC Services

As described in Chapter 3, the evidence presented in this report reinforces earlier findings that volunteers for Stage 2 of the demonstration assigned to the offset plus EWIC (T22 subjects) received more counseling services than those randomly assigned to the offset plus WIC (T21 subjects). This is as intended by the study design. Although WIC services were designed to be comparable to the certified work incentive counseling (CWIC) services provided to C2 subjects under WIPA, Derr et al. (2015) noted two differences: 1) during the period when WIPA was not funded, CWIC services might have been

⁶⁹ Derr et al. (2015), Gubits et al. (2013), and Wittenburg et al. (2012) report findings from earlier process and participation analyses.

unavailable to some C2 subjects; and 2) WIC staff provided post-entitlement services, which took time away from benefits counseling for T21 subjects in a way not present for C2 cases receiving WIPA services.⁷⁰ The current report presents evidence that assignment to T21 increased use of counseling services relative to C2 (by 6 percentage points relative to the 15 percent of C2 subjects who received benefits counseling in the year after study enrollment). Counseling services are defined as talking to someone about how work and earnings affect Social Security benefits and assistance from other programs. For services and supports that are intended to lead to employment or increased earnings (such as help finding a job or training to learn a new job or new skill), the differences between T21 and C2 were either quite small (2 percentage points at the largest) or not detectable. These small differences indicate that, in terms of their receipt of services and supports which are potentially most consequential to future employment, T21 and C2 subjects are largely equivalent. Again, this is as intended by the design.

As anticipated in planning, the percent of treatment group beneficiaries receiving work incentive counseling services declined in 2014 in comparison to a year prior, and this decline was proportionally larger for beneficiaries assigned to WIC services than beneficiaries assigned to EWIC services. At the same time, caseloads for both EWIC and WIC counselors decreased slightly at all sites, and beneficiaries receiving services reported high levels of satisfaction with counseling services.

Beyond counseling receipt, this chapter presents evidence that assignment to T22 relative to T21 increased beneficiary activities that are intended to lead to employment or higher earnings, including use of vocational rehabilitation services, completion of a work or job assessment, receipt of assistance in finding a job, enrollment in school or classes, and receipt of advice about modifying a job or work place to accommodate a disability. These differences are smaller in magnitude (13 percentage points or less), than the large differential in benefits counseling received (30 percentage points).

9.2. Awareness of How Earnings Affect Calculation of Benefits

Chapter 4 presents follow-up results showing that only about half of the Stage 2 treatment subjects provide responses consistent with an accurate understanding of how the benefit offset works—in particular, how earnings affect SSDI benefits under the offset. The results also show that the responses of T22 subjects are only slightly more accurate than those of T21 subjects. In addition, treatment beneficiaries who were working when they enrolled in the study do not provide responses consistent with an accurate understanding of the offset rules at a higher rate than those not working, despite the potentially high salience of the opportunity for combining work and benefits in the former group. The results for the control group subjects also show a relatively high amount of confusion about the standard SSDI program rules and the relationship of earnings to benefits. Only about half of control group subjects provide responses consistent with an accurate understanding of the current-law rules that apply to them.

Understanding the offset is presumably a crucial prerequisite for a behavioral response to the offset. These results suggest that this prerequisite has only been satisfied for about half of the treatment subjects. In addition, the results do not show that EWIC has substantially improved beneficiary understanding of the offset offer, one of its key purposes.

⁷⁰ In December 2013, the BOND Implementation Team removed post-entitlement work from the responsibilities of WIC providers in seven of the 10 sites.

9.3. Use of the Benefit Offset

Benefit adjustment under the offset has steadily increased for Stage 2 treatment subjects from 431 beneficiaries (5.4 percent) at the end of Stage 2 enrollment in September 2012 to 762 beneficiaries (9.5 percent) at the end of 2014. The proportion of the offset-plus-EWIC group with at least one month of offset use (11.0 percent) was 2.2 percentage points greater than the corresponding proportion for the offset-plus-WIC group (8.8 percent). It is unknown to what extent this differential reflects increased beneficiary reporting of earnings (i.e., front-door entry into the offset) on the part of the EWIC group rather than an actual difference in earnings behavior.

Several beneficiary characteristics associate with benefit adjustment in or before December 2014. Young age, good or very good general health status, back or musculoskeletal disorders and injuries as primary impairments, baseline employment, high baseline average indexed monthly earnings, and residence in a county with a high employment rate among people with disabilities were all associated with an increased likelihood of benefit adjustment, holding other characteristics constant.

We expect the percentage with a benefit offset adjustment to increase in the future, above the 9.5 percent figure for December 2014 because not all beneficiaries eligible for the offset *for* 2014 received offset payment *in* 2014. Beneficiaries who entered the adjustment process more recently likely will experience shorter delays in the administrative process than their predecessors due to recent process improvements, including a change in guidance on when to initiate work CDRs, a more effective method for identifying beneficiaries in need of a work CDR, and improvements in the quality of AEEs.

9.4. Delays with Benefit Adjustment and Improper Payments

The rate of benefit adjustment is dependent on establishment of offset eligibility as well as on administrative processes that support adjustment. By design, beneficiaries must work at a substantial level for more than 12 months (the nine-month TWP months and three-month GP) before they are eligible for offset use. Of course, the establishment of offset eligibility depends on working beneficiaries' ability and willingness to engage in sustained SGA-level work. Conditional on ability and willingness, it may still take time before a beneficiary is able to use the offset, for a variety of reasons. First, those interested, but unprepared, may require time to obtain counseling or employment-related services, address a variety of personal issues, or find an SGA-level job. Second, those who, in response to BOND, promptly initiated SGA-level work may need time to work towards total annual earnings that exceed BYA.

Once offset eligibility is established, benefit adjustment may be delayed for various reasons. We have identified three main sources of delays in proper benefit adjustment. First, the failure of some offset-eligible beneficiaries to report their earnings delays the start of the benefit adjustment process. Second, delays in processing work CDRs, once SSA recognizes the need for a work CDR, delay the determination of when eligibility for payment under the offset begins and provide a longer time frame over which the beneficiary may accrue improper payments. These lags are the result of (1) a tiered workflow for processing work CDRs between SSA field offices or processing centers and the ORDES work unit and (2) high work CDR caseload responsibility at the work unit. Finally, BSAS deficiencies have caused substantial delays in automated reconciliation, thereby delaying benefit adjustment for some first-time offset users. As a result, for many users, benefit adjustment is delayed.

The qualitative data suggest that offset use frequently leads to the occurrence of improper payments. There are a number of reasons for this. First, the retroactive nature of most initial benefit adjustments means that improper payments are often made during the first months of offset use. Second, ongoing offset users incur improper payments of some amount whenever their actual earnings differ from the amount they estimated at the beginning of the year. Finally, BSAS errors in benefit calculation have resulted in some payment errors to offset users.

As discussed in Section 4.1, the BOND logic model posits that beneficiaries need to understand the benefit offset in order to change their behavior in response to the incentive. Delays in delivery of the incentive via the adjustment process and the occurrence of improper payments are noteworthy because they may weaken beneficiary understanding of how the offset works.

9.5. Impact of the Offset and EWIC on Earnings

Chapters 7 and 8 of this report present results of the impact analysis of 2013 administrative data and 12-month interim survey data. We do not find evidence of effects on 2013 earnings for either the offset-plus-WIC or the offset-plus-EWIC groups when compared to the current-law control group. We also do not find an effect on earnings for the offset-plus-EWIC group compared to the offset-plus-WIC group.

Theory predicts that the offset will have 1) a positive average effect on earnings for those who do not engage in SGA under current law and 2) a negative average effect on earnings for those who do engage in SGA under current law. At the exploratory standard of evidence, we find some evidence consistent with these theoretical predictions. During calendar year 2013, we find exploratory evidence of a 3 percentage point employment effect for the offset-plus-WIC group compared to current law. In addition, there is exploratory evidence that the offset increased the proportion of beneficiaries with 2013 earnings above BYA for both of the offset treatment groups, by about 2 percentage points in each instance. We find similar evidence in the administrative and survey data that the offset has increased the proportion with current weekly earnings above the weekly equivalent of BYA for both offset groups.

The survey data also provide exploratory evidence that the benefit offset increased work effort and pay at the point of the follow-up survey interviews. Both the offset-plus-WIC and offset-plus-EWIC groups had higher average weekly hours and higher average weekly earnings when compared to the current-law control group.

There is no evidence in the administrative data or the survey data that EWIC differentially affects earnings or employment outcomes compared to less intensive services provided under WIC.

9.6. Impact of the Offset and EWIC on SSDI Benefits

The offset's effect on benefits is expected to arise from two sources. Theory predicts that the offset will have 1) a negative average effect on benefits for those who do not engage in SGA under current law and 2) a positive average effect on benefits for those who do engage in SGA under current law. In order to generate a reduction in benefits paid across the whole sample, the offset's reduction in benefits paid to those who become eligible for the offset but would otherwise receive full benefits under current law would have to be larger than the offset's resulting increase in benefits paid to those who would otherwise not receive benefits due to benefits suspense under current law.

Chapter 7 presents impact analysis results on receipt of SSDI benefits. We do not find evidence that the benefit offset increased or decreased SSDI benefits paid in 2013 for either the offset-plus-WIC or the offset-plus-EWIC treatment groups. We also do not find evidence of an effect of EWIC compared to WIC on SSDI benefits paid in 2013. The relatively small share of beneficiaries induced by the offset to increase earnings above BYA (about 2 percent of all treatment subjects) compared to the much larger proportion of subjects with earnings above BYA under current law (about 8 percent of all control group subjects) suggests that the offset should be having a *positive* effect on SSDI benefits for 2013, for both T21 and T22.⁷¹ Data on SSDI benefits paid for 2013 were not available at the time of this report, but we will examine that outcome in the Stage 2 final report.

At the exploratory level of evidence, we find the expected positive effects on number of months of SSDI receipt in 2013 for the two offset groups compared to current law. In contrast, we do not find exploratory evidence of an effect of EWIC compared to WIC on the number of months of SSDI receipt in 2013.

9.7. Taking Stock

What do we learn from the process, participation, and impact evidence presented in this report? Through 2014, the EWIC and WIC services appear to have been implemented generally in accord with the study's design. The difference between EWIC and WIC is greatest in the amount of benefits counseling provided and is narrower in the amount of employment services ultimately received by study subjects. The limited availability of employment services in the study sites (Section 2.6 of Derr et al. 2015) may contribute to the lack of a substantial differential in employment supports and job search assistance. Although the difference in counseling provided is quite large, this counseling surprisingly had only a very small effect on the understanding of how earnings affect benefits. Only about half of treatment subjects provided survey responses consistent with a correct understanding of the offset.

As of the end of 2014, only about 10 percent of Stage 2 treatment subjects had used at least one month of the offset. Although this amount is expected to rise in the next year (based upon the number of subjects undergoing a work CDR and the past number of subjects with back-door entry to the offset), this proportion still appears relatively small given that all Stage 2 subjects volunteered for the study presumably because they had some interest in using the offset.

The impact analysis provides no evidence of an effect of the offset on the primary outcome of 2013 earnings. It is unknown whether the offset truly has no effect on earnings or if it has a positive effect that is of a magnitude simply too small to be detected with our confirmatory statistical methods. We do find some exploratory evidence consistent with the economic theory underlying the offset (positive effects on employment and proportion earning above BYA). We also find no evidence of an effect of EWIC compared to WIC on the key outcomes of earnings and benefits in 2013. The next Stage 2 snapshot report will examine how the offset and EWIC affect earnings, employment, and benefits in 2014.

⁷¹ The emergence this effect is also dependent on the relative proportions of treatment and control subjects who have had cessation dates determined. Similar proportions or a greater proportion of control subjects would facilitate the emergence of a positive effect on benefits. The more rapid completion of work CDRs for control subjects thus seems likely to have increased the size of the positive measured effect of the offset on SSDI benefits paid.

References

- Abt Associates. "Application to Provide Work Incentives Counseling (WIC) and/or Enhanced Work Incentives Counseling (EWIC) Under the Benefit Offset National Demonstration." Submitted to the Social Security Administration (Contract SS00-10-60011), Office of Program Development & Research. Cambridge, MA: Abt Associates, June 2010.
- Bell, Stephen H., Daniel Gubits, David Stapleton, David Wittenburg, Michelle Derr, Arkadipta Ghosh, and Sara Ansell. "BOND Implementation and Evaluation. Evaluation Analysis Plan." Final Report Submitted to Social Security Administration. Cambridge, MA: Abt Associates, March 2011.
- Bloom, Howard B., and Charles Michalopoulos. "When Is the Story in the Subgroups? Strategies for Interpreting and Reporting Intervention Effects for Subgroups." *Prevention Science*, vol. 14, no. 2, 2013, pp. 179-188.
- Cain, Glen G, and Douglas Wissoker. "A Reanalysis of Marital Stability in the Seattle-Denver Income-Maintenance Experiment." *American Journal of Sociology*, vol. 95, no. 5, 1990, pp. 1235-1269.
- Derr, Michelle, Denise Hoffman, Jillian Berk, Ann Person, David Stapleton, Sarah Croake, Christopher Jones, and Jonathan McCay. "BOND Implementation and Evaluation: Process Study Report." Submitted to the Social Security Administration (contract deliverable 24d under Contract SS00-10-60011), Office of Research, Demonstrations, and Employment Support. Cambridge, MA: Abt Associates, February 2015.
- Gubits, Daniel, Michelle Derr, Jillian Berk, Ann Person, David Stapleton, Denise Hoffman, Stephen Bell, Rachel Cook, and David Wittenburg. "BOND Implementation and Evaluation: Stage 2 Early Assessment Report." Submitted to the Social Security Administration. Cambridge, MA: Abt Associates, and Washington, DC: Mathematica Policy Research, 2013.
- Gubits, Daniel, Winston Lin, and Stephen Bell. "BOND Implementation and Evaluation: First- and Second-Year Snapshot of Earnings and Benefit Impacts for Stage 2." Submitted to the Social Security Administration. Cambridge, MA: Abt Associates, and Washington, DC: Mathematica Policy Research, 2014.
- Hannan, Michael T. and Nancy Brandon Tuma "A Reassessment of the Effect of Income Maintenance on Marital Dissolution in the Seattle-Denver Experiment." *American Journal of Sociology*, vol. 95, no. 5, March 1990, pp. 1270-1298.
- Hannan, Michael T., Nancy B. Tuma, and Lyle P. Groeneveld. "Income and Marital Events: Evidence from an Income-Maintenance Experiment." *American Journal of Sociology*, vol. 82, 1977, pp. 1186-1211.
- Hanson, Robert Harold. *The Current Population Survey—Design and Methodology*, Technical Paper 40, U.S. Bureau of the Census, Washington, D.C. 1978.

Internal Revenue Service. Topic 751- Social Security and Medicare Withholding Rates.

<https://www.irs.gov/taxtopics/tc751.html>, Accessed 9/19/16

Kehn, Matthew, and Jody Schimmel. "Enrollment, Employment, and Earnings in the Medicaid Buy-In Program, 2010." Submitted to the Centers for Medicare & Medicaid Services Disabled and Elderly Health Programs Group. Washington, DC: Mathematica Policy Research, May 2013.

Livermore, Gina A., Denise Hoffman, and Maura Bardos. "Ticket to Work Participant Characteristics and Outcomes Under the Revised Regulations." Submitted to the Social Security Administration, Office of Retirement and Disability Policy. Washington, DC: Mathematica Policy Research, 2012.

O'Day, Bonnie, and David Vandergoot. "Benefit Counseling Specifications and Costs." Report submitted to Social Security Administration (contract deliverable 10.1). Cambridge, MA: Abt Associates, July 2010.

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

Schimmel, Jody, David Stapleton, and Jea Song. "How Common Is 'Parking' Among Social Security Disability Insurance Beneficiaries? Evidence from the 1999 Change in the Earnings Level of Substantial Gainful Activity." *Social Security Bulletin*, vol. 71, no. 4, 2011.

Schochet, Peter Z. "An Approach for Addressing the Multiple Testing Problem in Social Policy Impact Evaluations." *Evaluation Review*, vol. 33, no. 6, 2009, pp. 539-567.

Social Security Administration. "The Social Security Administration's Report on Continuing Disability Reviews Resulting from Work Activity." Report to House Appropriations Committee. Washington, DC: SSA, January 26, 2011.

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

Social Security Administration. Annual Statistical Supplement to the Social Security Bulletin, 2015

Social Security Administration. *Social Insurance Programs*.

<https://www.ssa.gov/policy/docs/progdesc/sspus/oasdi.pdf> ; Accessed 9/19/2016.

Stapleton, David C., Stephen H. Bell, David C. Wittenburg, Brian Sokol, and Debi McInnis. "BOND Implementation and Evaluation: BOND Final Design Report." Submitted to the Social Security Administration (contract deliverable 2.2 under Contract SS00-10-60011), Office of Program Development & Research. Cambridge, MA: Abt Associates, December 2010.

Stapleton, David, David Wittenburg, Daniel Gubits, David Judkins, David Mann, and Andrew McGuirk. "BOND Implementation and Evaluation: First-Year Snapshot of Earnings and Benefit Impacts

- for Stage 1.” Cambridge, MA: Abt Associates, and Washington, DC: Mathematica Policy Research, 2013.
- Stapleton, David, David Wittenburg, David R. Mann, Denise Hoffman, and Andrew McGuirk. “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 Program Development & Research. Washington, DC: Mathematica Policy Research, 2014.
- Stapleton, David, David Wittenburg, David R. Mann, Denise Hoffman, and Andrew McGuirk. “BOND Implementation and Evaluation: Third-Year Snapshot of Earnings and Benefit Impacts for Stage 1.” Submitted to the Social Security Administration. Cambridge, MA: Abt Associates, and Washington, DC: Mathematica Policy Research, 2015.
- Weathers, Robert, and Michelle Stegman Bailey. “The Impact of Rehabilitation and Counseling Services on the Labor Market Activity of Social Security Disability Insurance (SSDI) Beneficiaries.” *Journal of Policy Analysis and Management*, vol. 33, no. 3, 2014, pp. 623–648.
- Weathers III, Robert R. and Jeffrey Hemmeter. “The impact of changing financial work incentives on the earnings of Social Security Disability Insurance (SSDI) beneficiaries.” *Journal of Policy Analysis and Management*, vol. 30, no. 4, 2011, pp. 708-728.
- Wittenburg, David, David Stapleton, Michelle Derr, Denise W. Hoffman, and David R. Mann. “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, May 2012.
- Wittenburg, David, David R. Mann, David Stapleton, Daniel Gubits, David Judkins and Andrew McGuirk. “BOND Implementation and Evaluation: Third-Year Snapshot of Earnings and Benefit Impacts for Stage 1.” Submitted to the Social Security Administration. Cambridge, MA: Abt Associates, and Washington, DC: Mathematica Policy Research, 2015.
- Wolter, Kirk. *Introduction to Variance Estimation*. New York: Springer-Verlag. 1985
- Wright, Deborah, Gina Livermore, Denise Hoffman, Eric Grau, and Maura Bardos. “2010 National Beneficiary Survey: Methodology and Descriptive Statistics.” Washington, DC: Mathematica Policy Research, 2012.

Appendix A. Impact Methodology

This appendix describes the impact estimation methodology used in this report. Specifically, this appendix describes the estimation procedure, the multiple comparisons adjustment, the covariates included in the estimation model, the analysis sample and weights, and the analysis of impact magnitudes by subgroup. The approach has some key methodological differences from that used in the 2014 Stage 2 *Snapshot Report*. In addition, we clarify that the administrative measure of total earnings used in all reports is total earnings that are taxable for Social Security. We discuss the methodological differences and outcome measure clarification at the outset of the appendix.

A.1 Changes from the 2014 Stage 2 *Snapshot Report*

The purpose of the most notable change of the estimation procedure from the 2014 Stage 2 *Snapshot Report* is to make the impact estimates generalizable to the national population of SSDI beneficiaries who would have volunteered for Stage 2 had they been offered the opportunity to enroll in the study. The 10 BOND sites were chosen at random from the universe of 53 geographic areas that collectively include all SSDI beneficiaries in the nation. For both Stage 1 and Stage 2, the BOND *Evaluation Analysis Plan* envisioned generalizing estimated effects beyond the 10 study sites to apply to the nation (Bell et al. 2011). This goal requires (1) the use of weights that account for the random selection of sites with varying probabilities and (2) appropriate standard error calculations. However, the impact estimates in the 2014 Stage 2 *Snapshot Report* were generalizable only to beneficiaries in the ten BOND sites rather than in the entire nation. This change from the original plan was made after a test of the planned method resulted in instances of widely varying estimated standard errors across the three pairwise comparisons used (T21 vs. C2, T22 vs. C2, and T22 vs. T21) for several outcomes. Upon investigation of the issue, the evaluation team determined that the instability in standard error estimates was related to having a small number of sites—only 10. Our simulations showed that an analysis with 10 sites was vulnerable to unstable variance estimates when true cross-site variance in impacts is relatively small. To address this issue, the evaluation team developed a method in the spirit of Hanson (1978) and Wolter (1985) in which techniques were used to “stabilize” estimates of variance. This method allows us to produce valid significance tests for national impact estimates.

This 2015 Stage 2 *Interim Report* (and we envision all future Stage 2 reports) uses the revised method to present impact estimates that generalize to the national population of beneficiaries who would have volunteered for Stage 2 had they been offered the opportunity to enroll in the study.⁷² The specific changes involved in generating national estimates are (1) estimating both “cluster-robust” and “robust, unclustered” standard errors and using the larger of the two standard errors for significance testing, (2) using 9 degrees of freedom in significance tests to reflect the 10 BOND sites, (3) using sampling weights that account for the random selection of study sites, and (4) making the refinement to the multiple comparisons procedure described in Section A.3.

In addition to the changes related to producing nationally generalizable impact estimates, we also made a correction to the “within-site” component of the analytic weight. During the preparation of this report, we

⁷² Appendix C of this report presents newly-estimated impacts for 2011 and 2012, previously presented in the 2014 Stage 2 *Snapshot Report*. The results in that report were generalizable only to the ten BOND sites, rather than to the entire nation.

discovered that the construction of the analytic weights used in the 2014 Stage 2 *Snapshot Report* did not fully capture the process by which the Implementation Team randomly selected subjects for the Stage 2 outreach waves. This discovery led us to revise the construction of the within-site weight component. Compared to the original method of weight construction, the new method:

- Accounts for the fact that probability of selection for Stage 2 outreach varied by SSDI entitlement month for subjects whose SSDI entitlement started between January 2008 and July 2009 (the most recently entitled subjects were more likely to be selected for outreach);⁷³ and,
- Accounts for some Stage 2 outreach waves being limited to certain geographical zones within sites.

We accounted for these two factors by calculating probabilities of selection to each of five groups of outreach waves rather than to each of three groups of waves. These changes in the construction of the within-site weight component resulted in a revised set of within-site weights that is similar to the original set of within-site weights. The correlation between the original and revised within-site weights is .885. We provide more details about the revised weights in Section A.6.

A.2 Social Security Earnings

The Social Security Administration made the Summary Segment of the Master Earnings File (MEF) available to this evaluation. The MEF is SSA's primary repository of earnings data for the US population. The MEF contains all information from the W-2 forms submitted annually by employers to SSA for each paid employee and the relevant information for calculating benefits from 1040-SE forms that self-employed individuals send to the IRS. The Summary Segment summarizes a limited set of data from the MEF. Therefore a limited set of information is available to the BOND evaluation. For example, the Summary Segment does not include total earnings subject to income tax. Rather, the Summary Segment contains data on annual earnings that are subject to Old-Age, Survivors, and Disability Insurance (OASDI) taxes, otherwise known as Social Security taxes. The revenue from OASDI taxes funds insurance benefit payments to retired workers and their spouses and children; survivors of deceased workers; and disabled workers and their spouses and dependent children (SSA Social Insurance Programs, accessed 9/19/2016). We next describe how Social Security taxes are reported to SSA.

The W-2 form lists several types of earnings amounts (Exhibit A-1 provides an image of the form). In Box 1 of the W-2 form, employers are required to report an employee's total wage, tips, and other compensation that is subject to income tax. Several types of wages are excluded from Box 1, such as payments to retirement accounts (401Ks). Employers are required to report social security taxable earnings in Box 3 ("Social security wages") and Box 7 ("Social security tips"); payments to retirement accounts are taxed, and therefore included. Social Security taxable earnings are capped at a maximum (IRS 2016). For 2013, the maximum was \$113,700.⁷⁴ Amounts above this maximum are not subject to Social Security taxes, and thus the sum of Box 3 and Box 7 will never exceed the maximum, regardless of

⁷³ The variation in selection probabilities by entitlement month was caused by the short/long entitlement duration status used for oversampling being determined as of the time of each Stage 2 outreach wave (i.e., the dividing line for short/long status was different for each outreach wave).

⁷⁴ The maximum social security taxable earnings varies each year, generally increasing with inflation. For 2011 and 2012, the maximums were \$106,800 and \$110,100, respectively.

what is reported in Box 1. The sum of Box 3 and Box 7 could be less than Box 1 (for example, because wages exceed the wage base limit). However, the sum of Box 3 and Box 7 could also be more than Box 1 (for example, payments to retirement accounts and dependent care accounts are taxable for Social Security in the year they are earned).

The Summary Segment of the MEF contains the summed total of the Social Security earnings amounts from all of the W-2 forms (Box 3 and Box 7) and the 1040-SE form posted to the MEF. Therefore, the summed totals of Social Security earnings amounts are the data available to the BOND evaluation. There are some disadvantages to relying on Social Security earnings as an overall earnings measure. Social Security earnings may be different from all employment income for the following reasons:

- (1) Not all jobs are covered by Social Security. Non-covered jobs include some state and local government positions and railroad workers. Only six percent of the US workforce does not participate in Social Security (Annual Statistical Supplement to the Social Security Bulletin, 2015). For example, teachers in some states do not pay Social Security taxes on their earnings. Of the BOND sites, teachers in Colorado, Maine, and Massachusetts fall into this category.
- (2) For each W-2 and 1040-SE form, Social Security earnings are capped at a maximum taxable amount, \$113,700 for 2013. However, very few have earnings at or above that amount (in 2013, 0.01 percent of all Stage 2 participants have earnings at or above the taxable amount). In addition, beneficiaries who are earning at or above that amount are unlikely to have a behavioral response to the offset.
- (3) Not all work and earnings will be reported on a W-2 or 1040-SE form (i.e. “under-the-table” earnings).

As the earnings data available on the Summary Segment of the MEF do not include all earnings countable towards SGA, our estimates of earnings, employment, and proportion working above BYA may have a small downward bias compared to measures defined by total earnings countable towards SGA.⁷⁵ In addition, the estimate of the impact of the offset on earnings, employment and proportion working above BYA may have a small downward bias if some who are encouraged to work choose jobs not covered by Social Security (item number one in the list, above). On the other hand, the estimate could have a slight upward bias due to the fact that the offset may induce some people with under the table earnings to report them. Measures of weekly earnings and employment taken from survey data should not be subject to the same source of bias (though they are subject to other biases; in particular, recall bias and non-response bias).

⁷⁵ Not available for this evaluation, the Social Security Administration also has records of Box 1 earnings in the Detailed Segment of the MEF. Still, Box 1 earnings data would not offer a complete picture of earnings countable towards SGA because not all work and earnings are reported on a W-2 or 1040-SE form.

Exhibit A-1. W-2 Wage and Tax Statement

22222		Void <input type="checkbox"/>	a Employee's social security number		For Official Use Only ▶ OMB No. 1545-0008	
b Employer identification number (EIN)			1 Wages, tips, other compensation		2 Federal income tax withheld	
c Employer's name, address, and ZIP code			3 Social security wages		4 Social security tax withheld	
			5 Medicare wages and tips		6 Medicare tax withheld	
			7 Social security tips		8 Allocated tips	
d Control number			9		10 Dependent care benefits	
e Employee's first name and initial		Last name	Suff.	11 Nonqualified plans		12a See instructions for box 12
f Employee's address and ZIP code			13 Statutory employee <input type="checkbox"/> Retirement plan <input type="checkbox"/> Third-party sick pay <input type="checkbox"/>		12b	
			14 Other		12c	
					12d	
15 State	Employer's state ID number	16 State wages, tips, etc.	17 State income tax	18 Local wages, tips, etc.	19 Local income tax	20 Locality name

Form W-2 Wage and Tax Statement 2013 Department of the Treasury—Internal Revenue Service
 Copy A For Social Security Administration — Send this entire page with Form W-3 to the Social Security Administration; photocopies are not acceptable.
 For Privacy Act and Paperwork Reduction Act Notice, see the separate instructions. Cat. No. 10134D

A.3 Estimation Procedure

A.3.1. Regression Model

As shown in Appendix B of the 2014 Stage 2 Snapshot Report, our basic impact estimation model is:

$$(1) \quad y_{ij} = \beta_0 + \beta_1 T_{21ij} + \beta_2 T_{22ij} + X_{ij} \Phi + \varepsilon_{ij}$$

where y_{ij} is an outcome measure for beneficiary i in site j ($j = 1, 2, \dots, 10$),

T_{21ij} = an indicator of whether beneficiary i in site j has been randomized into the T21 group (= 1 if so, = 0 if in T22 or C2 groups),

T_{22ij} = an indicator of whether beneficiary i in site j has been randomized into the T22 group (= 1 if so, = 0 if in T21 or C2 groups),

X_{ij} = a vector of baseline characteristics (listed in Exhibits C-1 and C-2) for individual i in site j ,

β_0 = the model intercept,

β_1 = the overall impact of the T21 treatment (vs. no treatment for the C2 group),

β_2 = the overall impact of the T22 treatment (vs. no treatment for the C2 group),

Φ = a vector of coefficients, and

ε_{ij} = an error term that is beneficiary- and site-specific (discussed below).

In this model, the incremental impact of the T22 treatment compared with the T21 treatment is the difference $\beta_2 - \beta_1$. We estimate model (1) by weighted least squares regression, using the SURVEYREG procedure in SAS.

The estimated standard errors are computed using a “variance stabilization” method, which errs on the side of a larger standard error when there is discrepancy between two different methods for estimating the standard error.⁷⁶ We first estimate the model while specifying that ε_{ij} is correlated within site and independent across sites (i.e., an “unconditional” standard error that treats the sites as clusters, sometimes called the “cluster-robust” standard error). Next, we estimate the model while specifying that ε_{ij} is independent between and within sites (i.e., a “conditional” standard error that treats the sites as strata, with sites entering into the model as dummy variables, sometimes called the “robust, unclustered” standard error).⁷⁷

These two standard errors are appropriate for different purposes. The unconditional standard errors are designed to support inferences about what would happen with a national implementation of one treatment variation or another. In contrast, the conditional standard errors are designed to support inferences about what would happen if one treatment variation or another were implemented throughout the 10 sites. Standard theoretical statistical analysis implies that the *true* unconditional standard errors are at least as large and usually larger (often considerably) larger than the conditional standard errors. This is because unconditional inference requires us to extrapolate from the 10 sites to the rest of the nation. However, the *estimated* (not true) unconditional standard errors are noisy (unstable) due to the fact that they use observed variation among a small number of sites—only 10. To address this issue due to the small number of sites, we stabilize the unconditional standard errors by replacing them with corresponding conditional standard errors whenever the unconditional standard error is smaller than the conditional standard error.⁷⁸ Both sets of standard errors are estimated using Taylor series linearization in the SURVEYREG procedure in SAS.

⁷⁶ The specific method described here is in the spirit of Hanson (1978) and Wolter (1985), where other variance stabilization methods were used.

⁷⁷ When estimating the unconditional standard error, the covariates omit site dummies. When estimating the conditional standard error, the covariates include site dummies.

⁷⁸ Our simulations have shown that the likelihood of the conditional standard error being larger than the unconditional standard error increases as the true cross-site variance of impacts decreases. In a simulation of very small true cross-site variance of impacts, we found that the 90% confidence interval contained the true effect 92.3% of the time. This result shows that when true cross-site variance is relatively small (and so occasionally the conditional standard error is larger than the unconditional standard error), the variance stabilization method is conservative, sacrificing some statistical power to avoid displaying grossly inconsistent variance estimates for pairs of statistics where generally similar variances are expected. Given the statistical issues, such conservative inference seems appropriate.

For both the unconditional and the conditional model, we compute estimated standard errors for the estimates of β_1 , β_2 , and $\beta_2 - \beta_1$ from the estimated variance-covariance matrix, using the ESTIMATE statement. Following the variance stabilization method, we report the standard error for each estimate that is the maximum of the conditional standard error and the unconditional standard error. In all cases, the impact estimate we report is from the unconditional model. The p-value for the t-statistic implied by the impact estimate and the reported standard error is calculated using 9 degrees of freedom, regardless of whether the reported standard error is the conditional or unconditional standard error.⁷⁹

The computational method that we use for estimation of Stage 2 impacts is somewhat simpler than the method used for Stage 1 because Stage 2 has many fewer subjects (about 13,000 rather than close to 1 million in Stage 1). The smaller computational burden allows estimation of the model (1) in one step for Stage 2, whereas for Stage 1 the greater computational burden led us to use multiple steps.

A.4 Multiple Comparisons

The 2015 Stage 2 *Interim Report* takes the same approach to the multiple comparisons problem as the 2014 Stage 2 *Snapshot Report*:

1. We separate the hypothesis tests into “confirmatory” and “exploratory” tests, as specified in the *Evaluation Analysis Plan*, prior to conducting the impact analysis. Only the two most important outcomes—annual earnings and SSDI benefits paid in a calendar year—are included in the confirmatory group, and in this report, only the impact estimates for those outcomes in 2013 are treated as confirmatory. All other impact estimates, including all estimates for subgroups and impact findings concerning other outcomes, are considered exploratory. Statistically significant findings from confirmatory analyses are interpreted as evidence that the benefit offset had impacts on earnings or SSDI benefits. In contrast, statistically significant findings from exploratory analyses are characterized as suggestive of demonstration impacts in other areas.
2. We implement a multiple comparisons adjustment procedure for our two confirmatory outcomes. The procedure controls the “familywise error rate”—the probability of rejecting at least one null hypothesis in a family of hypothesis tests when all null hypotheses are true.

As was done in the 2014 Stage 2 *Snapshot Report*, we handle the pairwise comparisons of T21 vs. C2 and T22 vs. C2 together in a multiple comparisons adjustment (total of four statistical tests). We handle the EWIC vs. WIC (T22 vs. T21) comparison separately (total of two statistical tests) from the other two comparisons.

The new method for estimating standard errors necessitates a corresponding refinement in the multiple comparisons procedure used to adjust the p-values of the two confirmatory significance tests. As in the 2014 Stage 2 *Snapshot Report*, we use the Westfall-Young permutation step-down method as the multiple comparisons procedure in this report. This approach involves generating a large number of re-randomized samples and comparing the p-values (for the two confirmatory outcomes of annual earnings and annual

⁷⁹ It is the national representativeness of the impact results that leads to the use of 9 degrees of freedom in the t-tests. Results that only generalize to the 10 BOND sites would use a number of degrees of freedom based on the number of study subjects in the impact comparison, rather than the number of study sites.

SSDI benefits) in each re-randomized sample with the p-values from the actual Stage 2 sample. Because the p-values from the actual Stage 2 sample are now derived from the larger of either unconditional or conditional standard errors, the p-values from the re-randomized samples need to be defined such that they use the same standard error chosen for each particular impact estimate in the actual Stage 2 sample. For example, suppose that in the actual Stage 2 sample, the unconditional standard error for the annual earnings impact for the T21 versus C2 comparison is larger than the conditional standard error, while the conditional standard error is larger for the annual SSDI benefits impact for the T21 versus C2 comparison. Then, in each re-randomized sample, the p-value for the earnings impact for that comparison needs to be derived from the unconditional standard error and the p-value for the benefits impact needs to be derived from the conditional standard error.⁸⁰

Exhibit A-2 shows both unadjusted and adjusted p-values for the 2013 confirmatory impact estimates in this report.

Exhibit A-2. Stage 2 Impact Estimates on Confirmatory Outcomes Illustrating the Multiple Comparison Adjustment on p-values

Comparison	Confirmatory Outcome	Impact Estimate (1)	p-value (Unadjusted) (2)	p-value (Multiple Comparisons Adjustment) (3)
First Multiple Comparison Procedure (4 hypothesis tests)				
T21 vs. C2	Total earnings in 2013	\$385 (\$179)	0.060	0.135
T21 vs. C2	Total SSDI benefits paid in 2013	\$357 (\$131)	0.023	0.104
T22 vs. C2	Total earnings in 2013	\$366 (\$284)	0.229	0.259
T22 vs. C2	Total SSDI benefits paid in 2013	\$374 (\$135)	0.022	0.104
Second Multiple Comparison Procedure (2 hypothesis tests)				
T22 vs. T21	Total earnings in 2013	\$-19 (\$198)	0.924	0.994
T22 vs. T21	Total SSDI benefits paid in 2013	\$16 (\$172)	0.926	0.994

Source: Analysis of SSA administrative records (from the MEF, BODS, MBR, and SSR) and the Stage 2 Baseline Survey.

Notes: See Chapter 6 for variable definitions. All earnings outcomes are based on a measure of earnings subject to Social Security taxes (see Appendix A.2 for further detail). Weights are used to ensure that the BOND subjects who met analysis criteria are representative of volunteers for offset participation in the nation. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: T21 = 4,854, T22 = 3,041, C2 = 4,849.

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.

⁸⁰ Our simulations show that this refinement controls the family-wise error rate at the expected level.

A.5 Covariates

Our estimation model includes almost all of the covariates used in the 2014 Stage 2 *Snapshot Report*. The only change in this set of covariates from the 2014 Stage 2 *Snapshot Report* is that we do not always use site dummies and never use covariates interacted with site. We include the site dummies when we are estimating the model with conditional standard errors; however, we do not include the site dummies when the model is estimated using unconditional standard errors. The reason for the omission of site dummies in unconditional (or, clustered) model is that fixed effects and random effects cannot be simultaneously identified.

Exhibits A-3 and A-4 show the covariates that we use, and distinguishes their origin: administrative data and baseline survey.

Exhibit A-3. Administrative-Data Covariates Included in Stage 2 Impact Regressions

Covariates (measured at baseline unless otherwise specified)
Age
Age (squared)
AIME (Average Indexed Monthly Earnings) as of May 2011
AIME (Average Indexed Monthly Earnings) as of May 2011 (squared)
AIME (Average Indexed Monthly Earnings) as of May 2011 are equal to zero
Any employment in 2010 ^a
County 2010 employment rate for people with a disability
County April 2011 unemployment rate
Dummy for missing 2010 unemployment rate and missing rural status
Dummy for missing employment rate for people with a disability
Earnings in 2010 ^a
Gender
Has a representative payee
Has SSDI start date on or after January 1, 2010 (very short-duration beneficiary)
Interaction of very short-duration x 2010 earnings ^a
Interaction of monthly benefit amount at baseline and AIME as of May 2011
Interaction of age and number of years receiving SSDI
Interaction of earnings in 2010 and randomly assigned in 2012 ^a
Interaction of earnings in 2011 and randomly assigned in 2012 ^a
Is a disabled adult child (DAC) beneficiary
Is a disabled widow(er) beneficiary (DWB)
Is a dually entitled DAC beneficiary
Is a dually entitled DWB
Monthly benefit amount (MBA) at baseline
Monthly benefit amount (MBA) at baseline is equal to zero
Number of years receiving SSDI
Number of years receiving SSDI (squared)
Primary impairment category: Neoplasms Mental disorders Back or other musculoskeletal Nervous system disorders Circulatory system disorders Genitourinary system disorders Injuries Respiratory Severe visual impairments Digestive system Other impairments Unknown impairments
Randomly assigned in 2012
Receives written beneficiary notices in Spanish
Rural area dummy
Short-duration SSDI receipt (36 months or fewer)
Site dummies ^b
SSI receipt dummy

^a Included in model for all earnings outcomes and total SSDI benefits only. Earnings data consist of Social Security taxable earnings as measured through administrative data (see Appendix A.2).

^b Included in estimation of conditional standard error only.

Exhibit A-4. Baseline Survey Covariates Included in Stage 2 Impact Regressions

Covariates
Marital status (married, widowed, divorced, separated, never married)
Cohabiting dummy
Education dummies (LT HS, HS/GED, Some college, 4yr college degree)
Child under age 18 in household ^b
Race/ethnicity (African American, Hispanic, White, Asian, Other)
Working at baseline (baseline survey)
Lives in non-group residence (single family home, regular apartment, or mobile home)
Enrolled in school or taking classes
Full-time student
Engaged in volunteer work
Health dummies (excellent, very good, good, fair, poor)
Personal goals include getting a job (if not working), moving up in a job, or learning new job skills
Health limits in moderate activities "a lot"
Health limits climbing several flights of stairs "a lot"
Emotional well-being (composite scale)
Stayed overnight in a hospital in 12 months before baseline
Needs the help of another to get around inside home
Needs the help of another to get around outside home
Earned \$12,000 or more in year before baseline
Change in health during past year (much better, somewhat better, about the same, somewhat worse, or much worse)
Not in the labor force
Self-employed
Employed at a steady job (neither temporary nor seasonal)
Employed full-time (35 or more hours per week)
Employed at a job with health insurance benefits
Employed at a job with many benefits (health insurance, paid sick days and vacation, long- and short-term disability benefits, and retirement benefits)
Able to drive a car
Has access to a car, truck, or van
Perceived barriers to employment (composite scale)
Able to do the same type of work as was doing when first became limited in the kind or amount of work or other daily activities one could do
Stayed in hospital more than 30 days in year before baseline
Body mass index 25 or higher
Emotional problems limited activities most or all of the time
Interaction of employed full-time and rural area
Interaction of employed full-time and 4-year college degree
Interaction of employed full-time and engaged in volunteer work
Interactions of employed full-time and health dummies
Interaction of employed full-time and self-employed
Interaction of employed full-time and job with health insurance
Interaction of employed full-time and job with many benefits
Interaction of employed full-time and able to do the same type of work as was doing when first became activity-limited
Interaction of employed full-time and access to a car, truck, or van
Interactions of change in health and earned \$12,000 or more in year before baseline
Interaction of not in the labor force and short-duration SSDI receipt
Interaction of not in the labor force and very short-duration beneficiary

Covariates
Interaction of not in the labor force and 2010 earnings ^a
Three-way interaction of not in the labor force, 2010 earnings, and very short-duration beneficiary ^a
Interaction of self-employed and county April 2011 unemployment rate
Interaction of self-employed and age
Interaction of self-employed and squared age
Interaction of self-employed and able to drive a car
Interactions of employed at a steady job and primary impairment category
Interaction of employed at a job with health insurance benefits and MBA at baseline
Interaction of employed at a job with health insurance benefits and very short-duration beneficiary
Interaction of employed at a job with health insurance benefits and 2010 earnings ^a
Three-way interaction of employed at a job with health insurance benefits, 2010 earnings, and very short-duration beneficiary ^a
Interaction of employed at a job with many benefits and county 2010 employment rate for people with a disability
Interaction of employed at a job with many benefits and dummy for missing employment rate for people with a disability
Interactions of employed at a job with many benefits and marital status
Interactions of employed at a job with many benefits and race/ethnicity
Interactions of employed at a job with many benefits and health dummies
Interactions of employed at a job with many benefits and health limits in moderate activities "a lot"
Interaction of perceived barriers to employment and earned \$12,000 or more in year before baseline
Interactions of primary impairment category and able to do the same type of work as was doing when first became activity-limited
Interactions of MBA and able to do the same type of work as was doing when first became activity-limited

^a Included in model for all earnings outcomes and total SSDI benefits only.

^b The 2014 Stage 2 *Snapshot Report* erroneously indicated the inclusion of three covariates in the impact model for that report: "Child under age 18 in household," "Number of months worked in previous 3 years," and "Square of number of months worked in previous 3 years." For this report, the "Child under age 18 in household" covariate is included in all impact analysis. The baseline survey did not collect sufficient information from subjects to create the "Number of months worked in previous 3 years" and "Square of number of months worked in previous 3 years" covariates. Therefore, these variables are not included in the impact analysis of this report (and will not be included in future Stage 2 reports).

A.6 Subgroup Analysis

The Stage 2 *Interim Report* presents impact estimates on administrative earnings and benefit outcomes for the full set of subgroups listed in the 2011 *Evaluation Analysis Plan*. These are:

<ul style="list-style-type: none"> • Short-duration beneficiaries (those receiving benefits 36 months or less when entering BOND) 	Vs.	<ul style="list-style-type: none"> • Longer-duration beneficiaries (those receiving benefits 37 months or more when entering BOND)
<ul style="list-style-type: none"> • Beneficiaries who are working at baseline 	Vs.	<ul style="list-style-type: none"> • Beneficiaries who are not working at baseline
<ul style="list-style-type: none"> • Beneficiaries with access to Medicaid buy-in programs 	Vs.	<ul style="list-style-type: none"> • Beneficiaries without access to Medicaid buy-in programs
<ul style="list-style-type: none"> • Younger beneficiaries (less than age 55) 	Vs.	<ul style="list-style-type: none"> • Older beneficiaries (age 55 and older)
<ul style="list-style-type: none"> • Beneficiaries with primary impairment of major affective disorder 	Vs.	<ul style="list-style-type: none"> • Beneficiaries with all other primary impairments
<ul style="list-style-type: none"> • Beneficiaries with primary impairment of back disorder 	Vs.	<ul style="list-style-type: none"> • Beneficiaries with all other primary impairments

We also examine an additional set of subgroups based on educational attainment, given the possibility of differential response to the offset due to (a) different earnings potential and (b) differing ability to understand the offset offer.

<ul style="list-style-type: none"> • Beneficiaries with any postsecondary degree 	Vs.	<ul style="list-style-type: none"> • Beneficiaries with H.S. degree or less (includes those with postsecondary classes but no degree)
---	-----	--

For each set of subgroups, we present four exhibits:

- T21 + T22 Vs. C2
- T21 Vs. C2
- T22 Vs. C2
- T22 Vs. T21

In addition to the three pairwise comparisons, we have added the comparison of T21 + T22 Vs. C2, which pools the two treatment groups together. We have added this comparison as a way to address the smaller sample size of some subgroups. Pooling the two treatment groups increases statistical power when looking for evidence of differential effects.

For subgroup analyses, we use the following extension of model (1) as explained in Appendix B of the 2014 Stage 2 *Snapshot Report*:

$$(2) \quad y_{ij} = \beta_0 + \beta_1 T_{21ij} + \beta_2 T_{22ij} + \beta_3 S_{ij} + \beta_4 (T_{21ij} S_{ij}) + \beta_5 (T_{22ij} S_{ij}) + X_{ij} \Phi + \varepsilon_{ij}$$

where

S_{ij} = 1 or 0 depending on which of two possible subgroups beneficiary i in site j belongs to,
 β_1 = the impact of the T21 treatment (vs. no treatment for the C2 group) for the subgroup with $S_{ij} = 0$,
 β_2 = the impact of the T22 treatment (vs. no treatment for the C2 group) for the subgroup with $S_{ij} = 0$,
 β_3 = the difference between the two subgroups' expected outcomes in the absence of treatment,
 β_4 = the difference between the two subgroups in the impacts of the T21 treatment,
 β_5 = the difference between the two subgroups in the impacts of the T22 treatment,

and the rest of the notation is as defined above. In this model, for the subgroup with $S_{ij} = 0$, the impact of T21 vs. C2 is β_1 , the impact of T22 vs. C2 is β_2 , and the incremental impact of T22 vs. T21 is the difference $\beta_2 - \beta_1$. For the subgroup with $S_{ij} = 1$, the impact of T21 vs. C2 is $\beta_1 + \beta_4$, the impact of T22 vs. C2 is $\beta_2 + \beta_5$, and the incremental impact of T22 vs. T21 is $\beta_2 + \beta_5 - \beta_1 - \beta_4$. The difference between the two subgroups in the incremental impact of T22 vs. T21 is $\beta_5 - \beta_4$. Similar to the estimation of model (1), we use the SURVEYREG procedure in SAS to compute weighted least squares estimates of all model parameters. As we did for the estimation of model (1), we follow the variance stabilization method: for each hypothesis test we use the maximum of the unconditional standard error and the conditional standard error. Similar to the estimation of model (1), we use the ESTIMATE statement to obtain estimated standard errors for parameter estimates and their sums and differences.

A.7 Analysis Weights and Sample Exclusions

A.7.1. Construction of Analysis Weights

This section describes the construction of analysis weights for the impact analysis presented in this *Interim Report*.

Analysis Weights for Administrative Outcomes

The construction of the analysis weights for administrative outcomes differs from that in the 2014 Stage 2 *Snapshot Report* in that (i) it includes a component that accounts for the random selection of sites into BOND and (ii) the within-site component has been revised.

The first component of the weight is necessary for impact estimates to be generalizable to the national population of Stage 2-eligible beneficiaries who would have volunteered for Stage 2 had they been asked. As explained in Stapleton et al. (2010), 10 SSA area offices were selected as sites for BOND from eight strata defined by census region (Northeast, Midwest, South, or West) and proportion of beneficiaries living in Medicaid buy-in states (low or high). With one exception, a single area office was selected from each stratum. The exception is that three area offices were selected from the low Medicaid Buy-in stratum

in the South region, a region with many more area offices and beneficiaries than the other strata.⁸¹ The area offices were selected in each stratum using probability proportional to size systematic sampling, in which size is defined as the number of SSDI beneficiaries served by the area office.

The second component of the weight, the within-site component, differs from that used in the 2014 Stage 2 *Snapshot Report* in that it better recovers the true probability of selection to the Stage 2 outreach. The random selection of Stage-2 eligible subjects into each outreach wave occurred in within-site strata defined by geographical zone and by the distinction between short- and long-duration entitlement *as of the mailing date* for the wave. (Initial outreach mailing began in January 2011 and ended in May 2012.) For logistical reasons, some outreach waves were limited to certain geographical zones within sites. Most waves oversampled short-duration subjects in order to ensure that at least half of Stage 2 volunteers were short-duration subjects (a goal of the Stage 2 design). The degree of oversampling differed across waves. Stage 2 eligibility for the first three outreach waves (known as the Stage 2 “pilot”) was based on the December 2010 sample file and eligibility for the remaining outreach waves was based on the subsequent April 2011 (for most subjects) and June 2011 (for newly added subjects)⁸² sample files. Probability for random selection to the Stage 2 outreach thus differed according to three factors:

- a) SSDI entitlement start month (which determined short-/long-duration status for each outreach wave),
- b) Geographical zone of residence, and
- c) Stage 2 eligibility status in December 2010, April 2011, and June 2011 files

There are four different possible Stage 2 eligibility statuses:

- 1 = beneficiary is Stage 2-eligible in December 2010 sample file only,
- 2 = beneficiary is Stage 2-eligible in December 2010 and April 2011 sample files,
- 3 = beneficiary is Stage 2-eligible in April 2011 sample file only, and
- 4 = beneficiary is Stage 2-eligible in June 2011 sample file of newly added subjects.

For administrative outcomes, each Stage 2 sample member is assigned an analysis weight given by:

$$w_{mkj\ell}^A = \left(\frac{N_m}{N_{mk}} \right) \left(\frac{1}{P_{mkj\ell}} \right)$$

where:

⁸¹ Because three area offices were selected from this stratum, the first component of all analysis weights for sample members from this stratum is $\left(\frac{N_m}{3N_{mk}} \right)$, rather than $\left(\frac{N_m}{N_{mk}} \right)$.

⁸² To insure that there would be sufficient short-duration subjects in the outreach waves (and thereby meet the goal of having at least half the volunteers be of short-duration), very recently-entitled Stage 2-eligible subjects were added in June 2011 to the pool of subjects from which outreach waves were drawn.

- $w_{mkj\ell}^A$ is the Stage 2 weight for administrative outcomes for a volunteer who is served by site k within site-selection stratum m and who is of longitudinal eligibility category j (one of the four Stage 2 eligibility statuses listed above) and within-site stratum ℓ (defined by SSDI entitlement start month and geographical zone)⁸³;
- N_m is the number of SSDI beneficiaries in site-selection stratum m ($m= 1, \dots, 8$);
- N_{mk} is the number of SSDI beneficiaries served by site k within site-selection stratum m ;
- $P_{mkj\ell}$ is the unconditional probability of random selection to any one of the outreach waves of the Stage 2 recruitment effort for a volunteer of site-selection stratum m , site k , longitudinal eligibility category j , and within-site stratum ℓ .

The first component of the weight is included so that the Stage 2 results represent Stage 2-eligible beneficiaries in the nation who would have volunteered had they been offered the opportunity to enroll in the study. The second component of the weight is the reciprocal of the probability of being selected into any outreach wave.

The unconditional probabilities of selection into any outreach wave differ for the four longitudinal eligibility categories of beneficiaries, and are based on conditional probabilities for inclusion into five sets of outreach waves: (a) the pilot waves (January to April 2011), (b) the June 2011 wave, (c) the July 2011 wave, (d) the August 2011 wave, and (e) all other later waves from September 2011 through May 2012. Subjects of longitudinal eligibility category $j=1$ were only eligible to be selected into the pilot waves, those of category $j=2$ were eligible for all waves, those of category $j=3$ were eligible beginning with the June 2011 wave, and those of category $j=4$ were eligible beginning with the July 2011 wave. The conditional probabilities for inclusion into each of these sets of waves are:

$$Pr(PILOT)^{mk\ell} = \left(\frac{N_{mk\ell}^{PILOT}}{N_{mk1\ell} + N_{mk2\ell}} \right)$$

$$Pr(JUNEWAVE|not previously selected)^{mk\ell} = \left(\frac{N_{mk\ell}^{JUNEWAVE}}{N_{mk2\ell} + N_{mk3\ell} - N_{mk2\ell}^{PILOT}} \right)$$

$$Pr(JULYWAVE|not previously selected)^{mk\ell} = \left(\frac{N_{mk\ell}^{JULYWAVE}}{N_{mk2\ell} + N_{mk3\ell} + N_{mk4\ell} - N_{mk2\ell}^{PILOT} - N_{mk\ell}^{JUNEWAVE}} \right)$$

⁸³ The number of possible within-site strata ℓ varies in the calculation of selection probabilities for different sets of outreach waves. For the pilot waves, there are 18 possible strata per site defined by SSDI entitlement start month (January 2008 and before, each month from February 2008 to May 2008, and June 2008 and after) and geographical zone (3 zones per site). For the June 2011 wave, there are 6 possible strata per site ($\{\text{July 2008 and before, August 2008 and after}\} \times \{3 \text{ zones per site}\}$). For the July 2011 wave, there are 6 possible strata per site ($\{\text{August 2008 and before, September 2008 and after}\} \times \{3 \text{ zones per site}\}$). For the August 2011 wave, there are up to 10 possible strata per site ($\{\text{September 2008 and before, October 2008 and after}\} \times \{\text{up to 5 zones per site}\}$). And for the remaining waves, there are up to 45 possible strata per site ($\{\text{October 2008 and before, each month from November 2008 to May 2009, and June 2009 and after}\} \times \{\text{up to 5 zones per site}\}$). The Greater Detroit and Greater Houston sites have 3 zones, the DC Metro site has 4 zones, and all other sites have 5 zones.

$$Pr(AUGWAVE|not\ previously\ selected)^{mk\ell} = \left(\frac{N_{mk\ell}^{AUGWAVE}}{N_{mk2\ell} + N_{mk3\ell} + N_{mk4\ell} - N_{mk2\ell}^{PILOT} - N_{mk\ell}^{JUNEWAVE} - N_{mk\ell}^{JULYWAVE}} \right)$$

$$Pr(LATERWAVES|not\ previously\ selected)^{mk\ell} = \left(\frac{N_{mk\ell}^{LATERWAVES}}{N_{mk2\ell} + N_{mk3\ell} + N_{mk4\ell} - N_{mk2\ell}^{PILOT} - N_{mk\ell}^{JUNEWAVE} - N_{mk\ell}^{JULYWAVE} - N_{mk\ell}^{AUGWAVE}} \right)$$

where:

- $Pr(\blacksquare)^{mk\ell}$ denotes the unconditional probability of being included in a particular set of outreach waves for a Stage 2 volunteer of site-selection stratum m , site k , and within-site stratum ℓ ;
- $Pr(\blacksquare|not\ previously\ selected)^{mk\ell}$ denotes the conditional probability of being included in a particular set of outreach waves conditional on not having been selected for a previous wave for a Stage 2 volunteer of site-selection stratum m , site k , and within-site stratum ℓ ;
- $N_{mk\ell}^{\blacksquare}$ denotes the number of subjects of national stratum m , site k , and within-site stratum ℓ (across all longitudinal eligibility categories j) who were selected into a particular set of outreach waves;
- $N_{mkj\ell}$ denotes the total number of Stage 2 eligible beneficiaries of site-selection stratum m , site k , longitudinal eligibility category j , and within-site stratum ℓ ;
- $N_{mkj\ell}^{\blacksquare}$ denotes the number of subjects of site-selection stratum m , site k , longitudinal eligibility category j , and within-site stratum ℓ who were randomly selected into a particular set of outreach waves; and
- the number of possible within-site strata ℓ differs according to the probability calculated.

Using the conditional probabilities for inclusion into particular sets of outreach waves, we can then define the unconditional probabilities for inclusion into particular sets of waves for the subjects in longitudinal eligibility categories $j = 2, 3$, and 4.

Longitudinal Eligibility Category 2:

$$Pr(JUNEWAVE)^{mk2\ell} = (1 - Pr(PILOT)^{mk\ell})Pr(JUNEWAVE|not\ previously\ selected)^{mk\ell}$$

$$Pr(JULYWAVE)^{mk2\ell} = (1 - Pr(PILOT)^{mk\ell} - Pr(JUNEWAVE)^{mk2\ell}) Pr(JULYWAVE|not\ previously\ selected)^{mk\ell}$$

$$\begin{aligned}
Pr(AUGWAVE)^{mk2\ell} &= (1 - Pr(PILOT)^{mk\ell} - Pr(JUNEWAVE)^{mk2\ell} \\
&\quad - Pr(JULYWAVE)^{mk2\ell}) Pr(AUGWAVE|not\ previously\ selected)^{mk\ell} \\
Pr(LATERWAVES)^{mk2\ell} &= (1 - Pr(PILOT)^{mk\ell} - Pr(JUNEWAVE)^{mk2\ell} - Pr(JULYWAVE)^{mk2\ell} \\
&\quad - Pr(AUGWAVE)^{mk2\ell}) Pr(LATERWAVES|not\ previously\ selected)^{mk\ell}
\end{aligned}$$

Longitudinal Eligibility Category 3:

$$Pr(JUNEWAVE)^{mk3\ell} = Pr(JUNEWAVE|not\ previously\ selected)^{mk\ell}$$

$$\begin{aligned}
Pr(JULYWAVE)^{mk3\ell} \\
= (1 - Pr(JUNEWAVE)^{mk3\ell}) Pr(JULYWAVE|not\ previously\ selected)^{mk\ell}
\end{aligned}$$

$$\begin{aligned}
Pr(AUGWAVE)^{mk3\ell} \\
= (1 - Pr(JUNEWAVE)^{mk3\ell} \\
- Pr(JULYWAVE)^{mk3\ell}) Pr(AUGWAVE|not\ previously\ selected)^{mk\ell}
\end{aligned}$$

$$\begin{aligned}
Pr(LATERWAVES)^{mk3\ell} \\
= (1 - Pr(JUNEWAVE)^{mk3\ell} - Pr(JULYWAVE)^{mk3\ell} \\
- Pr(AUGWAVE)^{mk3\ell}) Pr(LATERWAVES|not\ previously\ selected)^{mk\ell}
\end{aligned}$$

Longitudinal Eligibility Category 4:

$$Pr(JULYWAVE)^{mk4\ell} = Pr(JULYWAVE|not\ previously\ selected)^{mk\ell}$$

$$Pr(AUGWAVE)^{mk4\ell} = (1 - Pr(JULYWAVE)^{mk4\ell}) Pr(AUGWAVE|not\ previously\ selected)^{mk\ell}$$

$$\begin{aligned}
Pr(LATERWAVES)^{mk4\ell} \\
= (1 - Pr(JULYWAVE)^{mk4\ell} \\
- Pr(AUGWAVE)^{mk4\ell}) Pr(LATERWAVES|not\ previously\ selected)^{mk\ell}
\end{aligned}$$

In the final step, using the unconditional probabilities for inclusion into particular sets of outreach waves, we can define the probability of inclusion into *any* outreach wave for the four categories of volunteers:

$$P_{mk1\ell} = Pr(PILOT)^{mk\ell}$$

$$\begin{aligned}
P_{mk2\ell} = Pr(PILOT)^{mk\ell} + Pr(JUNEWAVE)^{mk2\ell} + Pr(JULYWAVE)^{mk2\ell} + Pr(AUGWAVE)^{mk2\ell} \\
+ Pr(LATERWAVES)^{mk2\ell}
\end{aligned}$$

$$\begin{aligned}
P_{mk3\ell} = Pr(JUNEWAVE)^{mk3\ell} + Pr(JULYWAVE)^{mk3\ell} + Pr(AUGWAVE)^{mk3\ell} \\
+ Pr(LATERWAVES)^{mk3\ell}
\end{aligned}$$

$$P_{mk4\ell} = Pr(JULYWAVE)^{mk4\ell} + Pr(AUGWAVE)^{mk4\ell} + Pr(LATERWAVES)^{mk4\ell}$$

It is important to note that the Stage 2 analysis weights *do not* align the weighted totals of the Stage 2 volunteers to represent all beneficiaries in the outreach waves. Only a small percentage (5 percent) of those solicited volunteered for the study. The sum of the Stage 2 analysis weights across the three assignment groups of Stage 2 subjects is an estimate of the number of Stage-2 eligible beneficiaries in the nation who would have volunteered had they been given the opportunity to enroll in the study.

Analysis Weights for Survey Outcomes

For estimating impacts on survey outcomes, each Stage 2 sample member is assigned an analysis weight that is the Stage 2 administrative weight adjusted for non-response to the survey. The survey non-response adjustment weights more heavily the respondents who are most similar to the non-respondents. In order to construct the non-response adjustment component we ran three predictive regressions: one for each of the Stage 2 assignment groups. This model regresses an indicator variable equal to one if a person responded to the survey, on all baseline characteristics used in the impact model.

$$(3) \quad R_{pkg} = X_{pkg}\Theta_g + \epsilon_{pkg}$$

where R_{pkg} is equal to one if participant p in site k in assignment group g responded to the survey, X_{pkg} = a vector of baseline characteristics (listed in Section C) for individual p in site k in assignment group g ,

Θ_g = a vector of coefficients for assignment group g , and

ϵ_{pkg} = an idiosyncratic error term for participant p , in site k , and assignment group g .

We use the coefficients from this predictive model to calculate each subject's propensity to respond to the survey ($X_{pkg}\Theta_g$ from regression), given their baseline characteristics. Then, we divide each assignment group into quintiles based on propensity to respond. The non-response adjustment component for each respondent is

$$nrw_{gq} = \left(\frac{NR_{gq} + R_{gq}}{R_{gq}} \right)$$

where:

- nrw_{gq} denotes the non-response weight component for a respondent in group g with response propensity quintile q .
- NR_{gq} denotes the *weighted* number of non-respondents in group g with response propensity quintile q where the weights are the analytical weights for administrative outcomes $w_{mkj\ell}^A$ (i.e., NR_{gq} is the sum of the administrative outcome weights for the non-respondents in group g with response propensity quintile q).
- R_{gq} denotes the *weighted* number of respondents in group g with response propensity quintile q where the weights are the analytical weights for administrative outcomes $w_{mkj\ell}^A$ (i.e., R_{gq} is the

sum of the administrative outcome weights for the respondents in group g with response propensity quintile q).

The analytical weight for survey outcomes is then given by:

$$w_{mkj\ell gq}^S = w_{mkj\ell}^A \times nrw_{gq}$$

where:

- $w_{mkj\ell gq}^S$ denotes the Stage 2 weight for survey outcomes for a volunteer of national stratum m , site k , category j , and stratum ℓ , assignment group g , and response propensity quintile q .

A.7.2. Sample Exclusions

The impact analysis chapters (Chapters 6, 7 and 8) of this *Interim Report* (and future Stage 2 reports) use the same sample as the 2014 Stage 2 *Snapshot Report*, with 12,744 subjects. This sample does not include volunteers with BOND-eligible relatives, who were excluded after random assignment symmetrically in the three Stage 2 random assignment groups. However, the excluded subset is only a small fraction of the original Stage 2 sample (1.6 percent unweighted; 1.7 percent using the analysis weights). (The 2014 Stage 2 *Snapshot Report* discusses this “contamination” issue and the sample exclusions as they affect Stage 2 analysis.) By contrast, the process analysis chapters (Chapters 3, 4, and 5) use the full sample of Stage 2 volunteers (12,954 subjects) because the full sample reflects the potential caseload of WIC and EWIC counselors.

Appendix B. Subgroup Exhibits for 2013 Earnings and Benefit Impacts

Exhibit B-1. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T21 Vs. C2) for Subgroups Defined by Duration of SSDI Receipt

Outcome	Short Duration			Long Duration			Estimated Difference in Impact (7)
	Average Outcome with Offset and WIC (T21) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and WIC (T21) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$4,525	\$4,020	\$505* (\$240)	\$4,133	\$3,835	\$298 (\$342)	\$207 (\$457)
Employment during year (%)	38.23	36.27	1.96 (1.78)	41.60	38.23	3.37** (1.45)	-1.41 (2.78)
Earnings above BYA (%)	11.31	8.88	2.42** (0.79)	10.97	9.26	1.71 (0.94)	0.71 (1.36)
Earnings above 2x BYA (%)	4.70	3.99	0.71 (0.52)	3.64	3.26	0.37 (0.64)	0.34 (0.78)
Earnings above 3x BYA (%)	2.18	1.98	0.20 (0.38)	1.08	1.30	-0.22 (0.37)	0.42 (0.53)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$13,377	\$13,110	\$267 (\$149)	\$12,351	\$11,926	\$425* (\$188)	\$-158 (\$243)
Number of months with SSDI payments	11.34	11.12	0.22** (0.07)	11.21	10.93	0.28** (0.11)	-0.06 (0.12)
Total SSI benefits paid	\$51	\$51	\$1 (\$13)	\$36	\$26	\$10 (\$13)	\$-9 (\$17)
Number of months with SSI payments	0.27	0.25	0.02 (0.05)	0.14	0.15	-0.01 (0.05)	0.03 (0.06)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Short Duration T21 = 3,125, Short Duration C2 = 3,102, Long Duration T21 = 1,729, Long Duration C2 = 1,747.

*/**/** 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 (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.

Exhibit B-2. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. C2) for Subgroups Defined by Duration of SSDI Receipt

Outcome	Short Duration			Long Duration			Estimated Difference in Impact (7)
	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and EWIC (T22) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$4,694	\$4,020	\$674* (\$329)	\$3,983	\$3,835	\$148 (\$411)	\$526 (\$503)
Employment during year (%)	39.13	36.27	2.86* (1.49)	38.57	38.23	0.34 (1.58)	2.53 (2.65)
Earnings above BYA (%)	12.54	8.88	3.66*** (0.89)	10.56	9.26	1.30 (1.33)	2.36 (1.78)
Earnings above 2x BYA (%)	4.92	3.99	0.93 (0.86)	2.91	3.26	-0.35 (0.65)	1.28 (1.00)
Earnings above 3x BYA (%)	2.28	1.98	0.30 (0.44)	1.23	1.30	-0.08 (0.42)	0.38 (0.56)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$13,430	\$13,110	\$320* (\$171)	\$12,338	\$11,926	\$412* (\$198)	-\$92 (\$263)
Number of months with SSDI payments	11.37	11.12	0.25*** (0.07)	11.21	10.93	0.28* (0.12)	-0.03 (0.14)
Total SSI benefits paid	\$45	\$51	-\$6 (\$15)	\$50	\$26	\$24 (\$19)	-\$29 (\$23)
Number of months with SSI payments	0.24	0.25	-0.00 (0.05)	0.23	0.15	0.09 (0.07)	-0.09 (0.08)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Short Duration T22 = 1,914, Short Duration C2 = 3,102, Long Duration T22 = 1,127, Long Duration C2 = 1,747.

*/**/** 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 (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.

Exhibit B-3. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. T21) for Subgroups Defined by Duration of SSDI Receipt

Outcome	Short Duration			Long Duration			Estimated Difference in Impact (7)
	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome with Offset and WIC (T21) (2)	Impact Estimate (3)	Average Outcome with Offset and EWIC (T22) (4)	Average Outcome with Offset and WIC (T21) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$4,694	\$4,525	\$169 (\$346)	\$3,983	\$4,133	-\$150 (\$341)	\$320 (\$565)
Employment during year (%)	39.13	38.23	0.91 (1.40)	38.57	41.60	-3.03*** (0.93)	3.94† (2.12)
Earnings above BYA (%)	12.54	11.31	1.24 (1.32)	10.56	10.97	-0.41 (1.08)	1.65 (2.03)
Earnings above 2x BYA (%)	4.92	4.70	0.22 (0.74)	2.91	3.64	-0.73 (0.54)	0.94 (1.18)
Earnings above 3x BYA (%)	2.28	2.18	0.10 (0.50)	1.23	1.08	0.14 (0.36)	-0.04 (0.73)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$13,430	\$13,377	\$53 (\$133)	\$12,338	\$12,351	-\$13 (\$252)	\$66 (\$252)
Number of months with SSDI payments	11.37	11.34	0.03 (0.04)	11.21	11.21	0.00 (0.08)	0.03 (0.08)
Total SSI benefits paid	\$45	\$51	-\$6 (\$18)	\$50	\$36	\$14 (\$15)	-\$20 (\$18)
Number of months with SSI payments	0.24	0.27	-0.03 (0.06)	0.23	0.14	0.09 (0.05)	-0.12† (0.05)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Short Duration T22 = 1,914, Short Duration T21 = 3,125, Long Duration T22 = 1,127, Long Duration T21 = 1,729.

*/**/** 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 (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.

Exhibit B-4. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 + T21 vs C2) for Subgroups Defined by Duration of SSDI Receipt

Outcome	Short Duration			Long Duration			Estimated Difference in Impact (7)
	Average Outcome with Offset and WIC and EWIC (T22 + T21) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and WIC and EWIC (T22 + T21) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$4,588	\$4,020	\$568** (\$212)	\$4,073	\$3,835	\$238 (\$331)	\$330 (\$389)
Employment during year (%)	38.56	36.27	2.29 (1.54)	40.40	38.23	2.16 (1.28)	0.13 (2.52)
Earnings above BYA (%)	11.77	8.88	2.89*** (0.67)	10.80	9.26	1.55 (0.98)	1.34 (1.18)
Earnings above 2x BYA (%)	4.79	3.99	0.79 (0.45)	3.35	3.26	0.08 (0.56)	0.71 (0.58)
Earnings above 3x BYA (%)	2.22	1.98	0.24 (0.32)	1.14	1.30	-0.16 (0.34)	0.40 (0.27)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$13,397	\$13,110	\$287* (\$134)	\$12,346	\$11,926	\$420** (\$165)	\$-133 (\$102)
Number of months with SSDI payments	11.35	11.12	0.23*** (0.06)	11.21	10.93	0.28** (0.11)	-0.05 (0.12)
Total SSI benefits paid	\$49	\$51	\$-2 (\$11)	\$42	\$26	\$15 (\$12)	\$-17 (\$17)
Number of months with SSI payments	0.26	0.25	0.01 (0.04)	0.18	0.15	0.03 (0.04)	-0.02 (0.04)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Short Duration T22 + T21 = 5,039, Short Duration C2 = 3,102, Long Duration T22 + T21 = 2,856, Long Duration C2 = 1,747.

*/**/** 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 (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.

Exhibit B-5. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T21 Vs. C2) for Subgroups Defined by Employment at Baseline

Outcome	Employed			Not Employed			Estimated Difference in Impact (7)
	Average Outcome with Offset and WIC (T21) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and WIC (T21) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$10,224	\$9,705	\$518 (\$491)	\$2,236	\$1,858	\$378* (\$194)	\$141 (\$516)
Employment during year (%)	76.67	77.08	-0.42 (1.93)	27.53	23.61	3.91*** (1.11)	-4.33† (2.22)
Earnings above BYA (%)	29.47	23.77	5.70** (1.90)	4.77	3.94	0.83 (0.65)	4.87†† (1.98)
Earnings above 2x BYA (%)	11.01	9.82	1.19 (1.31)	1.66	1.30	0.36 (0.37)	0.83 (1.34)
Earnings above 3x BYA (%)	3.91	4.11	-0.20 (0.87)	0.71	0.64	0.07 (0.20)	-0.27 (0.89)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$12,351	\$11,218	\$1,132*** (\$306)	\$12,951	\$12,843	\$108 (\$129)	\$1,024†† (\$330)
Number of months with SSDI payments	10.64	9.83	0.82*** (0.22)	11.48	11.43	0.05 (0.06)	0.76††† (0.23)
Total SSI benefits paid	\$8	\$15	-\$7 (\$9)	\$54	\$44	\$10 (\$12)	-\$17 (\$14)
Number of months with SSI payments	0.07	0.09	-0.02 (0.06)	0.23	0.22	0.01 (0.04)	-0.02 (0.06)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Employed T21 = 1,207, Employed C2 = 1,187, Not Employed T21 = 3,610, Not Employed C2 = 3,627.

*/**/** 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 (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.

Exhibit B-6. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. C2) for Subgroups Defined by Employment at Baseline

Outcome	Employed			Not Employed			Estimated Difference in Impact (7)
	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and EWIC (T22) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$10,471	\$9,705	\$766 (\$589)	\$2,149	\$1,858	\$291 (\$217)	\$475 (\$599)
Employment during year (%)	76.57	77.08	-0.51 (2.17)	25.63	23.61	2.01 (1.26)	-2.53 (2.88)
Earnings above BYA (%)	29.04	23.77	5.27** (2.24)	5.36	3.94	1.43* (0.78)	3.85 (2.33)
Earnings above 2x BYA (%)	10.27	9.82	0.44 (1.49)	1.54	1.30	0.24 (0.41)	0.20 (1.53)
Earnings above 3x BYA (%)	4.46	4.11	0.35 (1.08)	0.73	0.64	0.09 (0.24)	0.26 (1.03)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$12,062	\$11,218	\$843** (\$291)	\$13,050	\$12,843	\$207 (\$152)	\$636† (\$328)
Number of months with SSDI payments	10.65	9.83	0.82*** (0.19)	11.50	11.43	0.07 (0.07)	0.75††† (0.19)
Total SSI benefits paid	\$32	\$15	\$17 (\$18)	\$54	\$44	\$10 (\$16)	\$7 (\$28)
Number of months with SSI payments	0.09	0.09	0.00 (0.07)	0.29	0.22	0.06 (0.07)	-0.06 (0.12)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Employed T22 = 702, Employed C2 = 1,187, Not Employed T22 = 2,317, Not Employed C2 = 3,627.

/ 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 (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.

Exhibit B-7. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. T21) for Subgroups Defined by Employment at Baseline

Outcome	Employed			Not Employed			Estimated Difference in Impact (7)
	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome with Offset and WIC (T21) (2)	Impact Estimate (3)	Average Outcome with Offset and EWIC (T22) (4)	Average Outcome with Offset and WIC (T21) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$10,471	\$10,224	\$247 (\$527)	\$2,149	\$2,236	-\$87 (\$175)	\$335 (\$525)
Employment during year (%)	76.57	76.67	-0.10 (2.27)	25.63	27.53	-1.90 (1.09)	1.81 (3.21)
Earnings above BYA (%)	29.04	29.47	-0.43 (1.34)	5.36	4.77	0.60 (0.65)	-1.02 (1.40)
Earnings above 2x BYA (%)	10.27	11.01	-0.75 (1.19)	1.54	1.66	-0.13 (0.24)	-0.62 (1.32)
Earnings above 3x BYA (%)	4.46	3.91	0.54 (0.88)	0.73	0.71	0.02 (0.16)	0.53 (0.93)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$12,062	\$12,351	-\$289 (\$218)	\$13,050	\$12,951	\$100 (\$214)	-\$389 (\$324)
Number of months with SSDI payments	10.65	10.64	0.01 (0.17)	11.50	11.48	0.02 (0.07)	-0.01 (0.21)
Total SSI benefits paid	\$32	\$8	\$24 (\$17)	\$54	\$54	-\$0 (\$17)	\$24 (\$23)
Number of months with SSI payments	0.09	0.07	0.02 (0.06)	0.29	0.23	0.06 (0.07)	-0.04 (0.09)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Employed T22 = 702, Employed T21 = 1,207, Not Employed T22 = 2,317, Not Employed T21 = 3,610.

*/**/** 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 (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.

Exhibit B-8. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 + T21 vs C2) for Subgroups Defined by Employment at Baseline

Outcome	Employed			Not Employed			Estimated Difference in Impact (7)
	Average Outcome with Offset and WIC and EWIC (T22 + T21) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and WIC and EWIC (T22 + T21) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$10,317	\$9,705	\$612 (\$445)	\$2,202	\$1,858	\$344* (\$184)	\$268 (\$241)
Employment during year (%)	76.64	77.08	-0.45 (1.73)	26.78	23.61	3.17** (0.98)	-3.61†† (1.59)
Earnings above BYA (%)	29.31	23.77	5.54** (1.71)	5.00	3.94	1.06 (0.63)	4.48††† (1.14)
Earnings above 2x BYA (%)	10.73	9.82	0.91 (1.16)	1.61	1.30	0.31 (0.37)	0.59 (0.77)
Earnings above 3x BYA (%)	4.12	4.11	0.01 (0.76)	0.71	0.64	0.07 (0.18)	-0.07 (0.62)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$12,242	\$11,218	\$1,023*** (\$260)	\$12,990	\$12,843	\$147 (\$117)	\$877††† (\$147)
Number of months with SSDI payments	10.65	9.83	0.82*** (0.19)	11.49	11.43	0.06 (0.05)	0.76††† (0.19)
Total SSI benefits paid	\$17	\$15	\$2 (\$10)	\$54	\$44	\$10 (\$10)	\$-8 (\$17)
Number of months with SSI payments	0.08	0.09	-0.01 (0.06)	0.25	0.22	0.03 (0.04)	-0.04 (0.07)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Employed T22 + T21 = 1,909, Employed C2 = 1,187, Not Employed T22 + T21 = 5,927, Not Employed C2 = 3,627.

*/**/** 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 (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.

Exhibit B-9. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T21 Vs. C2) for Subgroups Defined by Access to Medicaid Buy-in Programs

Outcome	Access to Medicaid Buy-in Programs			No Access to Medicaid Buy-in Programs			Estimated Difference in Impact (7)
	Average Outcome with Offset and WIC (T21) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and WIC (T21) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$4,372	\$4,200	\$172 (\$212)	\$4,094	\$3,138	\$956 (\$523)	\$-784 (\$554)
Employment during year (%)	41.46	39.87	1.59 (1.18)	36.66	30.75	5.90** (2.21)	-4.32 (2.61)
Earnings above BYA (%)	11.41	9.60	1.81** (0.76)	10.29	7.75	2.54* (1.31)	-0.73 (1.48)
Earnings above 2x BYA (%)	4.02	3.91	0.12 (0.51)	4.23	2.67	1.56 (0.97)	-1.45 (0.96)
Earnings above 3x BYA (%)	1.50	1.72	-0.22 (0.32)	1.68	1.24	0.44 (0.74)	-0.66 (0.81)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$12,999	\$12,493	\$506*** (\$156)	\$12,207	\$12,236	\$-29 (\$219)	\$535† (\$253)
Number of months with SSDI payments	11.29	10.96	0.34*** (0.08)	11.18	11.16	0.03 (0.12)	0.31†† (0.13)
Total SSI benefits paid	\$44	\$38	\$7 (\$11)	\$38	\$34	\$4 (\$13)	\$3 (\$17)
Number of months with SSI payments	0.20	0.19	0.01 (0.04)	0.19	0.20	-0.01 (0.05)	0.02 (0.07)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Access to Medicaid Buy-in Programs T21 = 3,276, Access to Medicaid Buy-in Programs C2 = 3,288, No Access to Medicaid Buy-in Programs T21 = 1,578, No Access to Medicaid Buy-in Programs C2 = 1,561.

*/**/*** 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 (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.

Exhibit B-10. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. C2) for Subgroups Defined by Access to Medicaid Buy-in Programs

Outcome	Access to Medicaid Buy-in Programs			No Access to Medicaid Buy-in Programs			Estimated Difference in Impact (7)
	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and EWIC (T22) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$4,549	\$4,200	\$349 (\$297)	\$3,551	\$3,138	\$413 (\$526)	-\$65 (\$544)
Employment during year (%)	40.53	39.87	0.66 (1.32)	34.06	30.75	3.31 (2.08)	-2.65 (2.24)
Earnings above BYA (%)	12.11	9.60	2.51** (0.89)	9.39	7.75	1.65 (1.51)	0.87 (1.57)
Earnings above 2x BYA (%)	4.06	3.91	0.15 (0.58)	2.92	2.67	0.26 (0.91)	-0.11 (1.01)
Earnings above 3x BYA (%)	1.80	1.72	0.08 (0.42)	1.32	1.24	0.08 (0.56)	0.01 (0.67)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$12,889	\$12,493	\$396** (\$168)	\$12,552	\$12,236	\$316 (\$214)	\$80 (\$272)
Number of months with SSDI payments	11.27	10.96	0.31*** (0.09)	11.29	11.16	0.13 (0.11)	0.18 (0.14)
Total SSI benefits paid	\$52	\$38	\$15 (\$16)	\$36	\$34	\$3 (\$26)	\$12 (\$29)
Number of months with SSI payments	0.27	0.19	0.09 (0.06)	0.14	0.20	-0.05 (0.06)	0.14 (0.08)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Access to Medicaid Buy-in Programs T22 = 2,067, Access to Medicaid Buy-in Programs C2 = 3,288, No Access to Medicaid Buy-in Programs T22 = 974, No Access to Medicaid Buy-in Programs C2 = 1,561.

*/**/*** 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 (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.

Exhibit B-11. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. T21) for Subgroups Defined by Access to Medicaid Buy-in Programs

Outcome	Access to Medicaid Buy-in Programs			No Access to Medicaid Buy-in Programs			Estimated Difference in Impact (7)
	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome with Offset and WIC (T21) (2)	Impact Estimate (3)	Average Outcome with Offset and EWIC (T22) (4)	Average Outcome with Offset and WIC (T21) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$4,549	\$4,372	\$177 (\$287)	\$3,551	\$4,094	-\$542 (\$384)	\$719 (\$558)
Employment during year (%)	40.53	41.46	-0.93 (0.80)	34.06	36.66	-2.60 (1.57)	1.67 (2.09)
Earnings above BYA (%)	12.11	11.41	0.70 (0.87)	9.39	10.29	-0.90 (1.13)	1.60 (1.60)
Earnings above 2x BYA (%)	4.06	4.02	0.03 (0.28)	2.92	4.23	-1.31* (0.61)	1.34† (0.72)
Earnings above 3x BYA (%)	1.80	1.50	0.30 (0.40)	1.32	1.68	-0.36 (0.42)	0.67 (0.73)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$12,889	\$12,999	-\$110 (\$202)	\$12,552	\$12,207	\$345 (\$218)	-\$455 (\$283)
Number of months with SSDI payments	11.27	11.29	-0.02 (0.06)	11.29	11.18	0.10 (0.09)	-0.13 (0.11)
Total SSI benefits paid	\$52	\$44	\$8 (\$18)	\$36	\$38	-\$1 (\$16)	\$9 (\$25)
Number of months with SSI payments	0.27	0.20	0.08 (0.07)	0.14	0.19	-0.05 (0.05)	0.12 (0.08)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Access to Medicaid Buy-in Programs T22 = 2,067, Access to Medicaid Buy-in Programs T21 = 3,276, No Access to Medicaid Buy-in Programs T22 = 974, No Access to Medicaid Buy-in Programs T21 = 1,578.

*/**/*** 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 (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.

Exhibit B-12. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 + T21 vs C2) for Subgroups Defined by Access to Medicaid Buy-in Programs

Outcome	Access to Medicaid Buy-in Programs			No Access to Medicaid Buy-in Programs			Estimated Difference in Impact (7)
	Average Outcome with Offset and WIC and EWIC (T22 + T21) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and WIC and EWIC (T22 + T21) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$4,440	\$4,200	\$240 (\$193)	\$3,886	\$3,138	\$748 (\$490)	-\$508 (\$477)
Employment during year (%)	41.09	39.87	1.22 (1.05)	35.66	30.75	4.91** (2.02)	-3.68 (2.25)
Earnings above BYA (%)	11.68	9.60	2.08** (0.69)	9.95	7.75	2.20 (1.28)	-0.12 (1.30)
Earnings above 2x BYA (%)	4.04	3.91	0.13 (0.45)	3.73	2.67	1.06 (0.90)	-0.93 (0.91)
Earnings above 3x BYA (%)	1.61	1.72	-0.10 (0.29)	1.54	1.24	0.30 (0.64)	-0.40 (0.67)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$12,957	\$12,493	\$463*** (\$137)	\$12,339	\$12,236	\$103 (\$170)	\$360† (\$166)
Number of months with SSDI payments	11.28	10.96	0.33*** (0.07)	11.22	11.16	0.07 (0.10)	0.26† (0.12)
Total SSI benefits paid	\$47	\$38	\$10 (\$10)	\$37	\$34	\$3 (\$17)	\$6 (\$17)
Number of months with SSI payments	0.23	0.19	0.04 (0.04)	0.17	0.20	-0.03 (0.05)	0.07 (0.04)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Access to Medicaid Buy-in Programs T22 + T21 = 5,343, Access to Medicaid Buy-in Programs C2 = 3,288, No Access to Medicaid Buy-in Programs T22 + T21 = 2,552, No Access to Medicaid Buy-in Programs C2 = 1,561.

*/**/** 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 (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.

Exhibit B-13. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T21 Vs. C2) for Subgroups Defined by Age at Baseline

Outcome	Age 49 or Less at Baseline			Age 50 or More at Baseline			Estimated Difference in Impact (7)
	Average Outcome with Offset and WIC (T21) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and WIC (T21) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$5,355	\$4,804	\$550* (\$277)	\$3,246	\$3,027	\$220 (\$208)	\$331 (\$346)
Employment during year (%)	45.92	43.71	2.20 (1.42)	34.45	31.14	3.31** (1.29)	-1.10 (1.91)
Earnings above BYA (%)	14.27	12.02	2.26** (0.98)	7.96	6.20	1.76** (0.75)	0.49 (1.24)
Earnings above 2x BYA (%)	5.79	4.85	0.94 (0.68)	2.38	2.30	0.08 (0.44)	0.86 (0.81)
Earnings above 3x BYA (%)	2.04	1.86	0.18 (0.43)	1.06	1.32	-0.26 (0.32)	-0.44 (0.53)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$11,646	\$11,337	\$309 (\$240)	\$13,909	\$13,504	\$405* (\$179)	-\$96 (\$289)
Number of months with SSDI payments	11.04	10.83	0.21 (0.14)	11.49	11.19	0.29** (0.10)	-0.08 (0.20)
Total SSI benefits paid	\$67	\$57	\$10 (\$17)	\$18	\$17	\$1 (\$9)	\$9 (\$18)
Number of months with SSI payments	0.25	0.29	-0.04 (0.05)	0.14	0.09	0.05 (0.03)	-0.09 (0.07)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Age 49 or Less at Baseline T21 = 2,407, Age 49 or Less at Baseline C2 = 2,385, Age 50 or More at Baseline T21 = 2,447, Age 50 or More at Baseline C2 = 2,464.

*/**/** 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 (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.

Exhibit B-14. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. C2) for Subgroups Defined by Age at Baseline

Outcome	Age 49 or Less at Baseline			Age 50 or More at Baseline			Estimated Difference in Impact (7)
	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and EWIC (T22) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$5,283	\$4,804	\$479 (\$301)	\$3,280	\$3,027	\$254 (\$350)	\$225 (\$382)
Employment during year (%)	45.61	43.71	1.89 (1.57)	32.00	31.14	0.86 (1.48)	1.04 (2.17)
Earnings above BYA (%)	14.61	12.02	2.60** (1.12)	8.16	6.20	1.96 (1.18)	0.64 (1.56)
Earnings above 2x BYA (%)	4.77	4.85	-0.08 (0.71)	2.72	2.30	0.43 (0.57)	-0.51 (0.89)
Earnings above 3x BYA (%)	2.23	1.86	0.37 (0.47)	1.12	1.32	-0.21 (0.40)	0.58 (0.56)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$11,565	\$11,337	\$228 (\$176)	\$14,022	\$13,504	\$518** (\$222)	-\$289 (\$317)
Number of months with SSDI payments	11.04	10.83	0.21 (0.14)	11.51	11.19	0.31*** (0.09)	-0.10 (0.18)
Total SSI benefits paid	\$66	\$57	\$10 (\$20)	\$30	\$17	\$13 (\$15)	-\$3 (\$27)
Number of months with SSI payments	0.31	0.29	0.02 (0.07)	0.17	0.09	0.07 (0.06)	-0.05 (0.09)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Age 49 or Less at Baseline T22 = 1,477, Age 49 or Less at Baseline C2 = 2,385, Age 50 or More at Baseline T22 = 1,564, Age 50 or More at Baseline C2 = 2,464.

*/**/** 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 (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.

Exhibit B-15. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. T21) for Subgroups Defined by Age at Baseline

Outcome	Age 49 or Less at Baseline			Age 50 or More at Baseline			Estimated Difference in Impact (7)
	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome with Offset and WIC (T21) (2)	Impact Estimate (3)	Average Outcome with Offset and EWIC (T22) (4)	Average Outcome with Offset and WIC (T21) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$5,283	\$5,355	-\$72 (\$204)	\$3,280	\$3,246	\$34 (\$274)	-\$106 (\$276)
Employment during year (%)	45.61	45.92	-0.31 (0.81)	32.00	34.45	-2.45 (1.52)	2.14 (2.23)
Earnings above BYA (%)	14.61	14.27	0.34 (0.90)	8.16	7.96	0.20 (0.85)	0.15 (1.23)
Earnings above 2x BYA (%)	4.77	5.79	-1.02* (0.55)	2.72	2.38	0.34 (0.47)	-1.36 (0.90)
Earnings above 3x BYA (%)	2.23	2.04	0.19 (0.35)	1.12	1.06	0.06 (0.33)	0.13 (0.52)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$11,565	\$11,646	-\$81 (\$241)	\$14,022	\$13,909	\$113 (\$232)	-\$194 (\$329)
Number of months with SSDI payments	11.04	11.04	0.01 (0.09)	11.51	11.49	0.02 (0.09)	-0.02 (0.14)
Total SSI benefits paid	\$66	\$67	-\$1 (\$21)	\$30	\$18	\$12 (\$18)	-\$13 (\$29)
Number of months with SSI payments	0.31	0.25	0.06 (0.05)	0.17	0.14	0.02 (0.06)	0.03 (0.04)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Age 49 or Less at Baseline T22 = 1,477, Age 49 or Less at Baseline T21 = 2,407, Age 50 or More at Baseline T22 = 1,564, Age 50 or More at Baseline T21 = 2,447.

*/**/** 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 (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.

Exhibit B-16. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 + T21 vs C2) for Subgroups Defined by Age at Baseline

Outcome	Age 49 or Less at Baseline			Age 50 or More at Baseline			Estimated Difference in Impact (7)
	Average Outcome with Offset and WIC and EWIC (T22 + T21) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and WIC and EWIC (T22 + T21) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$5,327	\$4,804	\$523* (\$248)	\$3,260	\$3,027	\$233 (\$230)	\$290† (\$153)
Employment during year (%)	45.80	43.71	2.08 (1.26)	33.49	31.14	2.35* (1.15)	-0.26 (1.43)
Earnings above BYA (%)	14.40	12.02	2.39** (0.87)	8.04	6.20	1.84** (0.80)	0.55 (1.00)
Earnings above 2x BYA (%)	5.40	4.85	0.55 (0.59)	2.51	2.30	0.22 (0.40)	0.33 (0.46)
Earnings above 3x BYA (%)	2.11	1.86	0.26 (0.38)	1.08	1.32	-0.24 (0.28)	0.50† (0.23)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$11,615	\$11,337	\$278 (\$169)	\$13,953	\$13,504	\$449** (\$161)	\$-171 (\$255)
Number of months with SSDI payments	11.04	10.83	0.21 (0.14)	11.50	11.19	0.30*** (0.09)	-0.09 (0.18)
Total SSI benefits paid	\$67	\$57	\$10 (\$14)	\$23	\$17	\$6 (\$8)	\$4 (\$14)
Number of months with SSI payments	0.27	0.29	-0.01 (0.05)	0.15	0.09	0.06* (0.03)	-0.07 (0.04)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Age 49 or Less at Baseline T22 + T21 = 3,884, Age 49 or Less at Baseline C2 = 2,385, Age 50 or More at Baseline T22 + T21 = 4,011, Age 50 or More at Baseline C2 = 2,464.

*/**/*** 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 (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.

Exhibit B-17. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T21 Vs. C2) for Subgroups Defined by Primary Impairment of Major Affective Disorder

Outcome	Primary Impairment of Major Affective Disorder			All Other Primary Impairments			Estimated Difference in Impact (7)
	Average Outcome with Offset and WIC (T21) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and WIC (T21) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$4,482	\$4,177	\$305 (\$422)	\$4,255	\$3,850	\$405* (\$188)	\$-100 (\$461)
Employment during year (%)	46.46	41.86	4.60* (2.35)	38.64	36.35	2.29* (1.04)	2.31 (2.56)
Earnings above BYA (%)	10.64	9.42	1.23 (1.51)	11.23	9.02	2.21*** (0.67)	-0.98 (1.66)
Earnings above 2x BYA (%)	4.07	3.81	0.25 (1.14)	4.09	3.51	0.58 (0.44)	-0.32 (1.15)
Earnings above 3x BYA (%)	1.48	1.61	-0.13 (0.60)	1.56	1.58	-0.02 (0.30)	-0.11 (0.67)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$12,268	\$12,049	\$220 (\$338)	\$12,904	\$12,513	\$391** (\$139)	\$-172 (\$357)
Number of months with SSDI payments	11.34	10.94	0.40* (0.21)	11.24	11.03	0.21** (0.07)	0.19 (0.22)
Total SSI benefits paid	\$48	\$31	\$18 (\$16)	\$41	\$38	\$3 (\$10)	\$15 (\$19)
Number of months with SSI payments	0.28	0.15	0.13* (0.07)	0.17	0.20	-0.02 (0.04)	0.15† (0.07)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Primary Impairment of Major Affective Disorder T21 = 902, Primary Impairment of Major Affective Disorder C2 = 876, All Other Primary Impairments T21 = 3,952, All Other Primary Impairments C2 = 3,973.

*/**/** 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 (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.

Exhibit B-18. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. C2) for Subgroups Defined by Primary Impairment of Major Affective Disorder

Outcome	Primary Impairment of Major Affective Disorder			All Other Primary Impairments			Estimated Difference in Impact (7)
	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and EWIC (T22) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$4,862	\$4,177	\$685 (\$594)	\$4,143	\$3,850	\$293 (\$247)	\$392 (\$513)
Employment during year (%)	45.74	41.86	3.87 (2.69)	37.14	36.35	0.79 (1.17)	3.08 (2.93)
Earnings above BYA (%)	13.48	9.42	4.06* (1.86)	10.89	9.02	1.87** (0.76)	2.18 (2.01)
Earnings above 2x BYA (%)	4.31	3.81	0.50 (1.20)	3.62	3.51	0.11 (0.56)	0.40 (1.29)
Earnings above 3x BYA (%)	1.59	1.61	-0.02 (0.72)	1.69	1.58	0.11 (0.42)	-0.13 (0.79)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$12,320	\$12,049	\$271 (\$345)	\$12,912	\$12,513	\$398** (\$152)	\$-127 (\$371)
Number of months with SSDI payments	11.30	10.94	0.36* (0.18)	11.27	11.03	0.24** (0.08)	0.12 (0.17)
Total SSI benefits paid	\$48	\$31	\$18 (\$26)	\$48	\$38	\$10 (\$14)	\$8 (\$29)
Number of months with SSI payments	0.25	0.15	0.10 (0.09)	0.24	0.20	0.04 (0.05)	0.06 (0.11)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Primary Impairment of Major Affective Disorder T22 = 499, Primary Impairment of Major Affective Disorder C2 = 876, All Other Primary Impairments T22 = 2,542, All Other Primary Impairments C2 = 3,973.

*/**/** 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 (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.

Exhibit B-19. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. T21) for Subgroups Defined by Primary Impairment of Major Affective Disorder

Outcome	Primary Impairment of Major Affective Disorder			All Other Primary Impairments			Estimated Difference in Impact (7)
	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome with Offset and WIC (T21) (2)	Impact Estimate (3)	Average Outcome with Offset and EWIC (T22) (4)	Average Outcome with Offset and WIC (T21) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$4,862	\$4,482	\$380 (\$538)	\$4,143	\$4,255	\$-112 (\$202)	\$492 (\$574)
Employment during year (%)	45.74	46.46	-0.73 (1.91)	37.14	38.64	-1.50** (0.64)	0.78 (2.26)
Earnings above BYA (%)	13.48	10.64	2.83 (1.77)	10.89	11.23	-0.34 (0.55)	3.17 (1.78)
Earnings above 2x BYA (%)	4.31	4.07	0.25 (1.04)	3.62	4.09	-0.47 (0.45)	0.72 (1.41)
Earnings above 3x BYA (%)	1.59	1.48	0.10 (0.63)	1.69	1.56	0.13 (0.26)	-0.02 (0.74)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$12,320	\$12,268	\$51 (\$345)	\$12,912	\$12,904	\$7 (\$154)	\$44 (\$289)
Number of months with SSDI payments	11.30	11.34	-0.04 (0.16)	11.27	11.24	0.03 (0.06)	-0.06 (0.19)
Total SSI benefits paid	\$48	\$48	\$0 (\$30)	\$48	\$41	\$7 (\$14)	\$-7 (\$30)
Number of months with SSI payments	0.25	0.28	-0.03 (0.09)	0.24	0.17	0.06 (0.06)	-0.09 (0.10)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Primary Impairment of Major Affective Disorder T22 = 499, Primary Impairment of Major Affective Disorder T21 = 902, All Other Primary Impairments T22 = 2,542, All Other Primary Impairments T21 = 3,952.

*/**/*** 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 (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.

Exhibit B-20. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 + T21 vs C2) for Subgroups Defined by Primary Impairment of Major Affective Disorder

Outcome	Primary Impairment of Major Affective Disorder			All Other Primary Impairments			Estimated Difference in Impact (7)
	Average Outcome with Offset and WIC and EWIC (T22 + T21) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and WIC and EWIC (T22 + T21) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$4,622	\$4,177	\$445 (\$381)	\$4,211	\$3,850	\$361* (\$182)	\$84 (\$263)
Employment during year (%)	46.20	41.86	4.33* (2.10)	38.05	36.35	1.70 (0.93)	2.63 (2.09)
Earnings above BYA (%)	11.69	9.42	2.27 (1.39)	11.10	9.02	2.08*** (0.59)	0.19 (0.88)
Earnings above 2x BYA (%)	4.16	3.81	0.35 (0.93)	3.91	3.51	0.39 (0.40)	-0.05 (0.75)
Earnings above 3x BYA (%)	1.52	1.61	-0.09 (0.54)	1.61	1.58	0.03 (0.30)	-0.12 (0.51)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$12,287	\$12,049	\$238 (\$297)	\$12,907	\$12,513	\$394** (\$123)	\$-156 (\$334)
Number of months with SSDI payments	11.32	10.94	0.39* (0.18)	11.25	11.03	0.22*** (0.06)	0.16 (0.18)
Total SSI benefits paid	\$48	\$31	\$18 (\$15)	\$44	\$38	\$6 (\$9)	\$12 (\$15)
Number of months with SSI payments	0.27	0.15	0.12* (0.06)	0.20	0.20	0.00 (0.03)	0.12† (0.06)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Primary Impairment of Major Affective Disorder T22 + T21 = 1,401, Primary Impairment of Major Affective Disorder C2 = 876, All Other Primary Impairments T22 + T21 = 6,494, All Other Primary Impairments C2 = 3,973.

*/**/*** 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 (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.

Exhibit B-21. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T21 Vs. C2) for Subgroups Defined by Primary Impairment of Back Disorder

Outcome	Primary Impairment of Back Disorder			All Other Primary Impairments			Estimated Difference in Impact (7)
	Average Outcome with Offset and WIC (T21) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and WIC (T21) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$3,563	\$3,062	\$501 (\$428)	\$4,409	\$4,039	\$370* (\$188)	\$132 (\$467)
Employment during year (%)	34.80	29.49	5.31* (2.69)	40.96	38.58	2.39** (1.03)	2.92 (2.88)
Earnings above BYA (%)	8.40	6.37	2.03 (1.41)	11.51	9.50	2.01** (0.67)	0.02 (1.56)
Earnings above 2x BYA (%)	3.85	3.20	0.64 (1.05)	4.12	3.63	0.50 (0.48)	0.15 (1.14)
Earnings above 3x BYA (%)	1.43	1.61	-0.18 (0.73)	1.57	1.59	-0.02 (0.29)	-0.16 (0.79)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$13,566	\$13,073	\$493 (\$338)	\$12,665	\$12,328	\$337** (\$138)	\$156 (\$362)
Number of months with SSDI payments	11.70	11.44	0.26* (0.13)	11.20	10.95	0.25*** (0.07)	0.01 (0.13)
Total SSI benefits paid	\$21	\$23	\$-1 (\$12)	\$45	\$39	\$7 (\$10)	\$-8 (\$15)
Number of months with SSI payments	0.13	0.17	-0.04 (0.07)	0.20	0.19	0.01 (0.04)	-0.05 (0.08)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Primary Impairment of Back Disorder T21 = 661, Primary Impairment of Back Disorder C2 = 682, All Other Primary Impairments T21 = 4,193, All Other Primary Impairments C2 = 4,167.

*/**/*** 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 (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.

Exhibit B-22. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. C2) for Subgroups Defined by Primary Impairment of Back Disorder

Outcome	Primary Impairment of Back Disorder			All Other Primary Impairments			Estimated Difference in Impact (7)
	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and EWIC (T22) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$2,520	\$3,062	\$-542 (\$449)	\$4,534	\$4,039	\$495 (\$291)	\$-1,037† (\$494)
Employment during year (%)	30.27	29.49	0.78 (2.81)	40.02	38.58	1.44 (1.16)	-0.66 (3.04)
Earnings above BYA (%)	6.34	6.37	-0.03 (1.58)	12.10	9.50	2.60*** (0.78)	-2.63 (1.76)
Earnings above 2x BYA (%)	2.20	3.20	-1.00 (1.14)	3.97	3.63	0.35 (0.56)	-1.35 (1.27)
Earnings above 3x BYA (%)	0.28	1.61	-1.33* (0.61)	1.87	1.59	0.29 (0.35)	-1.62†† (0.69)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$13,079	\$13,073	\$7 (\$328)	\$12,751	\$12,328	\$424** (\$147)	\$-417 (\$360)
Number of months with SSDI payments	11.60	11.44	0.16 (0.14)	11.23	10.95	0.28*** (0.08)	-0.12 (0.16)
Total SSI benefits paid	\$45	\$23	\$22 (\$20)	\$49	\$39	\$10 (\$14)	\$13 (\$25)
Number of months with SSI payments	0.29	0.17	0.12 (0.12)	0.23	0.19	0.04 (0.05)	0.08 (0.14)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Primary Impairment of Back Disorder T22 = 424, Primary Impairment of Back Disorder C2 = 682, All Other Primary Impairments T22 = 2,617, All Other Primary Impairments C2 = 4,167.

*/**/** 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 (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.

Exhibit B-23. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. T21) for Subgroups Defined by Primary Impairment of Back Disorder

Outcome	Primary Impairment of Back Disorder			All Other Primary Impairments			Estimated Difference in Impact (7)
	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome with Offset and WIC (T21) (2)	Impact Estimate (3)	Average Outcome with Offset and EWIC (T22) (4)	Average Outcome with Offset and WIC (T21) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$2,520	\$3,563	\$-1,043*** (\$254)	\$4,534	\$4,409	\$125 (\$190)	\$-1,168††† (\$160)
Employment during year (%)	30.27	34.80	-4.53** (1.89)	40.02	40.96	-0.94* (0.48)	-3.59 (1.97)
Earnings above BYA (%)	6.34	8.40	-2.06 (1.29)	12.10	11.51	0.59 (0.65)	-2.65† (1.37)
Earnings above 2x BYA (%)	2.20	3.85	-1.64 (1.05)	3.97	4.12	-0.15 (0.23)	-1.50 (1.05)
Earnings above 3x BYA (%)	0.28	1.43	-1.15** (0.46)	1.87	1.57	0.30 (0.18)	-1.46††† (0.29)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$13,079	\$13,566	\$-487 (\$271)	\$12,751	\$12,665	\$87 (\$169)	\$-573†† (\$226)
Number of months with SSDI payments	11.60	11.70	-0.10 (0.18)	11.23	11.20	0.03 (0.06)	-0.13 (0.21)
Total SSI benefits paid	\$45	\$21	\$24 (\$19)	\$49	\$45	\$3 (\$15)	\$21 (\$25)
Number of months with SSI payments	0.29	0.13	0.16 (0.12)	0.23	0.20	0.03 (0.06)	0.13 (0.14)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Primary Impairment of Back Disorder T22 = 424, Primary Impairment of Back Disorder T21 = 661, All Other Primary Impairments T22 = 2,617, All Other Primary Impairments T21 = 4,193.

*/**/*** 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 (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.

Exhibit B-24. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 + T21 vs C2) for Subgroups Defined by Primary Impairment of Back Disorder

Outcome	Primary Impairment of Back Disorder			All Other Primary Impairments			Estimated Difference in Impact (7)
	Average Outcome with Offset and WIC and EWIC (T22 + T21) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and WIC and EWIC (T22 + T21) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$3,156	\$3,062	\$94 (\$381)	\$4,458	\$4,039	\$419* (\$211)	-\$325 (\$382)
Employment during year (%)	33.02	29.49	3.53 (2.34)	40.60	38.58	2.02* (0.92)	1.51 (2.56)
Earnings above BYA (%)	7.60	6.37	1.23 (1.26)	11.74	9.50	2.24*** (0.60)	-1.01 (1.19)
Earnings above 2x BYA (%)	3.20	3.20	-0.00 (0.91)	4.07	3.63	0.44 (0.50)	-0.44 (1.05)
Earnings above 3x BYA (%)	0.98	1.61	-0.63 (0.62)	1.69	1.59	0.10 (0.27)	-0.73 (0.40)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$13,376	\$13,073	\$303 (\$282)	\$12,699	\$12,328	\$371** (\$119)	-\$68 (\$285)
Number of months with SSDI payments	11.66	11.44	0.22* (0.11)	11.21	10.95	0.26*** (0.07)	-0.04 (0.08)
Total SSI benefits paid	\$31	\$23	\$8 (\$12)	\$47	\$39	\$8 (\$9)	-\$0 (\$12)
Number of months with SSI payments	0.19	0.17	0.02 (0.08)	0.22	0.19	0.02 (0.03)	-0.00 (0.08)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Primary Impairment of Back Disorder T22 + T21 = 1,085, Primary Impairment of Back Disorder C2 = 682, All Other Primary Impairments T22 + T21 = 6,810, All Other Primary Impairments C2 = 4,167.

*/**/** 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 (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.

Exhibit B-25. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law T21 Vs. C2) for Subgroups Defined by Education at Baseline

Outcome	Less than Associate's Degree			Any Postsecondary Degree			Estimated Difference in Impact (7)
	Average Outcome with Offset and WIC (T21) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and WIC (T21) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$3,910	\$3,278	\$632*** (\$183)	\$5,047	\$5,109	-\$63 (\$364)	-\$695 (\$405)
Employment during year (%)	39.13	34.38	4.75*** (1.16)	42.12	43.53	-1.41 (1.72)	-6.16†† (2.07)
Earnings above BYA (%)	10.19	7.76	2.43*** (0.71)	12.95	11.65	1.30 (1.19)	-1.12 (1.39)
Earnings above 2x BYA (%)	3.37	2.53	0.84 (0.46)	5.46	5.47	-0.00 (0.84)	-0.84 (0.95)
Earnings above 3x BYA (%)	1.04	0.98	0.06 (0.28)	2.53	2.69	-0.16 (0.59)	-0.22 (0.65)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$11,944	\$11,718	\$226 (\$155)	\$14,534	\$13,888	\$646** (\$257)	\$420 (\$289)
Number of months with SSDI payments	11.23	11.05	0.18* (0.10)	11.34	10.96	0.37*** (0.11)	0.19 (0.13)
Total SSI benefits paid	\$48	\$49	-\$1 (\$14)	\$32	\$12	\$19 (\$15)	\$20 (\$24)
Number of months with SSI payments	0.22	0.25	-0.03 (0.04)	0.13	0.07	0.06 (0.05)	0.10 (0.07)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Less than Associate's Degree T21 = 3,290, Less than Associate's Degree C2 = 3,224, Any Postsecondary Degree T21 = 1,524, Any Postsecondary Degree C2 = 1,586.

*/**/** 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 (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.

Exhibit B-26. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. C2) for Subgroups Defined by Education at Baseline

Outcome	Less than Associate's Degree			Any Postsecondary Degree			Estimated Difference in Impact (7)
	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and EWIC (T22) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$3,612	\$3,278	\$335 (\$209)	\$5,658	\$5,109	\$549 (\$549)	\$214 (\$472)
Employment during year (%)	36.20	34.38	1.81 (1.56)	43.93	43.53	0.40 (3.08)	-1.41 (4.34)
Earnings above BYA (%)	10.05	7.76	2.30** (0.82)	14.17	11.65	2.52 (1.38)	0.23 (1.61)
Earnings above 2x BYA (%)	2.63	2.53	0.09 (0.44)	6.12	5.47	0.65 (1.13)	0.56 (1.10)
Earnings above 3x BYA (%)	0.98	0.98	0.00 (0.30)	3.11	2.69	0.42 (0.71)	0.42 (0.72)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$12,063	\$11,718	\$346** (\$151)	\$14,305	\$13,888	\$417 (\$342)	\$71 (\$457)
Number of months with SSDI payments	11.31	11.05	0.26* (0.12)	11.22	10.96	0.25* (0.13)	-0.01 (0.20)
Total SSI benefits paid	\$61	\$49	\$13 (\$16)	\$22	\$12	\$9 (\$19)	-\$3 (\$24)
Number of months with SSI payments	0.32	0.25	0.07 (0.06)	0.07	0.07	0.01 (0.05)	-0.07 (0.08)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Less than Associate's Degree T22 = 2,035, Less than Associate's Degree C2 = 3,224, Any Postsecondary Degree T22 = 981, Any Postsecondary Degree C2 = 1,586.

*/**/** 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 (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.

Exhibit B-27. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 Vs. T21) for Subgroups Defined by Education at Baseline

Outcome	Less than Associate's Degree			Any Postsecondary Degree			Estimated Difference in Impact (7)
	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome with Offset and WIC (T21) (2)	Impact Estimate (3)	Average Outcome with Offset and EWIC (T22) (4)	Average Outcome with Offset and WIC (T21) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$3,612	\$3,910	\$-298 (\$204)	\$5,658	\$5,047	\$612 (\$377)	\$909†† (\$387)
Employment during year (%)	36.20	39.13	-2.94** (1.14)	43.93	42.12	1.81 (2.60)	4.75 (3.56)
Earnings above BYA (%)	10.05	10.19	-0.13 (0.71)	14.17	12.95	1.22 (0.67)	1.35† (0.62)
Earnings above 2x BYA (%)	2.63	3.37	-0.74** (0.31)	6.12	5.46	0.66 (0.87)	1.40 (1.07)
Earnings above 3x BYA (%)	0.98	1.04	-0.06 (0.24)	3.11	2.53	0.58 (0.52)	0.64 (0.59)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$12,063	\$11,944	\$119 (\$157)	\$14,305	\$14,534	\$-229 (\$315)	\$-349 (\$320)
Number of months with SSDI payments	11.31	11.23	0.08 (0.05)	11.22	11.34	-0.12 (0.14)	-0.20 (0.16)
Total SSI benefits paid	\$61	\$48	\$13 (\$21)	\$22	\$32	\$-10 (\$18)	\$-23 (\$30)
Number of months with SSI payments	0.32	0.22	0.10 (0.07)	0.07	0.13	-0.06 (0.06)	-0.16† (0.08)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Less than Associate's Degree T22 = 2,035, Less than Associate's Degree T21 = 3,290, Any Postsecondary Degree T22 = 981, Any Postsecondary Degree T21 = 1,524.

*/**/** 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 (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.

Exhibit B-28. Estimated Impacts on 2013 Outcomes of the Offset Compared to Current Law (T22 + T21 vs C2) for Subgroups Defined by Education at Baseline

Outcome	Less than Associate's Degree			Any Postsecondary Degree			Estimated Difference in Impact (7)
	Average Outcome with Offset and WIC and EWIC (T22 + T21) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and WIC and EWIC (T22 + T21) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2013)							
Total earnings	\$3,796	\$3,278	\$519** (\$162)	\$5,293	\$5,109	\$184 (\$375)	-\$335 (\$296)
Employment during year (%)	38.02	34.38	3.63*** (1.03)	42.85	43.53	-0.68 (1.78)	-4.32† (2.34)
Earnings above BYA (%)	10.14	7.76	2.38*** (0.64)	13.44	11.65	1.79 (1.06)	-0.58 (0.73)
Earnings above 2x BYA (%)	3.09	2.53	0.55 (0.39)	5.73	5.47	0.26 (0.83)	-0.29 (0.79)
Earnings above 3x BYA (%)	1.02	0.98	0.04 (0.24)	2.77	2.69	0.08 (0.55)	0.04 (0.55)
Benefit Outcomes (January–December 2013)							
Total SSDI benefits paid	\$11,990	\$11,718	\$272* (\$134)	\$14,442	\$13,888	\$554** (\$225)	\$282 (\$297)
Number of months with SSDI payments	11.26	11.05	0.21* (0.10)	11.29	10.96	0.33*** (0.10)	0.11 (0.13)
Total SSI benefits paid	\$53	\$49	\$4 (\$11)	\$28	\$12	\$15 (\$13)	\$11 (\$16)
Number of months with SSI payments	0.26	0.25	0.01 (0.04)	0.11	0.07	0.04 (0.04)	0.03 (0.05)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Less than Associate's Degree T22 + T21 = 5,325, Less than Associate's Degree C2 = 3,224, Any Postsecondary Degree T22 + T21 = 2,505, Any Postsecondary Degree C2 = 1,586.

*/**/** 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 (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. 2011 & 2012 Earnings and Benefits Impacts

Exhibit C-1. Estimated Impacts on 2012 Earnings and Benefits of Stage 2 Volunteers: Confirmatory Results, All Policy Comparisons

Outcome	Average Outcome with Offset and WIC (T21) (1)	Average Outcome with Offset and EWIC (T22) (2)	Average Outcome under Current Law (C2) (3)	Estimated Impact of Offset + WIC vs Current Law (T21 vs. C2) (4)	Estimated Impact of Offset + EWIC vs Current Law (T22 vs. C2) (5)	Estimated Impact of EWIC instead of WIC Given Offset (T22 vs. T21) (6)
Total earnings	\$3,949	\$4,000	\$3,638	\$312 ^a (\$156)	\$363 ^a (\$258)	\$51 (\$194)
Total SSDI benefits paid	\$12,766	\$12,791	\$12,551	\$216 ^b (\$128)	\$241 ^b (\$122)	\$25 (\$152)

Source: Analysis of SSA administrative records (from the MEF, BODS, MBR, and SSR), with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: T21 = 4,854, T22 = 4,849, C2 = 3,041

###,### 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.

^a The impact estimates for total earnings for T21 vs. C2 and T22 vs. C2 had p-values after multiple comparison adjustments of 0.302, and hence do not provide evidence of an impact. Prior to multiple comparison adjustments, the p-value for the impact estimate for T21 vs. C2 was 0.078, providing some evidence of impact using an exploratory standard of evidence. Prior to multiple comparison adjustments, the p-value for the impact estimate for T22 vs. C2 was 0.194, failing to provide evidence of impact using an exploratory standard of evidence.

^b The impact estimate for total SSDI benefits paid for T21 vs. C2 and T22 vs. C2 had p-values after multiple comparison adjustments of 0.302 and hence do not provide evidence of a proven impact. Prior to multiple comparison adjustments, the p-values for the impact estimate for T21 vs. C2 was 0.126, failing to provide evidence of impact using an exploratory standard of evidence. Prior to multiple comparison adjustments, the p-value for the impact estimate for T22 vs. C2 was 0.080, providing some evidence of impact using an exploratory standard of evidence.

**Exhibit C-2. Estimated Impacts on 2012 Earnings and Benefits of Stage 2 Volunteers:
Exploratory Results, All Policy Comparisons**

Outcome	Average Outcome with Offset and WIC (T21) (1)	Average Outcome with Offset and EWIC (T22) (2)	Average Outcome under Current Law (C2) (3)	Estimated Impact of Offset + WIC vs Current Law (T21 vs. C2) (4)	Estimated Impact of Offset + EWIC vs Current Law (T22 vs. C2) (5)	Estimated Impact of EWIC instead of WIC Given Offset (T22 vs. T21) (6)
Earnings Outcomes (January–December 2012)						
Employment during year (%)	41.22	40.81	39.62	1.60* (0.86)	1.20 (1.09)	-0.41 (1.32)
Earnings above BYA (%)	10.46	10.18	9.12	1.34** (0.58)	1.06 (0.79)	-0.28 (0.67)
Earnings above 2x BYA (%)	3.25	2.77	2.71	0.54 (0.35)	0.05 (0.37)	-0.48 (0.39)
Earnings above 3x BYA (%)	1.16	1.42	1.02	0.14 (0.21)	0.40 (0.36)	0.27 (0.25)
Benefit Outcomes (January–December 2012)						
Number of months with SSDI payments	11.45	11.46	11.36	0.09 (0.06)	0.10 (0.07)	0.01 (0.05)
Total SSI benefits paid	\$53	\$54	\$38	\$14 (11)	\$15 (14)	\$1 (16)
Number of months with SSI payments	0.18	0.19	0.17	0.02 (0.03)	0.03 (0.04)	0.01 (0.04)

Source: Analysis of SSA administrative records (from the MEF, BODS, MBR, and SSR), with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: T21 = 4,854, T22 = 3,041, C2 = 4,849

*/**/** 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 (and with no multiple comparisons adjustment).

Exhibit C-3. Estimated Impacts on 2012 Outcomes of the Offset Plus WIC Compared to Current Law (T21 Vs. C2) for Subgroups Defined by Duration of SSDI Receipt

Outcome	Short Duration			Long Duration			Estimated Difference in Impact (7)
	Average Outcome with Offset and WIC (T21) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and WIC (T21) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2012)							
Total earnings	\$4,005	\$3,513	\$492** (\$184)	\$3,905	\$3,728	\$177 (\$225)	\$315 (\$267)
Employment during year (%)	39.22	37.82	1.40 (1.07)	42.67	40.92	1.75 (1.28)	-0.35 (1.66)
Earnings above BYA (%)	10.00	8.76	1.25 (0.68)	10.80	9.38	1.42 (0.87)	-0.17 (1.11)
Earnings above 2x BYA (%)	3.71	3.27	0.43 (0.43)	2.92	2.31	0.61 (0.52)	-0.18 (0.69)
Earnings above 3x BYA (%)	1.74	1.35	0.39 (0.30)	0.73	0.78	-0.06 (0.29)	0.45 (0.42)
Benefit Outcomes (January–December 2012)							
Total SSDI benefits paid	\$13,466	\$13,329	\$138 (\$215)	\$12,260	\$11,986	\$274 (\$157)	\$-137 (\$243)
Number of months with SSDI payments	11.56	11.51	0.05 (0.05)	11.37	11.25	0.12 (0.09)	-0.07 (0.09)
Total SSI benefits paid	\$74	\$68	\$6 (\$19)	\$37	\$16	\$20 (\$13)	\$-14 (\$23)
Number of months with SSI payments	0.24	0.23	0.00 (0.04)	0.14	0.12	0.03 (0.04)	-0.02 (0.06)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Short Duration T21 = 3,125, Short Duration C2 = 3,102, Long Duration T21 = 1,729, Long Duration C2 = 1,747.

*/**/** 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 (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.

Exhibit C-4. Estimated Impacts on 2012 Outcomes of the Offset Plus EWIC Compared to Current Law (T22 Vs. C2) for Subgroups Defined by Duration of SSDI Receipt

Outcome	Short Duration			Long Duration			Estimated Difference in Impact (7)
	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and EWIC (T22) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2012)							
Total earnings	\$4,016	\$3,513	\$503* (\$258)	\$3,989	\$3,728	\$261 (\$345)	\$242 (\$350)
Employment during year (%)	38.92	37.82	1.10 (1.28)	42.19	40.92	1.27 (1.81)	-0.16 (2.31)
Earnings above BYA (%)	10.23	8.76	1.48 (0.83)	10.15	9.38	0.77 (1.07)	0.71 (1.25)
Earnings above 2x BYA (%)	2.90	3.27	-0.38 (0.41)	2.66	2.31	0.36 (0.56)	-0.73 (0.69)
Earnings above 3x BYA (%)	1.89	1.35	0.54 (0.33)	1.09	0.78	0.30 (0.47)	0.23 (0.47)
Benefit Outcomes (January–December 2012)							
Total SSDI benefits paid	\$13,508	\$13,329	\$180 (\$148)	\$12,270	\$11,986	\$284 (\$180)	\$-105 (\$234)
Number of months with SSDI payments	11.61	11.51	0.10* (0.06)	11.35	11.25	0.10 (0.10)	0.00 (0.10)
Total SSI benefits paid	\$85	\$68	\$16 (\$25)	\$31	\$16	\$15 (\$17)	\$2 (\$31)
Number of months with SSI payments	0.23	0.23	0.00 (0.05)	0.16	0.12	0.04 (0.05)	-0.04 (0.07)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Short Duration T22 = 1,914, Short Duration C2 = 3,102, Long Duration T22 = 1,127, Long Duration C2 = 1,747.

*/**/** 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 (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.

Exhibit C-5. Estimated Impacts on 2012 Outcomes of the Offset Plus EWIC Compared to the Offset Plus WIC (T22 Vs. T21) for Subgroups Defined by Duration of SSDI Receipt

Outcome	Short Duration			Long Duration			Estimated Difference in Impact (7)
	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome with Offset and WIC (T21) (2)	Impact Estimate (3)	Average Outcome with Offset and EWIC (T22) (4)	Average Outcome with Offset and WIC (T21) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2012)							
Total earnings	\$4,016	\$4,005	\$11 (\$211)	\$3,989	\$3,905	\$84 (\$299)	-\$73 (\$364)
Employment during year (%)	38.92	39.22	-0.30 (1.62)	42.19	42.67	-0.49 (2.09)	0.19 (2.78)
Earnings above BYA (%)	10.23	10.00	0.23 (0.85)	10.15	10.80	-0.65 (0.89)	0.88 (1.27)
Earnings above 2x BYA (%)	2.90	3.71	-0.81** (0.32)	2.66	2.92	-0.25 (0.58)	-0.56 (0.79)
Earnings above 3x BYA (%)	1.89	1.74	0.14 (0.30)	1.09	0.73	0.36 (0.27)	-0.22 (0.41)
Benefit Outcomes (January–December 2012)							
Total SSDI benefits paid	\$13,508	\$13,466	\$42 (\$142)	\$12,270	\$12,260	\$10 (\$191)	\$32 (\$163)
Number of months with SSDI payments	11.61	11.56	0.05 (0.04)	11.35	11.37	-0.02 (0.06)	0.07 (0.05)
Total SSI benefits paid	\$85	\$74	\$10 (\$29)	\$31	\$37	-\$6 (\$10)	\$16 (\$27)
Number of months with SSI payments	0.23	0.24	-0.00 (0.06)	0.16	0.14	0.02 (0.04)	-0.02 (0.08)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: Short Duration T22 = 1,914, Short Duration T21 = 3,125, Long Duration T22 = 1,127, Long Duration T21 = 1,729.

*/**/** 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 (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.

Exhibit C-6. Estimated Impacts on 2012 Outcomes of the Offset Plus WIC Compared to Current Law (T21 Vs. C2) for Subgroups Defined by Year of Study Enrollment

Outcome	2011 Enrollees			2012 Enrollees			Estimated Difference in Impact (7)
	Average Outcome with Offset and WIC (T21) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and WIC (T21) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2012)							
Total earnings	\$3,876	\$3,638	\$239 (\$238)	\$3,905	\$3,638	\$268* (\$144)	\$29 (\$264)
Employment during year (%)	41.64	39.12	2.52 (1.40)	40.70	40.02	0.69 (1.05)	-1.83 (1.75)
Earnings above BYA (%)	10.42	8.90	1.52 (0.93)	10.28	9.30	0.99 (0.70)	-0.53 (1.16)
Earnings above 2x BYA (%)	3.26	2.87	0.40 (0.58)	3.07	2.59	0.48 (0.40)	0.08 (0.71)
Earnings above 3x BYA (%)	1.15	1.07	0.08 (0.33)	1.06	0.98	0.07 (0.26)	-0.01 (0.42)
Benefit Outcomes (January–December 2012)							
Total SSDI benefits paid	\$12,452	\$12,273	\$179 (\$169)	\$13,028	\$12,773	\$255* (\$139)	\$76 (\$218)
Number of months with SSDI payments	11.37	11.22	0.15 (0.10)	11.53	11.47	0.06 (0.07)	-0.09 (0.11)
Total SSI benefits paid	\$51	\$44	\$7 (\$17)	\$53	\$33	\$20 (\$15)	\$13 (\$25)
Number of months with SSI payments	0.23	0.20	0.03 (0.05)	0.15	0.14	0.01 (0.04)	-0.02 (0.07)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: 2011 Enrollees T21 = 1,948, 2011 Enrollees C2 = 1,941, 2012 Enrollees T21 = 2,906, 2012 Enrollees C2 = 2,908.

*/**/** 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 (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.

Exhibit C-7. Estimated Impacts on 2012 Outcomes of the Offset Plus EWIC Compared to Current Law (T22 Vs. C2) for Subgroups Defined by Year of Study Enrollment

Outcome	2011 Enrollees			2012 Enrollees			Estimated Difference in Impact (7)
	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome under Current Law (C2) (2)	Impact Estimate (3)	Average Outcome with Offset and EWIC (T22) (4)	Average Outcome under Current Law (C2) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2012)							
Total earnings	\$3,978	\$3,638	\$340 (\$247)	\$3,748	\$3,638	\$111 (\$153)	-\$230 (\$286)
Employment during year (%)	39.92	39.12	0.80 (1.74)	41.03	40.02	1.01 (1.22)	0.21 (2.02)
Earnings above BYA (%)	10.27	8.90	1.37 (1.23)	9.56	9.30	0.27 (0.77)	-1.10 (1.38)
Earnings above 2x BYA (%)	3.02	2.87	0.15 (0.62)	2.12	2.59	-0.47 (0.39)	-0.62 (0.74)
Earnings above 3x BYA (%)	1.44	1.07	0.38 (0.46)	1.12	0.98	0.14 (0.39)	-0.24 (0.66)
Benefit Outcomes (January–December 2012)							
Total SSDI benefits paid	\$12,671	\$12,273	\$399* (\$205)	\$12,915	\$12,773	\$142 (\$146)	-\$256 (\$252)
Number of months with SSDI payments	11.39	11.22	0.17 (0.10)	11.54	11.47	0.08 (0.07)	-0.09 (0.11)
Total SSI benefits paid	\$56	\$44	\$12 (\$32)	\$52	\$33	\$18 (\$17)	\$7 (\$38)
Number of months with SSI payments	0.24	0.20	0.05 (0.08)	0.15	0.14	0.01 (0.05)	-0.04 (0.10)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: 2011 Enrollees T22 = 1,212, 2011 Enrollees C2 = 1,941, 2012 Enrollees T22 = 1,829, 2012 Enrollees C2 = 2,908.

*/**/** 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 (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.

Exhibit C-8. Estimated Impacts on 2012 Outcomes of the Offset Plus EWIC Compared to the Offset Plus WIC (T22 Vs. T21) for Subgroups Defined by Year of Study Enrollment

Outcome	2011 Enrollees			2012 Enrollees			Estimated Difference in Impact (7)
	Average Outcome with Offset and EWIC (T22) (1)	Average Outcome with Offset and WIC (T21) (2)	Impact Estimate (3)	Average Outcome with Offset and EWIC (T22) (4)	Average Outcome with Offset and WIC (T21) (5)	Impact Estimate (6)	
Earnings Outcomes (January–December 2012)							
Total earnings	\$3,978	\$3,876	\$102 (\$216)	\$3,748	\$3,905	-\$157 (\$159)	-\$259 (\$247)
Employment during year (%)	39.92	41.64	-1.72 (1.70)	41.03	40.70	0.33 (1.83)	2.04 (2.55)
Earnings above BYA (%)	10.27	10.42	-0.15 (0.99)	9.56	10.28	-0.72 (0.50)	-0.57 (1.08)
Earnings above 2x BYA (%)	3.02	3.26	-0.25 (0.53)	2.12	3.07	-0.95* (0.43)	-0.70 (0.64)
Earnings above 3x BYA (%)	1.44	1.15	0.29 (0.36)	1.12	1.06	0.07 (0.38)	-0.23 (0.69)
Benefit Outcomes (January–December 2012)							
Total SSDI benefits paid	\$12,671	\$12,452	\$220 (\$184)	\$12,915	\$13,028	-\$112 (\$138)	-\$332 ^{†††} (\$100)
Number of months with SSDI payments	11.39	11.37	0.02 (0.09)	11.54	11.53	0.02 (0.03)	-0.00 (0.10)
Total SSI benefits paid	\$56	\$51	\$5 (\$26)	\$52	\$53	-\$2 (\$19)	-\$6 (\$33)
Number of months with SSI payments	0.24	0.23	0.02 (0.06)	0.15	0.15	0.00 (0.04)	-0.02 (0.07)

Source: SSA administrative records, from the MEF, BODS, MBR, and SSR, with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: 2011 Enrollees T22 = 1,212, 2011 Enrollees T21 = 1,948, 2012 Enrollees T22 = 1,829, 2012 Enrollees T21 = 2,906.

*/**/** 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 (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.

Exhibit C-9. Estimated Impacts on 2011 Earnings and Benefits of Stage 2 Volunteers: All Policy Comparisons

Outcome	Average Outcome with Offset and WIC (T21) (1)	Average Outcome with Offset and EWIC (T22) (2)	Average Outcome under Current Law (C2) (3)	Estimated Impact of Offset + WIC vs Current Law (T21 vs. C2) (4)	Estimated Impact of Offset + EWIC vs Current Law (T22 vs. C2) (5)	Estimated Impact of EWIC instead of WIC Given Offset (T22 vs. T21) (6)
Earnings Outcomes (January–December 2011)						
Total earnings	\$3,157	\$3,247	\$2,951	\$206* (\$98)	\$296 (\$171)	\$90 (\$120)
Employment during year (%)	37.66	37.15	36.82	0.85 (1.03)	0.34 (1.15)	-0.51 (0.91)
Earnings above BYA (%)	7.44	7.63	7.32	0.11 (0.49)	0.30 (0.71)	0.19 (0.55)
Earnings above 2x BYA (%)	1.85	2.52	1.92	-0.07 (0.27)	0.60 (0.39)	0.67* (0.32)
Earnings above 3x BYA (%)	0.81	0.89	0.60	0.21 (0.17)	0.29 (0.18)	0.08 (0.20)
Benefit Outcomes (January–December 2011)						
Total SSDI benefits paid	\$13,091	\$13,128	\$12,942	\$149 (\$122)	\$187 (\$131)	\$37 (\$181)
Number of months with SSDI payments	11.28	11.31	11.28	0.00 (0.06)	0.03 (0.06)	0.03 (0.05)
Total SSI benefits paid	\$263	\$267	\$201	\$62* (\$29)	\$66 (\$42)	\$4 (\$46)
Number of months with SSI payments	0.19	0.20	0.18	0.01 (0.02)	0.02 (0.03)	0.01 (0.03)

Source: Analysis of SSA administrative records (from the MEF, BODS, MBR, and SSR), with covariates from Stage 2 baseline survey and baseline SSA administrative data used in impact analysis regression equations.

Notes: See Chapter 6 for variable definitions. Weights reflecting sample selection are used to ensure that the BOND subjects who met analysis criteria are representative of the national population of SSDI-only beneficiaries who would volunteer for study enrollment. Standard errors are in parentheses. Means and impact estimates are regression-adjusted for baseline characteristics.

Unweighted sample sizes: T21 = 4,854, T22 = 3,041, C2 = 4,849

*/**/** 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 (and with no multiple comparisons adjustment).